

तथ्याङ्क र तथ्याङ्कीय साक्षरता

विकास प्रशासनका लागि नीतिनिर्माता, योजनाकार, विकास साभेदार,
अध्येता एवं पत्रकारहरूको तथ्याङ्कीय प्रयोगात्मक हाते पुस्तिका



नेपाल सरकार

राष्ट्रिय योजना आयोग

सिंहदरबार, काठमाण्डौं

तथ्याङ्क र तथ्याङ्कीय साक्षरता

विकास प्रशासनका लागि नीतिनिर्माता, योजनाकार, विकास साभेदार,
अध्येता एवं पत्रकारहरूको तथ्याङ्कीय प्रयोगात्मक हाते पुस्तिका

तथ्याङ्क र तथ्याङ्कीय साक्षरता

विकास प्रशासनका लागि नीतिनिर्माता, योजनाकार, विकास साभेदार,
अध्येता एवं पत्रकारहरूको तथ्याङ्कीय प्रयोगात्मक हाते पुस्तिका



नेपाल सरकार
राष्ट्रिय योजना आयोग
सिंहदरबार, काठमाण्डौ

प्रकाशक: नेपाल सरकार
राष्ट्रिय योजना आयोग
सिंहदरबार, काठमाण्डौं
फोन: ०१-४२१११३२, ४२१११३३
फ्याक्स: ०१-४२११७००
इमेल: npcs@npc.gov.np
वेबसाइट: www.npc.gov.np

डिजाइन: नेपाल इन डाटा, विकास उद्यमीको पहल
बालुवाटार, काठमाण्डौं
फोन: ०१-४४४२०१४
इमेल: nepalindata@bikasudhyami.com
वेबसाइट: www.nepalindata.com

प्रकाशन वर्ष: २०७८ चैत्र

प्रथम संस्करण: २००० प्रति

ISBN 978-9937-1-1482-0

मुद्रक: विजय अफसेट
धपागल-१६, ललितपुर

शुभकामना मन्तव्य

विकासको मापन तथ्याङ्कबाट गरिन्छ । वैज्ञानिक विधि प्रयोग गरी उत्पादन भएको सही तथ्याङ्कले समाजको ऐनाको रूपमा काम गर्दछ । तसर्थ, गुणस्तरीय तथ्याङ्कले नागरिकको वास्तविक अवस्था प्रस्तुत गरी नीति, योजना र कार्यक्रम निर्माण गर्दा नागरिकका अपेक्षाहरूलाई सम्बोधन गर्नुपर्दछ । अहिले सूचना प्रविधिको विकास र प्रयोगसँगै तथ्याङ्कको महत्त्व अत्यन्त बढेको छ । विकसित राष्ट्रमा प्रत्येक निर्णयहरू तथ्याङ्कमा आधारित भएर गर्ने प्रचलन सामान्य भइसकेको परिप्रेक्ष्यमा नेपालमा पनि तथ्याङ्कको प्रयोग र माग दिनानुदिन बढ्दै गइरहेको छ ।

तथ्याङ्क कुनै सीमित प्राविधिक विषयवस्तु मात्र नभई यो एउटा सार्वजनिक महत्त्वको विषय पनि हो । त्यसैकारणले गर्दा हाल विश्वभर तथ्याङ्कलाई सार्वजनिक वस्तु वा सम्पत्तिको रूपमा हेरिन्छ । तथ्याङ्कको महत्त्वलाई उजागर गर्ने र यसको सही प्रयोगको बारेमा जानकारी दिने उपयुक्त माध्यम भनेको तथ्याङ्क साक्षरता हो । सबै नागरिकलाई तथ्याङ्कको बारेमा साक्षर बनाउनु राज्यको एउटा महत्त्वपूर्ण जिम्मेवारी पनि हो ।

प्रस्तुत यथार्थतालाई मनन गरी राष्ट्रिय योजना आयोगले सबै नागरिकमा तथ्याङ्कसम्बन्धी साक्षरताको अभिवृद्धि गर्ने अभिप्रायले यो “तथ्याङ्क र तथ्याङ्कीय साक्षरता” पुस्तक प्रकाशन गरेको छ । तथ्याङ्क सबैको चासो र महत्त्वको विषयका साथै प्राविधिक विषय पनि हो । यस्तो प्राविधिक विषय सामान्य नागरिकले बुझ्ने भाषामा प्रस्तुत गर्ने कार्य चुनौतीपूर्ण नै हुन्छ, तथापि यस पुस्तकले तथ्याङ्कसम्बन्धी विषयहरूलाई सकेसम्म सरल रूपमा प्रस्तुत गर्ने जमर्को गरेको छ । राष्ट्रिय योजना आयोगको यो पहिलो प्रयास हो । तथ्याङ्क साक्षरता एउटा राष्ट्रिय अभियानसमेत भएको हुनाले आउँदा दिनहरूमा यस पुस्तकलाई अभै परिष्कृत तथा परिमार्जन गर्दै लगिनुका साथै सबै नागरिकसमक्ष तथ्याङ्क सम्बन्धी सचेतना विस्तार गर्ने कार्यलाई निरन्तरता दिइनेछ ।

यो “तथ्याङ्क र तथ्याङ्कीय साक्षरता” नामक पुस्तकको मस्यौदा तयारीका कार्यमा केन्द्रीय तथ्याङ्क विभागको महत्त्वपूर्ण भूमिका रहेको छ । अतः उक्त विभागका महानिर्देशक नेबिन लाल श्रेष्ठ, उपमहानिर्देशक डा. हेमराज रेग्मी साथै यो पुस्तकको मस्यौदा तयार पारी राष्ट्रिय योजना आयोगलाई हस्तान्तरण गरी सहयोग गर्नुहुने कार्यदलका सदस्यद्वय निर्देशक ऋषि राम सिग्देल र निर्देशक राजन सिलवालप्रति हार्दिक धन्यवाद प्रकट गर्दछु । त्यसरी नै एसिया फाउन्डेसन र विकास उद्यमीको नेपाल इन डाटा टिमको सहयोगप्रति पनि आभार व्यक्त गर्दछु ।

चैत्र, २०७८
काठमाण्डौं



केवल प्रसाद भण्डारी
सदस्य सचिव
राष्ट्रिय योजना आयोग

Foreword

The importance of using data in decision-making cannot be over-stated, especially when it comes to navigating Nepal's roadmap to development. Using data in planning and policy-making is a tried and tested method of ensuring efficient and effective delivery.

According to Nepal's 2015 Constitution, data-related functions are managed by provincial and local governments. This puts you on the driving seat in taking crucial decisions to bring meaningful change to the lives of Nepali people. Understanding and learning how to mobilize the power of good data will not only help you do this better, but also help you achieve multiple objectives. For example, with the right data, used in the right way, you can contribute to Nepal's development and do so in a sustainable, climate-resilient way.

The UK's Evidence for Development Programme has helped facilitate the use of data and evidence in crucial development work, including strengthening sub-national governance. This handbook is a part of that ambition – to ensure government officials have access to a useful guide that can train them and build their capacity to work with data.

This handbook will help you understand and work with data, statistics, and analysis irrespective of your position, or your level of understanding of these topics. The last chapter is on digital literacy, which explains the use of computerized systems for data and statistics management.

Under the leadership of the National Planning Commission (NPC) and the Central Bureau of Statistics (CBS), this handbook has the potential to bridge the gap in knowledge and improve coordination between, national and sub-national governments on data use.

The launch of this handbook is an important milestone on the path to modernising Nepal's data ecosystem and encouraging wider evidence-based decision making.

On behalf of the British Embassy, I would like to thank the NPC for leading the development of this handbook, the CBS for its technical inputs, The Asia Foundation for facilitating the process and Bikas Udhyami for the coordination work.

I wish you all the very best in making good use of this handbook and look forward to hearing about the changes that it brings to evidence-based decision making for sub-national governance.

Nicola Pollitt
British Ambassador to Nepal

Message

Statistical literacy is important in today's world fuelled by data; it has become mandatory for tomorrow's future. The need to interpret, cross verify, and communicate data is rapidly increasing. Marketing, social media, advertising, private and public sector decisions and community mobilizing are all increasingly relying on vast amounts of data. Data literacy will reach far into the future for Nepal's youth as they prepare for their careers. The country needs a more data literate population along side the state owned data and statistical system. The Constitution of Nepal has given rights to all three spheres to government to collect and manage their own data. Based on this, governments at all three levels use data generated from introduction of various information systems which is a foundation for an ever-more robust and diverse evidence-based against which decisions are made. We hear this first and foremost from those on the front-lines of government service delivery, eager for this diversified evidence base and the human resources to analyze and interpret. This holds potential for deep impact on all aspects of our lives including health, education, economy, disaster management, public service, and public participation in governing systems.

The scope of statistical literacy has grown beyond statisticians conducting surveys, gathering data, sharing data, analysing data, and using findings. These are the foundations of strong statistical systems. Even stronger data ecosystems include statistical literacy to apply in planning and decision making for polished public administration. We regularly hear in the Asia Foundation's work that the more public service providers exercising skills to make sense of data, facilitate demand, supply, and use of data for planning and decision making, the state will better thrive at its mandates.

Driven by these observations and requests to strengthen local data and information ecosystems, the Asia Foundation has remained committed the past five years to partner with public service representatives in their roles and functions to best integrate data and statistics into their functions assigned by the Constitution. This handbook was co-designed by the users themselves directly informed by the statistical expertise of the National Planning Commission and Central Bureau of Statistics. And that's what makes this handbook unique – the core approach is meeting the reader/student from an applied approach with relatable, real life examples from Nepal's public service, using existing data from the government. It seeks to be accessible to the wide range of individuals within public service with the goal of direct applicability.

I would like to thank National Planning Commission and Central Bureau of Statistics for providing us the vote of confidence to be a part of the development of this handbook. I am grateful to team members of The Asia Foundation, Bikas Udhyami, and all who are involved in the development of

the book. Collectively, our vision is that this handbook will lead to more data users among the public service sector. Lastly, I also want to extend my best wishes to the readers of this handbook for your contribution towards Nepal's unrelenting dedication to progress.

A handwritten signature in black ink, appearing to read 'Meghan T. Nalbo'. The signature is fluid and cursive, with the first name 'Meghan' being more prominent than the last name 'Nalbo'.

Meghan T. Nalbo
Country Representative
The Asia Foundation

दुई शब्द

सूचना प्रविधिको आधुनिक युगमा मानिसहरूको दैनिक जीवन र व्यवसायमा तथ्याङ्क बढ्दो रूपमा प्रयोग हुँदै आइरहेको छ । प्रमाणमा आधारित नीति, योजना तथा कार्यक्रमहरू तय गर्न तथा सोको मूल्याङ्कन गर्नका लागि निर्णयकर्ताको भरपर्दो साधन तथ्याङ्क हुन्छ । तथ्याङ्कले मानव आवश्यकताहरू परिपूर्ति गर्न आवश्यक सूचनाहरू उपलब्ध गराउने भएकाले यसलाई ज्ञान र शक्तिको स्रोतको रूपमा पनि लिने गरिन्छ । यथेष्ट तथ्याङ्कीय ज्ञानविना यसको प्रयोग गर्दा गलत व्याख्या वा सूचनाहरूको दुरुपयोग हुने सम्भावना हुन्छ । तथ्याङ्कप्रतिका यस किसिमका गलत व्याख्या वा सूचनाहरूको दुरुपयोग हटाउनका लागि पनि तथ्याङ्कीय साक्षरताको विस्तार गर्नु अपरिहार्य हुन्छ ।

नेपालको संविधानको अनुसूचीमा उल्लेखित संघ, प्रदेश र स्थानीय तहका अधिकार सूचीको कार्य विस्तृतीकरण प्रतिवेदनअनुसार केन्द्रीय तथ्याङ्कको राष्ट्रिय एवं अन्तर्राष्ट्रिय मानक र गुणस्तर निर्धारण गर्ने जिम्मेवारी संघीय सरकारअन्तर्गत राखिएको छ । तथ्याङ्क ऐन, २०१५ ले तथ्याङ्कीय क्रियाकलापको अग्रणी भूमिका निर्वाह गर्ने अधिकार केन्द्रीय तथ्याङ्क विभागलाई सुम्पेको छ । त्यसैगरी तथ्याङ्क प्रणाली विकासको लागि राष्ट्रिय रणनीति, २०७५ मा तथ्याङ्कको गुणस्तर सुनिश्चित गर्ने विधि अवलम्बन गर्ने रणनीति लिइएको छ । संयुक्त राष्ट्रसंघीय आधिकारिक तथ्याङ्कका आधारभूत सिद्धान्तहरूले पनि तथ्याङ्कको त्रुटिपूर्ण व्याख्या एवं दुरुपयोग रोक्नका लागि आधिकारिक तथ्याङ्कीय निकायहरू जिम्मेवार हुनुपर्दछ भनी उल्लेख गरेको छ । यिनै नीतिगत पृष्ठभूमिका आधारमा तथ्याङ्कसम्बन्धी आधारभूत विषयहरू समेटेर तथ्याङ्कको महत्त्व, प्रयोग र सचेतना अभिवृद्धि गर्नु तथ्याङ्क प्रणालीका आधिकारिक निकायहरूको महत्त्वपूर्ण जिम्मेवारी हुन्छ ।

केन्द्रीय तथ्याङ्क विभागले विभिन्न गणना र सर्वेक्षणहरू मार्फत् सामाजिक, आर्थिक, जनसांख्यिक, लैङ्गिक तथा वातावरण तथ्याङ्क उपलब्ध गराउँदै आएको छ । यसले नियमित रूपमा राष्ट्रिय लेखा अनुमानहरू पनि प्रकाशन गर्दै आएको छ । तथ्याङ्क र तथ्याङ्कीय साक्षरता नामक यस पुस्तकको अध्ययनले नेपालको तथ्याङ्कीय प्रणाली, विभिन्न निकायहरूबाट उत्पादन हुने तथ्याङ्क र अन्य विषयगत सैद्धान्तिक तथ्याङ्कलाई सरल र सहज रूपमा बुझ्न र त्यसको आधारभूत ज्ञान हासिल गर्न मद्दत पुग्नेछ भन्ने विश्वास लिएको छु । तसर्थ, नीतिनिर्माता, योजनाकार, तथ्याङ्ककर्मी, विकास साभेदार, पत्रकार तथा विद्यार्थी सबैले यसको अधिकतम प्रयोग गर्न पनि म हार्दिक अनुरोध गर्दछु ।

राष्ट्रिय योजना आयोग अन्तर्गतको केन्द्रीय तथ्याङ्क विभाग संघीय सरकारको प्रमुख तथ्याङ्कीय निकायको रूपमा रहेको छ । राष्ट्रिय योजना आयोगको अग्रसरतामा यो पुस्तक प्रकाशित भएकोमा मलाई धेरै खुसी लागेको छ । यो पुस्तक प्रकाशन गर्न योगदान दिनुहुने सबैमा कृतज्ञतासहित हार्दिक आभार व्यक्त गर्दछु ।

चैत्र, २०७८
काठमाण्डौं


नेबिन लाल श्रेष्ठ
महानिर्देशक
केन्द्रीय तथ्याङ्क विभाग

कृतज्ञता

राष्ट्रिय योजना आयोगको मार्गनिर्देशन तथा केन्द्रीय तथ्याङ्क विभागको नेतृत्वमा दि एशिया फाउण्डेसन र विकास उद्यमीको सहकार्यमा यस हाते पुस्तिका तथ्याङ्क तथा तथ्याङ्कीय साक्षरता तयार भएको छ । यस हाते पुस्तिकाको संयोजन र तयारीको लागि दि एशिया फाउण्डेसनद्वारा सञ्चालित विकासको लागि तथ्याङ्क कार्यक्रम (Data for Development Programme) मार्फत् बेलायत सरकार (UK Aid) को आर्थिक सहयोग प्राप्त भएको हो ।

पुस्तिकाको सामग्री तयार पार्न तथा यसलाई मूर्तरूप दिन सहयोग गर्नुहुने केन्द्रीय तथ्याङ्क विभागका महानिर्देशक श्री नेबिन लाल श्रेष्ठ, उपमहानिर्देशक डा. हेमराज रेग्मी, निर्देशक श्री ऋषि राम सिन्देल, र निर्देशक श्री राजन सिलवालप्रति विशेष अनुगृहीत छौं । आफ्नो अमूल्य समय दिएर यस पुस्तिकाको हरेक भागलाई अन्तिम रूप दिन विशेष सहयोग पुऱ्याउनु भएकै कारण पुस्तिकाले यो परिष्कृत रूप लिन सफल भएको छ ।

पुस्तिकाको विस्तृत अध्ययन गरेर यसलाई परिष्कृत बनाउन प्रतिक्रिया तथा सुझाव दिनुहुने दि एशिया फाउण्डेसनका कार्यक्रम निर्देशक श्री सजना महर्जन अमात्य, कार्यक्रम प्रबन्धक श्री प्रणय रत्न स्थापित, कार्यक्रम अधिकृत श्री क्षितिज खनाल र श्री अर्नभ उपाध्यायप्रति कृतज्ञ छौं । समेटिएका सुझावले पुस्तिकामा भएका विषयवस्तुको दायरा फराकिलो र सन्दर्भबद्ध हुन मद्दत पुगेको छ ।

पुस्तिकाको ढाँचा तयार गर्न तथा भाषिक कोणलाई सुक्ष्म ढंगले अध्ययन गरी सम्पादन गर्नुहुने विकास उद्यमीको प्रमुख कार्यकारी अधिकृत श्री सन्तोष गर्तौला तथा उक्त संस्थाका श्री सजनी लामा, श्री अनुज तामाङ, श्री इन्द्र लाल शाक्य, श्री पेमा तामाङ, र श्री ज्ञानेन्द्र न्यौपानेप्रति हार्दिक आभार व्यक्त गर्दछौं । नेपाली भाषाको मर्म, सरलता, र पठनीयतालाई ध्यान दिई अन्तिम रूप दिइएको यस पुस्तिका आगामी दिनहरूमा तथ्याङ्कीय ज्ञान हासिल गर्न चाहने व्यक्ति र जिज्ञासुहरूलाई निकै सहयोगी हुनेछ भन्ने विश्वास लिएका छौं ।

‘विकासको लागि तथ्याङ्क’ कार्यक्रमका परियोजना सल्लाहकार समितिका सबै पदाधिकारी तथा सदस्यहरूलाई यसको कार्यान्वयनमा पुऱ्याउनु भएको रचनात्मक र मूल्यवान सहयोगको लागि हार्दिक आभार व्यक्त गर्दछौं ।

परिचय

तथ्याङ्क र तथ्याङ्कीय साक्षरता पुस्तिका विकास प्रशासनका लागि नीतिनिर्माता, योजनाकार, विकास साभेदार, अध्येता एवं पत्रकारहरूको तथ्याङ्कीय प्रयोगात्मक हाते पुस्तिका हो । समकालीन समयमा सूचना तथा प्रविधिले निकै ठुलो फड्को मारेको र अधिकांश सूचनाहरू, तथ्याङ्क व्यवस्थापनका उपज हुने हुनाले यससँग सम्बन्धित सरोकारवालाहरूलाई तथ्याङ्कसम्बन्धी आधारभूत जानकारी गराउनु यस पुस्तिकाको प्रमुख उद्देश्य रहेको छ । तथ्याङ्कका अधिकांश पुस्तकहरू अंग्रेजी संस्करणमा पाइने र उपलब्ध पुस्तकहरूमा तथ्याङ्कका सैद्धान्तिक पक्षलाई बढी महत्त्व दिइने हुनाले तथ्याङ्कका आम सरोकारवालाहरूलाई सैद्धान्तिक तथा व्यावहारिक पक्षलाई जोड्ने सेतुको रूपमा नेपाली संस्करणमा यो पुस्तिका तयार गरिएको छ । यस पुस्तिकामा तथ्याङ्कका आधारभूत सैद्धान्तिक पक्षलाई समेटेर यसका परिचयात्मक अवधारणासहित आधिकारिक तथ्याङ्कको प्रयोग भल्काउने गरी विषयवस्तुहरू प्रस्तुत गरिएका छन् । यसबाट पाठकवर्गमा तथ्याङ्कलाई व्यावहारिक रूपमा कसरी प्रयोग गर्न सकिन्छ भन्ने विश्वासमा वृद्धि हुनेछ भन्ने विश्वास लिइएको छ ।

विकास प्रशासनसँग सम्बन्धित योजना तर्जुमा, नीति निर्माण तथा कार्यक्रमहरूको तयारीका कार्यहरू तथ्यमा आधारित हुनुपर्दछ भन्ने सर्वमान्य मान्यतालाई आत्मसात गर्दै त्यस्ता विकासका प्रतिफल मापन गर्ने कतिपय विकास सूचकहरू तथ्याङ्कको सहायताबाट निर्माण गरिन्छन् । यसका अतिरिक्त विकास कार्यक्रमको अनुगमन तथा मूल्याङ्कन गर्ने भरपर्दो औजार पनि तथ्याङ्क नै हो । प्रमाणमा आधारित निर्णय प्रक्रियाको आधार स्तम्भ पनि तथ्याङ्क नै हो । यसले समाजको यथार्थतालाई पारदर्शी तरिकाले सम्बन्धित सरोकारवालाहरू समक्ष प्रस्तुत गर्दछ । फलस्वरूप सुशासन कायम गर्नदेखि कल्याणकारी कार्यक्रम सञ्चालन गर्न तथ्याङ्क नभई हुँदैन । तथ्याङ्कको महत्त्वको साथसाथै यसका उत्पादक, व्यवस्थापक तथा प्रयोगकर्तामा यससम्बन्धी जानकारी हुनु जरूरी हुन्छ । नेपालको संविधानले निर्दिष्ट गरेको जनताको सुसूचित हुने अधिकारको सम्मान गर्न तथा गोपनीयता समेतको रक्षा गर्न तथ्याङ्कसँग सम्बन्धित सरोकारवालाहरूले विशेष ध्यान पुऱ्याउनु पर्ने हुन्छ । यसका लागि तथ्याङ्कीय साक्षरता अपरिहार्य हुन्छ । सवैधानिक व्यवस्था बमोजिम सञ्चालित राज्यका तिनवटै तहका सरकारमा आधिकारिक तथ्याङ्क व्यवस्थापन र प्रयोग गर्नका लागि सहज होस् भन्ने ध्येयका साथ यो पुस्तिका तयार गरिएको छ । तथ्याङ्क साक्षर भएका वा नभएका जुनसुकै तह र क्षेत्रमा रहेका व्यक्तित्वहरूका लागि सरल र सहज रूपले आधारभूत तथ्याङ्कीय ज्ञान यस पुस्तिकाले प्रदान गर्नेछ । तसर्थ, विद्यार्थी, पत्रकार, नीतिनिर्माता, योजनाकार तथा विकास साभेदारहरू यस पुस्तिकाका प्रत्यक्ष प्रयोगकर्ता हुनेछन् ।

यस पुस्तिकालाई मुख्य रूपमा पाँच भाग अन्तर्गत २३ वटा पाठमा विभाजन गरिएको छ । पहिलो भागमा नेपालको तथ्याङ्कीय प्रणाली र तथ्याङ्कीय साक्षरतासम्बन्धी विषयहरूको चर्चा गरिएको छ । यस भागअन्तर्गत जम्मा ३ वटा पाठहरू रहेका छन् । दोस्रो भागमा तथ्याङ्कसम्बन्धी आधारभूत जानकारी समेटेर ७ वटा पाठहरू समावेश गरिएका छन् । तेस्रो भागमा वर्णनात्मक र अनुमानात्मक तथ्याङ्कअन्तर्गत ८ वटा पाठहरू प्रस्तुत गरिएका छन् । यसैगरी चौथो भागमा

तथ्याङ्कीय क्रियाकलापमा थप विषयवस्तु समेटेर ३ वटा पाठ राखिएका छन् । अन्तिम भागमा डिजिटल साक्षरताअन्तर्गत तथ्याङ्क व्यवस्थापनमा कम्प्युटरको प्रयोगसम्बन्धी जानकारी गराउने उद्देश्यले २ वटा पाठ समावेश गरिएका छन् । यी पाठहरूका बारेमा संक्षिप्त जानकारी तलको अनुच्छेदहरूमा दिइएको छ ।

यस पुस्तिकाको पहिलो पाठमा नेपालको तथ्याङ्क प्रणाली कस्तो छ भन्ने बारेमा चर्चा गरिएको छ । तथ्याङ्कीय गतिविधिमा नेतृत्वदायी भूमिका निर्वाह गर्ने निकाय तथा आधिकारिक तथ्याङ्क उत्पादनमा संलग्न निकायहरूको परिचय तथा समग्र तथ्याङ्क प्रणालीका अङ्गहरूको जानकारी प्रदान गर्नु यस पाठको उद्देश्य रहेको छ । यसका अतिरिक्त यस पाठमा नेपालको तथ्याङ्क प्रणालीका सबल पक्ष, कमजोर पक्ष, अवसर तथा चुनौतीहरूका बारेमा पनि लेखाजोखा गरिएको छ ।

दोस्रो पाठ तथ्याङ्कसम्बन्धी नीति तथा संस्थागत व्यवस्थामा केन्द्रित रहेको छ । जसअनुसार तथ्याङ्कसम्बन्धी कानुनी प्रावधानहरू, आवधिक योजनाहरूमा तथ्याङ्कको परिदृश्य, तथ्याङ्कसम्बन्धी अन्तर्राष्ट्रिय पहलजस्ता विषयहरू समावेश गरिएका छन् । यसका अतिरिक्त यस पाठमा तथ्याङ्क व्यवस्थापनका लागि केन्द्रदेखि स्थानीय तहसम्मको संस्थागत व्यवस्थाको बारेमा पनि उल्लेख गरिएको छ ।

पुस्तिकाको पाठ ३ अन्तर्गत तथ्याङ्कीय साक्षरतासम्बन्धी अवधारणा प्रस्तुत गरिएको छ । खासगरी तथ्याङ्कीय साक्षरता भनेको के हो, यसको के महत्त्व छ, विकास प्रशासनमा तथ्याङ्कीय साक्षरताले कस्तो प्रभाव पारेको छ भन्ने जस्ता विषयहरूका बारेमा यस पाठमा चर्चा गरिएको छ ।

चौथो पाठमा तथ्याङ्कको आधारभूत जानकारी प्रदान गरिएको छ । जसअन्तर्गत तथ्याङ्कको परिभाषा, महत्त्व, दायरा, सीमा तथा यसप्रतिको अविश्वासजस्ता विषयहरू पर्दछन् । यस पाठको अध्ययनबाट तथ्याङ्कलाई सरल तरिकाले बुझ्न सकिन्छ ।

पाँचौँ पाठमा मापन एकाइका बारेमा चर्चा गरिएको छ । मापनको अर्थ, महत्त्व, प्रयोगजस्ता कुराहरू यस पाठमा समावेश गरिएका छन् । यस पाठबाट तथ्यहरूको मापन गर्दा प्रयोग गरिने एकाइ किन महत्त्वपूर्ण छ भन्ने कुराको जानकारी प्राप्त हुन्छ ।

त्यसैगरी छैटौँ पाठमा तथ्याङ्क प्राप्तिका स्रोत र तथ्याङ्कको व्यवस्थापनसँग सम्बन्धित विषयहरू समावेश गरिएका छन् । मुख्यतया तथ्याङ्कका प्राथमिक तथा माध्यमिक स्रोतका बारेमा चर्चा गरेर गणना, नमुना सर्वेक्षण, प्रशासनिक अभिलेख तथा बिग डाटाको बारेमा चर्चा गरिएको छ । तथ्याङ्कको स्रोतका बारेमा चर्चा गरिसकेपछि तथ्याङ्क संकलन, प्रशोधनसम्बन्धी विषयमा प्रकाश पारिएको छ । यसै पाठमा चरसम्बन्धी अवधारणा, यसका प्रकारजस्ता विषय समावेश गरेर तथ्याङ्क प्रस्तुतिको एउटा विधि तालिकीकरणका बारेमा व्याख्या गरिएको छ । साथै, आउटलायर्सको परिचय पनि यसै पाठमा दिइएको छ ।

पाठ ७ मा तथ्याङ्क र सूचकसम्बन्धी अवधारणा प्रस्तुत गरिएको छ । तथ्याङ्कलाई सूचनाको किन कच्चा पदार्थ भनिएको हो, सूचकको महत्त्व के छ, तथ्याङ्क विश्लेषण र सूचकको सम्बन्ध के हो भन्ने जस्ता विषयहरू यस पाठमा समावेश गरिएका छन् ।

पाठ ८ मा तथ्याङ्क प्रस्तुतिकरणका विविध विधिहरूका बारेमा चर्चा गरिएको छ । खासगरी यस पाठमा तथ्याङ्क प्रस्तुतिका ग्राफिकल र चित्रात्मक विधिको व्याख्या एवं विश्लेषण गरिएको छ । तथ्याङ्क प्रस्तुतिको महत्त्व, प्रकारका बारेमा चर्चा गर्दै यसको प्रयोग र व्याख्या गर्ने तरिकासमेत उल्लेख गरिएको छ ।

पाठ ९ मा तथ्याङ्क विश्लेषणमा अत्याधिक प्रयोग हुने केही महत्त्वपूर्ण गणितीय क्रियाकलाप जस्तै: दर, अनुपात, प्रतिशत, समानुपातका बारेमा व्याख्या गरिएको छ । यी विषयहरूको अर्थ, निकाल्ने तरिका र व्याख्या गर्दा ध्यान दिनुपर्ने विषयहरूका बारेमा यस पाठमा उल्लेख गरिएको छ ।

पाठ १० मा सूचकाङ्क र यसको प्रयोगसम्बन्धी विषयहरू समावेश छन् । सूचकाङ्कका प्रकार र नेपालमा निर्माण गरिने सूचकाङ्कहरूका बारेमा विस्तृत रूपमा जानकारी दिइएको छ । सूचकाङ्कको व्याख्या र प्रयोग कसरी गर्ने भन्ने सन्दर्भमा पनि यस पाठमा चर्चा गरिएको छ । यसका अतिरिक्त सूचकाङ्कको आधार वर्ष परिवर्तन गर्ने, सूचकाङ्कको विभाजन गर्ने तथा डिफ्लेटिड गर्ने विधिका बारेमा जानकारी दिइएको छ । साथै, जीविकोपार्जन सूचकाङ्कको निर्माण र प्रयोगका बारेमा पनि यसै पाठमा उल्लेख गरिएको छ ।

पाठ ११ मा तथ्याङ्कीय कार्यकलाप अन्तर्गतको महत्त्वपूर्ण विषय, केन्द्रीय प्रवृत्ति मापनका बारेमा जानकारी दिइएको छ । जसअन्तर्गत मध्यक, मध्यिका, बहुलक तथा विभाजकहरूका बारेमा जानकारी दिनुका साथै यस्ता सूचकहरू निर्माण गर्ने विधिका बारेमा पनि चर्चा गरिएको छ । केन्द्रीय प्रवृत्तिको गणना, व्याख्या र प्रयोग यसका गुण र अवगुण, यिनीहरूको छनोट र सीमाका बारेमा यस पाठमा विस्तृत चर्चा गरिएको छ ।

पाठ १२ मा तथ्याङ्क विश्लेषणको अर्को महत्त्वपूर्ण विषय, विचलनको मापनसम्बन्धी विषय समावेश छ । यस पाठमा विचलनका गुण, निरपेक्ष र सापेक्ष मापन, व्याख्या, प्रयोग जस्ता विषयहरू समावेश छन् । विचलनको मापन अन्तर्गत विस्तार, मध्य विचलन, स्तरीय भिन्नता, भेरियन्स, जेड स्कोर, लरेन्ज बक्र तथा स्तरीय भिन्नताको मानाङ्क जस्ता विषयहरू समावेश छन् ।

त्यसैगरी पाठ १३ मा सहसम्बन्ध र यसको प्रयोगका बारेमा चर्चा गरिएको छ । यस पाठमा सहसम्बन्धका प्रकार र गुणहरू, सहसम्बन्धको अध्ययन, गणना, व्याख्या र प्रयोगका बारेमा चर्चा गरिएको छ ।

पाठ १४ मा समय श्रृङ्खला विश्लेषणका आधारभूत विधिहरू प्रस्तुत गरिएको छ । यस पाठमा विभिन्न तथ्याङ्क (समय श्रृङ्खला, क्रस-सेक्शन, प्यानल डाटा) का परिचय दिनुका साथै वृद्धिदरका बारेमा पनि उल्लेख गरिएको छ । इन्टरपोलेसन, एक्सट्रापोलेसन, पूर्वानुमान र प्रक्षेपणजस्ता प्राविधिक विषय पनि यही पाठमा समावेश गरिएका छन् ।

पाठ १५ मा अविच्छिन्न चरले पढ्नुअघि सैद्धान्तिक डिस्ट्रिब्युसन जस्तै: नर्मल, स्टाण्डर्ड नर्मल डिस्ट्रिब्युसनका बारेमा चर्चा गरिएको छ । अनुमानात्मक तथ्याङ्कीय कार्य तथा निर्णय प्रक्रियामा नर्मल डिस्ट्रिब्युसनको प्रयोग हुने भएकोले यस पाठमा यो विषयका बारेमा चर्चा गरिएको हो ।

पाठ १६ मा अनुमानका सिद्धान्त, यसका गुणका साथै स्याम्प्लड डिस्ट्रिब्युसन अफ स्टाटिस्टिक्स, सेन्ट्रल लिमिट थेोरम र यिनका प्रयोगसम्बन्धी विषयमा व्याख्या गरिएको छ । अनुमान अन्तर्गत बिन्दु र अन्तराल अनुमानको परिचयात्मक टिप्पणी गरिएको छ । त्यसैगरी विश्वसनीय अन्तराल निर्माण र नमुना छनोटको आकार निर्धारण गर्ने विधिका बारेमा यस पाठमा उल्लेख गरिएको छ ।

यसैगरी पाठ १७ मा परिकल्पना परीक्षणका बारेमा व्याख्या गरिएको छ । यससँग सम्बन्धी अवधारणा र परिभाषा उल्लेख गरी परिकल्पना परीक्षणका चरणहरूका बारेमा विस्तृत व्याख्या गरिएको छ ।

पाठ १८ मा प्रतिगमन विश्लेषणसम्बन्धी विषयहरू समावेश गरिएका छन् । यही पाठमा स्वतन्त्र र निर्भर चरका बारेमा पनि प्रकाश पारिएको छ ।

पाठ १९ मा तथ्याङ्कका गुणहरूका बारेमा चर्चा गरिएको छ । त्यसैगरी यस पाठमा तथ्याङ्कको उपलब्धता र पहुँचका बारेमा पनि चर्चा गरिएको छ । पाठ २० मा खोजमूलक परियोजना सञ्चालन गर्दा अनुसरण गर्नुपर्ने प्रमुख चरणहरूका बारेमा व्याख्या गरिएको छ । यसका अतिरिक्त यस पाठमा तथ्याङ्कमा गोपनीयता र नैतिकताको बारेमा पनि चर्चा गर्दै तथ्याङ्कीय कार्यकलापमा कानुनी प्रावधानको पक्षलाई पनि प्रस्तुत गरिएको छ । पाठ २१ मा विभिन्न सर्वेक्षणका तथ्याङ्कीय प्रतिवेदनहरू अध्ययन गर्ने तरिकाका बारेमा चर्चा गरिएको छ । यो पाठ खासगरी परियोजना कार्यका लागि तयार गरिएको हो । तसर्थ, पाठकवर्गलाई यो पाठको अधीनमा रहेर विभिन्न सर्वेक्षण प्रतिवेदनहरू अध्ययन गरी संक्षिप्त प्रतिवेदन तयार गर्ने प्रेरित गरिएको छ । यस पाठमा सर्वेक्षण प्रतिवेदन अध्ययन गर्दा ध्यान दिनुपर्ने कुराहरू जस्तै: सर्वेक्षणको उद्देश्य, सर्वेक्षणको दायरा र कभरेज, लक्षित जनसंख्या र नमुना चयन, प्रश्नावली तयार गर्न प्रयोग गरिएको दृष्टिकोण, सर्वेक्षण कार्यान्वयन प्रक्रिया, अनुत्तरित प्रश्न र यसको व्यवस्थापन, परिणाम र तिनीहरूको प्रयोगजस्ता विषयहरूमा प्रकाश पारिएको छ ।

पाठ २२ र २३ मूलतः डिजिटल साक्षरतासँग सम्बन्धित विषयमा केन्द्रित छन् । जसअनुसार पाठ २२ मा एक्सेलसम्बन्धी र पाठ २३ मा स्टेटासम्बन्धी आधारभूत विषय समावेश गरिएको छ । यी दुई पाठबाट पाठकवर्गले तथ्याङ्क व्यवस्थापनमा कम्प्युटरको महत्त्व बुझ्नका साथै यसको प्रयोग गर्न सक्षम हुने विश्वास लिइएको छ ।

यस पुस्तकमा २३ वटा विभिन्न पाठका अतिरिक्त अनुसूचीमार्फत् अभ्यासका लागि राखिएका वस्तुगत प्रश्नहरूका समाधान, तथ्याङ्कीय क्रियाकलापमा आवश्यक पर्ने सांख्यिकीय तालिका, अवधारणा र परिभाषाजस्ता विषयहरू समावेश गरिएका छन् । समग्र पुस्तिकालाई प्रयोगात्मक बनाउने ध्येयका साथ संलग्न उदाहरणहरूमा सकभर आधिकारिक तथ्याङ्कको प्रयोग गरिएको छ । तसर्थ, प्राज्ञिक क्षेत्रका अतिरिक्त अन्य व्यावहारिक क्षेत्रमा समेत यस पुस्तिकाको प्रयोग उपयुक्त हुनेछ । यस पुस्तिकाले समाजका विभिन्न क्षेत्रमा तथ्याङ्कीय साक्षरता सुदृढ गराउनमा थप टेवा पुऱ्याउने विश्वास लिइएको छ ।

विषयसूची

भाग १ – नेपालको तथ्याङ्क प्रणाली र तथ्याङ्कीय साक्षरता

पाठ १	नेपालको तथ्याङ्क प्रणाली	०३
	१.१ परिचय	
	१.२ राष्ट्रिय तथ्याङ्क कार्यालय र यसको भूमिका	
	१.३ तथ्याङ्क उत्पादन र प्रयोगकर्ताका रूपमा रहेका प्रमुख संस्थाहरू	
	१.४ नेपालको तथ्याङ्क प्रणाली र यसका अङ्गहरू	
	१.५ नेपालको तथ्याङ्क प्रणालीका सबल पक्ष, कमजोर पक्ष, अवसर तथा चुनौतीहरू	
	१.६ अभ्यासका लागि प्रश्नहरू	
	१.७ सन्दर्भ सामग्री	
पाठ २	तथ्याङ्कसम्बन्धी नीति तथा संस्थागत व्यवस्था	१३
	२.१ परिचय	
	२.२ तथ्याङ्कीय नीतिहरू	
	२.२.१ नेपालको संविधान	
	२.२.२ तथ्याङ्क ऐन, २०१५	
	२.३ आवधिक योजनाहरूमा तथ्याङ्क	
	२.४ तथ्याङ्कसम्बन्धी अन्तर्राष्ट्रिय पहल	
	२.४.१ संयुक्त राष्ट्र संघको आधिकारिक तथ्याङ्कसम्बन्धी आधारभूत सिद्धान्तहरू	
	२.४.२ तथ्याङ्कमा प्राविधिक सहयोग सम्बन्धी राम्रा अभ्यासहरूको घोषणा	
	२.४.३ अन्तर्राष्ट्रिय तथ्याङ्क सञ्चालन गर्ने सिद्धान्तहरू	
	२.४.४ PARIS 21 संघको स्थापना	
	२.४.५ तथ्याङ्कको लागि माराकेच कार्ययोजना	
	२.५ तथ्याङ्क व्यवस्थापनका लागि केन्द्रदेखि स्थानीय तहसम्मको संस्थागत व्यवस्था	
	२.६ अभ्यासका लागि प्रश्नहरू	
	२.७ सन्दर्भ सामग्री	
पाठ ३	तथ्याङ्कीय साक्षरता तथा डिजिटल साक्षरतासम्बन्धी अवधारणा	३१
	३.१ परिचय	
	३.२ तथ्याङ्कीय साक्षरता (Statistical Literacy)	
	३.३ तथ्याङ्कीय साक्षरताका मापदण्ड	
	३.३.१ तथ्याङ्क बोध (Data Awareness)	
	३.३.२ तथ्याङ्कीय अवधारणाहरू बुझ्ने क्षमता	
	३.३.३ तथ्याङ्कीय जानकारीको विश्लेषण, व्याख्या र मूल्याङ्कन गर्ने क्षमता	
	३.३.४ तथ्याङ्कीय जानकारी र बुझाइ सञ्चार गर्ने क्षमता	
	३.४ तथ्याङ्कीय साक्षरताको महत्त्व	
	३.५ विकास प्रशासनमा तथ्याङ्कीय साक्षरताको प्रभाव	

- ३.६ डिजिटल साक्षरता (Digital Literacy)
- ३.७ निर्णय प्रक्रियामा डिजिटल साक्षरता
- ३.८ अभ्यासका लागि प्रश्नहरू
- ३.९ सन्दर्भ सामग्री

भाग २ – तथ्याङ्कसम्बन्धी आधारभूत जानकारी

पाठ ४	तथ्याङ्कको परिचय	४३
	४.१ तथ्याङ्कको परिभाषा	
	४.२ तथ्याङ्कको महत्त्व तथा दायरा	
	४.३ तथ्याङ्कका सीमा	
	४.४ तथ्याङ्क प्रतिको अविश्वास	
	४.५ अभ्यासका लागि प्रश्नहरू	
	४.६ सन्दर्भ सामग्री	
पाठ ५	चरको परिचय तथा मापन एकाइ सम्बन्धी अवधारणा	४९
	५.१ चरको परिचय	
	५.२ चरका प्रकार	
	५.२.१ गुणात्मक चर	
	५.२.२ परिमाणात्मक चर	
	५.३ मापन भनेको के हो ?	
	५.३.१ नाममात्रको तहका तथ्याङ्क (Nominal - Level Data)	
	५.३.२ क्रमिक तहका तथ्याङ्क (Ordinal - Level Data)	
	५.३.३ अन्तराल तहका तथ्याङ्क (Interval - Level Data)	
	५.३.४ अनुपात तहका तथ्याङ्क (Ratio - Level Data)	
	५.४ मापनको महत्त्व	
	५.५ मापनको प्रयोग	
	५.६ अभ्यासका लागि प्रश्नहरू	
	५.७ सन्दर्भ सामग्री	
पाठ ६	तथ्याङ्कका स्रोत तथा तथ्याङ्क व्यवस्थापन	५७
	६.१ परिचय	
	६.२ तथ्याङ्क स्रोत	
	६.२.१ गणना (Census)	
	६.२.२ नमुना सर्वेक्षण (Sample Survey)	
	६.२.३ प्रशासनिक अभिलेख (Administrative Records)	
	६.२.४ बिग डाटा (Big Data)	
	६.३ तथ्याङ्क संकलन	
	६.३.१ प्रत्यक्ष व्यक्तिगत अन्तरवार्ता (Direct Personel Interview)	
	६.३.२ अप्रत्यक्ष व्यक्तिगत अन्तरवार्ता (Indirect Personel Interview)	

- ६.३.३ स्थानीय प्रतिनिधि/संवाददाताहरूबाट तथ्याङ्क संकलन
- ६.३.४ प्रश्नावली मेल गर्ने विधि (Mailed Questionnaire)
- ६.३.५ सेड्युल्सको माध्यमबाट गणकद्वारा तथ्याङ्क संकलन
- ६.३.६ इन्टरनेटको प्रयोग गरी तथ्याङ्क संकलन
- ६.४ तथ्याङ्क संकलनमा प्रयोग हुने प्रश्नावली/सेड्युल्स निर्माण गर्दा ध्यान दिनुपर्ने कुराहरू
- ६.५ तथ्याङ्क वर्गीकरण र यसका नियम
- ६.६ आवृत्ति वितरण
- ६.७ द्विविभाजन आवृत्ति वितरण
- ६.८ तालिकीकरण: अर्थ, महत्त्व तथा प्रकार
- ६.९ आउटलायर्सको परिचय
- ६.१० अभ्यासका लागि प्रश्नहरू
- ६.११ सन्दर्भ सामग्री
- ६.१२ पदावलीहरू र तिनीहरूको अर्थ

पाठ ७ तथ्याङ्क र सूचक

७९

७.१ परिचय

- ७.१.१ व्यक्तिगत डेटा
- ७.१.२ खण्डित डेटा
- ७.१.३ अविच्छिन्न डेटा
- ७.२ सूचकका प्रकार (Types of Indicators)
 - ७.२.१ लागत सूचक (Cost Indicator)
 - ७.२.२ प्रक्रियागत सूचक (Process Indicator)
 - ७.२.३ प्रतिफलमापक सूचक (Output Indicator)
 - ७.२.४ प्रभाव वा असरमापक सूचक (Outcome Indicator)
- ७.३ सूचनाको कच्चा पदार्थका रूपमा आँकडा
- ७.४ सूचकको महत्त्व तथा प्रयोग
- ७.५ असल सूचकमा हुनुपर्ने गुणहरू
- ७.६ तथ्याङ्क विश्लेषण र सूचक निर्माण
- ७.७ तथ्याङ्क विश्लेषणका उद्देश्य
- ७.८ सूचक सन्दर्भ पाना
- ७.९ अभ्यासका लागि प्रश्नहरू
- ७.१० शब्दावली

पाठ ८ तथ्याङ्क प्रस्तुतिकरण

९३

८.१ परिचय

- ८.२ तथ्याङ्क प्रस्तुतिको महत्त्व
- ८.३ तथ्याङ्क प्रस्तुतिका प्रकार
- ८.४ तथ्याङ्क प्रस्तुतिकरण गर्दा ध्यान दिनुपर्ने कुराहरू

- ८.५ चित्रात्मक प्रस्तुतिकरण
 - ८.५.१ स्तम्भचित्र
 - ८.५.२ वृत्तचित्र (Pie chart)
 - ८.५.३ नक्सा
 - ८.५.४ हिस्टोग्राम (Histogram)
 - ८.५.५ समय श्रृङ्खलाको ग्राफ
- ८.६ डायग्राम र ग्राफको प्रयोगबाट तथ्याङ्कको विश्लेषण तथा व्याख्या
- ८.७ अभ्यासका लागि प्रश्नहरू
- ८.८ सन्दर्भ सामग्री

पाठ ९ तथ्याङ्क विश्लेषणका लागि आधारभूत गणितीय सूचक

११३

- ९.१ परिचय
- ९.२ संख्याको सामान्य ज्ञान
- ९.३ भिन्नको सामान्य ज्ञान
- ९.४ अनुपात, समानुपात, दर तथा प्रतिशतको व्याख्या र प्रयोग
 - ९.४.१ अनुपात (Ratio)
 - ९.४.२ समानुपात (Proportion)
 - ९.४.३ दर (Rate)
 - ९.४.४ प्रतिशत (Percentage)
 - ९.४.५ प्रतिशत बिन्दु (Percentage Point)
- ९.५ दर, अनुपात, समानुपात, प्रतिशत, प्रतिशत बिन्दु प्रयोग गर्दा ध्यान दिनुपर्ने कुराहरू
- ९.६ अभ्यासका लागि प्रश्नहरू
- ९.७ सन्दर्भ सामग्री
- ९.८ शब्दावली

पाठ १० सूचकाङ्क र यसको प्रयोग

१२५

- १०.१ परिचय
- १०.२ सूचकाङ्कका प्रकार र नेपालमा निर्माण गरिने सूचकाङ्कहरू
 - १०.२.१ सूचकाङ्कका प्रकार
 - १०.२.२ नेपालमा निर्माण गरिने सूचकाङ्कहरू
- १०.३ सूचकाङ्क निर्माणका तरिका तथा सूचकाङ्क निर्माणमा आइपर्ने समस्या
- १०.४ सूचकाङ्कको व्याख्या एवं प्रयोग
- १०.५ सूचकाङ्क प्रयोग गर्दा ध्यान दिनुपर्ने कुराहरू
- १०.६ आधार समय परिवर्तन र सूचकाङ्कको स्प्लाइसिड तथा डिफ्लेटीड
- १०.७ जीविकोपार्जन सूचकाङ्क: निर्माण र प्रयोग
- १०.८ सूचकाङ्कका सीमाहरू
- १०.९ अभ्यासका लागि प्रश्नहरू
- १०.१० सन्दर्भ सामग्री

पाठ ११ केन्द्रीय प्रवृत्ति मापन

१४३

- ११.१ परिचय
- ११.२ असल केन्द्रीय प्रवृत्तिका गुणहरू
- ११.३ केन्द्रीय प्रवृत्ति मापन
 - ११.३.१ मध्यक (Mean)
 - ११.३.२ मध्यिका (Median)
 - ११.३.३ बहुलक (Mode)
- ११.४ मध्यक, मध्यिका र बहुलक बिचको सम्बन्ध
- ११.५ केन्द्रीय प्रवृत्तिको गुण, अवगुण तथा प्रयोगका क्षेत्र
- ११.६ केन्द्रीय प्रवृत्ति मापनको छनोट
- ११.७ मध्यिका र बहुलकको ग्राफिकल मापन
- ११.८ औसतका सीमा
- ११.९ विभाजक गर्ने मानअभ्यासका लागि प्रश्नहरू
 - ११.९.१ चतुर्थांश (Quartile)
 - ११.९.२ दशमक (Decile)
 - ११.९.३ प्रतिशतक
 - ११.९.४ विभाजन गर्ने मानहरूको प्रयोग तथा महत्त्व
- ११.१० अभ्यासका लागि प्रश्नहरू
- ११.११ सन्दर्भ सामग्री

पाठ १२ विचलनको मापन

१७९

- १२.१ परिचय र अर्थ
- १२.२ असल विचलन मापनका गुणहरू
- १२.३ विचलनको निरपेक्ष र सापेक्ष मापन
- १२.४ विचलनको विभिन्न मापन
 - १२.४.१ स्क्वाटर डायग्राम
 - १२.४.२ विस्तार
 - १२.४.३ मध्य विचलन
 - १२.४.४ स्तरीय भिन्नता
 - १२.४.५ स्तरीय भिन्नताको मानाङ्क
 - १२.४.६ भेरियन्स
- १२.५ विचलन मापनका विशेषताहरू
- १२.६ जेड-स्कोर (z-score)
- १२.७ लरेन्ज बक्ररेखा र जिनीको गुणाङ्क (Lorenz Curve and Gini Coefficient)
 - १२.७.१ लरेन्ज बक्ररेखा बनाउने तरिका
 - १२.७.२ जिनीको गुणाङ्क

१२.८ अभ्यासका लागि प्रश्नहरू

१२.९ सन्दर्भ सामग्री

पाठ १३ सहसम्बन्ध र यसको प्रयोग

२०९

१३.१ परिचय

१३.२ सहसम्बन्धका प्रकार र यसका विशेषता

१३.२.१ सकारात्मक र नकारात्मक सहसम्बन्ध

१३.२.२ रेखीय तथा गैर-रेखीय सहसम्बन्ध

१३.२.३ सम्बन्धको स्तर (डिग्री)

१३.३ सहसम्बन्ध अध्ययन गर्ने विधि

१३.३.१ स्क्याटर चित्रको माध्यमबाट सहसम्बन्ध अध्ययन

१३.३.२ सहसम्बन्ध गुणाङ्कको गणना

१३.४ सहसम्बन्ध गुणाङ्कको व्याख्या र प्रयोग

१३.५ कोअफिसियन्ट अफ डिटरमिनेशन

१३.६ अभ्यासका लागि प्रश्नहरू

१३.७ सन्दर्भ सामग्री

पाठ १४ समय श्रृङ्खला विश्लेषण

२२१

१४.१ परिचय

१४.२ समय श्रृङ्खलाका अवयवहरू

१४.३ समय श्रृङ्खलाको विश्लेषण तथा महत्त्व

१४.४ समय श्रृङ्खला विश्लेषणको लागि गणितीय मोडेलहरू

१४.५ समय श्रृङ्खलाको विश्लेषण तथा महत्त्व

१४.५.१ प्रवृत्ति (Trend = T)

१४.५.२ मौसमी विचलन (Seasonal Variation = S)

१४.५.३ चक्रीय तथा अनियमित विचलन

(Seasonal and Irregular Variation = S)

१४.६ समय श्रृङ्खलामा अनुमान

१४.६.१ समय श्रृङ्खलाको सिधा रेखीय प्रक्षेपण

१४.७ वृद्धिदर

१४.८ अभ्यासका लागि प्रश्नहरू

१४.९ सन्दर्भ सामग्री

पाठ १५ नर्मल डिस्ट्रिब्यूसन

२५३

१५.१ परिचय

१५.२ नर्मल डिस्ट्रिब्यूसन बक्ररेखाका विशेषता र समीकरण

१५.३ स्टान्डर्ड नर्मल डिस्ट्रिब्यूसन

१५.४ स्टान्डर्ड नर्मल सम्भाव्य बक्ररेखा अन्तर्गतको क्षेत्रफल

१५.५ नर्मल डिस्ट्रिब्यूसनको महत्त्व र प्रयोग

१५.६ अभ्यासका लागि प्रश्नहरू

१५.७ सन्दर्भ सामग्री

पाठ १६ अनुमान तथा नमुनासम्बन्धी आधारभूत जानकारी

२६५

१६.१ अनुमानको सिद्धान्त

१६.२ अनुमानका गुणहरू

१६.३ स्याम्प्लिङ डिस्ट्रिब्यूशन अफ स्टाटिस्टिक्स

१६.४ सेन्ट्रल लिमिट सिद्धान्त र यसको प्रयोग

१६.४.१ सेन्ट्रल लिमिट सिद्धान्तको महत्त्व तथा प्रयोग

१६.४.२ सेन्ट्रल लिमिट सिद्धान्तले प्रयोग गर्नका लागि ध्यान दिनुपर्ने कुराहरू

१६.५ अनुमानका प्रकारहरू

१६.६ विश्वस्तता अन्तराल

१६.७ नमूना छनोटसम्बन्धी अवधारणा

१६.७.१ नमूना छनोट अध्ययनका फाइदा तथा बेफाइदा

१६.७.२ नमूना छनोटका प्रकार

१६.७.३ नमूनाको आकार (Sample Size)

१६.८ अभ्यासका लागि प्रश्नहरू

१६.९ सन्दर्भ सामग्री

१६.१० पदावलीहरू र तिनीहरूको अर्थ

पाठ १७ परिकल्पना परीक्षण

२८१

१७.१ परिचय

१७.२ तथ्याङ्कीय परिकल्पना र यसका प्रकार

१७.३ परिकल्पना परीक्षणमा हुने त्रुटि र तिनीहरूको प्रकार

१७.४ लेभल अफ सिग्निफिक्यान्स

१७.५ क्रिटिकल मान तथा क्रिटिकल क्षेत्र

१७.६ टेष्ट स्टाटिस्टिक्स र यसको पी-मान

१७.७ परिकल्पना परीक्षणका प्रमुख चरणहरू

१७.८ एकल मध्यक परीक्षण गर्ने उदाहरण

१७.९ दुई मध्यकबिचको फरक परीक्षण गर्ने उदाहरण

१७.१० अभ्यासका लागि प्रश्नहरू

१७.११ सन्दर्भ सामग्री

पाठ १८ प्रतिगमन विश्लेषण

२९९

१८.१ परिचय

१८.२ सहसम्बन्ध र प्रतिगमन बिचको भिन्नता

१८.३ स्वतन्त्र तथा निर्भर चरहरू

१८.४ सिधा रेखीय प्रतिगमन विश्लेषण

१८.५ गुणाङ्कहरूको अनुमान तथा व्याख्या

- १८.६ स्पेसिफिकेशन र डाटासम्बन्धी सवाल
- १८.७ हिटेरोस्केडारिस्टिसटी
- १८.८ आर स्क्वायर्ड र गुडनेस अफ फिट
- १८.९ अभ्यासका लागि प्रश्नहरू
- १८.१० सन्दर्भ सामग्री

भाग ४ – तथ्याङ्कीय क्रियाकलापमा थप विषयको अध्ययन

- पाठ १९ तथ्याङ्कका गुण** ३१५
- १९.१ परिचय
 - १९.२ तथ्याङ्क गुणका आयाम
 - १९.३ तथ्याङ्कको उपलब्धता र पहुँच
 - १९.४ अभ्यासका लागि प्रश्नहरू
 - १९.५ सन्दर्भ सामग्री
- पाठ २० प्रायोगिक परियोजना सञ्चालनका प्रमुख चरणहरू तथा तथ्याङ्कमा गोपनीयता र नैतिकता** ३२१
- २०.१ परिचय
 - २०.२ प्रायोगिक परियोजना सञ्चालनका प्रमुख चरणहरू
 - २०.३ तथ्याङ्कमा गोपनीयता र नैतिकता
 - २०.४ गोपनीयता सम्बन्धी कानुनी प्रावधान
 - २०.५ अभ्यासका लागि प्रश्नहरू
 - २०.६ सन्दर्भ सामग्री
- पाठ २१ सर्वेक्षण प्रतिवेदन अध्ययन** ३२७
- २१.१ परिचय
 - २१.२ सर्वेक्षण प्रतिवेदन भनेको के हो?
 - २१.३ सर्वेक्षण प्रतिवेदन अध्ययन गर्दा ध्यान दिनुपर्ने कुराहरू
 - २१.४ परियोजना कार्य
 - २१.५ सन्दर्भ सामग्री

भाग ५ – तथ्याङ्कीय क्रियाकलापमा कम्प्युटरको प्रयोग

- पाठ २२ एक्सेलसम्बन्धी आधारभूत ज्ञान** ३३५
- २२.१ परिचय
 - २२.१.१ मेनुबार (Menu Bar)
 - २२.१.२ हरफ, महल र कोठा (Row, Column, Cell)
 - २२.१.३ फिल्टर, सर्टिंग र फिल्टर निष्क्रिय गन
 - २२.१.४ सिट (थप्ने, नाम परिवर्तन गर्ने, नक्कल गर्ने र हटाउने)
 - २२.१.५ अक्षरहरूको अवस्थिति मिलान (Alignment), कोठाहरू मर्ज (एकीकरण) र अक्षरलाई हरफीकरण (Wrap) गर्ने

- २२.१.६ कपी गरेका कुरा कतै लगेर टाँस (Paste) गर्ने र फर्मेट गर्ने
- २२.१.७ हरफ र महल थप्ने
- २२.१.८ महल वा हरफको शीर्षक राखेर अरू भाग लुकाउने (Freeze Panes)
- २२.१.९ एक्सेलमा तथ्याङ्क व्यवस्थापन
- २२.१.१० तथ्याङ्क व्यवस्थापन र निर्णय निर्माणमा एक्सेल
- २२.२ तालिका तथा यसका ढाँचा
- २२.३ एक्सेलमा सूत्रको प्रयोग
- २२.४ एक्सेलमा पिभोट टेबल बनाउने विधि
- २२.५ एक्सेलमा विभिन्न ग्राफ तथा चित्र बनाउने विधि
 - २२.५.१ बार डायग्राम बनाउने विधि
- २२.६ अभ्यासका लागि प्रश्नहरू
- २२.७ सन्दर्भ सामग्री

पाठ २३ स्टेटा सम्बन्धी आधारभूत ज्ञान ३६३

- २३.१ STATA को परिचय र यसका आधारभूत विशेषताहरू
 - २३.१.१ बार डायग्राम बनाउने विधि
- २३.२ स्टेटामा डाटा लोड गर्ने तरिका, तथ्याङ्कको आयात तथा निर्यात
- २३.३ मेनुको परिचय
- २३.४ STATA कमाण्डहरू, "Do" फाइल र लग फाइल
- २३.५ STATA मा डाटा व्यवस्थापन
- २३.६ डाटा संक्षेपीकरण
- २३.७ डाटा भिजुअलाइजेशन
- २३.८ परियोजना कार्य
- २३.९ अभ्यासका लागि प्रश्नहरू
- २३.१० सन्दर्भ सामग्री

अनुसूचीहरू

- अनुसूची १ वस्तुगत प्रश्न समाधानका लागि सुझाव ३७९
- अनुसूची २ साँख्यिकीय तालिकाहरू ३९९

भाग १

नेपालको तथ्याङ्क प्रणाली
र तथ्याङ्कीय साक्षरता

नेपालको तथ्याङ्क प्रणाली

१.१ परिचय

राष्ट्रिय तथ्याङ्क प्रणाली भन्नाले कुनै देशको आधिकारिक तथ्याङ्कसम्बन्धी कानुनी प्रावधान, पूर्वाधार तथा संस्थागत व्यवस्थाको समष्टिगत रूपलाई बुझिन्छ । स्वभावतः यसमा तथ्याङ्क उत्पादक, प्रदायक, प्रयोगकर्ता र शोध एवम् अनुसन्धानसँग सम्बन्धित पक्षहरूको अन्तरआबद्धता रहेको हुन्छ । यसले तथ्याङ्क संकलन, व्यवस्थापन, वितरण तथा उपयोगको खाका निर्धारण गरेको हुन्छ । अर्गनाइजेसन फर इकोनोमिक कोअपरेसन एण्ड डेभेलपमण्ट (OECD) ले राष्ट्रिय तथ्याङ्क प्रणालीलाई निम्नानुसार परिभाषित गरेको पाइन्छ¹:

“राष्ट्रिय तथ्याङ्क प्रणाली (National Statistical System) कुनै पनि देश भित्रको तथ्याङ्कीय संगठन र एकाइहरूको समूह हो, जसले राष्ट्रिय सरकारको तर्फबाट आधिकारिक तथ्याङ्कहरू संयुक्त रूपमा सङ्कलन, व्यवस्थापन र प्रसार गर्दछ” (OECD, 2004) ।

सामान्यतया, तथ्याङ्क प्रणाली केन्द्रीकृत (जस्तै: क्यानडा, अष्ट्रेलिया) र विकेन्द्रित (जस्तै: अमेरिका, बेलायत, चीन, फिलिपिन्स, फ्रान्स) गरी दुई प्रकारको हुन्छ । तथापि, समग्रमा कुनै पनि तथ्याङ्क प्रणालीलाई पूर्णरूपमा केन्द्रीकृत र विकेन्द्रित भनेर छुट्याउन भने कठिन नै हुन्छ । नेपालको सन्दर्भमा कुरा गर्दा देशको तथ्याङ्क प्रणाली सुरुका दिनहरूमा परम्परागत केन्द्रीकृत मोडेलमा आधारित थियो भने हाल विकेन्द्रित मोडेलमा आधारित रहेको पाइन्छ । नेपालको तथ्याङ्क प्रणाली औपचारिक रूपमा वि.सं. १९६८ मा सम्पन्न पहिलो जनगणनाबाट सुरु भएको मानिन्छ । तथ्याङ्क प्रणालीलाई थप सुदृढ गर्दै लैजाने क्रममा वि.सं. १९९५ मा जनसंख्या गोश्वारा स्थापना गरिएको थियो । यस अड्डाले नेपालको चौथो जनगणना (वि.सं. १९९८) सञ्चालन गरेको थियो भने पछि यस अड्डालाई औद्योगिक व्यापारिक समाचार संग्रह अड्डासँग गाभेर संख्या विभाग (वि.सं. २००८) को स्थापना गरिएको थियो । यो विभागले नेपालको पाँचौँ जनगणना (वि.सं. २००९/११) सम्पन्न गरेको थियो ।

तथ्याङ्क प्रणालीको विकासकै क्रममा वि.सं. २०१५ मा तथ्याङ्क ऐन जारी भई केन्द्रीय तथ्याङ्क विभागको स्थापना भएको थियो । देशको केन्द्रीय तथ्याङ्क निकायको रूपमा यस

1 OECD. Retrieved on November 14, 2021 from: <https://stats.oecd.org/glossary/detail.asp?ID=1726>

राष्ट्रिय तथ्याङ्क प्रणाली भन्नाले कुनै देशको आधिकारिक तथ्याङ्कसम्बन्धी कानुनी प्रावधान, पूर्वाधार तथा संस्थागत व्यवस्थाको समष्टिगत रूपलाई बुझिन्छ । स्वभावतः यसमा तथ्याङ्क उत्पादक, प्रदायक, प्रयोगकर्ता र शोध एवम् अनुसन्धानसँग सम्बन्धित पक्षहरूको अन्तरआबद्धता रहेको हुन्छ । यसले तथ्याङ्क संकलन, व्यवस्थापन, वितरण तथा उपयोगको खाका निर्धारण गरेको हुन्छ । अर्गनाइजेसन फर इकोनोमिक कोअपरेसन एण्ड डेभेलपमण्ट (OECD) ले राष्ट्रिय तथ्याङ्क प्रणालीलाई निम्नानुसार परिभाषित गरेको पाइन्छ¹:

“राष्ट्रिय तथ्याङ्क प्रणाली (National Statistical System) कुनै पनि देश भित्रको तथ्याङ्कीय संगठन र एकाइहरूको समूह हो, जसले राष्ट्रिय सरकारको तर्फबाट आधिकारिक तथ्याङ्कहरू संयुक्त रूपमा सङ्कलन, व्यवस्थापन र प्रसार गर्दछ” (OECD, 2004) ।

सामान्यतया, तथ्याङ्क प्रणाली केन्द्रीकृत (जस्तै: क्यानडा, अष्ट्रेलिया) र विकेन्द्रित (जस्तै: अमेरिका, बेलायत, चीन, फिलिपिन्स, फ्रान्स) गरी दुई प्रकारको हुन्छ । तथापि, समग्रमा कुनै पनि तथ्याङ्क प्रणालीलाई पूर्णरूपमा केन्द्रीकृत र विकेन्द्रित भनेर छुट्याउन भने कठिन नै हुन्छ । नेपालको सन्दर्भमा कुरा गर्दा देशको तथ्याङ्क प्रणाली सुरुका दिनहरूमा परम्परागत केन्द्रीकृत मोडेलमा आधारित थियो भने हाल विकेन्द्रित मोडेलमा आधारित रहेको पाइन्छ । नेपालको तथ्याङ्क प्रणाली औपचारिक रूपमा वि.सं. १९६८ मा सम्पन्न पहिलो जनगणनाबाट सुरु भएको मानिन्छ । तथ्याङ्क प्रणालीलाई थप सुदृढ गर्ने लैजाने क्रममा वि.सं. १९९५ मा जनसंख्या गोश्वारा स्थापना गरिएको थियो । यस अङ्गले नेपालको चौथो जनगणना (वि.सं. १९९८) सञ्चालन गरेको थियो भनेपछि यस अङ्गलाई औद्योगिक व्यापारिक समाचार संग्रह अङ्गसँग गाभेर संख्या विभाग (वि.सं. २००८) को स्थापना गरिएको थियो । यो विभागले नेपालको पाँचौँ जनगणना (वि.सं. २००९/११) सम्पन्न गरेको थियो ।

तथ्याङ्क प्रणालीको विकासकै क्रममा वि.सं. २०१५ मा तथ्याङ्क ऐन जारी भई केन्द्रीय तथ्याङ्क विभागको स्थापना भएको थियो । देशको केन्द्रीय तथ्याङ्क निकायको रूपमा यस विभागले हालसम्म तथ्याङ्कीय क्रियाकलापहरू सञ्चालन गर्ने नेतृत्वदायी भूमिका निर्वाह गर्दै आइरहेको छ ।

१.२ राष्ट्रिय तथ्याङ्क कार्यालय र यसको भूमिका

वि.सं. २०१५ पुस १७ गते केन्द्रीय तथ्याङ्क विभागको स्थापना भएपश्चात् राष्ट्रिय तथ्याङ्क कार्यालयको रूपमा यस विभाग क्रियाशील हुँदै आइरहेको तथ्य सर्वविदितै छ । तथ्याङ्क ऐन, २०१५ बमोजिम यस विभागले तथ्याङ्क सङ्कलन, प्रशोधन, विश्लेषण, प्रकाशन र

प्रसारका लागि केन्द्रीय निकायको रूपमा भूमिका निर्वाह गर्दै आएको छ । यस विभागले राष्ट्रिय योजना आयोग अन्तर्गतको विभागीय संगठनको रूपमा कार्य गर्दछ । यसले मुख्यतया जनगणना, विभिन्न सर्वेक्षण तथा अन्य तथ्याङ्कीय अध्ययनका आधारमा निश्चित समयभित्र भरपर्दो आर्थिक तथा सामाजिक तथ्याङ्क उत्पादन तथा वितरण गर्दछ ।

विभाग स्थापना गर्नुको मुख्य उद्देश्य योजना, नीति तथा कार्यक्रमहरू तर्जुमा गर्नुका साथै राष्ट्रिय योजना आयोग तथा अन्य सरकारी निकायहरूलाई आधिकारिक तथ्याङ्क उपलब्ध गराउनु र देशको आर्थिक वृद्धि मापन गर्न राष्ट्रिय लेखा अनुमानहरू उत्पादन गर्नु हो । तथ्याङ्क ऐन, २०१५ को प्रस्तावनामा केन्द्रीय तथ्याङ्क विभाग स्थापना गर्नुको उद्देश्य निम्नानुसार उल्लेख गरिएको छ ।²

“जनतालाई अधिक मात्रामा ठोस हित हुने योजना तर्जुमा गर्न, नीति निर्धारण गर्न र मुलुकको आर्थिक सामाजिक गतिविधिबारे यथार्थ ज्ञान प्राप्त गर्न सरकारलाई भरपर्दो आधार जुटाउन सामयिक तथा गुणस्तरयुक्त तथ्याङ्क उपलब्ध गराउने”
(प्रस्तावना: तथ्याङ्क ऐन, २०१५) ।

केन्द्रीय तथ्याङ्क विभागले देशको सामाजिक आर्थिक अवस्थाको मूल्याङ्कन गर्न नियमित रूपमा जनगणना गर्नुका साथै अन्य विभिन्न सर्वेक्षणहरू गर्दछ । विभागले गरिरहेको प्रमुख कार्यहरूमा जनगणना, कृषिगणना, औद्योगिक गणना, आर्थिक गणना, राष्ट्रिय लेखा अनुमान, मूल्य सूचकाङ्क अनुमान, जीवनस्तर सर्वेक्षण, श्रमशक्ति सर्वेक्षण, वार्षिक घरपरिवार सर्वेक्षण, बहुसूचक सर्वेक्षण तथा अन्य विविध सर्वेक्षणहरू पर्दछन् । यसरी विभिन्न माध्यमबाट संकलन गरिएका तथ्याङ्कहरूको प्रशोधन, विश्लेषण, प्रकाशन, वितरण तथा भण्डारण गर्ने प्रमुख जिम्मेवारी यस विभागले वहन गर्दै आएको छ । विभागको समष्टिगत जिम्मेवारीलाई देहायबमोजिम प्रस्तुत गर्न सकिन्छ ।

- विभिन्न गणना, सर्वेक्षण तथा समष्टिगत आर्थिक तथ्याङ्कको व्यवस्थापन गर्ने,
- तथ्याङ्कीय क्रियाकलापमा समन्वय, स्तरीकरण, गुणस्तर व्यवस्थापन गर्ने,
- राष्ट्रिय तथा अन्तर्राष्ट्रिय विकास एजेण्डालाई आवश्यक पर्ने तथ्याङ्कको आपूर्ति गर्ने,
- अन्तर्राष्ट्रियस्तरका तथ्याङ्कीय गतिविधिमा प्रतिनिधित्व तथा सहकार्य गर्ने,
- तथ्याङ्कीय जनशक्ति एवं तथ्याङ्कीय पूर्वाधारको विकास गर्ने,
- राष्ट्रिय तथ्याङ्क प्रणालीको विकास गर्ने,
- तथ्याङ्कसम्बन्धी नीति तथा रणनीतिहरूको विकास गर्ने,

2 Nepal Law Commission. (2015). *Statistics Act 2015*. Kathmandu, Nepal; Nepal Law Commission. Retrieved on November 14, 2021 from: <https://bit.ly/3CqBLpO>

- नेपालको मुख्य तथ्याङ्कीय निकायको रूपमा गर्नुपर्ने अन्य कार्यहरू गर्ने ।

राष्ट्रिय तथ्याङ्क प्रणालीको समग्र रूपमा प्रभावकारी सञ्चालनका लागि केन्द्रीय तथ्याङ्क विभागको भूमिका सधैं महत्त्वपूर्ण तथा निर्णायक रहेको छ । यसले प्राज्ञिक समुदायका सदस्यहरू, तथ्याङ्क उत्पादक तथा प्रयोगकर्ताहरू बिच समन्वयात्मक तथा सहयोगी अनुसन्धान प्रयासहरूलाई अभिवृद्धि गरिरहेको हुन्छ । तसर्थ, देशमा तथ्याङ्क प्रणालीको विकास र तथ्याङ्कीय स्तर कायम राख्नमा यसको भूमिका निकै महत्त्वपूर्ण रहेको छ ।

१.३ तथ्याङ्क उत्पादन र प्रयोगकर्ताका रूपमा रहेका प्रमुख संस्थाहरू

नेपालमा तथ्याङ्क उत्पादनमा संलग्न प्रमुख संस्थाका रूपमा केन्द्रीय तथ्याङ्क विभाग रहेको कुरा माथि नै चर्चा गरिसकिएको छ । यस विभागका अतिरिक्त विभिन्न क्षेत्रका तथ्याङ्क उत्पादन, प्रशोधन, विश्लेषण, प्रकाशन, वितरण तथा भण्डारण जस्ता तथ्याङ्कीय क्रियाकलापमा अन्य सरकारी तथा अर्धसरकारी संघ संस्थाहरू संलग्न रहेका छन् । उदाहरणका रूपमा शिक्षा मन्त्रालयले शैक्षिक तथ्याङ्क, स्वास्थ्य मन्त्रालयले स्वास्थ्यसम्बन्धी तथ्याङ्क, कृषि मन्त्रालयले कृषि तथ्याङ्क, नेपाल राष्ट्र बैंकले मूल्य तथ्याङ्क, शोधनान्तर स्थिति तथा मुद्रा र बैंकसम्बन्धी तथ्याङ्क आदि उत्पादन तथा व्यवस्थापन गर्दछन् । त्यसैगरी तथ्याङ्क प्रयोगकर्ताका रूपमा राष्ट्रिय योजना आयोग, विषयगत मन्त्रालयहरू, विभिन्न सरकारी, अर्धसरकारी, गैर सरकारी तथा निजी क्षेत्रका संघ संस्थाहरू प्रमुख रहेका छन् । कतिपय संस्थाहरूले तथ्याङ्क उत्पादन र प्रयोग दुवै कार्य गर्दछन् भने कतिपय निकायहरूले तथ्याङ्क उत्पादनभन्दा प्रयोग बढी मात्रामा गरिरहेका हुन्छन् । नेपालका प्रमुख तथ्याङ्क उत्पादकहरू र तिनले उत्पादन गर्ने तथ्याङ्क तालिका १.१ मा प्रस्तुत गरिएको छ ।

तालिका १.१: नेपालका प्रमुख तथ्याङ्क उत्पादक तथा उत्पादित तथ्याङ्कको क्षेत्र

क्र.सं.	तथ्याङ्क उत्पादकको नाम	तथ्याङ्क उत्पादन विधि	तथ्याङ्कको प्रमुख क्षेत्र
१	केन्द्रीय तथ्याङ्क विभाग	गणना तथा सर्वेक्षण	आर्थिक तथा सामाजिक तथ्याङ्क
२	स्वास्थ्य तथा जनसंख्या मन्त्रालय	स्वास्थ्य व्यवस्थापन सूचना प्रणाली तथा विभिन्न सर्वेक्षण	सामाजिक तथ्याङ्क (स्वास्थ्यसम्बन्धी)
३	शिक्षा, विज्ञान तथा प्रविधि मन्त्रालय	शिक्षा व्यवस्थापन सूचना प्रणाली तथा विभिन्न सर्वेक्षण	सामाजिक तथ्याङ्क (स्वास्थ्यसम्बन्धी)

४	श्रम, रोजगार तथा सामाजिक सुरक्षा मन्त्रालय	श्रम बजार सूचना प्रणाली, वैदेशिक रोजगार सूचना प्रणाली तथा विभिन्न सर्वेक्षण	आर्थिक तथा सामाजिक तथ्याङ्क (श्रमशक्ति, रोजगारी तथा सामाजिक सुरक्षासम्बन्धी)
५	महिला, बालबालिका तथा ज्येष्ठ नागरिक मन्त्रालय	सूचना व्यवस्थापन प्रणाली तथा विभिन्न सर्वेक्षण	सामाजिक तथ्याङ्क (लैङ्गिक तथ्याङ्क)
६	निर्वाचन आयोग, नेपाल	सूचना प्रणाली, मतदाता नामावली	सामाजिक तथा राजनीतिक तथ्याङ्क
७	संघीय मामिला तथा सामान्य प्रशासन मन्त्रालय	व्यक्तिगत विवरण वैयक्तिक सूचना प्रणाली तथा अन्य सूचना प्रणाली	सरकारी कर्मचारीसम्बन्धी तथ्याङ्क, अन्य आर्थिक तथा सामाजिक तथ्याङ्क
८	गृह मन्त्रालय	व्यक्तिगत घटना दर्ता तथा सामाजिक सुरक्षा व्यवस्थापन सूचना प्रणाली (VERSP-MIS)	जनसांख्यिक तथ्याङ्क (जन्म, मृत्यु, विवाह, बसाईसराइ, सम्बन्ध विच्छेद, सामाजिक सुरक्षा आदि); सार्वजनिक सुरक्षा र अपराधसम्बन्धी
१०	कृषि तथा पशुपन्छी विकास मन्त्रालय	कृषि व्यवस्थापन सूचना प्रणाली तथा विभिन्न सर्वेक्षण	आर्थिक तथ्याङ्क (कृषि)
११	नेपाल राष्ट्र बैंक	राजस्व व्यवस्थापन सूचना प्रणाली (RMIS), वित्तीय सूचना एकाइ (FIU) तथा विभिन्न सर्वेक्षण	आर्थिक तथ्याङ्क (मूल्य सूचकाङ्क, शोधनान्तर स्थिति तथा मुद्रा र बैंकसम्बन्धी तथ्याङ्क)
१२	उद्योग, वाणिज्य तथा आपूर्ति मन्त्रालय	औद्योगिक सूचना प्रणाली, अन्य सूचना प्रणाली तथा सर्वेक्षणहरू	आर्थिक तथ्याङ्क (उद्योग प्रतिष्ठान, रोजगार र तिनीहरूको लागत, उत्पादनसम्बन्धी)
१३	संस्कृति, पर्यटन तथा नागरिक उड्डयन मन्त्रालय	पर्यटन व्यवस्थापन सूचना प्रणाली तथा विभिन्न सर्वेक्षणहरू	आर्थिक तथ्याङ्क (पर्यटनसम्बन्धी)
१४	विभिन्न प्रतिष्ठानहरू	लेखा तथा प्रतिवेदनहरू	आर्थिक तथ्याङ्क (उत्पादनसम्बन्धी)
१५	ऊर्जा, जलस्रोत तथा सिंचाइ मन्त्रालय	जलवायु तथ्याङ्कसम्बन्धी पोर्टल	वातावरण तथ्याङ्क (तापक्रम, वर्षा आदि)
१६	वन तथा वातावरण मन्त्रालय	वातावरण व्यवस्थापन सूचना प्रणाली	वातावरण तथ्याङ्क (प्रदूषण, वन क्षेत्रफल आदि)

१७	विभिन्न विश्वविद्यालय तथा क्याम्पसहरू	विभिन्न सर्वेक्षण, अध्ययन अनुसन्धान	आर्थिक, सामाजिक, राजनीतिक, वातावरणसम्बन्धी तथ्याङ्क
१८	अदालत	प्रशासनिक अभिलेख	मुद्दा मामिलासम्बन्धी
१९	प्रदेश सरकार	प्रशासनिक अभिलेख तथा सर्वेक्षण	प्रदेश तथ्याङ्क
२०	स्थानीय तह	प्रशासनिक अभिलेख तथा स्थानीय तहका प्रोफाइल	व्यक्तिगत घटनासम्बन्धी, स्थानीय तहको साधन स्रोतसम्बन्धी, रोजगारीसम्बन्धी र अन्य स्थानीयस्तरका आर्थिक र सामाजिक तथ्याङ्क

नोट: माथि उल्लेख गरिएका संघ संस्था र ती मातहतका संगठनहरू बाहेक अन्य कतिपय निकायहरूले पनि तथ्याङ्क उत्पादन प्रशोधन तथा प्रसार गर्दछन् तर ती सबैको सूची यहाँ प्रस्तुत गरिएको छैन। तसर्थ, माथि उल्लेख नगरिएका तर तथ्याङ्कीय क्रियाकलापमा संलग्न केही संघ संस्थाहरूको पहिचान गर्ने अभिरूचि र जिम्मा पाठक वर्गको काँधमा छोडिएको छ। साथै, प्रदेश तथा स्थानीय तहका तथ्याङ्कीय क्रियाकलापका बारेमा मन्थन गर्नका लागि समेत आग्रह गरिएको छ।

१.४ नेपालको तथ्याङ्क प्रणाली र यसका अङ्गहरू

सुरूका दिनहरूमा परम्परागत केन्द्रीकृत प्रणालीबाट नेपालको तथ्याङ्क प्रणाली स्थापित भएको भए तापनि हाल नेपालको संविधानले स्पष्ट रूपमा संघ, प्रदेश र स्थानीय तहको अधिकार सूचीमा छुट्टाछुट्टै तथ्याङ्कीय क्रियाकलापको जिम्मेवारी प्रदान गरेबमोजिम विकेन्द्रित रूपमा तथ्याङ्कीय गतिविधिहरू सञ्चालनमा रहेका छन्। तसर्थ, हाल नेपालको तथ्याङ्क प्रणाली तेर्सो (Horizontal) र ठाडो (Vertical) दुवै किसिमले विकेन्द्रित रूपमा सञ्चालित रहेको छ। सोही बमोजिम तथ्याङ्क ऐन, २०१५ लाई समयसापेक्ष परिमार्जन गर्नका लागि तथ्याङ्क ऐनको नयाँ मस्यौदा तयार भई राष्ट्रिय सभाबाट पारित समेत भइसकेको अवस्था छ।^३

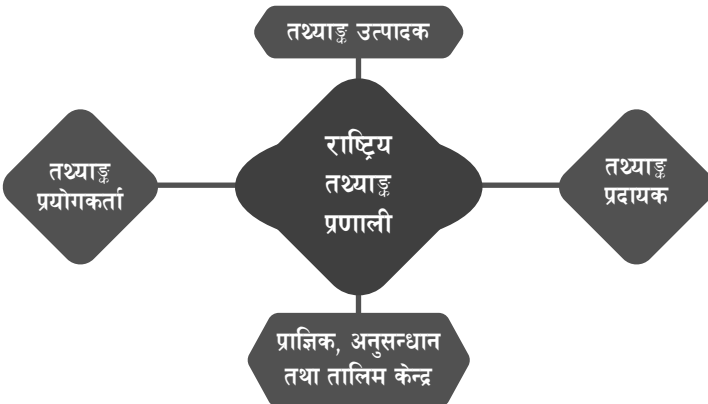
नेपालको संविधानको अनुसूची ५, ६ र ८ मा क्रमशः संघ, प्रदेश र स्थानीय तहको तथ्याङ्कीय अधिकारहरूको सूची तोकिएको छ।^४ यस अवस्थामा तिनैवटा सरकारमा

3 National Assembly. Retrieved on November 14, 2021 from: <https://na.parliament.gov.np/np/bills/F43qwywR>

4 Nepal Law Commission. (2015). *Constitution of Nepal*. Kathmandu, Nepal; Nepal Law Commission. Retrieved on November 14, 2021 from: <https://bit.ly/3Hzi664>

तथ्याङ्कीय गतिविधि विस्तार हुँदै गएको र सरकारी, अर्धसरकारी, गैर सरकारी, निजी क्षेत्र तथा अनुसन्धानकर्ता, सञ्चार जगत, नागरिक समाजलगायत सर्वाङ्गीण खाँचो र चासोको विषय भएकोले तथ्याङ्कको माग दिनानुदिन बढ्दै गइरहेको छ । यस परिप्रेक्ष्यमा राष्ट्रिय तथ्याङ्क प्रणालीलाई थप प्रभावकारी र परिष्कृत बनाउँदै सबल र सक्षम प्रणालीको रूपमा स्थापित गर्न पर्ने अवस्था रहेको छ । यसका लागि यस प्रणालीका अभिन्न अङ्गहरूका बारेमा सामान्य जानकारी राख्नु सान्दर्भिक हुन्छ । चित्र नं. १.१ मा प्रस्तुत गरिएबमोजिम राष्ट्रिय तथ्याङ्क प्रणालीका मुख्यगरी चारवटा प्रमुख अङ्गहरू रहेका हुन्छन् । तथ्याङ्क प्रदायकहरू तथ्याङ्कका श्रोत हुन् । जस्तै: व्यक्ति, घरपरिवार, प्रतिष्ठान, संघ संस्था आदिले आफू र आफूसँग सम्बन्धित विवरणमार्फत् तथ्याङ्क प्रदान गरिरहेका हुन्छन् भने यस्ता तथ्याङ्क प्रदायकबाट सूचना संकलन गरेर तथ्याङ्क उत्पादकहरूले तथ्याङ्क उत्पादन गरिरहेका हुन्छन् । यसरी उत्पादित तथ्याङ्कको प्रयोग सम्बन्धित प्रयोगकर्ताहरूले गर्दछन् । सरकारी तथा अर्धसरकारी निकाय जस्तै: राष्ट्रिय योजना आयोग, विभिन्न विषयगत मन्त्रालय तथा मातहतका संस्थाहरू, संवैधानिक निकाय, स्थानीय सरकारलगायत राष्ट्रिय तथा अन्तर्राष्ट्रिय संघ संस्था, निजी क्षेत्र, अनुसन्धानकर्ता, अध्येता, विद्यार्थी, पत्रकार, नागरिक समाज र सर्वसाधारण नागरिकसमेत तथ्याङ्कका प्रमुख प्रयोगकर्ता हुन् । यसैगरी तथ्याङ्क प्रणालीको अर्को प्रमुख अङ्गका रूपमा विश्वविद्यालय, क्याम्पस, अनुसन्धान तथा तालिम केन्द्रलाई लिन सकिन्छ । यस्ता निकाय तथा केन्द्रहरूले तथ्याङ्क विकासका लागि अध्ययन अनुसन्धानमार्फत् तथ्याङ्कको उत्पादन, व्यवस्थापन, प्रसार तथा प्रयोगका विविध विषयमा निरन्तर पृष्ठपोषण गरिरहेका हुन्छन् । यसरी तथ्याङ्क प्रणालीका यी चारवटै अङ्गहरूले एक आपसमा परस्पर अन्तरसम्बन्धित र परिपूरक हुँदै बलियो तथ्याङ्क प्रणालीको स्वरूप निर्धारण गरेका हुन्छन् ।

चित्र नं. १.१: तथ्याङ्क प्रणाली र यसका अङ्गहरू



१.५ नेपालको तथ्याङ्क प्रणालीका सबल पक्ष, कमजोर पक्ष, अवसर तथा चुनौतीहरू औपचारिक रूपमा वि.सं. १९६८ देखि नेपालमा तथ्याङ्क प्रणालीको सुरुवात भएको र वि.सं. २०१५ देखि तथ्याङ्क ऐन जारी गरी विधिवत् रूपमा तथ्याङ्क प्रणालीको स्थापना गरेसँगै हालसम्म आइपुग्दा यसले धेरै अनुभव, समस्या तथा चुनौतीहरू साक्षात्कार गरेको छ। यस प्रणालीलाई थप सुदृढ गर्दै बढ्दो तथ्याङ्क प्रयोगकर्ताको मागबमोजिम गुणस्तरीय तथ्याङ्क उत्पादन तथा प्रसार गर्नका लागि यस प्रणालीका सबल पक्ष, कमजोर पक्ष, अवसर तथा चुनौतीहरूको बारेमा चर्चा गर्नु अति नै महत्त्वपूर्ण हुन्छ।

सबल पक्ष

- उच्चस्तरीय समन्वय संयन्त्रको रूपमा राष्ट्रिय तथ्याङ्क परिषद्को व्यवस्था भएको,
- उच्च शैक्षिक योग्यता सहितको दक्ष जनशक्तिको उपलब्धता,
- तथ्याङ्क संकलन, प्रशोधन, विश्लेषण, प्रकाशन, वितरण तथा भण्डारणमा एक सय वर्षभन्दा बढीको अनुभव रहेको,
- विभिन्न गणना तथा नमुना सर्वेक्षणहरू नियमित रूपमा सञ्चालन भएका,
- अधिकांश निकायहरूमा प्रशासनिक अभिलेखहरूको व्यवस्थित उपयोग गरी तथ्याङ्क उत्पादन गरिएको,
- तथ्याङ्कीय कार्यको केन्द्रीय निकायको रूपमा केन्द्रीय तथ्याङ्क विभाग र ३३ वटा जिल्लामा तथ्याङ्क कार्यालयहरू रहेको,
- स्थानीय तहमा तथ्याङ्क गतिविधि सञ्चालनमा रहेको।

कमजोर पक्ष

- समयसापेक्ष बढ्दो तथ्याङ्कको माग सम्बोधन गर्न आवश्यक साधन र श्रोतको कमी छ,
- तथ्याङ्क व्यवस्थापनमा आधुनिक प्रविधिको अधिकतम उपयोग गर्न सकिएको छैन,
- केन्द्रीय तथ्याङ्क विभागको संगठन संरचना (एउटा विभाग तथा ३३ जिल्ला तथ्याङ्क कार्यालय) ले बढ्दो तथ्याङ्कको माग र आपूर्तिबिच सन्तुलन कायम गर्न कठिन छ,
- आधिकारिक तथ्याङ्कसम्बन्धी प्राविधिक तालिम दिने संस्थाको व्यवस्था छैन,
- नीति निर्माणदेखि योजना र्जुमा र कार्यान्वयनको चरणसम्म तथ्याङ्कीय साक्षरताको कमी छ,
- तथ्याङ्कीय क्रियाकलापमा संलग्न सरोकारवालाहरूले प्रयोग गर्ने अवधारणा तथा परिभाषामा एकरूपता कायम हुन सकेको छैन,
- कतिपय तथ्याङ्क प्रदायकहरू तथ्याङ्क संकलन कार्यमा कर्तव्यविमुख छन्,
- तथ्याङ्क उत्पादनमा दोहोरोपनाका कारण कतिपय सर्वेक्षणहरूले उत्तरदातालाई थप बोझ सृजना गरेका छन्,

- तथ्याङ्क प्रणालीका प्रमुख अङ्गहरू बिचको समन्वय कमजोर छ,
- सबै क्षेत्रमा खुला तथ्याङ्कको अवधारणा लागु गर्न सकिएको छैन,
- कतिपय तथ्याङ्क प्रयोगकर्तामैत्री रूपमा उपलब्ध छैनन्,
- स्थानीय तहमा तथ्याङ्क जनशक्तिको अभाव छ,
- तथ्याङ्क सेवामा कार्यरत कर्मचारीको वृत्ति विकास अत्यन्तै न्यून छ ।

अवसर

- नेपालको संविधानले संघ, प्रदेश र स्थानीय तह गरी तिनै तहका सरकारको तथ्याङ्कीय अधिकार सूची प्रदान गरेको,
- कानुनी प्रावधानबमोजिम तथ्याङ्क ऐन लगायत अन्य ऐनहरूले तथ्याङ्कीय क्रियाकलापका बारेमा मार्गदर्शन गरेको,
- नेपाल सरकारले तथ्याङ्क विकासका लागि राष्ट्रिय रणनीति (२०७५/७६-२०७९/८०) लागु गरेको छ,
- तथ्याङ्क ऐनको नयाँ मस्यौदा तयार भई राष्ट्रिय सभाबाट पारित भएको छ,
- नयाँ संगठन संरचनाका लागि संगठन तथा व्यवस्थापनसम्बन्धी अध्ययन प्रतिवेदन तयार भएको छ,
- तथ्याङ्क क्षेत्रमा अन्तर्राष्ट्रिय संघ संस्थाहरूको चासो र सहयोग छ,
- तथ्याङ्कप्रति सम्बन्धित सरोकारवाला तथा सर्वसाधारणको माग र चासोमा वृद्धि भएको छ,
- बढ्दो प्रविधिको प्रयोगसँगै समकालीन समयमा विकसित “बिग डाटा” प्रयोगको सम्भावना बढेको छ ।

चुनौती

- तथ्याङ्क क्षेत्र अन्य क्षेत्रको तुलनामा कम प्राथमिकतामा परेको छ,
- तथ्याङ्क व्यवस्थापन र समन्वयका लागि कानुनी रूपमै बलियो संयन्त्रको अभाव छ,
- तथ्याङ्क प्रणालीअन्तर्गत ठाडो र तेर्सो समन्वय गर्ने संयन्त्र तथा विधिको अभाव छ,
- आकस्मिक रूपमा आइपर्ने प्राकृतिक प्रकोप तथा महामारीका अवस्थामा तथ्याङ्कीय गतिविधि सञ्चालन गर्ने संरचना तथा विधिको अभाव,
- परम्परागत गणना तथा सर्वेक्षणको प्रयोगबाट मात्रै तथ्याङ्क उत्पादनको लागत कम गर्नु तथा उत्तरदाताको बोझलाई कम गर्नु चुनौतीपूर्ण छ ।

नोट: माथि उल्लेख गरिएबाहेक नेपालको राष्ट्रिय तथ्याङ्क प्रणालीका सबल पक्ष, कमजोर पक्ष, समस्या तथा चुनौती के के हुन सक्लान् ? छलफल गर्नुहोस् ।

१.६ अभ्यासका लागि प्रश्नहरू

- क) नेपालको पहिलो जनगणना कहिले सञ्चालन भएको थियो ?
- ख) राष्ट्रिय तथ्याङ्क प्रणाली भन्नाले के बुझिन्छ ? यसका प्रमुख अङ्गहरू के के हुन् ?
- ग) नेपालको राष्ट्रिय तथ्याङ्क प्रणाली केन्द्रीकृत वा विकेन्द्रित के हो ? तार्किक व्याख्या गर्नुहोस् ।
- घ) नेपालको राष्ट्रिय तथ्याङ्क प्रणालीका प्रमुख समस्याहरू उल्लेख गर्दै समाधानका उपायहरूका बारेमा चर्चा गर्नुहोस् ।
- ङ) केन्द्रीय तथ्याङ्क विभागका कामहरू के के हुन् ? चर्चा गर्नुहोस् ।

१.७ सन्दर्भ सामग्री

नेपालको राष्ट्रिय तथ्याङ्क प्रणाली विस्तृत रूपमा अध्ययन गर्नका लागि देहायबमोजिमका सामग्री उपयोगी छन् ।

१. केन्द्रीय तथ्याङ्क विभाग । *केन्द्रीय तथ्याङ्क विभागको कार्यविधि* । काठमाण्डौं, नेपाल: केन्द्रीय तथ्याङ्क विभाग ।
२. केन्द्रीय तथ्याङ्क विभाग (२०७५) । *तथ्याङ्क प्रणाली विकासको लागि राष्ट्रिय रणनीति* । काठमाण्डौं, नेपाल: केन्द्रीय तथ्याङ्क विभाग ।
३. वास्तोला, तुङ्ग शिरोमणि (२०६६) । नेपालको तथ्याङ्क प्रणाली । केन्द्रीय तथ्याङ्क विभाग, *नेपालमा आधिकारिक तथ्याङ्क प्रणाली तथा विधि* (पृष्ठ १-१०) । काठमाण्डौं, नेपाल: केन्द्रीय तथ्याङ्क विभाग ।
४. सुवाल, रूद्र (२०१७) । आधिकारिक तथ्याङ्कको परिचय र यसको प्रणाली । केन्द्रीय तथ्याङ्क विभाग, *नेपालको राष्ट्रिय तथ्याङ्क प्रणाली, एक संग्रह (अंग्रेजी संस्करण)* (पृष्ठ १०-२०) । काठमाण्डौं, नेपाल: केन्द्रीय तथ्याङ्क विभाग ।
५. श्रेष्ठ, नेबिन लाल (२०१७) । नेपालको राष्ट्रिय तथ्याङ्क प्रणाली । केन्द्रीय तथ्याङ्क विभाग, *नेपालको राष्ट्रिय तथ्याङ्क प्रणाली, एक संग्रह (अंग्रेजी संस्करण)* (पृष्ठ १-९) । काठमाण्डौं, नेपाल: केन्द्रीय तथ्याङ्क विभाग ।

पाठ २

तथ्याङ्कसम्बन्धी नीति तथा संस्थागत व्यवस्था

२.१ परिचय

राष्ट्रिय तथ्याङ्क प्रणालीलाई व्यवस्थित रूपमा सञ्चालन एवं विकास गर्न यससँग सम्बन्धित ऐन तथा नियमहरूको आवश्यकता हुन्छ । सरकारलाई तथ्याङ्क गतिविधि सञ्चालन गर्नका लागि कानुनी प्रावधानको आवश्यकता पर्दछ । कानुन कार्यान्वयनका लागि थप नीति तथा संस्थागत संरचनाहरूको निकै महत्त्वपूर्ण भूमिका रहेको हुन्छ । कानुनी व्यवस्थाका सन्दर्भमा नेपालमा सुरुका चारवटा जनगणना गर्दासम्म जनगणना गर्ने कानुनी अधिकार र आधारका लागि पटके इस्तहार जारी गरेको पाइन्छ । वि. सं. १९७६, १९९८ तथा २००७ का इस्तहारहरू यसका उदाहरण हुन् । पाँचौँ जनगणना सञ्चालन गर्नका लागि भने जनगणना ऐन, २००९ जारी भएको पाइन्छ । वि.सं. २०१३ सालमा जनगणना ऐन, २०१३ जारी भएको र वि.सं. २०१५ मा तथ्याङ्क ऐन जारी भएपछि उक्त जनगणना ऐन खारेज भएको पाइन्छ । हाल नेपालमा तथ्याङ्कीय क्रियाकलाप सञ्चालन गर्नका लागि तथ्याङ्क ऐनको अतिरिक्त, नेपाल राष्ट्र बैंक ऐन, २०५८, स्थानीय सरकार सञ्चालन ऐन, २०७४ लगायतका विषयगत ऐन तथा नियमहरू कार्यान्वयनमा छन् । यस पाठको अर्को खण्डमा हामी नेपालको संविधान, तथ्याङ्क ऐन २०१५, तथ्याङ्क नियम २०४१, नेपाल राष्ट्र बैंक ऐन २०५८, स्थानीय सरकार सञ्चालन ऐन २०७४, तथ्याङ्क विकासका लागि राष्ट्रिय रणनीति २०७५ मा भएका प्रमुख विशेषताहरूका बारेमा बेग्लाबेग्लै चर्चा गर्नेछौं ।

२.२ तथ्याङ्कीय नीतिहरू

तथ्याङ्क व्यवस्थापनका लागि विद्यमान संविधान, ऐन, नियम, नीति तथा रणनीतिहरूले के कस्ता व्यवस्था गरेका छन् भनी चर्चा गर्नु यस खण्डको प्रमुख उद्देश्य रहेको छ । विद्यमान कानुनी दस्तावेजहरूले नेपालको तथ्याङ्क प्रणालीलाई कसरी निर्देशित गरेको छ भनेर लेखाजोखा गर्नु समग्र तथ्याङ्क प्रणाली विकासका लागि अत्यन्तै महत्त्वपूर्ण विषय हो । यद्यपि सबै ऐन, नियम, नीति तथा रणनीतिहरूको समीक्षा नगरी यहाँ तथ्याङ्कसँग प्रत्यक्ष सरोकार राख्ने दस्तावेजहरूको मात्र चर्चा गरिएको छ ।

२.२.१ नेपालको संविधान

वि.सं. २०७२ मा जारी गरिएको नेपालको संविधानले नेपालको राष्ट्रिय तथ्याङ्क प्रणालीलाई स्पष्ट रूपमा विकेन्द्रित तथ्याङ्क प्रणालीको रूपमा स्थापित गरेको छ । यसका अतिरिक्त संविधानका विभिन्न धाराहरूमा उल्लेख भएका तथ्याङ्क सम्बन्धित विषयहरू देहायबमोजिम रहेका छन् ।

- धारा २८ मा तथ्याङ्कको गोपनीयतासम्बन्धी विषय उल्लेख छ ।
- राज्यका नीतिहरू (धारा ५१) को (च) को विकाससम्बन्धी नीतिको बुँदा नं. ८ मा जनसांख्यिक तथ्याङ्कलाई अद्यावधिक गर्ने विषय उल्लेख छ ।
- प्रतिनिधि सभाको गठन (धारा ८४), प्रदेश सभाको गठन (धारा १७६) मा जनसंख्यालाई निर्वाचन क्षेत्र निर्धारणको आधार मानिएको छ ।
- राष्ट्रिय समावेशी आयोगको काम, कर्तव्य र अधिकार (धारा २५९) को १ (छ) मा राष्ट्रिय जनगणना उल्लेख गरिएको छ ।
- विशेष अधिकारको समीक्षा तथा पुनरावलोकन (धारा २८१) मा राष्ट्रिय जनगणना उल्लेख गरिएको छ ।
- निर्वाचन क्षेत्र निर्धारण आयोग (धारा २८६) निर्वाचन क्षेत्र निर्धारण गर्दा जनसंख्यालाई आधार मान्ने विषय उल्लेख छ ।
- अनुसूची ८ मा स्थानीय तहको अधिकार सूचीअन्तर्गत स्थानीय तथ्याङ्क र अभिलेख संकलन, बेरोजगार तथ्याङ्क संकलनसम्बन्धी विषय उल्लेख छ ।
- अनुसूची ९ मा संघ, प्रदेश र स्थानीय तहको अधिकारको साभा सूचीअन्तर्गत व्यक्तिगत घटना, जन्म, मृत्यु, विवाह र तथ्याङ्कसम्बन्धी विषय उल्लेख छ ।

नोट: संविधानको अनुसूची ५, ६, ८ र ९ मा उल्लेख भएका अधिकार सूचीहरूलाई नेपाल सरकार (मन्त्रपरिषद्) को मिति २०७३ माघ १८ को निर्णयबमोजिम विस्तृतीकरण गरी कार्यान्वयनमा आएको छ ।⁵

२.२.२ तथ्याङ्क ऐन, २०१५

तथ्याङ्कसम्बन्धी सबै क्रियाकलापहरूलाई व्यवस्थित रूपमा सञ्चालन गर्न वि.सं. २०१५ मंसिर १० गते तथ्याङ्क ऐन जारी भई हालसम्म पनि पटक पटक संशोधन हुँदै कार्यान्वयनमा रहेको छ । यस ऐनका प्रमुख विशेषताहरू देहायबमोजिम रहेका छन् ।

- केन्द्रीय तथ्याङ्क विभागको गठन यसै ऐनको व्यवस्थाअन्तर्गत भएको हो ।
- ऐनले सरकारलाई सूचित आदेश प्रकाशित गरी जुनसुकै प्रकारको तथ्याङ्क संकलन गर्ने अधिकार दिएको छ ।
- ऐनले सरकारलाई देशका जुनसुकै वर्ग वा समुदायका व्यक्तिसँग तथ्याङ्क माग गर्न सक्ने अधिकार दिएको छ ।
- ऐनले उत्तरदातासँग लिइएका विवरण वा तथ्याङ्कहरूको गोपनीयता कायम राख्नुपर्ने व्यवस्था गरेको छ ।

5 Office of the Prime Minister and Council of Ministers. Retrieved on November 14, 2021 from: <https://www.opmcm.gov.np/federalism-admin/>

- ऐनले अन्य निकायले तथ्याङ्क संकलन र प्रकाशन गर्न केन्द्रीय तथ्याङ्क विभागको स्वीकृति लिनुपर्ने व्यवस्था गरेको छ ।
- ऐनले कानुन विपरीत गोपनीयता भङ्ग गर्ने, उपलब्ध गराउनु पर्ने सूचना उपलब्ध नगराउने वा अन्य यो ऐन विपरीत कार्य गर्ने जो कोहीलाई कानुनी कारवाहीको व्यवस्था गरेको छ ।

सरकारले तथ्याङ्क ऐन, २०१५ मा भएका व्यवस्थाहरूलाई कार्यान्वयन गर्न सहज बनाउने उद्देश्यले वि.सं. २०४१ मा तथ्याङ्क नियम जारी गरी कार्यान्वयनमा ल्याएको छ ।^६

तथ्याङ्क ऐन तथा नियमका अतिरिक्त अन्य ऐन तथा नियमहरू पनि नेपालको राष्ट्रिय तथ्याङ्क प्रणालीसँग सम्बन्धित रहेका छन् । उदाहरणका लागि नेपाल राष्ट्र बैंक ऐन, २०५८; स्थानीय सरकार सञ्चालन ऐन, २०७४; जन्म, मृत्यु, तथा अन्य व्यक्तिगत घटना (दर्ता गर्ने) ऐन, २०३३; जन्म, मृत्यु, तथा अन्य व्यक्तिगत घटना (दर्ता गर्ने) नियमावली, २०३४; राष्ट्रिय सूचना प्रविधि विकास समिति (गठन) आदेश, २०५८; मतदाता नामावली सम्बन्धी ऐन, २०७३ आदि रहेका छन् । यी ऐन कानुनहरूमा समावेश भएका कतिपय प्रावधानहरू एक आपसमा सामञ्जस्यता कायम गर्न सकेका छैनन् । यसका अतिरिक्त विद्यमान तथ्याङ्क ऐनले समेत समकालीन तथ्याङ्कीय गतिविधिहरूलाई पूर्ण रूपमा मार्गनिर्देशन गर्न सक्षम नभएका कारण नयाँ कानुन तर्जुमा गर्ने उद्देश्यले तथ्याङ्क विधेयक, २०७६ राष्ट्रिय सभाबाट पारित भई प्रतिनिधि सभामा विचाराधीन अवस्थामा रहेको छ । ऐन नियमका अतिरिक्त हाल तथ्याङ्कसँग प्रत्यक्ष सम्बन्धित नीतिको रूपमा तथ्याङ्क विकासका लागि राष्ट्रिय रणनीति, (२०७५/७६-२०७९/८०) क्रियाशील रहेको छ ।

२.३ आवधिक योजनाहरूमा तथ्याङ्क^७

नेपालमा विकास प्रशासनको अवधारणा सुरु भएसँगै योजनावद्ध विकासको थालनी भएको विश्वास गरिन्छ । यसैअनुरूप वि.सं. २०१३ मा पहिलो पटक पञ्चवर्षीय योजना (२०१३-२०१८) तर्जुमा गरिएको थियो । नेपालमा हालसम्म पन्द्रहटा आवधिक योजनाहरू तर्जुमा गरिसकिएको छ र ती योजनाहरूमा तथ्याङ्कसम्बन्धी विषयलाई के कसरी राखिएको छ भनी संक्षेपमा चर्चा गर्नु यस खण्डको उद्देश्य हो ।

प्रथम पञ्चवर्षीय योजना (२०१३-२०१८): देशमा पहिलो पटक योजना तर्जुमा गर्दा

6 Office of the Prime Minister and Council of Ministers. Retrieved on November 14, 2021 from: <https://www.opmcm.gov.np/federalism-admin/>

7 National Planning Commission. *Periodic Plans of Nepal*. Retrieved on November 15, 2021 from: https://npc.gov.np/np/category/periodic_plans;page=2

तत्कालीन अवस्थामा तथ्याङ्कको निकै अभाव महसुस गरिएको देखिन्छ । सोही कारणले यस योजनामा तथ्याङ्कको क्षेत्र विस्तार गर्ने, केन्द्रीय तथ्याङ्क विभाग र यसका शाखा कार्यालयहरूको स्थापना गर्ने विषयहरू उल्लेख छ । यसका अतिरिक्त, यस योजनाले व्यक्तिगत घटनासम्बन्धी तथ्याङ्क, स्वास्थ्य तथ्याङ्क, कृषि तथ्याङ्क, मूल्य तथ्याङ्क, व्यापार तथ्याङ्क (आयात, निर्यात), जलवायु तथा मौसमसम्बन्धी तथ्याङ्कलगायत विभिन्न क्षेत्रका तथ्याङ्कको आवश्यकता अनुभूत गरी आगामी दिनहरूमा त्यस्ता तथ्याङ्क आपूर्ति गर्दै जाने लक्ष्य राखेको पाइन्छ ।

दोस्रो त्रिवर्षीय योजना (२०१९-२०२२): यस योजनाको प्राथमिकतामा नै विभिन्न क्षेत्रका तथ्याङ्क प्राप्त गर्ने कुरा उल्लेख गरिएको छ । कुनै योजनामा हात हाल्नु अगाडि तथ्याङ्क प्राप्त गर्न सर्वेक्षण गर्ने भन्ने विषय प्राथमिकताका साथ यस योजनामा उल्लेख गरिएको छ । यस योजनाले विभिन्न क्षेत्रका तथ्याङ्क अभावका कारण ती क्षेत्रका बारेमा चर्चा गर्न सम्भव नभएको भन्ने विषय उल्लेख गरेको छ । यस योजनामा तथ्याङ्कसम्बन्धी विषयमा उल्लेख भएको एउटा महत्त्वपूर्ण वाक्य जस्ताको तस्तै यसप्रकार रहेको छ:

“जनगणना, कृषिगणना र साधारण तथ्याङ्क संकलन, राष्ट्रिय उत्पादन आर्थिक साधन, जनशक्ति, प्राकृतिक साधन, यातायात शिक्षा, स्वास्थ्य, बन्दव्यापार, उद्योगधन्दा रोजगार आदिको सम्बन्धमा तथ्याङ्क संकलन गर्न र कृषिगणना गर्नेतर्फ प्रयास जारी नै राखिएको छ ।”

यसका अतिरिक्त, यस योजनामा सर्वे र तथ्याङ्क भन्ने छुट्टै शीर्षक सहितको योजनाको एउटा खण्डमा विभिन्न क्षेत्रहरू जस्तै: वन, खनिज पदार्थ, माटो, औद्योगिक कच्चा वस्तु तथा जलशक्ति आदिको योजना बनाउन आवश्यक तथ्याङ्क उपलब्ध गराउन साधारण सर्वेक्षण गर्ने भन्ने विषय उल्लेख छ । साथै, आर्थिक सामाजिक क्षेत्रका तथ्याङ्क, पूर्वाधारसम्बन्धी तथ्याङ्क उपलब्ध गराउने गरी सर्वेहरू गर्ने र जनगणनाको तथ्याङ्क प्रशोधन गरी प्रतिवेदन तयार गर्ने विषय यस योजनामा उल्लेख भएको पाइन्छ ।

तेस्रो पञ्चवर्षीय योजना (२०२२-२०२७): तथ्याङ्कसम्बन्धी एउटा छुट्टै पाठसहित यस योजनामा तथ्याङ्क विकासका लागि विभिन्न लक्ष्यहरू समावेश गरिएको थियो । यस योजनाले मूलतः औद्योगिक सर्वेक्षण गर्ने, आन्तरिक व्यापार तथ्याङ्क तथा व्यक्तिगत घटनासम्बन्धी तथ्याङ्क उत्पादनका लागि कार्य गर्ने र केन्द्रीय तथ्याङ्क विभागको प्रशासनिक सुधारका कार्यहरू गर्ने भन्ने विषयहरू प्राथमिकताका साथ उल्लेख गरेको पाइन्छ ।

चौथो पञ्चवर्षीय योजना (२०२७-२०३२): यस योजनाले समन्वय, मानकीकरण र तथ्याङ्कको क्षेत्रमा यान्त्रीकरण गर्नुपर्ने आवश्यकता अनुभूत गरेको पाइन्छ । यो योजनाले तथ्याङ्कको

दोहोरोपनाबाट प्रयोगकर्तामा हुने भ्रमलाई नियन्त्रण गर्न केन्द्रीय तथ्याङ्क विभागलाई समन्वयकारी भूमिका निर्वाह गर्ने जिम्मेवारी प्रदान गरेको देखिन्छ । योजनाको प्रस्तावित कार्यक्रममा जनगणना गर्ने र विगतको भन्दा बढी सूचना संकलन गर्ने तथा अन्तर्राष्ट्रियस्तरमा तुलनायोग्य बनाउने लक्ष्य लिइएको पाइन्छ । यसका अतिरिक्त, कृषिगणना, जनसंख्या सर्वेक्षण, पशु धन सर्वेक्षण, कृषि लगानी सर्वेक्षण, आन्तरिक व्यापार सर्वेक्षण, औद्योगिक गणना, माछा खेती सर्वेक्षण, पारिवारिक सर्वेक्षण, राष्ट्रिय उत्पादनको वार्षिक अनुमान जस्ता कार्यहरू गर्ने लक्ष्य राखिएको पाइन्छ । तथ्याङ्क व्यवस्थापनलाई छरितो र भरपर्दो बनाउन केन्द्रीय तथ्याङ्क विभागमा आधुनिक उपकरण जडान गर्ने विषयसमेत उल्लेख भएको छ ।

पाँचौँ पञ्चवर्षीय योजना (२०३२-२०३७): यस योजनामा तथ्याङ्क विकासका लागि निम्नानुसारका कार्यक्रमहरू तय गरिएको थियो ।

- जनसंख्या सर्वेक्षण गर्ने,
- प्रयोगात्मक रूपमा जन्म, मृत्यु दर्ता गर्ने,
- आगामी जनगणनाका लागि अध्ययन तथा परीक्षण गर्ने,
- तेस्रो औद्योगिक गणना गर्ने,
- वि.सं. २०२८ को कृषिगणनाको प्रतिवेदन निकाल्ने तथा आगामी कृषिगणनाका लागि प्रारम्भिक कार्य गर्ने,
- विभागलाई तथ्याङ्क व्यवस्थापनमा सहयोग गर्न यान्त्रिकीकरण गर्ने र कम्प्युटरसम्बन्धी दक्ष जनशक्ति उत्पादन गर्ने,
- व्यापार सर्वेक्षण गर्ने,
- रोजगारीसम्बन्धी सर्वेक्षण गर्ने,
- राष्ट्रिय आय, बचत र लगानीसम्बन्धी अध्ययन गर्ने,
- राष्ट्रिय आय अनुमानका लागि थप अध्ययन सर्वेक्षण गर्ने, र
- घरेलु तथा साना उद्योगसम्बन्धी सर्वेक्षण गर्ने ।

छैटौँ पञ्चवर्षीय योजना (२०३७-२०४२): यस योजनाले पनि तथ्याङ्क विकासका लागि महत्त्वपूर्ण नीति अवलम्बन गरेको पाइन्छ । खासगरी यस योजनामा तथ्याङ्कीय क्रियाकलापलाई निरन्तरता दिने; तथ्याङ्कको असंगति र दोहोरोपना हटाउन तथ्याङ्क ऐनबमोजिम केन्द्रीय तथ्याङ्क विभागको भूमिकालाई थप प्रभावकारी बनाउने; जनगणना, कृषिगणना, औद्योगिक गणनाजस्ता महत्त्वपूर्ण कार्यक्रमलाई प्राथमिकतामा राख्ने र अन्य तथ्याङ्कसम्बन्धी क्रियाकलापलाई पनि निरन्तर सञ्चालन गर्ने; तथ्याङ्क व्यवस्थापनमा कम्प्युटरको क्षमता अभिवृद्धि गर्ने; राष्ट्रिय कम्प्युटरको

प्रयोगलाई सरकारी क्षेत्रमा मात्र सीमित नराखी गैरसरकारी क्षेत्रको उपभोगको लागि समेत उपलब्ध गराउने नीति अनुकूल आवश्यकता अनुसार तथ्याङ्क कोष, मोडेलिङ, समीक्षा, टेप प्रशोधन, कम्प्युटरसम्बन्धी तालिम कार्यक्रमहरू पनि सञ्चालन गराउँदै लैजाने जस्ता नीतिहरू अवलम्बन गरेको पाइन्छ । उल्लेखित नीति बमोजिम योजनाले तथ्याङ्कसम्बन्धी विविध कार्यक्रम तथा बजेट तर्जुमा गरेको थियो । जसअनुसार प्राथमिकता प्राप्त प्रमुख तथ्याङ्कीय कार्यक्रमहरूमा राष्ट्रिय जनगणना २०३८, राष्ट्रिय कृषिगणना २०३८, औद्योगिक गणना २०३९, राष्ट्रिय लेखा र डिप्लेटर विकास, कृषि र औद्योगिक उत्पादन सूचकाङ्क निर्माण, अन्तरउद्योग विश्लेषण, घरेलु तथा साना उद्योग सर्वेक्षण, कम्प्युटर सुविधा र क्षमता अभिवृद्धिलगायत अन्य कार्यक्रमहरू रहेका थिए ।

सार्तौँ पञ्चवर्षीय योजना (२०४२-२०४७): यस योजनामा तथ्याङ्कसम्बन्धी तय गरिएका नीति तथा कार्यक्रमहरू यसप्रकार रहेका छन् ।

नीति

- प्राविधिक सहयोग र समन्वयका लागि केन्द्रीय तथ्याङ्क विभागलाई सक्रिय र प्रभावकारी बनाउने,
- राष्ट्रिय महत्त्वका तथ्याङ्कहरूलाई प्राथमिकता दिई अन्य महत्त्वपूर्ण क्षेत्रमा पनि तथ्याङ्क सङ्कलन गराई प्रशोधन र विश्लेषण कार्यमा असंगति, दोहरोपन हटाई शुद्ध र विश्वसनीय तथ्याङ्क उपलब्ध गराउने,
- कम्प्युटरको क्षमता अभिवृद्धि गर्ने ।

कार्यक्रम

- औद्योगिक गणना, २०४४ सञ्चालन गरी प्रतिवेदन प्रकाशन गर्ने,
- घरेलु तथा साना उद्योग सर्वेक्षण सञ्चालन गरी प्रतिवेदन प्रकाशन गर्ने,
- सामाजिक तथ्याङ्कको विकास गर्ने,
- तथ्याङ्क समन्वय, सुदृढीकरण तथ्याङ्क विधि विकास एवं प्रमापन गर्ने,
- तथ्याङ्क प्रकाशन तथा पुस्तकालयको विकास गर्ने,
- कृषि तथा औद्योगिक उत्पादन सूचकाङ्क निर्माण गर्ने,
- राष्ट्रिय लेखा विकासका लागि विभिन्न कार्यक्रम गर्ने,
- कृषिगणनाका प्रतिवेदन प्रकाशन गर्ने,
- जनसंख्या सर्वेक्षण तथा जनसंख्या तथ्याङ्क विश्लेषण गर्ने,
- राष्ट्रिय जनगणना, २०४८ को तयारीका कार्य गर्ने ।

आठौँ पञ्चवर्षीय योजना (२०४९-२०५४): आठौँ पञ्चवर्षीय योजनाले पनि तथ्याङ्क विकासका लागि निकै महत्त्वपूर्ण नीति तथा कार्यक्रमहरू तय गरेको पाइन्छ । यस योजनाले अवलम्बन गरेका तथ्याङ्कसम्बन्धी प्रमुख नीतिहरू यसप्रकार रहेका छन् ।

- राष्ट्रिय तथ्याङ्क परिषद्लाई प्रभावकारी तुल्याउने,
- केन्द्रमा एक वैज्ञानिक तथा प्रभावकारी तथ्याङ्क भण्डारको व्यवस्था गर्ने,
- तथ्याङ्क सेवाको विकास गर्ने,
- केन्द्रीय तथ्याङ्क विभागको सङ्गठनात्मक स्वरूपमा सुधार ल्याउने,
- तथ्याङ्कसँग सम्बन्धित विभिन्न निकायहरू बिच समन्वय कायम राख्न आर्थिक र सामाजिक तथ्याङ्कमा संलग्न प्रमुख निकायहरूलाई सम्पर्क निकायको रूपमा विकसित गराउने ।

त्यसैगरी यस योजनाले तथ्याङ्कसम्बन्धी तय गरेका कार्यक्रमहरू यसप्रकार छन् ।

- तथ्याङ्क प्रकाशन र पुस्तकालय विकास,
- सामाजिक तथ्याङ्क विकास,
- आर्थिक सूचकाङ्क निर्माण,
- राष्ट्रिय लेखा विकास,
- राष्ट्रिय जनगणना २०४८, जनसंख्या विश्लेषण र डेमोग्राफिक स्याम्पल सर्वे,
- राष्ट्रिय कृषिगणना,
- औद्योगिक गणना, दूला तथा साना उद्योग प्रतिष्ठान सर्वेक्षण,
- पारिवारिक सर्वेक्षण विकास,
- पञ्जीकरण कार्यक्रम ।

नवौँ पञ्चवर्षीय योजना (२०५४-२०५९): यस योजनामा प्रत्येक विषयगत क्षेत्रहरूसँगै सम्बन्धित तथ्याङ्कको बारेमा नीति तथा कार्यक्रमहरू तर्जुमा गरिएको थियो । अघिल्ला योजनामा उल्लेख भएका तथ्याङ्क कार्यक्रम प्रगतिको मूल्याङ्कन गर्दै भरपर्दो र गुणस्तरीय तथ्याङ्क उत्पादन तथा प्रयोग सम्बन्धमा विभिन्न विषयसँग सम्बन्धित भएर यस योजनामा तथ्याङ्कका कार्यक्रमहरू तयार गरिएको थियो ।

दशौँ पञ्चवर्षीय योजना (२०५९-२०६४): यस योजनाले केन्द्रीय तथ्याङ्क विभागको संस्थागत विकासमा विशेष जोड दिएको थियो । यसले उपयोगी, समयसापेक्ष, भरपर्दो र योजना प्रक्रियाको लागि आवश्यक उच्च गुणस्तरको तथ्याङ्क उत्पादनको आवश्यकतालाई जोड दिएको पाइन्छ । साथै, यसले एकीकृत राष्ट्रिय तथ्याङ्क योजना (CNSP) कार्यान्वयन गर्न र नेपाल आर्थिक योजना तथा तथ्याङ्क समूह अन्तर्गतको आर्थिक योजना समूहको

सक्रिय बनाउन सिफारिस गरेको थियो ।

एघारौँ त्रिवर्षीय योजना (२०६४/६५-२०६६/२०६७): यो योजनाले नेपालको राष्ट्रिय तथ्याङ्क प्रणालीलाई सुदृढ गर्ने विषय उल्लेख गरेको पाइन्छ । यस योजनाले राष्ट्रिय लेखा प्रणाली सुदृढीकरण गर्नका लागि कार्यक्रम तर्जुमा गर्नुका साथै समावेशी तथ्याङ्क उत्पादनसम्बन्धी नीति अवलम्बन गरेको देखिन्छ ।

बाह्रौँ त्रिवर्षीय योजना (२०६७/६८-२०६९/७०): यस योजनाले देशको नयाँ संरचनाअनुसार भरपर्दो राष्ट्रिय तथ्याङ्क प्रणाली निर्माणमा जोड दिएको थियो । योजना र तथ्याङ्कसम्बन्धी कार्यमा संलग्न निकाय र एकाइहरूको संस्थागत सुदृढीकरण एवं क्षमता अभिवृद्धि गर्ने रणनीति लिइएको थियो । राष्ट्रिय योजना आयोगमा भू-सूचना प्रणाली (Geographic Information System – GIS) मा आधारित तथ्याङ्क बैंक स्थापना गर्ने योजना पनि यसै त्रिवर्षीय योजनाले तर्जुमा गरेको थियो । तथ्याङ्क अभिलेख संग्रहको विकास गर्नुका साथै वडा तहसम्मको तथ्याङ्क उत्पादन गर्ने कार्यनीति यस योजनाले लिएको थियो । यस योजनाले तथ्याङ्क संकलन, प्रशोधन तथा विश्लेषणका लागि आधुनिक प्रविधिको प्रयोगमा जोड दिएको थियो ।

तेह्रौँ त्रिवर्षीय योजना (२०७०/७१-२०७२/७३): तथ्याङ्क विकासमा यो योजना निकै महत्त्वपूर्ण योजनाको रूपमा लिन सकिन्छ । यस योजनाले खासगरी तथ्याङ्क प्रणाली विकासको लागि राष्ट्रिय रणनीति तयार गर्ने, सर्वेक्षण अनुमति प्रणाली (Survey Clearance System) कार्यान्वयन गर्ने, योजना तथा अनुगमनका लागि विश्वसनीय र अन्तर्राष्ट्रियस्तरमा तुलनायोग्य तथ्याङ्क उत्पादन गर्ने र तथ्याङ्क विकासका लागि द्विपक्षीय र बहुपक्षीय समन्वय सुदृढीकरण गर्ने जस्ता महत्त्वपूर्ण तथ्याङ्कसम्बन्धी योजना बनाएको थियो । यस योजनाको प्रमुख कार्यक्रमहरू निम्नानुसार रहेका छन् ।

- खुला तथ्याङ्क (Open Data) प्रणाली कार्यान्वयनमा ल्याइने,
- राष्ट्रिय तथ्याङ्क विकास रणनीति तयार गरी कार्यान्वयनमा ल्याउने,
- त्रैमासिक राष्ट्रिय लेखा तथ्याङ्कको संस्थागत विकास गर्ने,
- विभिन्न सर्वेक्षणहरूका लागि आवश्यक पर्ने Master Sample Frame तयार गर्ने,
- तथ्याङ्कको सेवाकालीन, अन्य राष्ट्रिय तथा अन्तर्राष्ट्रियस्तरका तालिम तथा तथ्याङ्कीय अनुसन्धानका लागि राष्ट्रिय तथ्याङ्क कार्यालयको मातहतमा रहने गरी राष्ट्रिय तथ्याङ्क अनुसन्धान तथा तालिम केन्द्रको स्थापना गर्ने,
- अन्तर्राष्ट्रिय औद्योगिक मापन वर्गीकरणको मुख्य विभाजनअन्तर्गत पर्ने सम्पूर्ण क्रियाकलापहरूलाई समेट्ने गरी प्रमुख राष्ट्रिय आर्थिक तथा सामाजिक सूचकहरू तयार गर्ने,

- सम्भाव्यता तथा महत्त्वका आधारमा पर्यटन, स्वास्थ्य, महिलाको राष्ट्रिय अर्थतन्त्रमा योगदान, अनौपचारिक आर्थिक क्रियाकलाप आदि क्षेत्रहरूको विस्तारित (स्याटलाइट) लेखा तयार गर्ने,
- केन्द्रीय तथ्याङ्क विभागको पुनःसंरचना तथा तथ्याङ्क कार्यालयको विस्तार र सुदृढीकरण गर्ने,
- तथ्याङ्कीय क्रियाकलाप गरिरहेका अन्य सरकारी संगठनहरूको तथ्याङ्क एकाइहरूको संस्थागत तथा संगठनात्मक विकास तथा पुनःसंरचना गर्ने,
- केन्द्रीय तथ्याङ्क विभागमा भौगोलिक सूचना प्रणालीलाई संस्थागत गर्ने,
- श्रमशक्ति सर्वेक्षण सञ्चालन गर्ने,
- गरिबी मापनका लागि वार्षिक घरपरिवार सर्वेक्षणलाई निरन्तरता दिने,
- केन्द्रीय तथ्याङ्क विभागबाट भएका गणना तथा सर्वेक्षणसँग सम्बन्धित तथ्याङ्क र तिनको प्रक्रियागत अभिलेखन गर्ने कार्यक्रम (National Data Archive – NADA) लाई आधिकारिक तथ्याङ्कका अन्य निकायहरूमा पनि क्रमशः विस्तार गर्ने ।

चौधौँ त्रिवर्षीय योजना (२०७३/७४-२०७५/७६): नेपालको संविधानले मार्गदर्शन गरे बमोजिमको तथ्याङ्क प्रणाली विकास गर्न यो योजनाले विभिन्न रणनीति, कार्यनीति तथा कार्यक्रम र बजेटको व्यवस्था गरेको पाइन्छ । यस योजनाका तथ्याङ्कसम्बन्धी प्रमुख तिनवटा रणनीतिहरू निम्नानुसार छन् ।

- क) उच्च गुणस्तरीय तथ्याङ्क उत्पादन, सुधार र आपूर्ति गर्ने ।
- ख) तथ्याङ्कका उत्पादनका सबै सम्भाव्य विधि र स्रोतहरूको उपयोग गर्ने ।
- ग) संघीय संरचनाअनुसार तथ्याङ्कको गुणस्तर सुनिश्चित गर्न संस्थागत क्षमता अभिवृद्धि गर्ने ।

उल्लिखित रणनीतिबमोजिम यस योजनाले विभिन्न कार्यनीति र कार्यक्रमहरू तर्जुमा गरेको छ । उपयुक्त कानुनी पूर्वाधार निर्माण गर्ने, तथ्याङ्क प्रणालीको व्यवस्थित रूपले सुदृढीकरण गर्ने, राष्ट्रिय तथा अन्तर्राष्ट्रिय विकास लक्ष्य मापन गर्न देशलाई आवश्यक तथ्याङ्कको उत्पादन गर्ने, केन्द्रीय तथ्याङ्क विभागको संगठन पुनःसंरचना गर्ने, तथ्याङ्कीय अवधारणा तथा परिभाषाहरूलाई एकीकृत गर्ने, राष्ट्रिय तथ्याङ्क प्रशिक्षण केन्द्र स्थापना गर्ने जस्ता कयौँ महत्त्वपूर्ण कार्यनीतिहरू यस योजनाले अवलम्बन गरेको छ । यस योजनाले यिनै रणनीति तथा कार्यनीतिका आधारमा नीतिगत, कानुनी र संस्थागत व्यवस्थासम्बन्धी कार्यक्रम, विभिन्न गणना तथा सर्वेक्षणसम्बन्धी कार्यक्रम, र अन्य कार्यक्रम अन्तर्गत राष्ट्रिय लेखाको आधार वर्ष परिवर्तन, आपूर्ति र उपयोग तालिका (Supply and Use Table – SUT) निर्माण, मूल्य सूचकाङ्क निर्माण, त्रैमासिक राष्ट्रिय लेखा अनुमान,

सरकारका सबै तहमा आवश्यक तथ्याङ्क उत्पादन, प्रशासनिक अभिलेखहरूको सुदृढीकरण गरी एकीकृत तथ्याङ्क व्यवस्थापन, तथ्याङ्कीय साक्षरता कार्यक्रम सञ्चालन जस्ता अत्यन्त महत्त्वपूर्ण कार्यक्रमहरू तय गरेको छ ।

पन्ध्रौँ पञ्चवर्षीय योजना (२०७६/७७-२०८०/८१): सबल, क्रियाशील र समन्वयात्मक राष्ट्रिय तथ्याङ्क प्रणालीको विकास भन्ने सोचका साथ चालु पन्ध्रौँ पञ्चवर्षीय योजनाले तथ्याङ्क प्रणाली सुदृढीकरण गर्नका लागि देहाय बमोजिम लक्ष्य, उद्देश्य तथा रणनीति तर्जुमा गरेको छ ।

लक्ष्य: समन्यायिक विकास र समृद्धिसम्बन्धी नीतिका लागि संघीय शासन प्रणाली सापेक्ष गुणस्तरीय तथ्याङ्क उत्पादन, व्यवस्थापन र आपूर्ति पद्धतिको विकास गर्ने ।

उद्देश्य:

- १) विश्वसनीय, गुणस्तरीय तथा तुलनायोग्य तथ्याङ्क उत्पादन गरी नियमित आपूर्ति गर्नु ।
- २) निकायगत तथा तहगत समन्वय स्थापना गरी संघीय संरचना अनुरूपको तथ्याङ्क प्रणाली विकास गर्नु ।

रणनीति:

- १) संघीय पद्धतिबमोजिम नीतिगत र संस्थागत व्यवस्था गर्ने ।
- २) सूचकहरूको पहिचान र प्राथमिकीकरण गरी प्रचलित विधि र गुणस्तर मापदण्ड अनुरूपको तथ्याङ्क उत्पादन र आपूर्ति गर्ने ।
- ३) तथ्याङ्क उत्पादन र आपूर्तिमा सबै तथ्याङ्कीय स्रोतको उपयोग गर्ने ।
- ४) तथ्याङ्कीय क्रियाकलापमा सूचना प्रविधि लगायतका नवीनतम प्रविधिको प्रयोग वृद्धि गर्ने ।
- ५) तथ्याङ्कसम्बन्धी संगठनात्मक व्यवस्था, क्षमता विकास, समन्वय र सञ्जालीकरण गर्ने ।

उल्लिखित नीतिबमोजिम यस योजनाले योजना अवधिभर ३ वटा राष्ट्रियस्तरका गणना र १६ वटा राष्ट्रियस्तरको सर्वेक्षण सम्पन्न गर्ने लक्ष्य लिएको छ । यी कार्यक्रमहरूका बारेमा योजनाको परिच्छेद १० अन्तर्गत तथ्याङ्क प्रणाली शीर्षकको पाठ खण्डमा विस्तृत रूपमा प्रस्तुत गरिएको छ । जसअनुसार, राष्ट्रिय जनगणना २०७८, राष्ट्रिय कृषिगणना २०७८, नेपाल जीवनस्तर सर्वेक्षण, नेपाल बहुसूचकाङ्क सर्वेक्षण, जनसांख्यिक तथा स्वास्थ्य सर्वेक्षण, पारिवारिक बजेट सर्वेक्षण, समय उपयोग सर्वेक्षण, महिलाविरुद्ध हुने हिंसासम्बन्धी सर्वेक्षण, नेपाल औद्योगिक सर्वेक्षण आदि प्रमुख गणना तथा सर्वेक्षणहरू रहेका छन् । यसका अतिरिक्त योजनाले तथ्याङ्कको स्तरीय मानक र कार्यविधि तयार हुने

अपेक्षा राखेको छ । यसैगरी सर्वेक्षण अनुमति प्रणाली र निर्धारित तथ्याङ्क प्रणाली लागु गर्नेतर्फ योजना अग्रसर देखिन्छ । दिगो विकास लक्ष्यका सूचकहरू मापन गर्न आवश्यक तथ्याङ्कको नियमित उत्पादन र आपूर्ति गर्ने, राष्ट्रिय लेखाको आधार वर्ष परिवर्तन गर्ने, राष्ट्रिय लेखाका अनुमानहरू वार्षिक तथा त्रैमासिक रूपमा नियमित प्रकाशन गर्ने, आपूर्ति उपयोग तालिका नियमित रूपमा तयार गर्ने, मूल्यसम्बन्धित तथ्याङ्क उत्पादनमा नियमितता दिने, राष्ट्रिय तथ्यगत विवरण तयार गर्ने, विभिन्न क्षेत्रमा खण्डीकृत तथ्याङ्क उत्पादन गर्ने जस्ता महत्त्वपूर्ण कार्यक्रमहरूको योजना तर्जुमा भएको छ । यस योजनाको सफल कार्यान्वयनबाट नेपालको तथ्याङ्क प्रणालीको विकासको एउटा महत्त्वपूर्ण फडको मार्ने प्रबल सम्भावना देखिन्छ ।

२.४ तथ्याङ्कसम्बन्धी अन्तर्राष्ट्रिय पहल

तथ्याङ्क विकासमा अन्तर्राष्ट्रिय भूमिका निकै महत्त्वपूर्ण मानिन्छ । खासगरी सन् १९८० पछि विश्वमा आएको परिवर्तनसँगै सरकारी तथ्याङ्कको परम्परागत भूमिकामा समेत परिवर्तनको महसुस गरिन थालियो । त्यसपछि तथ्याङ्कसम्बन्धी साभा मापदण्डहरूको खोजी हुन थाल्यो र तथ्याङ्कीय क्रियाकलापमा अन्तर्राष्ट्रिय सहयोग तथा समन्वयको विस्तार हुन थाल्यो । फलस्वरूप आधिकारिक तथ्याङ्कका दशवटा सिद्धान्तहरू अगाडि सारिए । सन् १९९२ मा युरोपका लागि आर्थिक आयोगले यी सिद्धान्त अपनाउने घोषणा गर्‍यो भनेपछि १९९४ मा संयुक्त राष्ट्र तथ्याङ्क आयोगले यी सिद्धान्तहरूलाई अनुमोदन गर्‍यो । त्यसपछि अन्तर्राष्ट्रियस्तरमा विभिन्न घोषणा, कार्ययोजनाहरू तयार गरी लागु गरियो । यिनै सिद्धान्त, घोषणा तथा कार्ययोजना सम्बन्धमा संक्षेपमा यस खण्डमा चर्चा गरिएको छ ।^८

२.४.१ संयुक्त राष्ट्र संघको आधिकारिक तथ्याङ्कसम्बन्धी आधारभूत सिद्धान्तहरू (Fundamental Principles of Official Statistics)

सन् १९९४ मा विश्वको सर्वोत्कृष्ट तथ्याङ्कीय प्राधिकारी मानिने संयुक्त राष्ट्र तथ्याङ्क आयोगले अनुमोदन गरी आफ्ना सदस्य राष्ट्रहरूलाई यी सिद्धान्त अपनाउन निरन्तर प्रोत्साहन गर्दै आएको छ । उक्त सिद्धान्तहरूलाई केन्द्रीय तथ्याङ्क विभागका तत्कालीन महानिर्देशक तुङ्ग शिरोमणि वास्तोलाले नेपालमा आधिकारिक तथ्याङ्क प्रणाली तथा विधि भन्ने पुस्तकमा आफ्नो लेख “नेपालको तथ्याङ्क प्रणाली” मार्फत संक्षेपमा प्रस्तुत गर्नुभएको सिद्धान्तको नेपाली अनुवाद तलको तालिका २.१ मा जस्ताको तस्तै प्रस्तुत

8 United Nations. *Fundamental Principles of Official Statistics*. Retrieved on November 15, 2021 from: https://unstats.un.org/unsd/dnss/hb/E-fundamental%20principles_A4-WEB.pdf

गरिएको छ ।

तालिका २.१ आधिकारिक वा सरकारी तथ्याङ्कका आधारभूत सिद्धान्तहरू

क्र.सं.	सिद्धान्त	सामान्य जानकारी
१	प्रासङ्गिकता, निष्पक्षता र सुलभता	लोकतान्त्रिक समाजको सूचना प्रणालीमा आधिकारिक तथ्याङ्क एक अपरिहार्य तत्त्व मानिन्छ । यसले सरकार, अर्थतन्त्र र आमजनतालाई आर्थिक, जनसांख्यिक, सामाजिक तथा वातावरणीय परिस्थितिसम्बन्धी सूचना उपलब्ध गराउँछ ।
२	व्यावसायिक मापदण्ड र आचार	लोकतान्त्रिक समाजमा सरकारी तथ्याङ्क निकायहरूले नागरिकको सार्वजनिक सूचना पाउने हकको सम्मान गर्नुपर्दछ । यसका लागि यी निकायहरूले व्यावसायिक उपयोगिता भएका आधिकारिक तथ्याङ्क निकालेर निष्पक्षताका साथ उपलब्ध गराउनु पर्दछ ।
३	उत्तरदायित्व र पारदर्शिता	तथ्याङ्क संकलन, प्रशोधन, भण्डारण तथा प्रस्तुतिका विधि एवं प्रक्रिया निर्धारण गर्दा तथ्याङ्क निकायहरूले वैज्ञानिक सिद्धान्त र व्यावसायिक आचारसंहिता नितान्त व्यावसायिक सोच विचारका साथ निर्णय लिनुपर्दछ । आधिकारिक वा सरकारी तथ्याङ्कमा विश्वास जगाई तथ्याङ्कमा आस्था कायम राख्न यस्तो प्रक्रिया महत्त्वपूर्ण हुन्छ ।
४	दुरुपयोग निवारण	तथ्याङ्कको सही अर्थ लगाउने काम सहज तुल्याउन तथ्याङ्क निकायहरूले तथ्याङ्कको स्रोत, विधि र प्रक्रियाको विवरण वैज्ञानिक मापदण्डअनुसार प्रस्तुत गर्नुपर्दछ ।
५	सरकारी तथ्याङ्कका स्रोतहरू	तथ्याङ्कीय सर्वेक्षण वा प्रशासनिक अभिलेखलगायत सबै प्रकारका स्रोतहरूबाट तथ्याङ्कीय प्रयोजनका लागि आँकडा वा विवरण लिन सकिन्छ । तथ्याङ्क निकायहरूले स्रोत चयन गर्दा गुणस्तर, सामयिकता, लागत एवं उत्तरदाताको बोभलाई विचार गरेर उपयुक्त स्रोत चयन गर्नुपर्दछ ।
६	गोपनीयता	तथ्याङ्क निकायहरूले तथ्याङ्क निकाल्नका लागि संकलन गरेका व्यक्तिगत विवरण वा आँकडाहरू प्राकृतिक वा कानुनी व्यक्ति जो जसको भए तापनि एकदम गोप्य राख्नुपर्दछ । संकलन गरिएका यस्ता व्यक्तिगत विवरणहरू तथ्याङ्कीय प्रयोजनका लागि मात्र प्रयोग गर्नुपर्दछ ।
७	कानुन	जुन ऐन, नियम र आदेशअन्तर्गत तथ्याङ्क प्रणाली सञ्चालन भएका छन् ती सबै कानुनी प्रावधानहरू सार्वजनिक हुनुपर्दछ ।
८	राष्ट्रिय समन्वय	तथ्याङ्क प्रणालीमा सामञ्जस्यता र दक्षता प्राप्त गर्न देशभित्रका सबै तथ्याङ्क निकायहरू बिच समन्वय अत्यावश्यक हुन्छ ।

९	अन्तर्राष्ट्रिय मापदण्डको उपयोग	प्रत्येक देशको तथ्याङ्कीय निकायहरूले अन्तर्राष्ट्रिय अवधारणा, वर्गीकरण तथा विधिहरूको प्रयोग गर्दा सबै सरकारी तहमा सामञ्जस्यता र दक्षताको वृद्धि हुन्छ।
१०	अन्तर्राष्ट्रिय सहयोग	तथ्याङ्कमा द्विपक्षीय एवं बहुपक्षीय सहयोगले सबै देशमा आधिकारिक तथ्याङ्क प्रणालीको सुधार गर्न सघाउ पुऱ्याउँदछ।

स्रोत: वास्तोला, तुङ्ग शिरोमणि (२०६६)। नेपालको तथ्याङ्क प्रणाली। नेपालमा आधिकारिक तथ्याङ्क प्रणाली तथा विधि (पृष्ठ ८)। काठमाण्डौं, नेपाल: केन्द्रीय तथ्याङ्क विभाग।

२.४.२ तथ्याङ्कमा प्राविधिक सहयोगसम्बन्धी राम्रा अभ्यासहरूको घोषणा^९

संयुक्त राष्ट्र तथ्याङ्क आयोगको सन् १९९९ को ३० औं सभाले अनुमोदन गरी अपनाएको तथ्याङ्कमा प्राविधिक सहयोगसम्बन्धी राम्रा अभ्यासहरूको घोषणाले अन्तर्राष्ट्रिय र बहुराष्ट्रिय निकायहरूमा कार्यरत तथ्याङ्ककर्मीहरूलाई तथ्याङ्कीय क्रियाकलाप सञ्चालन गर्न स्पष्ट मार्गदर्शन गर्दछन्। कुल दशवटा सिद्धान्तान्तर्गत व्याख्या गरिएको राम्रा अभ्यासका प्रावधानहरूले अन्तर्राष्ट्रिय निकायका व्यावसायिक तथ्याङ्ककर्मीले पारदर्शी, व्यावसायिक एवं वैज्ञानिक तथ्याङ्कीय विधिहरूको प्रयोग गर्ने मार्गदर्शन प्रदान गरेको छ। यी अभ्यासहरूको प्रयोगले अन्तर्राष्ट्रिय तथ्याङ्क प्रणालीलाई प्रवर्द्धन गर्दछ। यी अभ्यासहरूले सर्वसाधारणको सूचनाको हकलाई सम्मान गर्दछन् र अन्तर्राष्ट्रिय मानव विकासमा सदस्य राष्ट्रहरूको सहभागितालाई जोड दिन्छन्। यसका अतिरिक्त यी राम्रा अभ्यासहरूले व्यक्तिगत सूचनाको गोपनीयता कायम राख्न र उपलब्ध तथ्याङ्कको सही सदुपयोग गर्नमा जोड दिएका छन्। साथै, तथ्याङ्क निकायहरू बिचको समन्वय तथा सहकार्यलाई व्यावसायिक वृद्धि सुधार गर्ने र अन्तर्राष्ट्रिय तथ्याङ्कको गुणस्तर, सुसंगतता र शासनलाई सुदृढ गर्ने माध्यमको रूपमा मान्यता दिएको छ।

२.४.३ अन्तर्राष्ट्रिय तथ्याङ्क सञ्चालन गर्ने सिद्धान्तहरू^{१०}

संयुक्त राष्ट्र तथ्याङ्क आयोगले आधिकारिक तथ्याङ्कसम्बन्धी आधारभूत सिद्धान्तहरू अनुमोदन गरेको दशौं वर्षमा अर्थात् सन् २००४ मा अन्तर्राष्ट्रियस्तरको तथ्याङ्कीय गतिविधिहरूको कामलाई मार्गदर्शन गर्न सक्ने समान प्रकृतिको सिद्धान्तहरूको सिर्जना गर्ने विचार विमर्श हुन थाल्यो। फलस्वरूप, तथ्याङ्कीय गतिविधि समन्वय समितिले

९ United Nations Statistics Division. Retrieved on November 15, 2021 from: https://unstats.un.org/unsd/methods/statorg/principles_stat_activities/principles_stat_activities.asp

१० United Nations Statistics Division. Committee for the Coordination of Statistical Activities - *Principles Governing International Statistical Activities*. Retrieved on November 15, 2021 from: https://unstats.un.org/unsd/ccsa/principles_stat_activities/

सन् २००५ मा अन्तर्राष्ट्रिय तथ्याङ्क सञ्चालन गर्ने सिद्धान्तहरू अनुमोदन गर्‍यो । यी सिद्धान्तहरू यसप्रकार छन्:

- सबैका लागि पहुँचयोग्य, उच्च गुणस्तरको अन्तर्राष्ट्रिय तथ्याङ्कहरू, विश्वव्यापी सूचना प्रणालीको आधारभूत तत्त्व हुन् ।
- अन्तर्राष्ट्रिय तथ्याङ्कहरूमा विश्वास कायम राख्न, तिनीहरूको उत्पादन निष्पक्ष र पूर्णरूपमा उच्चतम व्यावसायिक मापदण्डहरूमा आधारित हुनुपर्छ ।
- संस्थाहरूको तथ्याङ्कीय कार्यको लागि प्राप्त अधिकारका बारेमा सर्वसाधारणलाई जानकारी पाउने अधिकार छ ।
- अन्तर्राष्ट्रिय तथ्याङ्कको उत्पादनमा प्रयोग गरिएका अवधारणाहरू, परिभाषाहरू, वर्गीकरणहरू, स्रोतहरू, विधिहरू र प्रक्रियाहरू व्यावसायिक र वैज्ञानिक मापदण्ड पूरा गर्ने गरी छनोट गरिन्छ र प्रयोगकर्ताहरूका लागि पारदर्शी बनाइन्छ ।
- समयबद्धता र गुणस्तरका अन्य पक्षहरू सुनिश्चित गर्न, लागतप्रभावी हुन र तथ्याङ्क प्रदायकहरूको बोझ कम गर्नको लागि तथ्याङ्क सङ्कलनका लागि स्रोतहरू र विधिहरू उपयुक्त रूपमा छनोट गरिन्छ ।
- प्राकृतिक व्यक्तिहरू र कानुनी संस्थाहरू, वा राष्ट्रिय गोपनीयता नियमहरूको अधीनमा रहेका साना समष्टिहरूको बारेमा सङ्कलन गरिएको व्यक्तिगत विवरण पूर्ण रूपमा गोप्य राखिनेछ र तथ्याङ्कीय प्रयोजनका लागि वा कानूनद्वारा निर्देशित उद्देश्यहरूको लागि मात्र प्रयोग गरिनेछ ।
- तथ्याङ्कको गलत व्याख्या र दुरुपयोगलाई तुरुन्तै उचित रूपमा सम्बोधन गरिनेछ ।
- राष्ट्रिय र अन्तर्राष्ट्रिय तथ्याङ्कका मानकहरू व्यावसायिक मापदण्डका आधारमा विकास गरिनेछ र यिनीहरूले व्यावहारिक उपयोगिता र सम्भाव्यताको परीक्षणलाई समेत आत्मसात् गरेका हुन्छन् ।
- गुणस्तर, सुसंगतता र सुशासनलाई सुदृढ गर्न अन्तर्राष्ट्रिय तथ्याङ्कीय कार्यक्रमहरूको समन्वय आवश्यक छ ।
- तथ्याङ्कमा द्विपक्षीय र बहुपक्षीय सहयोगले संलग्न तथ्याङ्कविद्हरूको व्यावसायिक वृद्धि गर्नुका साथै संघ संस्था र देशहरूमा तथ्याङ्कको सुधारमा योगदान पुऱ्याउँछ ।

२.४.४ PARIS 21 को स्थापना¹¹

PARIS 21 भनेको २१ औँ शताब्दीमा विकासको लागि तथ्याङ्कमा साभेदारी गर्ने प्रमुख अन्तर्राष्ट्रिय संघ हो । यसका प्रमुख उद्देश्यहरू निम्नानुसार रहेका छन् ।

11 PARIS 21. Retrieved on November 15, 2021 from: <https://paris21.org/>

PARIS 21 भनेको २१ औँ शताब्दीमा विकासको लागि तथ्याङ्कमा साभेदारी गर्ने प्रमुख अन्तर्राष्ट्रिय संघ हो । यसका प्रमुख उद्देश्यहरू निम्नानुसार रहेका छन् ।

१) नीति निर्माण र अनुगमनको लागि प्रमाणमा आधारित हुने संस्कृतिको विकास गर्ने,
२) उपलब्ध स्रोतहरू प्रभावकारी रूपमा प्रयोग गर्दै, राम्रोसँग व्यवस्थित तथ्याङ्कीय प्रणाली विकास गर्ने ।

निर्णय प्रक्रियामा भरपर्दो तथ्याङ्कको एकीकरणलाई प्रवर्द्धन गरेर विकासोन्मुख देशहरूमा गरिबी घटाउने र सुशासनमा सुधार गर्ने लक्ष्यसहित द्रुत रूपमा परिवर्तन भइरहेको संसारमा भरपर्दो, उच्च गुणस्तरको आधिकारिक तथ्याङ्कहरू उत्पादन गर्ने र उपलब्ध गराउने चुनौतीहरूलाई सम्बोधन गर्न अन्तर्राष्ट्रिय संघ PARIS 21 को स्थापना भएको हो । यसले विकासोन्मुख देशहरू, दातृ राष्ट्रहरू र बहुपक्षीय र अन्तर्राष्ट्रिय संस्थाहरूलाई एक ठाउँमा ल्याउँछ । सन् १९९९ को नोभेम्बरमा यस संघको स्थापना भएको हो र सन् २००० को जुनदेखि यसका बैठकहरू हुँदै आएका छन् । सन् २००९ को नोभेम्बरमा सेनेगलको डकारमा भएको बैठकले तथ्याङ्क विकासका लागि डकार घोषणा गरेको थियो ।¹²

२.४.५ तथ्याङ्कको लागि माराकेस कार्ययोजना¹³

सहस्राब्दी विकास लक्ष्य अनुगमन गर्न आवश्यक तथ्याङ्क सुधार गर्ने र विकासशील देशहरूको तथ्याङ्कीय क्षमता सबल र सक्षम बनाउने मूल उद्देश्यसहित तथ्याङ्कको लागि माराकेस कार्ययोजना सन् २००४ मा तयार गरिएको थियो । यसले खासगरी प्रमुख सूचकहरूको उपलब्धता सुधार गर्न तथ्याङ्क विकासका लागि राष्ट्रिय रणनीति तयार गर्ने मिति तय गर्दछ । यसका साथै, यस योजनाले सहस्राब्दी विकास लक्ष्य प्रगतिको अनुगमन गर्ने क्षमता आँकलन गर्दै उक्त रणनीतिमा केन्द्रित हुन प्रेरित गर्दछ ।

माथि उल्लेख गरिएका बाहेक तथ्याङ्क विकासका लागि अन्तर्राष्ट्रियस्तरमा अन्य धेरै प्रयासहरू भएका छन् । उदाहरणका लागि तथ्याङ्कको लागि बुसान कार्ययोजना¹⁴, संयुक्त राष्ट्र संघको विश्व तथ्याङ्क फोरम आदिलाई लिन सकिन्छ ।

12 PARIS 21, 2009. *Declaration on the Development of Statistics*. Retrieved from: <https://www.paris21.org/sites/default/files/DDDS-en.pdf>

13 IEG (Independent Evaluation Group). 2011. *Marrakech Action Plan for Statistics, Partnership in Statistics for Development in the 21st Century, and Trust Fund for Statistical Capacity Building*. Global Program Review Vol. 5, Issue 3.

14 PARIS21. (2011). *Statistics for Transparency, Accountability, and Results: A Busan Action Plan for Statistics*. Retrieved from: https://www.paris21.org/sites/default/files/Busanactionplan_nov2011.pdf

२.५ तथ्याङ्क व्यवस्थापनका लागि केन्द्रदेखि स्थानीय तहसम्मको संस्थागत व्यवस्था

संघीय संरचनामा प्रवेश गरिसकेपछि तथ्याङ्कीय क्रियाकलापहरू पनि सबै तहका सरकारको अधिकार क्षेत्रभित्र प्रवेश गरिसकेको छ। संविधानप्रदत्त अधिकारको कार्यान्वयन गर्न केन्द्रीयस्तरमा रहेको केन्द्रीय तथ्याङ्क विभागलाई राष्ट्रिय तथ्याङ्क कार्यालयको रूपमा स्तरोन्नति गर्ने तथा समयसापेक्ष गुणस्तरीय तथ्याङ्क उत्पादन, प्रशोधन, विश्लेषण, प्रसार गर्ने उद्देश्यबमोजिम संगठन तथा व्यवस्थापन सर्वेक्षण अध्ययनको चरणमा रहेको छ। त्यसैगरी संघीय संरचना सापेक्ष कानुनी पूर्वाधारका लागि तथ्याङ्क विधेयक राष्ट्रिय सभाबाट पारित भएर प्रतिनिधि सभामा विचाराधीन अवस्थामा छ। यद्यपि विद्यमान अवस्थामा तथ्याङ्कीय कार्य गर्ने नेतृत्वदायी भूमिकामा केन्द्रीय तथ्याङ्क विभाग क्रियाशील छ भने यस विभाग मातहत ३३ वटा जिल्लामा तथ्याङ्क कार्यालयहरू सञ्चालनमा रहेका छन्। यसका अतिरिक्त कतिपय संवैधानिक निकाय, विषयगत मन्त्रालय, आयोग, विभाग र अन्य सरकारी तथा अर्धसरकारी निकायहरूमा तथ्याङ्क व्यवस्थापनका लागि छुट्टै शाखा वा एकाइ क्रियाशील रहेका छन्। उदाहरणका लागि लोकसेवा आयोग; अख्तियार दुरुपयोग अनुसन्धान आयोग; निर्वाचन आयोग; राष्ट्रिय प्राकृतिक स्रोत तथा वित्त आयोग; स्वास्थ्य तथा जनसंख्या मन्त्रालय; अर्थ मन्त्रालय; कृषि तथा पशुपन्छी विकास मन्त्रालय; उद्योग, वाणिज्य तथा आपूर्ति मन्त्रालय; संस्कृति, पर्यटन तथा नागरिक उड्डयन मन्त्रालय; ऊर्जा, जलस्रोत तथा सिँचाई मन्त्रालय; वन तथा वातावरण मन्त्रालय; राष्ट्रिय योजना आयोग आदिमा तथ्याङ्क शाखा रहेको र ती शाखाहरूमा तथ्याङ्क सेवाका कर्मचारीहरू कार्यरत रहेका छन्। त्यसैगरी कतिपय विभागहरूमा पनि तथ्याङ्क शाखा र तथ्याङ्क सेवाका कर्मचारीहरू कार्यरत छन्।

प्रदेशस्तरको तथ्याङ्कीय संस्थागत व्यवस्थाको सन्दर्भमा कुरा गर्दा लुम्बिनी प्रदेशले तथ्याङ्क ऐन, २०७५ तथा तथ्याङ्क नियमावली, २०७८ जारी गरेको तर सोबमोजिम संस्थागत व्यवस्था गरी नसकेको अवस्था छ भने अन्य प्रदेशहरूले तथ्याङ्क ऐन नै बनाई नसकेको अवस्था छ। यद्यपि प्रदेशका केही मन्त्रालय वा मातहत निकायहरूमा तथ्याङ्क सेवाका कर्मचारी कार्यरत रहेको अवस्था छ। त्यसैगरी प्रदेश मातहतका स्वास्थ्य कार्यालयहरूमा पनि तथ्याङ्क एकाइ रहेको र सो एकाइमा तथ्याङ्क सेवाका कर्मचारी कार्यरत रहेका छन्। स्थानीय तहको हकमा महानगरपालिका तथा उपमहानगरपालिकाहरूमा तथ्याङ्क सेवाका कर्मचारी कार्यरत रहेका छन् भने अन्य स्थानीय तहमा तथ्याङ्क उत्पादन हुने गरेको तर यथोचित रूपमा तथ्याङ्क एकाइ र तथ्याङ्क सेवाका कर्मचारीको व्यवस्था नभएको अवस्था छ। समग्रमा हेर्दा, प्रदेश तथा स्थानीय तह दुवैमा तथ्याङ्क व्यवस्थापन गर्ने छुट्टै आधिकारिक निकाय नभएको र भएका सरकारी निकायहरूमध्ये केहीमा तथ्याङ्क सेवाका कर्मचारी

कार्यरत भए तापनि कुनै पनि निकायमा तथ्याङ्कको छुट्टै महाशाखा वा शाखा नभएको अवस्था छ ।

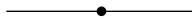
२.६ अभ्यासका लागि प्रश्नहरू

- क) नेपालको संविधानबमोजिम संघ, प्रदेश, र स्थानीय तहको तथ्याङ्कसम्बन्धी अधिकार के के छन् ? उल्लेख गर्नुहोस् ।
- ख) तथ्याङ्क ऐन, २०१५ का प्रमुख विशेषताहरू के के हुन् ? समयसापेक्ष तथ्याङ्क व्यवस्थापन गर्न के कस्ता कानुनी प्रावधानहरू यस ऐनले पूरा गर्न सकेको छैन, व्याख्या गर्नुहोस् ।
- ग) नेपाल राष्ट्र बैंक ऐन, २०५८ ले राष्ट्र बैंकलाई के कस्ता तथ्याङ्कीय क्रियाकलाप गर्ने अधिकार प्रदान गरेको छ, चर्चा गर्नुहोस् ।
- घ) प्रदेश तथा स्थानीय तहमा तथ्याङ्कीय क्रियाकलाप गर्ने छुट्टै सरकारी निकाय नहुँदा के कस्ता समस्याहरू सृजना भएका छन्, विवेचना गर्नुहोस् ।
- ङ) संयुक्त राष्ट्र संघको आधिकारिक तथ्याङ्कसम्बन्धी आधारभूत सिद्धान्तहरू उल्लेख गर्दै नेपालले यस सिद्धान्तलाई कुन हदसम्म अपनाएको छ, टिप्पणी गर्नुहोस् ।

२.७ सन्दर्भ सामग्री

तथ्याङ्कसम्बन्धी राष्ट्रिय नीति सम्बन्धमा विस्तृत रूपमा अध्ययन गर्नका लागि निम्न सामग्रीहरू उपयोगी छन् ।

१. सिग्देल, ऋषि राम (२०१७) । नेपालमा राष्ट्रिय लेखा र तथ्याङ्कीय नीतिहरू । एसिया प्यासिफिक आर्थिक तथ्याङ्क सप्ताह, <https://communities.unescap.org/econ-statistics/apes-2017-papers-presentations>. Retrieved on November 15, 2021.
२. सिलवाल, राजन (२०१७) । नेपालको तथ्याङ्कीय योजना तथा फ्रेमवर्क । केन्द्रीय तथ्याङ्क विभाग, *नेपालको राष्ट्रिय तथ्याङ्क प्रणाली, एक संग्रह* (अंग्रेजी संस्करण) (पृष्ठ २१-२९) । काठमाण्डौं, नेपाल: केन्द्रीय तथ्याङ्क विभाग ।



तथ्याङ्कीय साक्षरता तथा डिजिटल साक्षरतासम्बन्धी अवधारणा

३.१ परिचय

नेपालको तथ्याङ्क प्रणाली र यससम्बन्धी नीतिगत व्यवस्थाका बारेमा हामीले पाठ १ र पाठ २ मा चर्चा गरिसकेका छौं। यस पाठमा हामीले तथ्याङ्कीय साक्षरता र डिजिटल साक्षरता के हो? यसको महत्त्व के छ? जस्ता विषयहरूमा छलफल गर्दछौं। तथ्याङ्कीय साक्षरता र डिजिटल साक्षरता बुझ्नका लागि साक्षरता भनेको के हो त्यसकै पुनरावलोकन गरौं। विभिन्न विद्वानहरूले साक्षरतालाई फरक-फरक तरिकाले परिभाषित गरेको पाइन्छ र यसको परिभाषा क्रमिक रूपमा विकसित भइरहेको छ। साक्षरताको सर्वमान्य परिभाषा भने पाइँदैन। उदाहरणका लागि, साक्षरताले कहिलेकाहीँ केवल पढाइ लेखाइलाई मात्र बुझाउँछ भने कहिलेकाहीँ यसका साथै अन्य क्षमतालाई समेत समेटेको पाइन्छ। अक्सफोर्ड एडभान्स लर्नर्स शब्दकोष (सातौँ संस्करण) ले पढ्न र लेख्न सक्ने क्षमतालाई साक्षरता भनी अर्थ लगाएको छ। साक्षरतालाई बृहत् रूपमा परिभाषित गर्दै संयुक्त राष्ट्र संघीय शैक्षिक, वैज्ञानिक र सांस्कृतिक संगठन (UNESCO) ले “पढ्ने, लेख्ने र गणना गर्ने सीपहरूको समूहको रूपमा यसको परम्परागत अवधारणा भन्दा बाहिर, साक्षरतालाई अब बढ्दो डिजिटल, पाठ-मध्यस्थता, सूचना-सम्पन्न र द्रुत-परिवर्तनशील संसारमा पहिचान, बुझाइ, व्याख्या, सिर्जना र सञ्चारको माध्यमको रूपमा बुझिन्छ” भनी व्याख्या गरेको छ। Progress in Reading Literacy Study – PIRLS ले समाज वा व्यक्तिलाई आवश्यक पर्ने लिखित भाषा बुझ्न र प्रयोग गर्न सक्ने क्षमतालाई पढ्न सक्ने क्षमता भनी परिभाषित गरेको छ।¹⁵ यी र यस्तै परिभाषाहरूलाई समायोजन गरी हेर्दा, कुनै पनि भाषामा पढ्न, लेख्न, बुझ्न, व्याख्या गर्न, सञ्चार गर्न सक्ने क्षमता नै साक्षरता हो भन्दा अतिशयोक्ति नहोला। यस अर्थमा तथ्याङ्कका भाषाहरू पढ्न, लेख्न, बुझ्न, व्याख्या गर्न, सञ्चार गर्न सक्ने क्षमतालाई तथ्याङ्कीय साक्षरता (Statistical Literacy) भन्न सकिने भयो। यसैगरी डिजिटल माध्यमबाट सूचना प्राप्त गर्ने, बुझ्ने, उत्पादन गर्ने, सञ्चार गर्ने जस्ता विविध क्रियाकलापलाई नै डिजिटल साक्षरता भन्न सकिने भयो।

15 Cambridge University Press and Assessment. (2013). *What is literacy? An investigation into definition of English as a subject and the relationship between English, literacy and 'being literate' – A research report commissioned by Cambridge Assessment*. United Kingdom; Cambridge University Press and Assessment.

३.२ तथ्याङ्कीय साक्षरता (Statistical Literacy)

तथ्याङ्कीय साक्षरता सम्बन्धमा विभिन्न विद्वान्हरूले विभिन्न तरिकाले परिभाषित गरेको पाइन्छ । केही विद्वानहरूको परिभाषालाई संक्षेपमा हेरौं ।

“तथ्याङ्कीय साक्षरता भनेको हाम्रो दैनिक जीवनमा व्याप्त तथ्याङ्कीय नतिजाहरूलाई बुझ्ने र आलोचनात्मक रूपमा मूल्याङ्कन गर्ने क्षमता हो । यसका साथै, तथ्याङ्कीय सोचले सार्वजनिक र निजी, व्यावसायिक र व्यक्तिगत निर्णयहरूमा गर्न सक्ने योगदानहरूको कदर गर्ने क्षमता राख्दछ ।” - वालम्यान (Wallman (1993, p.1)

“तथ्याङ्कीय साक्षरतालाई “ट्रान्सन्युमेरेटिभ सोच (transnumerative thinking)” को रूपमा हेर्न सकिन्छ । जसको अर्थ हो, अध्येताहरूले तथ्याङ्कको विविध पक्षको प्रयोग गरी वास्तविक संसारको अर्थ बताउन सक्षम हुनेछन् भन्ने हो ।” - चिक, फ्वानकुच र वाट्सन (Chick, Pfannkuch, and Watson (2005)

“तथ्याङ्कीय साक्षरताले तथ्याङ्कको आधारभूत भाषा र उपकरणहरू बुझ्ने र प्रयोग गर्ने क्षमता बुझाउँदछ । अर्को अर्थमा, आधारभूत तथ्याङ्कीय अवधारणाहरूको अर्थ के हो भनेर जान्नु, साधारण तथ्याङ्कीय संकेतहरूको प्रयोग बुझ्ने र पहिचान गर्ने तथा तथ्याङ्कको विभिन्न प्रस्तुतिको व्याख्या गर्न सक्षम हुनु नै तथ्याङ्कीय साक्षरता हो ।”- गार्फिल्ड, डेलमास र जिफ्लर (Garfield, delMas, and Zieffler - 2010)

“तथ्याङ्कीय सूचना तथा तथ्याङ्क सम्बन्धित तर्कहरू व्याख्या गर्ने र आलोचनात्मक रूपमा मूल्याङ्कन गर्ने मानिसहरूको क्षमतालाई तथ्याङ्कीय साक्षरता भनिन्छ ।”- गल (Gal -2004, p= 49)

गलका अनुसार तथ्याङ्कीय साक्षर हुनका लागि निम्नानुसारका क्षेत्रमा ज्ञान हासिल गरेको हुनुपर्दछ:

- | | | |
|-------------------|----------------------|--------------------|
| क) गणितको ज्ञान, | ख) तथ्याङ्कको ज्ञान, | ग) सन्दर्भको ज्ञान |
| घ) साक्षरता कौशल, | ड) आलोचनात्मक प्रश्न | |

संक्षेपमा भन्नुपर्दा, तथ्याङ्कीय साक्षरता भनेको तथ्याङ्कलाई बुझ्ने र तर्क गर्ने क्षमता हो । यस्तो क्षमता तथ्याङ्कसँग सम्बन्धित सबै सरोकारवालाहरूसँग भएमा मात्र तथ्याङ्कको सही उत्पादन तथा प्रयोग भई प्रमाणमा आधारित निर्णय प्रक्रियाको पूर्ण प्रयोग हुन्छ र प्रमाणकै आधारमा नीति, योजना तथा कार्यक्रमहरू तय भई जनताको अधिकतम हित हुन सक्दछ ।

३.३ तथ्याङ्कीय साक्षरताका मापदण्ड

तथ्याङ्कीय रूपले साक्षर हुनका लागि बढी मात्रामा तथ्याङ्कको प्रयोग गरेर मात्र पुग्दैन ।

कतिपय मानिसहरूलाई उनीहरूलाई आवश्यक परेका सूचनाको स्रोत थाहा नहुन सक्दछ । सूचना पाइहाले पनि त्यसमा अन्तरनिहित अवधारणा तथा परिभाषा थाहा नहुन सक्दछ । साथै, सूचनामा इञ्जित गरिएका संख्याहरूले बोल्ने भाषा बुझ्ने क्षमताको विश्वास पनि नहुन सक्दछ । तथ्याङ्गीय साक्षरता भनेको संख्याहरू सम्बद्ध पद्धतिलाई सूत्रबद्ध गर्ने वा बुझ्ने मात्र नभई संख्याहरू व्याख्या गर्ने र यसमा समावेश जानकारीलाई प्रभावकारी रूपमा सञ्चार गर्ने क्षमता पनि हो । तसर्थ, तथ्याङ्गीय रूपले साक्षर हुनका लागि निम्नानुसारका चारवटा मापदण्ड पूरा भएको हुनुपर्दछ ।

३.३.१ तथ्याङ्क बोध (Data Awareness)

उपलब्ध तथ्याङ्क सान्दर्भिक र उपयुक्त छ वा छैन, तथ्याङ्कको स्रोत के हो र कसरी संकलन गरिएको हो तथा उद्देश्यानुरूप उक्त तथ्याङ्क उपयुक्त हुन्छ वा हुँदैन भन्ने जस्ता कुराहरू यकिन गर्न सक्नुपर्दछ ।

३.३.२ तथ्याङ्गीय अवधारणाहरू बुझ्ने क्षमता (Ability to Understand Statistical Concepts)

तथ्याङ्कलाई तालिका, चित्र वा नक्सामा प्रस्तुत गरिएका हुन्छन् । यस्ता तथ्याङ्कहरू प्रतिशत, अनुपात वा दरका रूपमा उल्लेख गरिएका हुन सक्दछन् । अभ्र परिस्कृत रूपमा औसत, विस्तार जस्ता सूचकहरू प्रयोग गरेर तथ्याङ्क प्रस्तुत गरिएको पनि हुन सक्दछ । यसका अतिरिक्त गणना, सर्वेक्षण, प्रवृत्ति जस्ता कैयौँ तथ्याङ्गीय प्राविधिक शब्दावलीहरू प्रयोग भएको हुन सक्दछ । तथ्यगत विवरणहरू जे जसरी प्रस्तुत गरिएको भए तापनि तिनीहरूको अवधारणा बुझ्न सक्ने क्षमता हुनुपर्दछ ।

३.३.३ तथ्याङ्गीय जानकारीको विश्लेषण, व्याख्या र मूल्याङ्कन गर्ने क्षमता (Ability to Analyze, Interpret and Evaluate Statistical Information)

तथ्याङ्गीय रूपले साक्षर हुनका लागि तथ्याङ्क कसरी व्यवस्थित गरिएको हुन्छ, व्यवस्थित रूपले प्रस्तुत गरिएको तथ्याङ्कले थप व्याख्या गर्न कसरी योगदान गरेको हुन्छ भन्ने कुरा बुझ्नुपर्दछ । उपलब्ध तथ्याङ्कको सन्दर्भ बुझेर तालिका वा ग्राफ वा चित्र जे जसरी तथ्याङ्क प्रस्तुत भएको भए पनि तिनीहरूको विश्लेषण, व्याख्या र मूल्याङ्कन गर्ने क्षमता हुनुपर्दछ । कतिपय अवस्थामा सर्वेक्षणानुसार परिभाषा फरक-फरक हुन सक्दछ । उदाहरणका लागि श्रमशक्ति उमेर समूह कसैले १० वर्षमाथि, कसैले १५ वर्षमाथि मानेका हुन सक्दछन् । त्यस्तै बालबालिकाको उमेर समूह पनि फरक-फरक परिभाषित गरिएको हुनसक्दछ । यस्ता विषयहरूलाई सूक्ष्म तरिकाले बुझ्नुपर्दछ ।

३.३.४ तथ्याङ्कीय जानकारी र बुझाइ सञ्चार गर्ने क्षमता (Ability to Communicate Statistical Information and Understandings)

तथ्याङ्कीय आँकडाहरू बुझेर मात्र तथ्याङ्कीय साक्षर हुनका लागि योग्य नहुन पनि सकिन्छ। यसका लागि बुझेका तथ्याङ्क सञ्चार गर्ने क्षमता हुनुपर्दछ। तसर्थ, तथ्याङ्क कसरी रिपोर्ट गरिएको छ भन्ने विषय महत्त्वपूर्ण हुन्छ। यसका साथै, तथ्याङ्कीय सञ्चार गर्ने सन्दर्भमा तथ्याङ्कको गोपनीयतालाई मूलभूत रूपमा अंगिकार गरिएको हुनुपर्दछ। प्रचलित नियम कानूनबमोजिम गोप्य राख्नुपर्ने सूचना तथा कानुनी वा प्राकृतिक व्यक्तिको व्यक्तिगत विवरण कानूनबमोजिम अति गोप्य राख्नुपर्ने कुरामा अत्यन्त संवेदनशील हुनुपर्दछ।

माथि उल्लेख गरिएका चारवटा मापदण्ड पूरा गर्ने व्यक्ति नै वास्तवमा तथ्याङ्कीय साक्षर व्यक्ति हो। जबसम्म तथ्याङ्कीय साक्षरता व्यापक हुँदैन तबसम्म यसको महत्त्व कमजोर भइरहन्छ। तथ्याङ्कीय प्रणाली सुधार गर्नेदेखि प्रमाणमा आधारित निर्णय प्रक्रिया र तथ्याङ्क पृष्ठपोषित नीति निर्माण, योजना तर्जुमा, अनुगमन तथा मूल्याङ्कन गर्ने परिपाटिको विकास गर्न तथ्याङ्कीय साक्षरता अपरिहार्य छ। नेपालको सन्दर्भमा देश संघीय संरचनामा गइसकेपछि नीति निर्माण, योजना तर्जुमा अनुगमन तथा मूल्याङ्कनका कार्य स्थानीय तहदेखि केन्द्रसम्म उत्तिकै महत्त्वपूर्ण सार्वजनिक क्रियाकलाप मानिन्छ। यसका लागि सबै तहमा तथ्याङ्कीय साक्षरता उत्तिकै महत्त्वपूर्ण रहेको छ। यो पुस्तकमा प्रस्तुत गरिएका अधिकांश विषयवस्तु प्रत्येक पाठकलाई तथ्याङ्कीय साक्षर बनाउने उद्देश्यले प्रेरित छन्।

३.४ तथ्याङ्कीय साक्षरताको महत्त्व

तथ्याङ्कीय साक्षरताको महत्त्वको चर्चा गर्दा अंग्रेज लेखक हर्वर्ट जर्ज वेल्स (H.G. Wells) का भनाईलाई उद्धृत गर्नु सान्दर्भिक हुन्छ। उनका अनुसार, “तथ्याङ्कीय सोच एक दिन कुशल नागरिकताको लागि पढ्न र लेख्न सक्ने क्षमताको लागि आवश्यक हुनेछ ... र त्यो दिन आइपुग्यो! उपलब्ध तथ्याङ्कले हाम्रो आर्थिक अवस्था, हाम्रो जनसंख्याको कल्याण र हाम्रो वातावरणको अवस्थाको बारेमा महत्त्वपूर्ण जानकारी प्रदान गर्दछन्। तिनीहरूले हाम्रो लोकतन्त्रको आधार बनाउन मद्दत गर्छन् र हामीलाई हाम्रो समाजको स्वास्थ्य र प्रगति मूल्याङ्कन गर्न आवश्यक ज्ञान प्रदान गर्छन्। हामी ती तथ्याङ्कहरू दृश्यात्मक, पहुँचयोग्य र बलियो भएकोमा भर पर्छौं र हामी तथ्याङ्कीय रूपमा साक्षर व्यक्तिहरूमा भर पर्छौं, जसले हाम्रो भविष्यको कार्य निर्धारण गर्न जानकारीको उत्कृष्ट प्रयोग गर्ने, स्पष्ट र विश्वस्त तर्कहरू प्रस्तुत गरेर र ‘प्रमाणमा आधारित नीति’ विकास गरेर हाम्रो निर्णय प्रक्रियामा मार्गनिर्देशन गर्दछन्।”

हर्वर्टका भनाइले तथ्याङ्कीय साक्षरता कति महत्त्वपूर्ण विषय हो भन्ने कुरा स्पष्ट पारेको

छ । तथ्याङ्गीय साक्षरताका प्रमुख महत्त्वहरूलाई निम्नानुसार बुँदागत रूपमा उल्लेख गर्न सकिन्छ ।

- तथ्याङ्गीय साक्षरताले हामी र हाम्रो समाजलाई बलियो बनाउँछ ।
- यसले अनुसन्धान, योजना र निर्णय गर्ने उद्देश्यहरूको लागि धेरै राम्रो ज्ञान प्रदान गर्दछ ।
- यसले सार्वजनिक नीति प्रक्रिया (स्थानीय तहदेखि केन्द्रीय तहसम्म) मा लिने निर्णयहरूको लागि आधार प्रदान गर्दछ, जस्तै: चुनावी सीमाहरू निर्धारण गर्ने; स्कूल, अस्पताल, बाटो, पुल, बिजुली, खानेपानी जस्ता पूर्वाधार कहाँकहाँ आवश्यक छन्, पत्ता लगाउने आदि ।
- यसले व्यवसाय बढाउन तिनीहरूको बजार जान्न, बुझ्न र तिनीहरूको बजारीकरण रणनीतिहरूमा सुधार गर्न सहयोग गर्दछ ।
- आजको सूचना-सम्पन्न समाजमा, तथ्याङ्गीय साक्षर हुनुले प्रत्येक व्यक्तिलाई समस्याको समाधान गर्ने आधार प्रदान गर्दछ ।
- तथ्याङ्गीय साक्षर व्यक्तिलाई उसका अन्य सहकर्मीभन्दा अब्बल सावित गर्दै रोजगारदाताका लागि आकर्षक बन्न सहयोग गर्दछ ।
- यसले प्रत्येक व्यक्तिलाई उसको अधिकार, शुद्धता र निष्ठाका साथ छलफल र निर्णय प्रक्रियामा संलग्न हुन सक्षम बनाउँछ ।

यसका अतिरिक्त दैनिक जीवनका सामान्य घटनाक्रमहरू जस्तै: समाचारपत्र, टेलिभिजन र इन्टरनेट आदिका प्रकाशनहरूमा प्रस्तुत गरिएका सामग्रीहरू बुझ्नको लागि तथ्याङ्क प्रयोग गर्ने तर्कहरू बुझ्न र तर्क गर्ने क्षमता आवश्यक पर्दछ । यस्तो क्षमता तथ्याङ्गीय साक्षरताले प्रदान गर्दछ ।

सञ्चार जगत तथ्याङ्क एवं सूचनाको महत्त्वपूर्ण प्रयोगकर्ता एवं सम्वाहक हो । सूचना प्रयोगमा सञ्चार जगत निकै संवेदनशील हुनुपर्दछ । आम मानिसमा पुग्ने सूचना तथ्याङ्गीय मर्मबमोजिम तथ्यमा आधारित हुनैपर्दछ । गलत सूचना सम्प्रेषण हुन गएमा तथ्याङ्क प्रतिको अविश्वास बढ्ने मात्र हैन, समग्र रूपमा व्यक्तिगत तथा सामाजिक जीवनमा ठुलो असर गर्दछ । तसर्थ, पत्रकार अथवा सञ्चारकर्मीले सूचना सम्प्रेषण गर्नका लागि न्यूनतम तथ्याङ्गीय अवधारणा बुझ्न आवश्यक हुन्छ । त्यसैगरी समाजमा सूचना सम्प्रेषण गर्ने अर्को वर्ग भनेको विद्यार्थी पनि हो । विद्यार्थीहरूले स्वयं ज्ञानको तह वृद्धि गर्न तथा आफ्नो घर छिमेकमा सही सूचना पुऱ्याउन तथ्याङ्गीय साक्षर हुनु आवश्यक हुन्छ । उच्च शिक्षा अध्ययनका क्रममा गरिने अनुसन्धानात्मक शैक्षिक क्रियाकलापको लागि शैक्षिक आधार तयार गर्नका लागि पनि विद्यालय स्तरदेखि नै तथ्याङ्कको आधारभूत जानकारी प्राप्त गरेको हुनुपर्दछ ।

संघीय संरचनामा सरकारका तिनवटै तहबाट नीतिनिर्माता, योजना तर्जुमा, अनुगमन तथा मूल्याङ्कनका कार्यहरू सञ्चालन हुन्छन् । हरेक तहमा नीतिनिर्माता तथा नीति पैरवी र पृष्ठपोषण गर्ने विज्ञहरू हुन्छन् । उनीहरूले प्रमाणमा आधारित निर्णय प्रक्रियाका लागि तथ्याङ्कको सही प्रयोग गर्नेपर्ने हुन्छ । यसका अतिरिक्त संघीय सरकारदेखि स्थानीय तहका सरकारले तथ्याङ्कसम्बन्धी विभिन्न क्रियाकलाप सञ्चालन गरेका हुन्छन् । नेपालको संविधानको अनुसूची ५, ६, ८ र ९ मा उल्लेख भएका अधिकार सूचीहरूलाई नेपाल सरकार (मन्त्रपरिषद्) को मिति २०७३ माघ १८ को निर्णयबमोजिम विस्तृतीकरण गरी कार्यान्वयनमा आइसकेको छ । यस सन्दर्भमा तिनै तहका सरकारले तथ्याङ्क संकलन, प्रशोधन, विश्लेषण, प्रस्तुति, व्याख्या, प्रसार तथा भण्डारण गर्दा तथ्याङ्कका मानक, सोसम्बन्धी अवधारणा, परिभाषा तथा अन्य प्राविधिक पक्षमा पूर्ण जानकारी राख्नुपर्दछ । तसर्थ, तथ्याङ्गीय साक्षरता तिनै तहका सरकारका लागि निकै महत्त्वपूर्ण छ ।

३.५ विकास प्रशासनमा तथ्याङ्गीय साक्षरताको प्रभाव

नेपालमा विकास प्रशासनको अवधारणा विधिवत् रूपमा वि.सं. २०१३ बाट सुरु भएको मानिन्छ । योजनावद्ध विकासका क्रममा हालसम्म तर्जुमा गरिएका १५ वटा आवधिक योजनाहरूमा तथ्याङ्कको उपस्थितिका बारेमा पाठ २ मा चर्चा गरिसकिएको छ । यसका अतिरिक्त योजना तर्जुमा गर्दा प्रत्येक विषय तथा क्षेत्रको वस्तुस्थिति मूल्याङ्कन गर्न; समस्या पहिचान गर्न; नीति, रणनीति, कार्यनीति, लक्ष्य तथा कार्यक्रम निर्धारण गर्न; नतिजा खाका तय गर्न तथ्याङ्कको अधिकतम प्रयोग भएको पाइन्छ ।

प्रायः सबै योजनाहरूले सम्बन्धित विषय क्षेत्रमा आवश्यक खण्डीकृत तथ्याङ्कको कुनै न कुनै स्तरमा अभाव रहेको विषय उल्लेख गरेका छन् । तसर्थ, नेपालमा राम्रा नीति बन्ने तर कार्यान्वयन नहुने भन्ने आरोप लाग्ने गरेको पनि पाइन्छ । यसको मुख्य कारणहरूमा तथ्याङ्कको अभावका साथसाथै तथ्याङ्गीय साक्षरताको कमी पनि हो । कतिपय तथ्याङ्कहरूको सही व्यवस्थापन गर्न नसक्दा गुणस्तरीय तथ्याङ्कको अभाव खाड्किएको छ । कतिपय सेवाप्रदायक निकायहरूले प्रशासनिक अभिलेखबाट तथ्याङ्क उत्पादन गर्न सकिरहेका छैनन् वा त्यस्ता अभिलेखहरूलाई व्यवस्थित गर्न सकेका छैनन् । उपलब्ध तथ्याङ्क पनि यथोचित रूपमा प्रयोग हुन सकेको छैन । फलस्वरूप योजना कार्यान्वयनको चरणमा हुने मूल्याङ्कन प्रक्रिया प्रभावित हुन पुगेको छ । अर्कोतर्फ प्रमाणका आधारमा नभई प्रभावका कारण योजना छनोट हुने र तथ्याङ्कविना नै मूल्याङ्कन हुने परिपाटीले गर्दा पनि विकासको अपेक्षित प्रतिफल प्राप्त गर्न कठिन परेको छ ।

विकास प्रशासनका लागि साधन र स्रोत परिचालन गर्दा केन्द्र सरकारले प्राकृतिक स्रोत तथा वित्त आयोगको सिफारिस कार्यान्वयन गरेको हुन्छ । वित्त आयोगले स्रोत

बाँडफाँटका लागि मूलभूत रूपमा तथ्याङ्कलाई नै आधार मानेको हुन्छ। विकास प्रशासन कार्यान्वयन गर्ने निकायहरूले पनि उपलब्ध साधन स्रोतको अधिकतम जनहितका लागि प्रयोग गर्दा तथ्याङ्ककै आधार लिएका हुन्छन्। विकासको प्रतिफलबाट प्रत्यक्ष लाभान्वित हुने वर्ग वा समुदाय पनि तथ्याङ्कको आधारमा विकास योजना माग गर्ने, सेवा प्रवाहमा कुनै पनि हिसाबले पूर्वाग्रह हुन नदिने तथा विकासमा समान र अर्थपूर्ण सहभागिताको पक्षपोषण गर्ने कार्य गरिरहेका छन्। त्यसो हुँदाहुँदै पनि जुन स्तरमा सरकारका सबै तहमा तथ्याङ्कीय साक्षरता विस्तार हुनुपर्ने हो, सोअनुसार हुन भने सकेको पाइँदैन।

विकास प्रशासनको अभ्यासकर्ताका रूपमा हाल नेपालमा केन्द्र सरकार, प्रदेश सरकार र स्थानीय सरकार क्रियाशील रहेका छन्। खासगरी प्रदेश र स्थानीय तहमा तथ्याङ्कीय साक्षर जनशक्तिको अभाव खड्किएको देखिन्छ। फलस्वरूप स्थानीय तहमा बनेका प्रोफाइल तथा विकास योजनाहरूमा अवधारणागत एकरूपता पाइँदैन। तर, कतिपय स्थानीय सरकारहरूले तथ्याङ्कीय क्षमता अभिवृद्धि गरी त्यसको प्रत्यक्ष प्रभाव विकासका कार्यहरूमा प्रदर्शन गरिरहेका पनि छन्। तथ्याङ्कीय साक्षरता अभिवृद्धि गर्ने सरकारहरू सफल भएका छन् भने अन्य सरकारहरू तुलनात्मक रूपमा पछि परेका देखिन्छन्। तसर्थ, सबै तहका सरकारले विकास प्रशासनलाई प्रभावकारी रूपमा सञ्चालन गर्न तथ्याङ्कीय साक्षरता न्यूनतम अत्यावश्यक पूर्वशर्त हो भन्नुमा अतिशयोक्ति नहोला।

३.६ डिजिटल साक्षरता (Digital Literacy)

डिजिटल साक्षरता भनेको आर्थिक र सामाजिक जीवनमा सहभागिताको लागि डिजिटल उपकरणहरू र नेटवर्क प्रविधिहरू मार्फत् सुरक्षित र उचित रूपमा सूचना सिर्जना गर्ने, सूचनाको पहुँचमा हुने, व्यवस्थान गर्ने, बुझ्ने, एकीकृत गर्ने, सञ्चार गर्ने र मूल्याङ्कन गर्ने क्षमता हो। यसले खासगरी विद्युतीय सूचनाको जानकारी, उपयोग र त्यसमा उपलब्ध हुने र हुनसक्ने सूचना तथा जानकारीहरूको आदानप्रदान गर्ने आधारभूत ज्ञानलाई बुझाउँछ। वर्तमान युग सूचना र प्रविधिको युग भएको हुँदा मानिसहरूलाई दैनिक व्यवहार गर्न परम्परागत साक्षरताले मात्र नपुग्ने अवस्था सृजना भएको छ। अधिकांश सरकारी, अर्धसरकारी, गैर सरकारी तथा निजी क्षेत्र डिजिटल युगमा प्रवेश गरिसकेका छन्। कतिपय कुराहरू कम्प्युटर र स्मार्ट फोनका माध्यमबाट मात्र प्राप्त गर्ने अवस्था रहेको छ। यसका अतिरिक्त डिजिटल साक्षरताले सेवा र सुविधाप्रदायक तथा प्रापक दुवैलाई सरल र सहज तुल्याइदिएको छ। पत्रपत्रिका पढ्नेदेखि बैङ्किङ कारोबार गर्नेसम्म डिजिटल प्रविधिको उपयोगमा बढोत्तरी हुँदै गएको छ। पत्राचार गर्नेदेखि मनोरञ्जनका लागि विविध माध्यमहरू अब डिजिटल प्रविधिमै आधारित हुन

पुगे । इन्टरनेटका माध्यमबाट दैनिक जीवनका कतिपय आवश्यकताहरू पूरा गर्न सकिन्छ । डिजिटल युगले कुनै पनि कुराको पहुँचलाई विश्वव्यापी बनाइदिएको छ । कतिपय कार्यालयका कामहरू डिजिटल प्रविधि प्रयोग गरेर घरबाटै गर्न सकिने भएको छ । तसर्थ, डिजिटल साक्षरता आधुनिक युगमा साधारण मानिसका लागि पनि आधारभूत योग्यताका रूपमा परिणत भएको छ ।

प्रोफेसर हिलर स्पाइर्सका अनुसार डिजिटल साक्षरतालाई तिन भागमा विभाजन गरेर हेर्न सकिन्छ ।

- १) डिजिटल सामग्री प्राप्त र उपयोग गर्ने,
- २) डिजिटल सामग्रीहरूको सिर्जना गर्ने र
- ३) डिजिटल सामग्रीहरूको सञ्चार वा वितरण गर्ने ।

यसरी डिजिटल साक्षरताले विद्युतीय उपकरण (कम्प्युटर, ट्याबलेट, मोबाइल फोन आदि) को उपयोग गर्दै दोस्रो पक्षबाट सिर्जित डिजिटल सामग्री प्राप्त गर्ने, स्वयं त्यस्ता सामग्रीहरू उत्पादन गर्ने, उत्पादित डिजिटल सामग्री दोस्रो वा तेस्रो पक्षसम्म सञ्चार वा सम्प्रेषण गर्नेसम्मका कार्यलाई समेट्दछ ।

समकालीन सूचनाको युगमा जीवनका प्रत्येक क्षेत्रमा डिजिटल साक्षरता अपरिहार्य बनिसकेको छ । व्यक्तिगत जीवनयापनका क्रियाकलापदेखि सार्वजनिक क्षेत्रका कार्यहरूमा समेत डिजिटल साक्षरताको प्रभावले हरेक काम लागतप्रभावी हुनुका साथै गुणस्तरीय हुन पुगेको छ । सार्वजनिक सेवा प्रवाहमा पनि डिजिटल साक्षरताको महत्त्व दिनानुदिन बढ्दै गइरहेको छ । कुनै पनि सेवा प्राप्तिका लागि अनलाइन आवेदन गर्नु, सेवा नतिजा अनलाइन मार्फत् नै प्राप्त गर्नु, सेवा प्राप्तिसँग सम्बन्धित गुनासो तथा उजुरीहरू अनलाइनमार्फत् सम्बन्धित निकायलाई प्रेषित गर्नु, आवश्यक सूचनाहरू अनलाइनमार्फत् प्राप्त गर्नु जस्ता कार्यहरूले सार्वजनिक सेवा प्रवाहलाई चुस्त दुरूस्त बनाउनुका साथै समय, पैसाजस्ता लागतका हिसाबले पनि कम खर्चिलो बनाएको छ । कोभिड-१९ जस्ता महामारीका समयमा समेत डिजिटल साक्षरताकै कारण कतिपय बालबालिका घरैबाट अनलाइन कक्षा अध्ययन गर्न सफल भए । तालिम, बैठक, सेमिनारहरू अनलाइन प्रविधिबाटै सञ्चालन भएका कारण मानिसको क्षमता र कार्यकौशलतामा ह्रास आउन पाएन ।

डिजिटल साक्षरताका असंख्य राम्रा पक्षहरूको बाबजुद यसका नकारात्मक पक्षहरूलाई बिर्सनु भने हुँदैन । बढ्दो डिजिटल प्रयोगले मानिसलाई नै यान्त्रिकीकरण गरेको जस्तो

अनुभूति हुन गइरहेको छ । कतिपय अनलाइन खेलका कारण बालबालिकाहरू कुलतमा फसेका प्रशस्तै उदाहरणहरू पनि छन् । डब्लु. एच. ओ. ले त अनलाइन खेलको अधिकतम प्रयोगबाट सीमित व्यक्तिहरूमा स्वास्थ्य समस्यासमेत सृजना हुनसक्ने जनाएको छ ।¹⁶ डिजिटल साक्षरताको नकारात्मक पक्षको एउटा उदाहरणका रूपमा अश्लील सामग्रीको अस्वाभाविक प्रयोगको जोखिम बढ्ने सम्भावनालाई पनि लिन सकिन्छ । यसका अतिरिक्त डिजिटल प्रविधि अन्तर्गत कमजोर सुरक्षा प्रणालीको फाइदा उठाएर साइबर अपराधीहरूले सूचना चोरी गर्ने, अनलाइन प्लेटफर्म ह्याक गर्ने तथा अन्य आपराधिक गतिविधि सञ्चालन गर्ने जोखिम पनि बढिरहेको छ । डिजिटल साक्षरताका यी र यस्ता अन्य कतिपय जोखिमका बाबजुद पनि यसको सही सदुपयोगबाट निजी जीवन, व्यावसायिक जीवन तथा सार्वजनिक जीवन सरल, सहज र सुन्दर बन्ने निश्चित छ ।

३.७ निर्णय प्रक्रियामा डिजिटल साक्षरता

सूचना प्रविधिको क्षेत्रमा भएको विकासले नीतिनिर्माता र निर्णयकर्ताहरूको भूमिकालाई सरल, सहज र छिटो बनाइदिएको छ । निर्णय प्रक्रियामा आवश्यक पर्ने कतिपय तथ्याङ्क र सूचनाका लागि निर्णयकर्ताहरू अरू मानिसहरू माथि भर पर्नुपर्ने अवस्थालाई डिजिटल साक्षरताले न्यूनीकरण गरिदिएको छ । यसको सबैभन्दा सुन्दर विशेषता के छ भने यो तत्काल उपलब्ध हुन्छ र किफायती पनि हुन्छ । डिजिटल साक्षरताले तथ्याङ्कका सामान्य प्रयोगकर्ताहरूको पहुँच र क्षमतालाई समेत विस्तार गरेको छ ।

आजको युग सूचना तथा प्रविधिको युग हो भन्ने गरिन्छ । तथ्याङ्क व्यवस्थापनको क्षेत्रमा कम्प्युटरको प्रयोग लगभग अनिवार्य शर्त नै भइसकेको छ । तथ्याङ्कको व्यवस्थापनका लागि विभिन्न कम्प्युटर सफ्टवेयरहरू विकास भइसकेका छन् । यस पृष्ठभूमिमा तथ्याङ्कका प्रमुख प्रयोगकर्ताहरू विद्यार्थी, पत्रकार तथा संघ, प्रदेश, र स्थानीय तहका नीतिनिर्माता एवं नीति पैरवी वा पृष्ठपोषण गर्ने जो कोहीले कम्प्युटरसम्बन्धी ज्ञान हासिल गर्नु निकै महत्त्वपूर्ण मानिन्छ । कतिपय स्थानमा कागजी कामकारबाहीलाई क्रमशः कम्प्युटरीकृत गर्दै लगेको छ । साथै, कतिपय सूचनाहरू अनलाइनमा प्राप्त हुन्छन् । यस्तो अवस्थामा उपलब्ध हुने सूचना दोहन गर्नका लागि डिजिटल साक्षर हुनैपर्दछ । प्रमाणमा आधारित भएर तर्क गर्न तथा निर्णय गर्नका लागि डिजिटल साक्षरता निकै महत्त्वपूर्ण मानिन्छ । अधिकांश मानिसको हातमा मोबाइल र इन्टरनेटको पहुँचमा विस्तार हुँदै गइरहेको अवस्थामा डिजिटल साक्षर व्यक्तिले चाहिएको बखत इन्टरनेटमार्फत् सूचना खोजेर प्रयोग गर्न सक्दछन् । यसले तुलनात्मक रूपमा लागत कम र उच्च प्रतिफल प्रदान गर्दछ ।

३.८ अभ्यासका लागि प्रश्नहरू

- क) तथ्याङ्कीय साक्षरता भनेको के हो, व्याख्या गर्नुहोस् ।
- ख) स्थानीय तहमा तथ्याङ्कीय साक्षरता किन आवश्यक छ, बुँदागत रूपमा उत्तर दिनुहोस् ।
- ग) डिजिटल साक्षरता भनेको के हो, यसका सकारात्मक तथा नकारात्मक पक्षहरूलाई केलाउनुहोस् ।
- घ) तथ्याङ्कीय साक्षर हुनका लागि के कस्ता मापदण्डहरू पूरा गर्नुपर्दछ, उल्लेख गर्नुहोस् ।
- ङ) विकास प्रशासनमा तथ्याङ्कीय साक्षरताले के कसरी सहयोग गर्दछ, व्याख्या गर्नुहोस् ।
- च) तलका भनाइहरू ठिक वा बेठिक छुट्याई कारण दिनुहोस् ।
- १) विद्यार्थीहरूले तथ्याङ्कसम्बन्धी जानकारी राख्नुपर्दछ ।
 - २) पत्रकारहरू तथ्याङ्कीय साक्षर नभए पनि हुन्छ ।
 - ३) स्थानीय तहमा कार्यरत कर्मचारीहरूलाई डिजिटल साक्षरताको त्यति आवश्यकता छैन ।
 - ४) तथ्याङ्कीय क्रियाकलाप संघीय सरकारको अधिकार क्षेत्रभित्र मात्र पर्ने विषय हुन् ।
 - ५) डिजिटल साक्षरताले मानिसको दक्षता वृद्धि गरेको छ ।

३.९ सन्दर्भ सामग्री

तथ्याङ्कीय साक्षरता सम्बन्धमा थप अध्ययन गर्नका लागि देहायबमोजिमका सामग्री उपयोगी छन् ।

१. गरफिल्ड, जे., डेलमास, आर., र जिफ्लर, ए. (२०१०) । तथ्याङ्कीय सोच मूल्याङ्कन । विडगुड, एन. हन्ट, र एफ. जोलिफ (इडिएस.) । *तथ्याङ्कीय शिक्षामा मूल्याङ्कन विधिहरू: अन्तर्राष्ट्रिय परिप्रेक्ष्य* (पृष्ठ १७५-१८६) । मिल्टन: जोन वाइले एण्ड सन्स । पाठ ११ ।
२. गल, आइ. (२००४) । तथ्याङ्कीय साक्षरता: अर्थ, तत्त्व, जिम्मेवारी । जे.वि. गरफिल्ड र डि. बेन्जभि (इडिएस.), *तथ्याङ्कीय साक्षरता, तर्क र सोच विकासको चुनौती* (पृष्ठ ४७-७८) । डोर्डेरच: क्लुवेर ।
३. चिक, एच., फ्वानकुच, एम., र वाट्सन, जे. (२००५) । *ट्रान्समुरेटिभ सोच: तथ्याङ्कभित्रका कथा पत्ता लगाउनु तथा बताउनु* । करिकुलम म्याटर्स, १, (पृष्ठ ८६-१०७) ।
४. वालम्यान, के. के. (१९९३) । *सांख्यिकीय साक्षरतामा अभिवृद्धि: हाम्रो समाजको समृद्धि* । अमेरिकी सांख्यिकी संघको जर्नल, ८८(४२१), १ ।
५. शशी, एस. (२०१७) । *तथ्याङ्कीय साक्षरताको परिभाषा र मोडेल: एक साहित्यिक समीक्षा, शैक्षिक अनुसन्धानको खुला समीक्षा* (पृष्ठ ११८-१३३)

भाग २

तथ्याङ्कसम्बन्धी
आधारभूत जानकारी

तथ्याङ्कको परिचय

४.१ तथ्याङ्कको परिभाषा

तथ्याङ्कसम्बन्धी साक्षर हुनका लागि तथ्याङ्कको अर्थ बुझ्नुपर्दछ । तथ्याङ्कका परिभाषामा विभिन्न विद्वानहरूको एकमत पाइँदैन । यद्यपि तथ्याङ्कलाई सरल तरिकाले बुझ्न पनि सकिन्छ । यस शब्दको बनोट तथ्य र अङ्क मिलेर बनेको छ । शाब्दिक हिसाबले हेर्दा तथ्यमा आधारित अङ्कले तथ्याङ्क बनेको छ । तर यतिले मात्र तथ्याङ्कको परिभाषा पूर्ण हुँदैन । अंग्रेजीमा डाटा (Data) र स्टाटिस्टिक्स (Statistics) दुईवटा शब्दहरू छन् जुन कतिपय मानिसहरूले पर्यायवाची शब्दका रूपमा बुझिरहेका र प्रयोग गरिरहेका हुन्छन् । वास्तवमा यी दुईमा तात्त्विक फरक छ । पहिलो शब्द डाटाले प्रशोधनपूर्वको कुनै खास विषयको अङ्कीय मान वा मानहरूको समूहलाई संकेत गर्दछ भने दोस्रो शब्द स्टाटिस्टिक्सले ती डाटाहरू प्रशोधन गरिसकेपछि अर्थपूर्ण रूपमा प्रस्तुत गरिने आँकडालाई बुझाउँछ । स्टाटिस्टिक्सलाई अध्ययनको एक विधाका रूपमा पनि बुझ्ने गरिन्छ । प्राध्यापक एस.सी. गुप्ताका अनुसार, “तथ्याङ्क भनेको तथ्याङ्कीय विधिहरूको त्यस्तो विज्ञान हो, जसले सांख्यिकीय आँकडाहरू संकलन, विश्लेषण गरेर अनुमान गर्दा प्रयोग हुने सिद्धान्त र प्रविधिहरूलाई प्रतिनिधित्व गर्दछ ।” अंग्रेज तथ्याङ्कविद् एम.जी. केन्डलका अनुसार, “तथ्याङ्क भनेको वैज्ञानिक विधिको एउटा त्यस्तो शाखा हो जुन प्राकृतिक घटनाको जनसंख्याको गुणहरू गणना वा मापन गरेर प्राप्त हुने आँकडासँग सम्बन्धित हुन्छ ।” सामान्य अर्थमा, तथ्याङ्क भनेको प्रभावकारी निर्णय प्रक्रियाका लागि तथ्यहरूको संकलन, संगठन, प्रशोधन, विश्लेषण, प्रस्तुति र व्याख्या गर्ने वैज्ञानिक विधा हो ।

४.२ तथ्याङ्कको महत्त्व तथा दायरा

तथ्याङ्क प्रयोगको क्षेत्र निकै व्यापक हुन्छ । दैनिक जीवनका विभिन्न क्रियाकलापदेखि सार्वजनिक क्षेत्रका विविध गतिविधिहरूमा तथ्याङ्क अपरिहार्य हुन्छ । स्थानीय तहदेखि केन्द्र तहसम्मका हरेक तहमा निर्णय प्रकृत्याका लागि तथ्याङ्क नभई हुँदैन । खासगरी नीति निर्माण, योजना तथा कार्यक्रम तर्जुमा, कार्यक्रम कार्यान्वयनको अनुगमन तथा मूल्याङ्कन गर्नमा तथ्याङ्कको प्रयोग भएको हुन्छ । यसका अतिरिक्त राष्ट्रिय तथा अन्तर्राष्ट्रिय विकास लक्ष्यहरूको प्रगति मापन गर्नका लागि तथ्याङ्ककै आवश्यकता हुन्छ । तथ्याङ्कको महत्त्वलाई निम्नानुसारका बुँदाहरूले स्पष्ट पार्दछन् ।

- क) तथ्याङ्क मानव कल्याणको अङ्कगणित हो ।
- ख) तथ्याङ्क विनाको योजना कल्पना गर्न पनि सकिँदैन ।

- ग) तथ्याङ्कले सम्बन्धित तथ्यहरू बिचको सम्बन्ध प्रस्तुत गर्दछ ।
घ) तथ्याङ्क प्रशासनको आँखा हो ।
ङ) अनुगमन तथा सुपरिवेक्षण कार्यलाई तथ्याङ्कले मद्दत गर्दछ ।
च) तथ्याङ्क व्यापार, व्यवसाय र वाणिज्यका लागि अमूल्य मानिन्छ ।
छ) तथ्याङ्क विगतको ज्ञान र अनुभवको अभिलेख हो ।
ज) विज्ञानका सबै क्षेत्रमा तथ्याङ्कीय विधिको प्रयोग हुन्छ ।
झ) नीति निर्माण तथा निर्णय प्रक्रियामा तथ्याङ्कले पृष्ठपोषकको कार्य गर्दछ ।

मानौं, तपाईं म्याग्दे गाउँपालिकाको आ.व. २०७७/७८ का लागि योजना तर्जुमा गर्दै हुनुहुन्छ । तपाईंसँग स्वास्थ्य क्षेत्रमा माग गरिएका २ वटा योजना: भाडापखाला नियन्त्रणका लागि सचेतना कार्यक्रम र श्वासप्रश्वाससम्बन्धी रोग नियन्त्रणका लागि सचेतना कार्यक्रममध्ये स्रोत र साधनको सीमितताको कारण एउटा मात्र छनोट गर्नुपर्ने बाध्यता छ । यस अवस्थामा कुन कार्यक्रम छनोट गर्ने भनेर प्राथमिकीकरण कसरी गर्नुहुन्छ ? किन ? यस्तो अवस्थामा निर्णय प्रक्रियाका लागि सर्वसम्मत हुने निर्विकल्प उपाय भनेको तथ्याङ्कको सहारा लिनु नै हो । उदाहरणका लागि तालिका ४.१ मा म्याग्दे गाउँपालिकाको भाडापखाला र श्वासप्रश्वाससम्बन्धी समस्याको वर्तमान अवस्था प्रस्तुत गरिएको छ । जसअनुसार उक्त गाउँपालिकामा भाडापखालाको तुलनामा श्वासप्रश्वासको समस्या तुलनात्मक रूपमा बढी गम्भीर देखिन्छ । तसर्थ, आगामी योजनामा श्वासप्रश्वास रोग नियन्त्रण कार्यक्रमलाई पहिलो प्राथमिकतामा पार्नुपर्ने हुन्छ । यस उदाहरणबाट योजना तर्जुमा तथा निर्णय प्रक्रियामा तथ्याङ्कको महत्त्वका बारेमा स्पष्ट हुन सकिन्छ ।

तालिका ४.१: म्याग्दे गाउँपालिकाको भाडापखाला तथा श्वासप्रश्वास रोगसम्बन्धी बिरामीहरूको संख्या

भाडापखाला र श्वासप्रश्वास रोगसम्बन्धी सूचक	२०७५/७६	२०७५/७६
५ वर्षमुनिका बच्चाको भाडापखाला संख्या	४१४	३९५
५ वर्षमुनिका बच्चाको जलवियोजन (सामान्य र कडा) भएको संख्या	१	०
बहिरङ्ग सेवामा आएका भाडापखाला संख्या	२२५	७१८
५ वर्षमुनिका बच्चाको सबै प्रकारका श्वासप्रश्वाससम्बन्धी रोग लागेको संख्या	६७३	७४२
५ वर्षमुनिका बच्चाको निमोनिया (सामान्य र कडा) संख्या	३४	१६

बहिरङ्ग सेवामा आएका सबै प्रकारका शवासप्रशवास सम्बन्धी रोग लागेको संख्या	७८७	१,४३८
--	-----	-------

स्रोत: स्वास्थ्य सेवा विभागको वार्षिक प्रतिवेदन, २०७५/७६ र २०७६/७७^१

तथ्याङ्कको दायरा निकै फराकिलो हुन्छ । योजना प्रक्रियामा, राज्यका प्रत्येक अंगहरूमा, राजनीतिमा, गणितमा, अर्थशास्त्रमा, व्यवसायमा, व्यवस्थापनमा, लेखामा र लेखापरीक्षणमा उद्योगमा, तथ्याङ्क अपरिहार्य मानिन्छ । त्यसैगरी बिमामा, खगोल विज्ञानमा, भौतिक तथा रसायन विज्ञानमा, वातावरण विज्ञानमा, सामाजिक विज्ञानमा, जीव विज्ञानमा, चिकित्सा विज्ञानमा, प्राकृतिक विज्ञानमा, मनोविज्ञानमा, शिक्षामा, युद्धमा तथ्याङ्कलाई अभिन्न अङ्गको रूपमा लिइन्छ । यसरी हेर्दा आर्थिक, सामाजिक, राजनीतिक, वातावरणीय, खगोलीय लगायत सम्पूर्ण क्षेत्रमा तथ्याङ्कको आवश्यकता पर्दछ । मानव जीवनका हरेक क्रियाकलापमा तथ्याङ्कको प्रयोग भएकै हुन्छ भन्दा अतिशयोक्ति हुँदैन । तथ्याङ्कलाई विकासको अभिन्न अङ्ग मानिन्छ । विकास योजना तर्जुमा गर्नका लागि वास्तविक संसारको यथार्थ चित्र आवश्यक हुन्छ र तथ्याङ्कले त्यो उपलब्ध गराउँदछ । उदाहरणका लागि कुनै स्थानमा पुल बनाउने वा नबनाउने विषयमा निर्णय गर्नका लागि सो स्थानमा पुलको प्रयोग गर्ने मानिसहरूको संख्या, अन्य वैकल्पिक पुल भए नभएको सूचना, पुल बनेपछि सम्बन्धित सेवाग्राहीहरूको आर्थिक, सामाजिक क्षेत्रमा हुने रूपान्तरणको सम्भावनाजस्ता विषयहरूमा सूचना आवश्यक हुन्छ र यस्ता सूचना तथ्याङ्कबाट मात्रै प्राप्त हुन्छ । तसर्थ, विकास साभेदारहरूलाई तथ्याङ्क नभई नहुने विषय हो । यसै सन्दर्भमा, विकास प्रशासनको अभ्यास गर्ने संघ, प्रदेश तथा स्थानीय सरकारले नीति निर्माण गर्न, योजनाहरूको छनोट, कार्यक्रमहरूको अनुगमन तथा मूल्याङ्कन गर्न विविध विषयमा तथ्याङ्कको आवश्यकता परेको हुन्छ । अर्कोतर्फ विकास कार्यका सूचनाहरू जनतामा सम्प्रेषण गर्ने सञ्चार जगतलाई पनि तथ्याङ्क अपरिहार्य नै हुन्छ । विकास प्रशासन मात्र नभई दैनिक प्रशासन एवं सार्वजनिक सरोकारका विषयमा निर्णय गर्न तथ्याङ्कको ठुलो महत्त्व छ । खासगरी विद्यार्थी जीवनदेखि नै सबै मानिसको तथ्याङ्क प्रयोग गर्ने बानीले वर्तमान तथा भविष्यमा बन्ने नीतिदेखि कार्यक्रमहरू आवश्यकतामा आधारित बनाउन मद्दत गर्दछ र विकासको

1 Department of Health Services. (2021). *Annual Report 2019/20*. Kathmandu, Nepal; Ministry of Health and Population.

Retrieved on November 17, 2021 from: <http://dohs.gov.np/annual-report-2076-77-2019-20/>; Department of Health Services (2020). *Annual Report (2018/19)*. Kathmandu, Nepal; Ministry of Health and Population. Retrieved on November 17, 2021 from: <http://dohs.gov.np/dohs-annual-report-fy-2076-77/>

प्रतिफल न्यायोचित रूपले वितरण हुन्छ ।

४.३ तथ्याङ्कका सीमा

निकै उपयोगी हुनुका बाबजुद पनि तथ्याङ्कका केही सीमाहरू छन् । तथ्याङ्कका केही प्रमुख सीमाहरू निम्नानुसार छन् ।

क) तथ्याङ्कले गुणात्मक प्रकृतिका घटनाको अध्ययन गर्दैन ।

ख) तथ्याङ्कले एकल व्यक्तिको अध्ययन गर्दैन ।

ग) तथ्याङ्कीय नियमहरू शतप्रतिशत सही हुँदैनन्, औसतमा मात्र सही हुन्छन् ।

घ) तथ्याङ्कको दुरुपयोग हुने खतरा हुन्छ । गलत निष्कर्षमा पुग्नाउने सम्भावना पनि हुन्छ ।

ङ) तथ्याङ्कको प्रयोग तथ्याङ्कीय साक्षर व्यक्तिहरूले मात्र गर्न सक्दछन् ।

च) तथ्याङ्कीय आँकडाहरूमा तुलनायोग्य रूपले एकरूपता हुन जरूरी हुन्छ ।

छ) तथ्याङ्कका परिणामहरूलाई सधैं कुनै पनि समूहको मूल्यको एकमात्र निर्धारकको रूपमा व्यवहार गरिनु हुँदैन ।

४.४ तथ्याङ्कप्रतिको अविश्वास

वैज्ञानिक अनुसन्धानमा तथ्याङ्कको ठुलो योगदान हुँदाहुँदै पनि यसप्रतिको अविश्वास पनि कायमै छ । कतिपय मानिसहरूले तथ्याङ्कले जे पनि प्रमाणित गरिदिन्छ भनेर आरोप लगाउने गर्दछन् । एउटै तथ्याङ्कलाई फरक-फरक व्यक्तिले फरक-फरक तवरले व्याख्या गरी निष्कर्ष निकाल्ने पनि गरेको पाइन्छ । केही मानिसले त तथ्याङ्कलाई भुटको संज्ञा पनि दिने गर्दछन् । यसरी तथ्याङ्कप्रति अविश्वास हुनुको प्रमुख कारण भनेको तथ्याङ्कीय साक्षर नभएका व्यक्तिले आफ्नो सीमित स्वार्थ पूर्तिका लागि तथ्याङ्कको गलत प्रयोग र गलत व्याख्या गर्नु नै हो । व्यावसायिक तथ्याङ्ककर्मीहरूले तथ्याङ्कप्रतिको अविश्वास तथा भ्रमहरूलाई चिर्दै व्यावसायिक मापदण्डबमोजिम तथ्याङ्कको सही सदुपयोग गर्नुपर्दछ । यसका लागि तथ्याङ्क प्रयोगकर्ता सधैं सजग र सवेदनशील हुनुपर्दछ । खासगरी तथ्याङ्क प्रयोगकर्ताहरूले सबैभन्दा ध्यान दिनुपर्ने कुरा भनेको तथ्याङ्कको आधिकारिकता हो । अर्थात् कुन निकायले कुन उद्देश्यका लागि तथ्याङ्क उत्पादन गरेको हो ? उसले अवलम्बन गरेको विधि कुन हो ? गणनाबाट आएको तथ्याङ्क हो वा सर्वेक्षणबाट आएको तथ्याङ्क हो ? सर्वेक्षणबाट आएको तथ्याङ्क हो भने नमुना छनौटको आधार र त्यसको पूर्णता के कस्तो हो ? सन्दर्भ समय के हो ? तथ्याङ्क प्रशोधन र विश्लेषणका सीमाहरू के के हुन् ? जस्ता विषयहरूमा तथ्याङ्क प्रयोगकर्ताले तथ्याङ्क प्रयोग गर्नुपूर्व जानकारी राख्नुपर्ने हुन्छ ।

४.५ अभ्यासका लागि प्रश्नहरू

- क) योजना तर्जुमा गर्दा तथ्याङ्कको प्रयोग कसरी गर्न सकिन्छ, उदाहरणसहित व्याख्या गर्नुहोस् ।
- ख) तथ्याङ्कीय साक्षर नभएको व्यक्तिको हातमा तथ्याङ्क पच्यो भने के कस्ता खतराहरू आइपर्न सक्दछन्, उदाहरणसहित व्याख्या गर्नुहोस् ।
- ग) तथ्याङ्कप्रति अविश्वास उत्पन्न हुनुको प्रमुख कारण के के हुन् ?
- घ) तथ्याङ्कका सीमाहरू के के छन्, उल्लेख गर्नुहोस् ।
- ङ) कुनै एउटा गाउँपालिकाको विद्यालयहरूको विद्यार्थी संख्या अध्ययन गरेर, कुन विद्यालयलाई धेरै अनुदान दिँदा न्यायोचित होला भनी निष्कर्ष निकाल्नुहोस् ।
- च) थप पुष्ट्याउँ गर्नुहोस् ।
- १) विद्यार्थीहरूलाई तथ्याङ्कको ठूलो महत्त्व हुन्छ ।
 - २) नीतिनिर्माताहरूले तथ्याङ्कको महत्त्व बुझेनन् भने सर्वाङ्गीण विकासमा बाधा पर्दछ ।
 - ३) स्थानीय तहमा तथ्याङ्कको महत्त्व दिनानुदिन बढ्दै गइरहेको छ ।

४.६ सन्दर्भ सामग्री

तथ्याङ्कसम्बन्धी आधारभूत जानकारीका लागि बजारमा धेरै सामग्रीहरू पाइन्छन् । जसरी सही तथ्याङ्कको प्रयोगले मात्र सही निर्णयमा पुग्न सकिन्छ, त्यसैगरी स्तरीय सामग्रीको अध्ययनबाट मात्रै तथ्याङ्कसम्बन्धी सही अवधारणा बुझ्न सकिन्छ । आजकल तथ्याङ्कसम्बन्धी अवधारणा बुझ्न अनलाइन विधिबाट गुगलमार्फत् खोजी गर्दा पनि धेरै सामग्री उपलब्ध हुन्छन् । तर अध्ययनकर्ताले सही सामग्री छनोट गरी अध्ययन गर्नु उपयुक्त हुन्छ । धेरै स्तरीय पुस्तकहरूमध्ये केही पुस्तकको जानकारी तल दिइएको छ ।

१. Statistical Techniques in Business & Economics by Douglas A. Lind, William G. Marchal & Samuel A. Wathen, Seventeenth Edition, McGraw-Hill Education, 2018
२. Fundamentals of Statistics by S.C. Gupta, Sixth Revised & Enlarged Edition, Himalaya Publishing House, 2009
३. Statistics and Probability, Concepts and Techniques by Hridaya B. Shrestha, Second Edition, Ekta Books Distributors Pvt. Ltd., 2008

पाठ ५

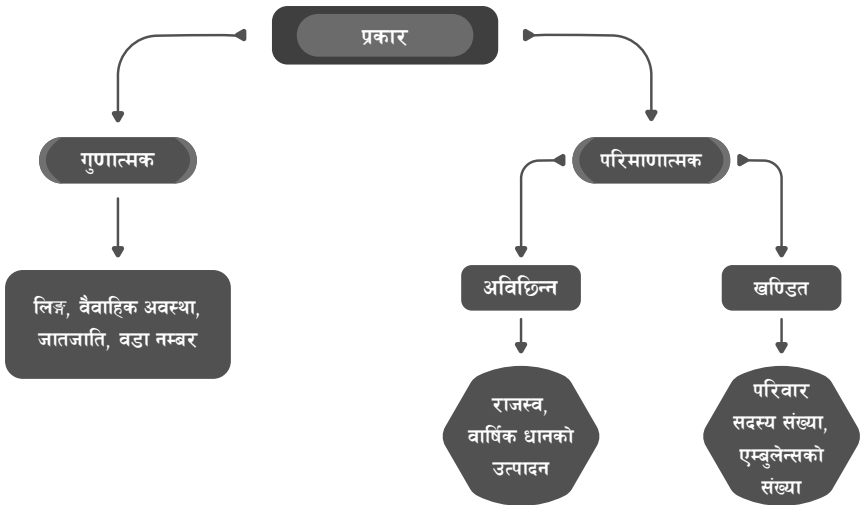
चरको परिचय तथा

मापन एकाइ सम्बन्धी अवधारणा

५.१ चर (Variable) को परिचय

हरेक वस्तु, व्यक्ति, समय, परिस्थिति, आदिअनुसार विभिन्न विशेषताहरू फरक-फरक हुन सक्छन् । यसरी फरक-फरक हुनसक्ने विशेषताहरूलाई चर (Variable) भनिन्छ । उदाहरणका लागि व्यक्तिगत घटना दर्ता गर्दा अभिलेख गरिएका विशेषताहरू जस्तै: जन्म मिति, लिंग, जन्मस्थान आदि चरहरू हुन् । चरलाई मुख्यतया दुई प्रकारमा विभाजन गर्न सकिन्छ । चरले कुनै पनि विशेषताको गुण वा परिमाणले केलाई प्रतिनिधित्व गरेको छ, सोही बमोजिम चरको प्रकार निर्धारण हुन्छ ।

चित्र नं. ५.१: चरका प्रकार



५.२ चरका प्रकार

कुनै पनि चरलाई कसरी मापन गरिएको छ भन्ने आधारमा चरको प्रकृति निर्धारण हुने र त्यसैअनुसार विश्लेषण विधिहरू फरक पर्ने हुनाले चरलाई मापनको आधारमा दुई भागमा वर्गीकरण गरिएको छ ।

क) गुणात्मक चर (Qualitative Variables)

ख) परिमाणात्मक चर (Quantitative Variables)

५.२.१ गुणात्मक चर

कुनै वस्तु वा व्यक्तिलाई गैर-संख्यात्मक विशेषताको रूपमा अवलोकन र अभिलेख

गरिन्छ भने त्यस्तो चरलाई गुणात्मक चर भनिन्छ । यस्ता चरहरू परिमाणात्मक हिसाबले मापन गर्न सकिँदैन । जस्तै: जातजाति, रङ्ग, लिङ्ग, आर्थिक स्थिति, आदि गुणात्मक चरका उदाहरण हुन् । नाम मात्रको मापन तह र क्रमिक तहमा मापन गरिएका चरहरू गुणात्मक चर हुन्छन् ।

५.२.२ परिमाणात्मक चर

कुनै पनि चरलाई यदि संख्यात्मक रूपमा अभिलेख गर्न सकिन्छ भने त्यसलाई परिमाणात्मक चर भनिन्छ । अर्थात्, परिमाण अथवा मानमा मापन गर्न सकिने चरलाई परिमाणात्मक चर भनिन्छ । जस्तै: लम्बाइ, चौडाइ, तौल, आमदानी, खर्च, क्षेत्रफल, आदि । अन्तराल (Interval) र अनुपात (Ratio) तहमा मापन गरिएका चरहरू परिमाणात्मक चर हुन्छन् । यस किसिमका चरहरूलाई खण्डित (Discrete) र अविच्छिन्न (Continuous) गरी पुनः दुई भागमा विभाजन गर्न सकिन्छ । खण्डित चरहरूले निश्चित मानहरू मात्र लिएका हुन्छन् र ती मानहरू बिच अन्तर हुन्छ । उदाहरणका लागि परिवार सदस्य संख्या, गाईको संख्या आदि । अविच्छिन्न चरले कुनै निश्चित दायराभित्र कुनै पनि मान लिन सक्छ । उदाहरणका लागि कुनै गाउँपालिकाको मासिक राजस्व संकलन, बढागत मकैको उत्पादन आदि । माथि चित्र नं. ५.१ मा चरका प्रकार र तिनीहरूका केही उदाहरण दिइएको छ ।

५.३ मापन भनेको के हो ?

तथ्याङ्कलाई विभिन्न स्तरमा मापन गरिएको हुन्छ र तिनीहरूलाई सोहीबमोजिम वर्गीकरण गर्न सकिन्छ । तसर्थ, मापन र मापनको तहका बारेमा जानकारी राख्नुपर्ने हुन्छ । कुनै पनि वस्तुको विशेषता जनाउन वा चिन्न वा अरूसँग तुलना गर्न सजिलो हुने हिसाबले विशेषताअनुसार वा अन्य कुनै आधारमा कुनै सो वस्तुलाई अङ्क, अक्षर वा चिन्ह उपलब्ध गराउनुलाई मापन भनिन्छ । कुनै व्यक्तिको उमेर २३ वर्ष छ भन्दाखेरि २३ ले उसको उमेर विशेषतालाई जनाएको छ । र, ऊ पुरुष हो भन्दा उसको लैङ्गिक विशेषतालाई जनाएको छ । मापन सँधै अङ्क वा मानमा मात्रै हुन्छ भन्ने हुँदैन । कोही व्यक्तिको नाम पनि उसलाई चिनाउने एउटा मापन नै हो । तसर्थ, कुनै विशेषताको परिमाण वा उपस्थितिलाई प्रतिनिधित्व गर्ने मान नै मापन हो । कुनै पनि चरको प्रकृति कस्तो हुने भन्ने कुरा त्यसलाई कसरी मापन गरिएको छ भन्ने कुरामा भर पर्दछ । तथ्याङ्कीय कार्यमा मापनको ठुलो महत्त्व हुन्छ किनभने चरहरूको मापनको तहअनुसार थप गणितीय क्रिया गर्न सकिने वा नसकिने भन्ने कुराको निर्धारण हुन्छ । त्यसको लागि पहिला मापन एकाइ (Measurement Scales) को बारेमा जानकारी राख्नु जरूरी हुन्छ ।

५.३.१ नाममात्रको तहका तथ्याङ्क (Nominal - Level Data)

मापनको नाममात्र तहमा अभिलेख गरिएको तथ्याङ्कलाई नदोहोरिने गरी लेबल (Labels) वा नामको रूपमा प्रतिनिधित्व गरिन्छ। उदाहरणका लागि सहरी क्षेत्र र ग्रामीण क्षेत्रमा पर्ने घरहरू, महिला, पुरुष र अन्य लिङ्गी मानिसको समूह, जातजातिको समूह आदि। यस किसिमको वर्गीकरणमा परिमाणको कुनै स्थान हुँदैन। यिनीहरूलाई सानो-ठुलो, राम्रो-नराम्रो भनेर भन्न सकिन्छ। तसर्थ, तिनीहरूको कुनै क्रमबद्धता हुँदैन। तिनीहरूलाई केवल वर्गीकरण र गणनामात्र गर्न सकिन्छ। यस्ता चरहरूलाई थप गणितीय क्रियाहरू गर्न पनि सकिँदैन। थप उदाहरणका लागि तल तालिका ५.१ मा ताप्लेजुङ जिल्लाका विभिन्न गाउँपालिकाहरूको कोड दिइएको छ। यी कोडहरू संख्यामा भए तापनि नाममात्रका हुन्छन्। यिनीहरूलाई सानो-ठुलो भन्न वा जोड, घटाउ, गुणा, भाग केही गर्न मिल्दैन।

तालिका ५.१: ताप्लेजुङ जिल्लाका विभिन्न गाउँपालिकाहरूको कोड

गाउँपालिकाको नाम	कोड
पक्ताङ्लुङ गाउँपालिका	१०,१०१
मिक्वाखोला गाउँपालिका	१०,१०२
मेरिङ्गदेन गाउँपालिका	१०,१०३
मैवाखोला गाउँपालिका	१०,१०४
आठराई त्रिवेणी गाउँपालिका	१०,१०५
फुङ्डीलङ नगरपालिका	१०,१०६
पाथीभरा याङवरक गाउँपालिका	१०,१०७
सिरीजङ्गा गाउँपालिका	१०,१०८
सिदिङ्गवा गाउँपालिका	१०,१०९

स्रोत: केन्द्रीय तथ्याङ्क विभाग^२

५.३.२ क्रमिक तहका तथ्याङ्क (Ordinal - Level Data)

मापनको क्रमिक तहमा अभिलेख गरिएको तथ्याङ्क परिभाषित विशेषता वा गुणात्मक चरमा आधारित वस्तुहरूको सापेक्ष श्रेणी वा मूल्याङ्कनमा आधारित हुन्छ। मापनको यस स्तरमा आधारित चरहरू मात्र क्रमबद्ध वा गणना गर्न सकिन्छ। उदाहरणका लागि उत्कृष्ट १० विद्यार्थीहरूको सूची, सापेक्ष गरिबीको तहअनुसार वडाहरूको क्रमबद्धता आदि। तसर्थ, क्रमिक स्केलमा मापन गरिएका चरहरूले सानो-ठुलो, तल-माथि जस्ता कुराहरू

2 Central Bureau of Statistics. *Geographical Codes*. Kathmandu, Nepal; Central Bureau of Statistic. Retrieved on November 18, 2021 from: <https://cbs.gov.np/geographical-codes/>

देखाउने गर्दछन् । यसबाट पनि कुनै परिमाणजन्य विशेषता हेर्न नसकिएता पनि तह छुट्टयाउन भने सकिन्छ । तालिका ५.२ मा दिइएको उपलब्धिको क्रममा उल्लेख भएका संख्याहरूले सम्बन्धित स्थानीय तहको क्रमबद्धता प्रतिनिधित्व गरेका छन् । यी अङ्कहरूको पनि जोड, घटाउ, गुणा, भाग गर्न सकिँदैन तर यिनीहरूले तुलनात्मक रूपमा एटा भन्दा अर्को अब्बल वा कमसल भनेर देखाउँदछन् ।

तालिका ५.२: दार्चुला जिल्ला भित्रका स्थानीय तहको विद्यालयको उपलब्धिको क्रम

स्थानीय तहको नाम	उपलब्धिको क्रम
महाकाली नगरपालिका	१
मालिकार्जुन गाउँपालिका	२
मार्मा गाउँपालिका	३
शैल्यशिखर नगरपालिका	४
ब्यास गाउँपालिका	५
दुहुँ गाउँपालिका	६
लेकम गाउँपालिका	७
नौगाड गाउँपालिका	८
अपिहिमाल गाउँपालिका	९

स्रोत: सामुदायिक माध्यमिक विद्यालय कार्यसम्पादन परीक्षण, २०७५/७६, शैक्षिक गुणस्तर परीक्षण केन्द्र, २०७७^३

५.३.३ अन्तराल तहका तथ्याङ्क (Interval - Level Data)

परिमाणमा मापन गर्न सकिने तर वास्तविक शून्य नभएका विशेषता मापन गर्न अन्तराल स्केलको प्रयोग गरिन्छ । जस्तै: तापक्रम । तापक्रमक वास्तविक शून्य हुँदैन । शून्य डिग्री सेल्सियस भनेको तापक्रम नै नभएको भन्ने होइन । यदि फरेनहाइट एकाइमा मापन गरियो भने शून्य डिग्री सेल्सियस बराबर ३२ डिग्री फरेनहाइट हुन्छ । दुईथरी एकाइमा मापन गर्दा दुईथरी शून्य फरक-फरक देखिएकोले वास्तविक शून्य नभएको हो । अन्तराल स्केलमा मापन गरिएको दुईवटा परिमाणलाई दौंज्दा अन्तराल (Interval/Difference) द्वारा मात्रै दौंज्न सकिन्छ । जस्तै: एक ठाउँको तापक्रम २० डिग्री सेल्सियस छ, त्यस्तै अर्को ठाउँको तापक्रम ४० डिग्री सेल्सियस छ भने, यी दुई ठाउँको तापक्रम दौंज्दा पहिलो

3 Education Review Office. (2020). *A Wholistic Report on Performance Audit of Community Schools, 2020*. Sanothimi, Bhaktapur; Education Review Office. Retrieved on November 18, 2021 from: https://www.ero.gov.np/post/6_60410f410d0d4

ठाउँको तुलनामा दोस्रो ठाउँको तापक्रम २० डिग्री सेल्सियसले बढी छ भनेर दाँजिन्छ तर पहिलो ठाउँको तुलनामा दोस्रो ठाउँको तापक्रम दुई गुणा छ भन्न मिल्दैन । यसबाट मापन गरिएका परिमाणहरूलाई जोड/घटाउ गर्न मिल्छ तर गुणन/भाग गर्न सकिन्न ।

५.३.४ अनुपात तहका तथ्याङ्क (Ratio - Level Data)

परिमाणमा मापन गर्न सकिने र वास्तविक शून्य भएका विशेषता मापन गर्न अनुपात स्केलको प्रयोग गरिन्छ । जस्तै: बिरामी संख्या, लम्बाइ, उचाइ, क्षेत्रफल, तौल, आमदानी, खर्च, आदि । यस स्केलमा अध्ययन गर्न लागिएका चरहरूलाई गुणन, भाग, जोड वा घटाऊ जेसुकै क्रिया गर्न पनि सकिन्छ । दुईवटा परिमाणहरूलाई दाँज्नु पर्दा अन्तराल वा अनुपात दुवैबाट दाँज्न सकिन्छ । जस्तै: कुनै एक व्यक्तिको मासिक आमदानी रू. १०,०००/- छ, त्यस्तै अर्को एक व्यक्तिको मासिक आमदानी रू. २०,०००/- छ भने, यी दुई व्यक्तिको आमदानी दाँज्नु पर्दा पहिलो व्यक्तिको तुलनामा दोस्रो व्यक्तिको मासिक आमदानी रू १०,०००/- ले बढी छ वा दुई गुणा छ भन्न सकिन्छ । तालिका ५.३ मा रामेछाप जिल्लाका स्थानीय तहमा बहिरंगमा आएका विभिन्न रोगीहरूको संख्या प्रस्तुत गरिएको छ । यी संख्याहरू पनि अनुपात तहका तथ्याङ्क हुन् । उक्त तालिकाबाट रामेछाप जिल्लामा आर्थिक वर्ष २०७६/७७ मा कुल हैजाका बिरामी कति आए भनेर निकाल्न सकिन्छ । त्यस्तै, उमाकुण्ड गाँउपालिकामा हैजाका भन्दा भाडापखालाका बिरामी ५३९ जनाले बढी छन् वा २७० गुणाभन्दा बढी छन् भनेर विश्लेषण गर्न सकिन्छ । त्यसै गरी, मन्थली नगरपालिकामा हैजा, जुका, भाडापखाला, टाईफाइड, निमोनियाका बिरामी अन्य स्थानीय तहमा भन्दा उल्लेख्य रूपमा बढी छन् भनेर तुलनात्मक अध्ययन पनि गर्न सकिन्छ । अनुपात तहका तथ्याङ्क तथ्याङ्कीय विश्लेषणमा सबैभन्दा बढी प्रयोग हुने तथ्याङ्कहरू हुन् ।

तालिका ५.३: रामेछाप जिल्लाका स्थानीय तहमा बहिरंगमा आएका विभिन्न बिरामीहरूको संख्या

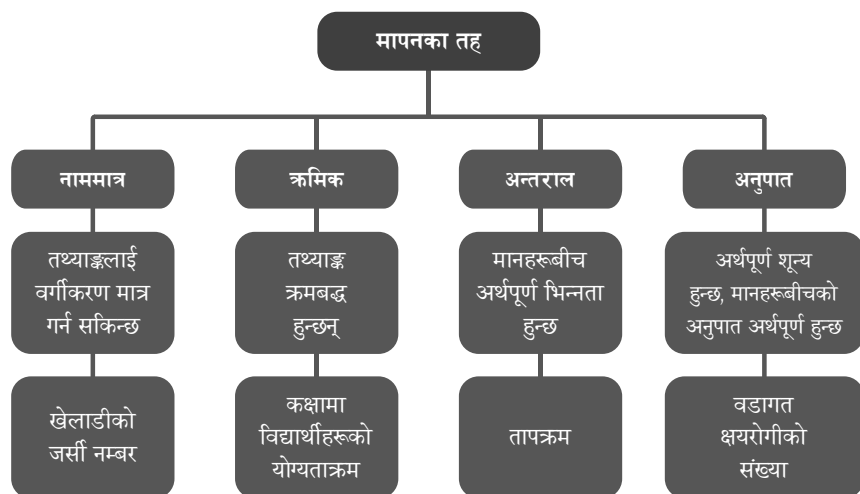
स्थानीय तहको नाम	हैजा	जुका	भाडापखाला	टाईफाइड	निमोनिया
उमाकुण्ड गाउँपालिका	२	१९२	५४१	२१३	१०६
गोकुलगङ्गा गाउँपालिका	७	१८७	५८७	३३०	१६७
लिखु तामाकोशी गाउँपालिका	३६	४४२	५६६	३८	८९
रामेछाप नगरपालिका	०	३४५	७८०	७६	१३७
मन्थली नगरपालिका	४२	१,०९७	१,६३८	९६४	३९५
खाँडादेवी गाउँपालिका	१७	५०५	९४७	११६	६६
दोरम्बा गाउँपालिका	४	३५०	६३८	४६९	२६३

सुनापती गाउँपालिका	१९	२१५	६६७	१२५	४४
--------------------	----	-----	-----	-----	----

स्रोत: वार्षिक प्रतिवेदन, स्वास्थ्य सेवा विभाग, २०७६/७७^४

संक्षेपमा मापनका चार तह, तिनको विशेषता, र उदाहरण चित्र ५.२ मा प्रस्तुत गरिएको छ । मापनको तहले चर विश्लेषण गर्न प्रयोग गर्न सकिने सांख्यिकीय विधिहरूको प्रकार निर्धारण गर्नेछ । यस पुस्तकका अन्य पाठहरूमा हामीले खासगरी अन्तराल र अनुपात तहका तथ्याङ्कहरूको बारेमा बढी चर्चा गर्नेछौं । नाममात्र तहका र क्रमिक तहका तथ्याङ्क विश्लेषणका लागि पाठकहरूलाई सन्दर्भ सामग्रीको क्र.सं. १ मा दिइएको पुस्तकको पाठ १५ र १६ अध्ययन गर्न आग्रह गरिन्छ ।

चित्र नं. ५.२: मापनको तहका लागि विशेषताहरूको सारांश र उदाहरणहरू



५.४ मापनको महत्त्व

कुनै पनि विशेषता भएका चरहरूका बारेमा अध्ययन गर्नुपूर्व ती चरहरू कुन एकाइमा मापन गरिएको छ भन्ने विषयमा पूर्ण रूपमा जानकारी राख्नुपर्दछ । तथ्याङ्क संकलनका लागि प्रश्नावली निर्माणदेखि तथ्याङ्क विश्लेषण र प्रस्तुतिसम्मका चरणमा मापन एकाइको विशेष महत्त्व हुन्छ । तथ्याङ्क विश्लेषणका लागि प्रयोग गरिने कम्प्युटर सफ्टवेयरहरूमा तथ्याङ्क प्रविष्टि गर्ने क्रममा मापन एकाइ निश्चित गर्नुपर्ने हुन्छ । मापन एकाइकै आधारमा विभिन्न चरहरूलाई कसरी प्रयोग गर्ने र के कस्ता गणितीय

4 Department of Health Services. (2021). *Annual Report 2019/20*. Kathmandu, Nepal; Ministry of Health and Population. Retrieved on November 18, 2021 from: <http://dohs.gov.np/annual-report-2076-77-2019-20/>

क्रियाद्वारा सूचकहरू निर्माण गर्न सकिन्छ भन्ने कुरा निर्धारण गर्दछ। तसर्थ, नीतिनिर्माता, योजनाकार, विकास साभेदार, विद्यार्थी एवं सञ्चार जगतले तथ्याङ्कको विश्लेषण र प्रयोग गर्दा सम्बन्धित तथ्याङ्क मापनको तहका बारेमा जानकारी राख्नुपर्दछ। उदाहरणका लागि तालिका ५.४ मा देखाइए अनुसार गत साल गाजरको मूल्य प्रति के.जी. ४० रूपैयाँले बढेको थियो। यस वर्ष गाजरको मूल्य प्रति के.जी. ६० रूपैयाँले बढ्यो। यस अवस्थामा एक के.जी. गाजरको मूल्य गत वर्षभन्दा यस वर्ष १.५ गुणाले बढी भयो भन्न मिल्दैन। किनभने एक के.जी. गाजरको मूल्य १.५ गुणा बढी भएको नभई मूल्य परिवर्तन चाहिँ १.५ गुणा बढी भएको हो। अर्थात्, गत वर्षको अन्तको मूल्य (रु ११०) भन्दा यस वर्षको अन्तको मूल्य (१८०) तुलना गर्दा रु. ७० ले बढी हो, १.५ गुणा बढी होइन। तर मूल्य वृद्धि भने पोहरभन्दा यस वर्ष ४० को तुलनामा ६० भनेको १.५ गुणाले बढी भएको हो।

तालिका ५.४: स्थानीय गाजरको मूल्य

साल (सन्)	वर्षको सुरुमा मूल्य	वर्षको अन्तमा मूल्य	परिवर्तन
२०२०	७०	११०	४०
२०२१	१२०	१८०	६०

स्रोत: कालिमाटी फलफूल तथा तरकारी बजार विकास समिति^५

५.५ मापनको प्रयोग

तथ्याङ्क संकलनका लागि तयार गरिने प्रश्नहरू तयार गर्दा नै मापनको प्रयोग गरिएको हुन्छ। गुणात्मक विशेषता (Qualitative Features) को अध्ययन गर्नुपर्दा नाम मात्रको वा क्रमिक तहका मापन एकाइ प्रयोग गरिन्छ भने परिमाणात्मक विशेषता (Quantitative Features) को अध्ययन गर्दा अन्तराल वा अनुपात तहका मापन एकाइ प्रयोग गरिन्छ। उदाहरणका लागि कुनै घटनाप्रतिको धारणा अध्ययन गर्दा क्रमिक तहका मापन एकाइ प्रयोग गरेर प्रश्नावली तयार गरिन्छ। जस्तै: वडा कार्यालयबाट प्रदान गरिने सेवाबाट कतिको सन्तुष्ट हुनुहुन्छ? यसका लागि क्रमिक तहका वैकल्पिक उत्तरहरू तयार गरिन्छ; जस्तै: १) पूर्ण सन्तुष्ट, २) सन्तुष्ट, ३) तटस्थ, ४) असन्तुष्ट, ५) निकै असन्तुष्ट।

5 Kalimati Fruits and Vegetables Market Development Board. *Kalimati Vegetable and Fruits Rate*. Kathmandu, Nepal; Kalimati Fruits and Vegetables Market Development Board. Retrieved on November 23, 2021 from: <https://kalimatimarket.gov.np/index.php/price>

त्यसैगरी तथ्याङ्क कम्प्युटरमा प्रविष्ट गर्ने बेला मापनको प्रयोग गरिन्छ भने तथ्याङ्क प्रशोधन तथा विश्लेषण गर्ने बेलामा पनि मापनको प्रयोग हुन्छ । तथ्याङ्कको प्रयोग गर्दा समेत मापन तहको प्रयोग गर्नुपर्दछ । तसर्थ, संघ, प्रदेश, स्थानीय तहका नीतिनिर्माता, योजनाकार, सरोकारवाला, विद्यार्थीदेखि सञ्चार जगतसम्मले तथ्याङ्क मापनको तहलाई विशेष ध्यान दिनुपर्दछ ।

५.६ अभ्यासका लागि प्रश्नहरू

क) खाली ठाउँ भर्नुहोस् ।

१) परिवारको मासिक आमदानीतहको मापन हो ।

२) चर प्रकारका हुन्छन् ।

३) गुणात्मक विशेषताको अध्ययन गर्नुपर्दा मापन एकाइ प्रयोग गरिन्छ ।

४) मापनका तह हुन्छन् ।

ख) निम्न शब्दावलीहरूलाई उदाहरणसहित परिभाषित गर्नुहोस् ।

१) चर

२) मापन

ग) क्रमिक तहका मापन जोडा मिलाउनुहोस् ।

नाम मात्र तहको मापन

विद्यार्थीको तौल

क्रमिक तहको मापन

फुटबल खेलाडीको जर्सी नम्बर

अन्तराल तहको मापन

मानव विकास सूचकाङ्कमा देशहरूको स्थान

अनुपात तहको मापन

बिरामीको ज्वरोको तापक्रम

घ) अविच्छिन्न तथ्याङ्क र खण्डित तथ्याङ्कबिच फरक छुट्याउनुहोस् ।

ङ) तथ्याङ्क व्यवस्थापनमा मापनको महत्त्व उल्लेख गर्नुहोस् ।

च) चरका प्रकार र मापनका तहबिचको सम्बन्धका बारेमा चर्चा गर्नुहोस् ।

छ) तथ्याङ्क व्यवस्थापनमा मापनको प्रयोग कसरी भएको हुन्छ, व्याख्या गर्नुहोस् ।

५.७ सन्दर्भ सामग्री

चर तथा मापनसम्बन्धी थप अध्ययनका लागि निम्नानुसार सामग्रीहरू उपयोगी छन् ।

१. Statistical Techniques in Business & Economics by

Douglas A. Lind, William G. Marchal & Samuel A. Wathen,
Seventeenth Edition, McGraw-Hill Education, 2018

२. Statistics and Probability, Concepts and Techniques by

Hridaya B. Shrestha, Second Edition, Ekta Books Distributors
Pvt. Ltd., 2008

तथ्याङ्कका स्रोत तथा तथ्याङ्क व्यवस्थापन

६.१ परिचय

तथ्याङ्क व्यवस्थापनअन्तर्गत तथ्याङ्क संकलन, एकीकरण, प्रशोधन, विश्लेषण, प्रस्तुति, व्याख्या, प्रसार तथा अभिलेख सम्मका कार्यहरू पर्दछन् । तथ्याङ्क संकलन कार्य तथ्याङ्क व्यवस्थापनको पहिलो र प्रमुख कार्य हो । गुणस्तरीय तथ्याङ्क प्राप्तिका लागि यसको संकलन सही तरिकाले हुनु जरूरी छ । जसका लागि उचित स्रोतहरूको पहिचान गर्नुपर्दछ । तथ्याङ्क संकलनको उद्देश्य तथा लागत दुवै पक्षलाई मध्यनजर गरेर सही स्रोतको पहिचान हुन सकेमा मात्र उद्देश्य अनुरूपको तथ्याङ्क प्राप्त हुन्छ र नीति निर्माणदेखि अनुगमन मूल्याङ्कन सम्मका कार्यहरू प्रभावकारी हुन्छन् । तसर्थ, यस पाठमा हामी तथ्याङ्कका स्रोत तथा तथ्याङ्क व्यवस्थापनका अन्य केही पक्षहरूका बारेमा चर्चा गर्नेछौं ।

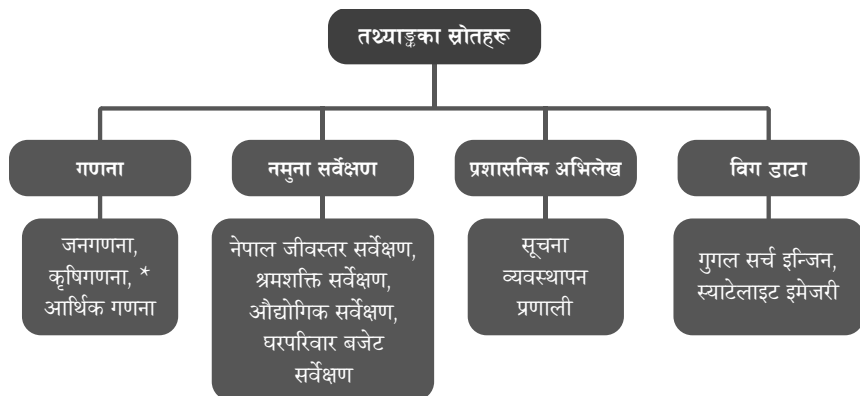
६.२ तथ्याङ्कका स्रोत

तथ्याङ्कका स्रोत भन्नाले तथ्याङ्क उपलब्ध हुने माध्यम वा स्थानलाई बुझाउँछ । तथ्याङ्क संकलनको उद्देश्य र क्षेत्र तय भइसकेपछिको अर्को चरणमा तथ्याङ्कको स्रोत पहिचान गरिन्छ । तथ्याङ्कप्राप्तिका हिसाबले तथ्याङ्कलाई मुख्यगरी दुई प्रकारमा विभाजन गर्न सकिन्छ: १) प्राथमिक तथ्याङ्क (Primary Data) र २) माध्यमिक तथ्याङ्क (Secondary Data) । अन्वेषक वा अनुसन्धानकर्ता आफैँले संकलन गरेको तथ्याङ्कलाई प्राथमिक तथ्याङ्क र स्रोतलाई प्राथमिक स्रोत भनिन्छ भने अरू कसैले संकलन गरेको तथ्याङ्क पुनः प्रयोगका लागि संकलन गरिएको तथ्याङ्कलाई माध्यमिक तथ्याङ्क र त्यस्ता स्रोतलाई माध्यमिक स्रोत भनिन्छ । प्रकाशित र अप्रकाशित गरी माध्यमिक तथ्याङ्क दुई प्रकारका हुन्छन् । उदाहरणका लागि केन्द्रीय तथ्याङ्क विभागले प्रकाशन गरेका राष्ट्रिय लेखा तथ्याङ्कहरू^६ अन्य कोही अन्वेषकले प्रयोग गर्दछन् भने सम्बन्धित अन्वेषकका लागि सो तथ्याङ्क प्रकाशित माध्यमिक तथ्याङ्क हुन्छ । यस्ता प्रकाशित आधिकारिक तथ्याङ्कका केही स्रोतहरू पाठ १ को तालिका १.१ का निकायहरूका प्रतिवेदनहरू वा तिनीहरूका वेबसाइटमा राखिएका तथ्याङ्कहरू हुन् । त्यसैगरी राष्ट्रिय तथा अन्तर्राष्ट्रिय संघसंस्थाहरूले पनि कतिपय तथ्याङ्कहरू प्रकाशन गरेर उपलब्ध गराएका हुन्छन् । त्यस्तै कतिपय तथ्याङ्क उत्पादकहरूले उत्पादन गरेका

6 Central Bureau of Statistics. *National Accounts*. Kathmandu, Nepal; Central Bureau of Statistics. Retrieved on November 18, 2021 from: <https://cbs.gov.np/national-accounts/>

तथ्याङ्क आन्तरिक प्रयोजनका लागि प्रकाशन नगरी राखेका हुन्छन्⁷ र त्यस्ता तथ्याङ्कलाई अन्य अन्वेषकले प्रयोग गर्दछन् भने त्यस्ता अन्वेषकका लागि उक्त तथ्याङ्क अप्रकाशित माध्यमिक तथ्याङ्क हुन्छ। तथ्याङ्कका स्रोतहरूका बारेमा चित्र ६.१ मा प्रस्तुत गरिएको छ।

चित्र नं. ६.१: तथ्याङ्कका स्रोतहरू



नोट: *नेपालको कृषिगणना तुलो स्केलको नमुना सर्वेक्षण नै हो तथापि कृषिगणना भन्ने गरिन्छ।

६.२.१ गणना (Census)

अध्ययन गर्न खोजिएको क्षेत्रभित्रका सम्पूर्ण वस्तु वा वस्तुहरूको समूहको विशेषता गणना गरिने कार्यलाई गणना भनिन्छ। जस्तै: नेपालमा अक्सर बसोबास गर्ने मानिसहरूको गणना गर्ने कार्यलाई जनगणना भनिन्छ, नेपालमा सञ्चालित सम्पूर्ण आर्थिक गतिविधि सञ्चालन गर्ने एकाइहरूको गणना गर्ने कार्यलाई आर्थिक गणना भनिन्छ। लागत (समय, पैसा, मानवस्रोत) का हिसाबले गणना निकै खर्चिलो हुन्छ, तर जनसंख्यासँग सम्बन्धित विशेषताहरू यथार्थमा आधारित हुन्छन्।

7 Sources for international statistics: The World Bank. *World Development Indicators*. Retrieved from: <https://databank.worldbank.org/source/world-development-indicators>; International Monetary Fund. *IMF DATA Access to Macroeconomic and Financial Data*. Retrieved from: <https://data.imf.org/?sk=388dfa60-1d26-4ade-b505-a05a558d9a42>; International Labour Organization. *Statistics and Databases*. Retrieved from: <https://www.ilo.org/global/statistics-and-databases/lang-en/index.htm>; CEIC. *Accurate Macro and Micro Economic Data*. Retrieved from: <https://www.ceicdata.com/en>; Economic Research Federal Reserve Bank of St. Louis. *FRED Economic Data*. Retrieved from: <https://fred.stlouisfed.org/>; University of Groningen. *Groningen Growth and Development Centre*. Retrieved from: <https://www.rug.nl/ggdc/productivity/pwt/>; Science direct. *Science Direct*. Retrieved from: <https://www.sciencedirect.com/>

६.२.२ नमुना सर्वेक्षण (Sample Survey)

अध्ययन गर्न खोजिएको जनसंख्यालाई उचित प्रतिनिधित्व गर्ने गरी सानो आकारको उपसमूह नमुना विधि (Sampling Method) बाट छनोट गरी जनसंख्याको विशेषताका बारेमा अनुमान गर्ने विधिलाई नमुना सर्वेक्षण भनिन्छ । हाम्रो दैनिक जीवनमा नमुना सर्वेक्षणको प्रयोग निकै भइरहेको हुन्छ । जस्तै: तरकारी पकाउँदा नुन लागे नलागेको तरकारीको सानो नमुना चाखेर, एक बोरा चामलको गुणस्तर एक अञ्जुली चामल जाँचेर आदि कार्यमा नमुना सर्वेक्षणको प्रयोग गरिरहेका हुन्छौं । तथ्याङ्कीय क्रियाकलापमा नेपालमा सञ्चालित केही नमुना सर्वेक्षणहरूको उदाहरणमा नेपाल जीवनस्तर सर्वेक्षण, श्रमशक्ति सर्वेक्षण, औद्योगिक सर्वेक्षण, घरपरिवार बजेट सर्वेक्षण, नेपाल स्वास्थ्य तथा जनसांख्यिक सर्वेक्षण आदि पर्दछन् । कृषिगणना पनि नमुना सर्वेक्षणमै आधारित सर्वेक्षण हो तथापि यसलाई गणना भन्ने गरिन्छ । लागतका हिसाबले गणनाभन्दा नमुना सर्वेक्षण निकै किफायती हुन्छ, तर जनसंख्यासँग सम्बन्धित विशेषताहरू अनुमानमा आधारित हुन्छन् । यस किसिमका अनुमानहरूमा नमुनासँग सम्बन्धित भ्रमांशहरू (Errors) समावेश भएका कारण जनसंख्याको विशेषतासँग नमुना सर्वेक्षणको अनुमानले दिएका विशेषता पूर्ण रूपमा मेल नखान सक्दछन् । कतिपय अध्ययन भने नमुना सर्वेक्षण गर्नुको विकल्प नै हुँदैन । जस्तै: हाम्रो शरीरमा भएको रगत परीक्षण गर्नुपर्‍यो भने नमुना विधिमा नै भर पर्नुपर्दछ ।

६.२.३ प्रशासनिक अभिलेख (Administrative Records)

स्थानीय तहदेखि प्रदेश, संघसम्मका विभिन्न सेवा प्रदायक निकायहरूले सेवा प्रवाह गरिरहँदा त्यस सम्बन्धमा कागजी वा विद्युतीय अभिलेख राखेका हुन्छन् । जस्तै: वडा कार्यालयमा जन्म दर्ताको काम हुँदा होस् वा मालपोत कार्यालयमा जग्गाको लिखत पास गर्दा त्यहाँ विभिन्न विवरणहरूको अभिलेख स्वतः सिर्जित हुन्छ । त्यसैगरी अस्पतालमा स्वास्थ्य सेवा प्रदान गर्दा, राजस्व कार्यालयले राजस्व संकलन गर्दा सेवाप्रदायक निकायले सेवा र सेवाग्राहीसँग सम्बन्धित विभिन्न विशेषताहरूको अभिलेख कायम गरेका हुन्छन् । यस्ता अभिलेखहरू तथ्याङ्क उत्पादनका लागि निकै महत्त्वपूर्ण स्रोत मानिन्छन् । उदाहरणका लागि व्यक्तिगत घटना दर्ताको अभिलेखबाट कुनै स्थानीय तहमा वर्ष भरिमा कति जना बच्चाको जन्मदर्ता भयो, कति जनाको बसाइँसराइ दर्ता भयो भन्ने जस्ता सूचनाहरू प्राप्तिका लागि कच्चा पदार्थको रूपमा तथ्याङ्क उपलब्ध हुन्छ । प्रशासनिक अभिलेखका तथ्याङ्कहरू सेवा प्रवाहसँग सम्बन्धित हुने हुनाले समुदायमा घटेका घटनाको पूर्ण रूपले जानकारी प्राप्त नहुन पनि सक्दछ । जस्तै: कुनै वडामा वर्षभरि जन्मिएका बच्चामध्ये दर्ता गर्नेहरूको संख्यामात्र प्रशासनिक अभिलेखबाट

प्राप्त हुने भयो । त्यस्तै, कुनै स्थानीय तहमा वर्षभरिमा कोभिडबाट मृत्यु हुनेको संख्यामध्ये स्वास्थ्य संस्थामा मृत्यु हुनेको संख्या मात्र स्वास्थ्य संस्थाको प्रशासनिक अभिलेखबाट प्राप्त हुने भयो । अस्पतालबाहेक समुदायमा कोभिडबाट मृत्यु हुनेको संख्या स्वास्थ्य संस्थाको प्रशासनिक अभिलेखबाट प्राप्त हुन सक्दैन । कतिपय तथ्याङ्कहरू जस्तै: राजस्व संकलन, सवारी साधनको कुल आयातजस्ता सूचनाहरू भने प्रशासनिक अभिलेखबाट पूर्णरूपमा प्राप्त हुन्छ । व्यक्तिगत घटना दर्ताको पूर्ण कार्यान्वयन, स्वास्थ्य सेवाको पहुँच तथा उपयोगमा विस्तारजस्ता कार्यले सम्बन्धित प्रशासनिक अभिलेखलाई भरपर्दो तथ्याङ्कको स्रोतका रूपमा स्थापित गराउन सकिन्छ । प्रशासनिक अभिलेखलाई सूचनाको स्रोतको रूपमा विकास गर्नका लागि कतिपय सेवासँग सम्बन्धि सूचना व्यवस्थापन प्रणाली सञ्चालनमा रहेका छन् । उदाहरणका लागि स्वास्थ्य व्यवस्थापन सूचना प्रणाली, शिक्षा व्यवस्थापन सूचना प्रणाली लगायतका सूचना प्रणालीहरूलाई तथ्याङ्कको वैकल्पिक स्रोतका रूपमा लिइन्छ ।

६.२.४ बिग डाटा (Big Data)

बिग डेटालाई तथ्याङ्कको वैकल्पिक स्रोतका रूपमा लिइन्छ । आयाम, विविधता र वेगका आधारमा अत्यन्त ठुलो र जटिल प्रकृतिको तथ्याङ्क भण्डारणलाई बिग डाटा भन्न गरिन्छ । सामान्य र चल्तीका कतिपय कम्प्युटर सफ्टवेयरहरूले समेत यस्ता तथ्याङ्कको प्रशोधन गर्न असमर्थ हुन्छन् । प्रविधिको प्रयोग जस्तै: टेलिफोन, सामाजिक सञ्जाल, इमेल, अनलाइन व्यापार, अनलाइन सर्च इन्जिन आदिको प्रयोग गर्दा ठुलो आयातनको, अति शीघ्र रूपमा र विविधतायुक्त (3 Vs: Volume, Velocity, and Variety) तथ्याङ्क उत्पादन भइरहेको हुन्छ । त्यसैगरी भू-उपग्रहबाट प्राप्त हुने नक्सा (Satellite Imagery) पनि बिग डाटाकै उदाहरण हो । तसर्थ, बिग डाटा संख्याका रूपमा, ध्वनिका रूपमा, नक्साका रूपमा वा अन्य यस्तै विविध रूपमा उपलब्ध हुन्छन् । यस्ता तथ्याङ्कहरूको उत्खनन, प्रशोधन, विश्लेषण, व्याख्या, एवं प्रसार गर्ने कार्य बिग डाटा अन्तर्गतका कार्य पर्दछन् । बिग डाटाको प्रयोगबाट पर्यटनसम्बन्धी, रोजगारीको अवस्था, उपभोक्ताको किनमेल सम्बन्धित बानी, पृथ्वीमा भएका विभिन्न गतिविधि, सवारी साधन सेयर गर्ने कम्पनीहरूका लागि प्रयोगकर्ताका माग, स्वास्थ्य अवस्था आदि सूचनाहरू प्राप्त गर्न सकिन्छ । उदाहरणका लागि पर्यटकहरूले गुगल सर्चबाट होटलहरूको जानकारी खोजी गर्दैगर्दा पर्यटकका विभिन्न विशेषता गुगल डाटा डिपोजिटरीमा जम्मा भइरहेको हुन्छ । त्यसबाट समग्रमा पर्यटकहरूले कुन स्थान रूचाउँछन् भन्ने जानकारी लिन सकिन्छ । त्यस्तै मानिसहरूले जागिर खोज्नका लागि गुगल सर्च गर्दै गर्दा उनीहरूको विशेषता गुगल डाटा डिपोजिटरीमा जम्मा भएको हुन्छ र यसबाट बेरोजगारीको अवस्था आँकलन गर्न सकिन्छ ।

यसका लागि उच्च दक्ष जनशक्ति र विकसित तथ्याङ्कीय पूर्वाधारको आवश्यकता पर्दछ । नेपालको सन्दर्भमा पनि बिग डाटालाई तथ्याङ्कको कम खर्चिलो वैकल्पिक स्रोतको रूपमा स्थापित गर्न सकिने प्रवल सम्भावना छ । त्यसका लागि बिग डाटासँग सम्बन्धित पूर्वाधार र जनशक्ति उत्पादनमा ध्यान दिनुपर्ने हुन्छ ।

६.३ तथ्याङ्क संकलन

सूचना चक्रको पहिलो कार्य भनेको तथ्याङ्क संकलन हो (चित्र नं. ६.२) । तथ्याङ्कका विभिन्न स्रोतबाट मापन गरिएका विवरणहरूलाई कागजी वा विद्युतीय माध्यममा स्थानान्तरण गर्ने कार्यलाई तथ्याङ्क संकलन भनिन्छ । तथ्याङ्क संकलनका लागि विभिन्न विधिहरू अबलम्बन गरिएको पाइन्छ । तीमध्ये केही महत्त्वपूर्ण विधिहरूका बारेमा यहाँ संक्षेपमा वर्णन गरिएको छ ।

६.३.१ प्रत्यक्ष व्यक्तिगत अन्तर्वार्ता (Direct Personal Interview)

यस विधिमा अन्वेषक स्वयंले सम्बन्धित उत्तरदातासँग प्रत्यक्ष व्यक्तिगत अन्तर्वार्ता गरेर आवश्यक विवरणहरू संकलन गर्दछन् । यस विधिबाट संकलित तथ्याङ्कमा अन्वेषक पूर्ण रूपमा विश्वस्त हुन सक्दछन् तर टूला प्रकृतिका सर्वेक्षणहरूमा यो विधि उपयुक्त हुँदैन । खासगरी शोधपत्र लेखिरहेका विद्यार्थीहरूले यो विधि प्रयोग गर्दछन् ।

६.३.२ अप्रत्यक्ष व्यक्तिगत अन्तर्वार्ता (Indirect Personal Interview)

विवरण उपलब्ध गराउनु पर्ने व्यक्तिको इच्छा नभएको अवस्था वा उक्त व्यक्तिले सही विवरण नदिने सम्भावना भएका सर्वेक्षणहरूमा यो विधि प्रयोग गरिन्छ । जस्तै: लागु पदार्थको प्रयोग गर्ने व्यक्ति, जुवातास खेल्ने व्यक्ति आदिले सोसम्बन्धी विवरण नदिन सक्दछन् । तसर्थ, सम्बन्धित व्यक्तिलाई नजिकबाट चिनेजानेको जिम्मेवार व्यक्तिबाट मौखिक अन्तर्वार्ता गरी तथ्याङ्क संकलन गर्ने प्रचलन रहेको छ । यसप्रकारको तथ्याङ्क संकलन विधिलाई अप्रत्यक्ष मौखिक अन्तर्वार्ता विधि भनिन्छ । यस विधिमा उत्तरदाताको पूर्वाग्रहले तथ्याङ्कको गुणस्तरलाई प्रत्यक्ष प्रभाव पारेको हुन्छ ।

६.३.३ स्थानीय प्रतिनिधि/संवाददाताहरूबाट तथ्याङ्क संकलन

यस विधिमा स्थानीयस्तरमा प्रतिनिधिहरू नियुक्त गरी विवरण संकलन गर्ने गरिन्छ । सञ्चारमाध्यमका संवाददाताहरूबाट विवरण संकलन गरिने कार्य यस विधिको एउटा उदाहरण हो ।

६.३.४ प्रश्नावली मेल गर्ने विधि (Mailed Questionnaire)

यस विधिमा विवरण संकलन गर्नका लागि प्रश्नावली तयार गरेर उत्तरदाताको ठेगानामा हुलाकबाट पठाइन्छ । उत्तरदाताले प्रश्नावलीका प्रश्नहरूको जवाफ लेखेर पुनः प्रेषकलाई

हुलाकबाटै फिर्ता गरिदिन्छन् । यसप्रकारको विधि शिक्षित उत्तरदाता भएको स्थानमा मात्र सम्भव हुन्छ । तथापि कतिपय उत्तरदाताले प्रश्नावलीलाई वेवास्ता गर्ने सम्भावना भने हुन्छ । नेपालमा यस विधिको खासै प्रयोग भएको पाइँदैन ।

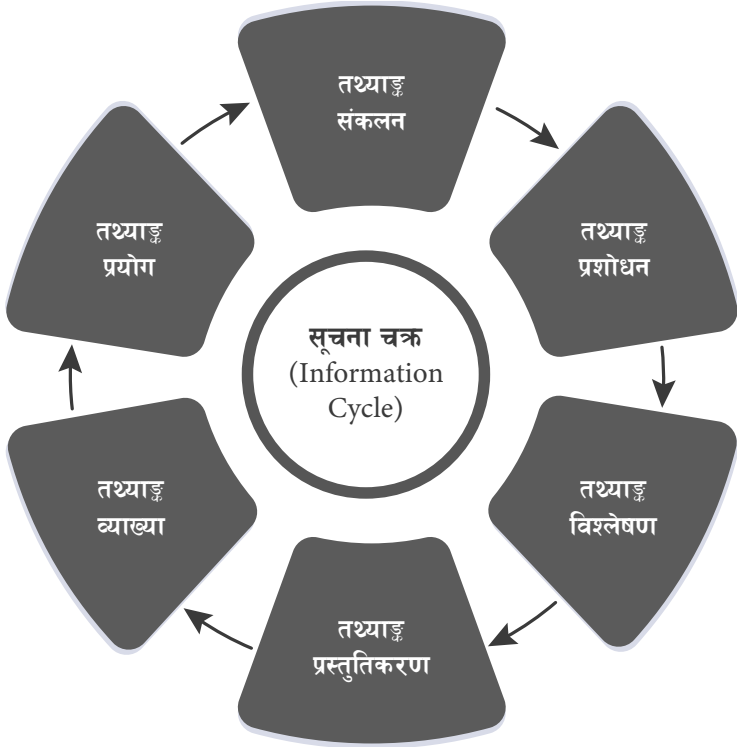
६.३.५ सेड्युल्सको माध्यमबाट गणकद्वारा तथ्याङ्क संकलन

यस विधिमा तथ्याङ्क संकलनका लागि प्रश्नावलीहरू तयार गरीएको हुन्छ । यस्ता प्रश्नावलीलाई सेड्युल्स (Schedules) भनिन्छ । तालिम प्राप्त गणकहरूद्वारा उत्तरदातालाई प्रत्यक्ष भेटेर व्यक्तिगत अन्तर्वार्ताका माध्यमबाट प्रश्नावलीमा भएबमोजिमका विवरण संकलन गर्दछन् । यस विधिमा सेड्युल्स कागजी वा विद्युतीय वा दुवै पनि हुन सक्दछन् । राष्ट्रिय जनगणना, २०७८ मा यस विधिको प्रयोग गरी काठमाण्डौँका केही स्थानमा ट्याबलेटको प्रयोगबाट र बाँकी स्थानमा कागजी प्रश्नावलीको प्रयोगबाट प्रत्येक परिवारको परिवारमूली वा परिवारको अधिक जानकारी राख्ने सदस्यसँग प्रत्यक्ष अन्तर्वार्ता गरी तथ्याङ्क संकलन गरिएको थियो । यो विधिको प्रयोग गर्दा गणकहरूले उत्तरदातालाई बुझाएर सही विवरण संकलन गर्न सक्दछन् भने अर्कोतर्फ गणकहरू पूर्वाग्रही भएमा गलत तथ्याङ्क संकलन हुने खतरा पनि रहन्छ । साथै, यो विधि लागतका हिसाबले निकै खर्चिलो विधि मानिन्छ ।

६.३.६ इन्टरनेटको प्रयोग गरी तथ्याङ्क संकलन

यस विधिमा प्रविधिको प्रयोग गरेर इन्टरनेटको माध्यमबाट प्रश्नावली भर्नका लागि उत्तरदातालाई अनुरोध गरिन्छ । उत्तरदाताले पनि तोकिएको ढाँचामा इन्टरनेटको माध्यमबाट आफ्ना विवरणहरू भरेर प्रेषण गरिदिन्छन् । तथ्याङ्क संकलनको सबैभन्दा सस्तो विधि भए तापनि इन्टरनेटको पहुँच पुगेका शिक्षित उत्तरदातासँग मात्र यो विधि प्रयोग गर्न सकिन्छ । नेपालको राष्ट्रिय जनगणना, २०७८ मा विदेशस्थित कुटनीतिक नियोगमा कार्यरत नेपाली कर्मचारीहरूको यही विधि (e-census) बाट तथ्याङ्क संकलन गरिएको थियो । आजकल कतिपय अध्येताहरूले यही विधिबाट अनुसन्धानमूलक कार्यका लागि तथ्याङ्क संकलन गरेको पनि पाइन्छ ।

चित्र नं. ६.२: सूचना चक्र



६.४ तथ्याङ्क संकलनमा प्रयोग हुने प्रश्नावली/सेड्युल निर्माण गर्दा ध्यान दिनुपर्ने कुराहरू तथ्याङ्क संकलन गर्ने विधिहरूका बारेमा हामीले माथिको छलफलबाट जानकारी प्राप्त गरिसकेका छौं । जुनसुकै विधिबाट तथ्याङ्क संकलन गर्नका लागि निश्चित ढाँचाको प्रश्नावली, सेड्युल वा अनुसूचीको आवश्यकता पर्दछ । यस किसिमका प्रश्नावली तयार गर्दा सर्वेक्षणको उद्देश्य प्राप्त हुने किसिमले विशेष ध्यान दिनुपर्ने हुन्छ । प्रश्नावली ज्यादै छोटो भएमा आवश्यक सूचनाहरू प्राप्त नहुन सक्दछन् भने ज्यादै लामो भएमा अनावश्यक खर्च हुने मात्र नभई समग्र तथ्याङ्कको गुणस्तरसमेत कमजोर हुन जान्छ । त्यसैगरी प्रश्नहरूको सिलसिलेवार नभिलेको खण्डमा प्रश्नकर्ता तथा उत्तरदाता दुवैमा द्विविधा उत्पन्न हुन्छ । यी र यस्तै प्रकारका विषयहरू समेटेर प्रभावकारी प्रश्नावली तयार गर्दा निम्नानुसारका कुराहरूमा विशेष ध्यान दिनुपर्दछ ।

क) तथ्याङ्क संकलन गर्नुको उद्देश्यप्रति अनुसन्धानकर्ता पूर्ण रूपमा स्पष्ट हुनुपर्दछ ।

ख) प्रश्नावली सीक्षित, सरल, स्वयं व्याख्यात्मक हुनुपर्दछ र यो भ्रम सिर्जना गर्ने खालको हुनुहुँदैन । प्रश्नहरू स्पष्ट, छोटो, सम्मानजनक, आदर्श प्रकृतिको हुनुपर्दछ । प्रश्नहरू

उत्तरदातालाई अनुमान गरेर जवाफ दिने सुविधारहित हुनुपर्दछ । बहुअर्थ लाग्ने किसिमका प्रश्नहरू राख्नुहुँदैन । उदाहरणका लागि तपाईंको पढाइ राम्रो छ ? भनेर सोध्ने किसिमका प्रश्न हुनुहुँदैन । सामान्यतया राम्रो, नराम्रो, दक्ष, प्रशस्त, समृद्ध, विरलै, बारम्बार, धनी, गरिबजस्ता भावार्थ आउने शब्दहरूको प्रयोग गर्नुहुँदैन । यस्ता शब्दहरू फरक-फरक व्यक्तिले फरक-फरक तरिकाले परिभाषित गर्ने गर्दछन् जसले गर्दा सर्वेक्षणको उद्देश्य बमोजिम उत्तर प्राप्त नहुन सक्दछ । कतिपय शब्दहरू जस्तै: मूल्य, सम्पत्ति, पूजी, आम्दानी, परिवार, लोकतन्त्र, समाजवादजस्ता शब्दको यथोचित परिभाषा उल्लेख गरेर मात्र प्रश्नहरूमा समावेश गर्नुपर्दछ ।

ग) संकलन गरिने विवरणबमोजिम खुला प्रश्न वा बन्द प्रश्न कस्तो प्रयोग गर्ने भन्ने विषयमा निक्यौल गर्नुपर्दछ । सामान्यतया तथ्याङ्क प्रशोधन तथा विश्लेषणका लागि बन्द प्रश्न प्रभावकारी मानिन्छ । उदाहरणका लागि “तपाईंको वैवाहिक अवस्था कस्तो हो ?” भन्ने प्रश्नलाई खुला छाड्नुभन्दा विकल्पहरू जस्तै: १) विवाहित, २) अविवाहित, ३) पारपाचुके, ४) छुट्टिएको, ५) अन्य भनेर प्रश्नावली तयार गर्दा तथ्याङ्क प्रशोधन कार्य सहज हुन्छ ।

घ) प्रश्नहरूको प्राकृतिक रूपमा सिसिलेवार मिलेको हुनुपर्दछ । उदाहरणका लागि पहिला “तपाईंको बच्चाबच्ची कति जना छन् ?” भनेर प्रश्न सोध्नु भन्दा पहिला “तपाईंको वैवाहिक अवस्था कस्तो हो ?” भनेर सोध्नु उपयुक्त हुन्छ । तपाईंसँग मोबाइल छ ? भन्ने प्रश्न नसोधी तपाईंले मोबाइल कति वर्ष पहिले किन्नु भएको हो ? भनेर प्रश्न गर्नु सान्दर्भिक मानिँदैन । यस्तो क्रम मिलाउनका लागि पहिला प्रश्नावलीको खेम्मा तयार गरेर परीक्षण गर्न सकिन्छ । परीक्षणका क्रममा प्राप्त हुने पृष्ठपोषण समेटेर प्रश्नावलीलाई अन्तिम रूप दिन सकिन्छ । साथै, पहिल्यै तयार भएका प्रश्नावलीहरू समेतलाई सन्दर्भ सामग्रीको रूपमा लिन सकिन्छ । उदाहरणका लागि परिवार सर्वेक्षण गर्नुपर्दा केन्द्रीय तथ्याङ्क विभागले सञ्चालन गरेका राष्ट्रिय जनगणना^८, नेपाल जीवनस्तर सर्वेक्षण, नेपाल श्रम शक्ति सर्वेक्षण लगायतका सर्वेक्षणहरूको प्रश्नावली हेर्न सकिन्छ ।

ङ) प्रश्नावली तयार गर्दा सही उत्तर दिन सजिलो हुने किसिमका प्रश्नहरू समावेश गर्नुपर्दछ । साथै, सामान्यतया गणितीय क्रियाको प्रयोग हुने जस्तै दर, प्रतिशत आदि सोध्नुहुँदैन ।

च) व्यक्तिगत संवेदनशील प्रश्नहरू सोध्नुहुँदैन । जस्तै: तपाईंले भाँडा माइने काम गर्नुहुन्छ ?

8 Central Bureau of Statistics. National Census 2021. Retrieved on February 2, 2021 from: <https://censusnepal.cbs.gov.np/Home/Details?tpid=6&dcid=68a1ef6-4ff6d-4c42-90b5-dfe561f251d0&tfsid=8>

भनेर प्रश्न गर्दा उत्तरदातालाई असहज हुनसक्दछ । कतिपय अवस्थामा यस किसिमका विवरण अप्रत्यक्ष रूपमा होसियारीपूर्वकसंकलन गर्नुपर्ने हुन्छ ।

छ) प्रश्नहरू प्रायः छोटो उत्तर आउने, सामान्य विकल्प छनोट गर्ने, बहुवैकल्पिक प्रश्न, र खुला प्रश्नहरू हुने गर्दछन् । कुन विवरणका लागि कुन किसिमको प्रश्न उपयुक्त हुन्छ भन्ने विषयमा विशेष ख्याल गर्नुपर्दछ ।

ज) अग्रणी प्रकृतिका प्रश्नहरू समावेश गर्नुहुँदैन । जस्तै: तपाईं साक्षर हो है ? भनेर प्रश्न सोध्नु हुँदैन । तपाईंले सामसुडको टिभी प्रयोग गर्नुहुन्छ है ? भनेर सोध्नुभन्दा तपाईंले कुन ब्राण्डको टिभी प्रयोग गर्नुहुन्छ ? भनेर सोध्नु उपयुक्त मानिन्छ ।

झ) प्रश्नावली तयार गर्दा ध्यान दिनुपर्ने अर्को कुरा के हो भने कतिपय प्रश्नहरू एक आपसमा सम्बन्धित हुन्छन् । यस्ता प्रश्नहरूले विवरणका असंगति परीक्षण गर्न सहयोग पुऱ्याउँदछन् । उदाहरणका लागि उत्तरदाताको उमेरसम्बन्धी प्रश्नका लागि निजको विवाह गर्दाको उमेर, निजको बच्चाको उमेर, निजको जन्म मिति आदि प्रश्नहरू सहयोगी प्रश्न हुन सक्दछन् । यस किसिमका प्रश्नहरू समावेश गर्नु तथ्याङ्कको गुणस्तर अभिवृद्धि गर्ने प्रयोजनका लागि राम्रो मानिन्छ ।

ञ) प्रश्नावलीको ड्राफ्ट तयार भएपछि यसको फिल्डमा गएर पूर्व परीक्षण गर्नुपर्दछ । प्रश्नावलीको पूर्व परीक्षणबाट यसमा भएका त्रुटिहरूको पहिचान गर्ने मात्र नभई समग्र फिल्ड कार्ययोजनाको व्यवस्थापन गर्न समेत सहयोग मिल्दछ ।

६.५ तथ्याङ्क वर्गीकरण र यसका नियम

संकलन गरिएका तथ्याङ्कको प्रशोधन तथा विश्लेषण गर्न सहज बनाउनका लागि विभिन्न आयामहरूमा तथ्याङ्कको वर्गीकरण गरिन्छ । तसर्थ, तथ्याङ्कलाई साभ्ना विशेषताका आधारमा श्रेणी वा समूहमा विभाजन गर्ने कार्य नै तथ्याङ्क वर्गीकरण हो । खासगरी भौगोलिक, समय शृङ्खला, गुण विशेष, संख्यात्मक समूह आदिका आधारमा तथ्याङ्कको वर्गीकरण गरिन्छ । उदाहरणका लागि तालिका ६.१ मा रौतहट जिल्लाको पर्रोहा नगरपालिकाको २०६८ सालको जम्मा जनसंख्या ३७,४५३ लाई वडा (भौगोलिक) र लिङ्ग (गुण विशेष) अनुसार वर्गीकरण गरेर प्रस्तुत गरिएको छ ।

तालिका ६.१: रौतहट जिल्ला, पर्रौहा नगरपालिकाको जनसंख्या, २०६८

वडा नं.	जनसंख्या		
	जम्मा	पुरुष	महिला
१	२,९२०	१,४८०	१,४४०
२	३,२५३	१,६०९	१,६४४
३	५,५८६	२,८७८	२,७०८
४	४,८३७	२,४८७	२,३५०
५	३,७०५	१,८९२	१,८१३
६	४,०२८	२,१३५	१,८९३
७	४,२२९	२,१७०	२,०५९
८	३,६६२	१,८७२	१,७९०
९	५,२३३	२,६८६	२,५४७
जम्मा	३७,४५३	१९,२०९	१८,२४४

स्रोत: केन्द्रीय तथ्याङ्क विभाग, राष्ट्रिय जनगणना, २०६८^९

तथ्याङ्क वर्गीकरणका लागि खास नियम नभए तापनि निम्नानुसारका विषयमा ध्यान दिनुपर्दछ ।

क) वर्गीकरणका श्रेणी वा समूहरू स्पष्ट रूपमा परिभाषित हुनुपर्दछ ।

ख) वर्गीकरणका श्रेणी वा समूहरू एक आपसमा नखर्चिने (Non-Overlapping) र पूर्ण हुनुपर्दछ ।

ग) वर्गीकरणका श्रेणी वा समूहरू स्थिर हुनुपर्दछ ।

घ) अध्ययनको उद्देश्य अनुरूपको वर्गीकरण हुनुपर्दछ ।

ङ) आवश्यकता अनुसार वर्गीकरणका आधारहरू परिवर्तनशील हुनुपर्दछ ।

६.६ आवृत्ति वितरण

व्यक्तिगत विशेषताका आधारमा संकलन गरिएका छरिएर बसेका आँकडाहरूलाई चरका कुनै निश्चित मान दोहोरिएको पटक गनेर आवृत्ति वितरण (Frequency Distribution) बमोजिम प्रस्तुत गर्न सकिन्छ । उदाहरणका लागि हुम्लाको नाम्खा गाउँपालिकाको ५ वर्षसम्मका बच्चाहरूको जनसंख्या तालिका ६.२ मा प्रस्तुत गरिएको छ । यस तालिकामा

9 Central Bureau of Statistics. *Population Census 2011*. Kathmandu, Nepal; Central Bureau of Statistics. Retrieved on November 20, 2021 from: <https://cbs.gov.np/population-2011/>

प्रस्तुत गरिएको वितरणलाई विच्छेदन आवृत्ति वितरण (Discrete or Ungrouped Frequency Distribution) भनिन्छ। संख्यात्मक विशेषता बमोजिमको तथ्याङ्कलाई समूहीकृत आवृत्ति वितरण (Grouped Frequency Distribution) अनुसार प्रस्तुत गर्न सकिन्छ। जस्तै: उमेर समूहअनुसार जनसंख्या वितरण गर्दा समूहीकृत आवृत्ति वितरणको प्रयोग गर्न सकिन्छ। यसबाट ठुलो आकारको प्रस्तुतिलाई सरलीकरण गरी सानो आकारमा परिणत गर्न सकिन्छ। उदाहरणका लागि नाम्खा गाउँपालिकाको उमेर अनुसारको जनसंख्यालाई तालिका ६.३ मा समूहीकृत आवृत्ति वितरण बमोजिम प्रस्तुत गरिएको छ।

तालिका ६.२: नाम्खा गाउँपालिकाको ५ वर्षसम्मका बच्चाहरूको संख्या

उमेर (वर्षमा)	जनसंख्या (आवृत्ति)
०	७१
१	५१
२	८८
३	८८
४	८९
५	९९
जम्मा	४८६

स्रोत: केन्द्रीय तथ्याङ्क विभाग, राष्ट्रिय जनगणना, २०६८¹⁰

अविच्छिन्न चरको आवृत्ति वितरण तयार गर्दा समावेशी भएको र नभएको गरी दुई तरिकाले तयार गर्न सकिन्छ। जस्तै: तालिका ६.३ को पहिलो दुई कोलम (Column) को आवृत्ति वितरण समावेशी हो किनभने प्रत्येक अन्तराल समूहका तल्लो र माथिल्लो सीमाका उमेरहरू सोही अन्तराल समूहमा समावेश भएका छन्। अर्थात्, ० वर्ष र ९ वर्ष उमेर समूहका मानिसको “० देखि ९” अन्तराल समूहमा समूहीकृत भएको छ। त्यस्तै, १० वर्ष र १९ वर्षका मानिसको पनि “१० देखि १९” अन्तराल समूहमा गणना भएको छ। अर्कोतर्फ सोही तालिकाको अन्तिम दुई कोलम (Column) को आवृत्ति वितरण समावेशी नभएको हो किनभने प्रत्येक अन्तराल समूहका माथिल्लो सीमाका उमेरहरू सोही अन्तराल समूहमा समावेश भएका छैनन्। अर्थात्, ० वर्षको मानिसको गणना “०

10 National Planning Commission and Central Bureau of Statistics. *National Data Profile*. Kathmandu, Nepal; National Planning Commission and Central Bureau of Statistics. Retrieved on November 20, 2021 from: <http://nationaldata.gov.np/Home/About>

देखि १०” अन्तराल समूहमा भएको छ भने १० वर्षका मानिसको गणना “० देखि १०” अन्तराल समूहमा नभएर “१० देखि २०” अन्तराल समूहमा भएको छ ।

तालिका ६.३: नाम्खा गाउँपालिकाको १० वर्ष उमेर समूहबमोजिम जनसंख्या वितरण

समावेशी भएको		समावेशी नभएको	
उमेर समूह (वर्षमा)	जनसंख्या (आवृत्ति)	उमेर समूह (वर्षमा)	जनसंख्या (आवृत्ति)
० देखि ९	७९३	० देखि १०	७९३
१० देखि १९	६८२	१० देखि २०	६८२
२० देखि २९	७३४	२० देखि ३०	७३४
३० देखि ३९	६३७	३० देखि ४०	६३७
४० देखि ४९	४६८	४० देखि ५०	४६८
५० देखि ५९	४२१	५० देखि ६०	४२१
६० देखि ६९	३३०	६० देखि ७०	३३०
७० देखि ७९	१३४	७० देखि ८०	१३४
८० देखि ८९	५१	८० देखि ९०	५१
जम्मा	४,२५०	जम्मा	४,२५०

स्रोत: केन्द्रीय तथ्याङ्क विभाग, राष्ट्रिय जनगणना, २०६८

६.७ द्विविभाजन आवृत्ति वितरण

दुई भिन्न चरहरूलाई एउटालाई ठाडो र अर्कोलाई तेर्सो लहरमा राखी तयार गरिएको आवृत्ति वितरणलाई द्विविभाजन आवृत्ति वितरण भनिन्छ । जस्तै: तालिका ६.१ मा ठाडो लहरमा वडा नं. र तेर्सो लहरमा लिङ्गअनुसार आवृत्ति वितरण तयार गरिएको छ । यो आवृत्ति वितरण दुई नाममात्र तहमा मापन गरिएका चरहरूको द्विविभाजन आवृत्ति वितरण हो । यसैगरी अनुपात तहमा मापन गरिएका दुई चरहरूको पनि द्विविभाजन आवृत्ति वितरण गर्न सकिन्छ । तालिका ६.४ मा उमेर र घरमा बसेको महिना संख्या अनुपात तहमा मापन गरिएका दुई चरहरूको द्विविभाजन आवृत्ति वितरण तालिका हो । यो तालिकाअनुसार कुल २८,६७० जनामध्ये २० वर्षभन्दा कम उमेरका मानिसहरू वर्षमा ७ महिनामात्र घरमा बस्ने जम्मा ६१ जना छन्, सोही उमेर समूहका ८ महिनामात्र घरमा बस्ने ६७ जना छन्,, एवं रीतले तथ्याङ्क प्रस्तुत गर्न सकिन्छ ।

तालिका ६.४: उमेर समूहअनुसार एक वर्षमा घरमा बस्ने अवधि महिनामा

उमेर समूह	घरमा बसेको जम्मा महिना						जम्मा
	७	८	९	१०	११	१२	
२० वर्षभन्दा कम	६१	६७	९२	११७	१३२	१२,९२१	१३,८५४
२० देखि ४० वर्ष	६४	९९	७९	१०३	१३०	६,९०५	७,६८८
४० देखि ६० वर्ष	१४	२९	३०	५२	४७	४,७०५	४,९३६
६० देखि ८० वर्ष	३	७	६	२१	२३	१,९२८	२,००६
८० वा सोभन्दा माथि	१	१	०	१	२	१८०	१८६
जम्मा	१४३	२०३	७	२९४	३३४	२६,६३९	२८,६७०

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

६.८ तालिकीकरण (Tabulation): अर्थ, महत्त्व तथा प्रकार

तथ्याङ्कको विश्लेषण तथा व्याख्या गर्नु पूर्व तथ्याङ्कलाई सरल र सहज तरिकाले बुझ्न र बुझाउनका लागि यसको व्यवस्थित प्रस्तुतिकरण गर्ने प्रचलन छ । तालिका, चित्र, रेखाचित्र, स्तम्भचित्र आदिका माध्यमबाट तथ्याङ्कको प्रस्तुतिकरण गरिन्छ । तथ्याङ्कलाई व्यवस्थित तरिकाले प्रस्तुति गर्ने विधिहरू मध्येको एउटा महत्त्वपूर्ण तथा धेरै प्रयोग हुने विधि तालिकीकरण विधि हो । तालिकीकरणले तथ्याङ्कलाई तुलना गर्न सहज तुल्याउँछ । माथि प्रस्तुत गरिएका तालिकाहरू ६.१, ६.२, ६.३, ६.४ तथ्याङ्कको तालिकीकरणका उदाहरणहरू हुन् । तालिकाका प्रमुख अङ्गहरूमा तालिका नम्बर, तालिका शीर्षक, क्याप्सन, स्तब, मुख्य भाग, नोट, तथ्याङ्कको स्रोतहरू पर्दछन् (तालिका ६.५ हेर्नुहोस्) ।

तालिका ६.५: तालिकाको शीर्षक

स्तब शीर्षक	(शीर्षक नोट)		
	क्याप्सन		
	कोलम शीर्षक	कोलम शीर्षक	कोलम शीर्षक
	आँकडा	आँकडा	आँकडा
	आँकडा	आँकडा	आँकडा
	आँकडा	आँकडा	आँकडा
	आँकडा	आँकडा	आँकडा

नोट: (थप स्पष्ट पारुपर्ने केही भएमा मात्र)

स्रोत: केन्द्रीय तथ्याङ्क विभाग, तथ्याङ्क तथा तथ्याङ्कीय साक्षरता पुस्तक, २०७८

तथ्याङ्क व्यवस्थापनमा तालिकीकरणको ठुलो महत्त्व छ । व्यापक रूपमा फैलिएका आँकडाहरूलाई एकत्रित गरी धेरै सूचनाहरूलाई एकैपटक प्रस्तुत गर्नका लागि तालिकीकरणले मद्दत गर्दछ । यसले तथ्याङ्कको विश्लेषण तथा व्याख्या गर्नका लागि पनि सहयोग गर्दछ । साथै, दुई वा दुईभन्दा बढी चरहरूबिच विविध विषयमा तुलना गर्न र एउटै चरको पनि विभिन्न विशेषताहरूको तुलना गर्न पनि तालिकीकरणले सहजीकरण गर्दछ । तालिकामा प्रस्तुत गरिएका तथ्याङ्कहरू आकर्षक र दर्शकहरूको ध्यान केन्द्रित गर्ने किसिमका हुन्छन् । यसका अतिरिक्त तालिकाका सूचनाहरू सम्झन सजिलो हुने हुँदा यसको महत्त्व निकै बढेको हो । तालिकाको अध्ययनबाट तथ्याङ्कमा भएका भ्रमांशहरू पनि पहिचान गर्न मद्दत मिल्दछ ।

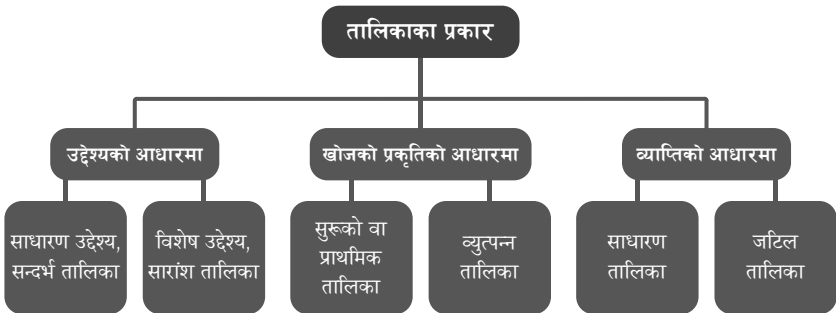
तथ्याङ्कीय तालिकाहरू विभिन्न तरिकाले तयार गरिन्छ । तालिकाको छोटो गर्दा खासगरी निम्न कुराहरूमा ध्यान दिनुपर्दछ र सोही बमोजिम तालिकाको प्रकारहरू निर्धारण गर्न सकिन्छ ।

क) खोज अनुसन्धानको उद्देश्य एवं दायरा

ख) खोज अनुसन्धानको प्रकृति (प्राथमिक वा माध्यमिक)

ग) खोज अनुसन्धानको व्याप्ति

चित्र नं. ६.३: तालिकाका प्रकार



उद्देश्यका आधारमा तालिका दुई प्रकारका हुन्छन् । साधारण उद्देश्यका सर्वेक्षण जस्तै: केन्द्रीय तथ्याङ्क विभागले जनगणना, कृषिगणनाका तथ्याङ्क प्रस्तुत गर्ने तालिकाहरू यस अन्तर्गत पर्दछन् । यस्ता तालिकाहरू अन्य अध्येताहरूले सन्दर्भ तालिकाको रूपमा धेरै प्रयोग गरेका हुन्छन् । उदाहरणका लागि राष्ट्रिय जनगणना, २०६८ को दोस्रो फारामबाट प्राप्त विवरणबमोजिम नतिजाको एउटा तालिकाको अंश तलको तालिका ६.६ मा दिइएको छ ।

तालिका ६.६: जन्मस्थान र लिङ्गानुसार जनसंख्या (संख्या लाखमा)

लिङ्ग	जम्मा जनसंख्या (लाखमा)	नेपालमै जन्मेका				विदेशमा जन्मेका	उल्लेख नभएका	
		नेपालमै जन्मेका जम्मा (लाखमा)	गृही जिल्लामा जन्मेका (लाखमा)	अन्य जिल्लामा जन्मेका				
				जम्मा (लाखमा)	ग्रामीण (लाखमा)			सहर
जम्मा	२६३	२५५	२१७	३८	३०	२	५	२
पुरुष	१२६	१२४	१०८	१६	१३	१	२	१
महिला	१३६	१३१	११०	२२	१७	१	३	१

नोट: संस्थागत परिवारको जनसंख्या समावेश छैन । स्रोत: केन्द्रीय तथ्याङ्क विभाग, राष्ट्रिय जनगणना, २०६८¹¹

11 Central Bureau of Statistics. *National Population and Housing Census 2011*. Kathmandu, Nepal; Central Bureau of Statistics. Retrieved on November 23, 2021 from: <https://cbs.gov.np/national-population-and-housing-census-2011/tables-from-form-ii/>

तालिका ६.७: उत्तरदाताको शैक्षिक स्तरबमोजिम घरपरिवारको वितरण

तथ्याङ्क विवरणको क्षेत्र	उत्तरदाताको शैक्षिक तह (%)									
	अनौपचारिक	प्रा.वि.	नि.मा.वि.	मा.वि.	एस.एल.सि.	१२ पास	स्नातक	स्नाकोत्तर	निरक्षर	जम्मा
सहरी	१७.७६	१२.४३	८.४९	६.९९	७.३८	५.५६	३.६५	१.८१	३५.९३	१००
ग्रामीण	२१.५८	१२.८१	५.७१	३.६९	४.३४	१.५५	०.७५	०.३८	४९.१९	१००
हिमाल	२६.६	९.६६	३.७३	२.९७	३.३	१.४२	०.५३	०	५१.७९	१००
पहाड	२५.२९	१६.०४	६.२८	४.९९	५.४५	२.८३	१.७१	१.०३	३६.३८	१००
तराई	१४.६२	९.९१	७.२४	४.६	५.३६	२.८४	१.६७	०.७	५३.०६	१००

स्रोत: केन्द्रीय तथ्याङ्क विभाग, राष्ट्रिय जलवायु परिवर्तनको प्रभाव सर्वेक्षण, २०१६¹²

12 Central Bureau of Statistics, *National Climate Change Impact Survey 2016*, Kathmandu, Nepal; Central Bureau of Statistics. Retrieved on November 23, 2021 from: https://cbs.gov.np/wp-content/uploads/2018/12/National-Climate-Change-Impact-Survey-2016-Report_final.pdf

विशेष उद्देश्यका सर्वेक्षणका तालिकाहरूमा विशेषगरी सूचनाका सारांशहरू प्रस्तुत गरिएको हुन्छ। यसप्रकारका तालिकाहरू विश्लेषणात्मक प्रकृतिका हुन्छन्। तुलनात्मक अध्ययन तथा चरहरूका सम्बन्ध अध्ययनका लागि यसप्रकारका तालिकाहरू प्रयोग गरिन्छ। यस किसिमका तालिकामा प्रस्तुत हुने तथ्याङ्क अनुपात, प्रतिशत आदिमा व्यक्त गरिएको हुन्छ। तालिका ६.७ विशेष उद्देश्यका सारांश तालिकाको उदाहरण हो।

त्यसैगरी खोज अनुसन्धानको आधारमा तयार गरिने तालिका पनि प्राथमिक र व्युत्पत्ति गरिएका (Derived) गरी दुई प्रकारका हुन्छन्। प्राथमिक तालिका तथ्यहरूलाई जस्ताको तस्तै प्रस्तुत गरिएको हुन्छ भने व्युत्पन्न तालिकामा सुरुका आँकडाहरूलाई गणितीय तथा तथ्याङ्कीय क्रियाको प्रयोग गरी दर, अनुपात, प्रतिशत, औसत, विस्तार आदि निकालेर प्रस्तुत गरिएको हुन्छ। कतिपय स्थानमा प्राथमिक र व्युत्पन्न दुवै मिश्रित भएको तालिका पनि प्रयोग गरिन्छ। अध्ययन क्षेत्रको व्याप्तिका आधारमा पनि तालिकाहरू साधारण र जटिल गरी दुई किसिमका हुन्छन्। चरको एउटा मात्र विशेषता प्रस्तुत गरिएको तालिका साधारण तालिका हो भने धेरै विशेषता प्रस्तुत गर्ने तालिका जटिल (Complex) तालिका हो। तालिका ६.८ साधारण तालिका हो भने तालिका ६.६, ६.७ जटिल तालिका हुन्।

तालिका ६.८: लिङ्गानुसार जनसंख्या

लिङ्ग	जनसंख्या
महिला	१,३६,४५,४६३
पुरुष	१,२८,४९,०४१
जम्मा	२,६४,९४,५०४

स्रोत: केन्द्रीय तथ्याङ्क विभाग, राष्ट्रिय जनगणना, २०६८

६.९ आउटलायर्सको परिचय

कुनै पनि तथ्याङ्कमा असामान्य रूपमा घटी वा बढी मान रहेको छ भने त्यसलाई हामी आउटलायर्स (Outliers) भन्दछौं। यस्ता असामान्य मानहरूले समग्र तथ्याङ्कको सूचनालाई प्रतिकूल प्रभाव पार्दछ। उदाहरणका लागि कुनै गाउँपालिकामा बसोबास गर्ने परिवारहरूको वार्षिक आमदानीको बारेमा अध्ययन गर्न संकलन गरिएको तथ्याङ्कमा सोही गाउँपालिकामा बसोबास गर्ने देशकै धनाढ्य व्यक्तिको पनि आमदानी संकलन भयो भने उक्त आमदानीको अङ्क अन्य परिवारको आमदानी अङ्कभन्दा निकै बढी हुन्छ र यसले तथ्याङ्कका अन्य विशेषताहरूलाई पूर्वाग्रही हिसाबले प्रभाव पार्दछ। यस किसिमका आउटलायर्सहरू पहिचान गरेर तथ्याङ्क विश्लेषण पूर्वनिश्चित विधिको प्रयोग गरी विस्थापन, प्रतिस्थापन वा अनुमानित मान राखी थप विश्लेषण गर्नुपर्दछ।

तथ्याङ्क विश्लेषणमा आउटलायर्सहरू हटाउँदा नमुना आकार (Sample Size) मा कुनै तात्विक असर पर्दैन भने हटाउन पनि सकिन्छ तर साधारणतया आउटलायर्स बाहेकका अन्य आँकडकहरूको औसत राखेर प्रतिस्थापन गर्ने चलन पनि छ । आउटलायर्सहरू पहिचान गर्नका लागि धेरै विधिहरू छन् । तीमध्ये स्क्याटर प्लट प्रयोग गरेर ग्राफिकल निरीक्षण गर्ने एउटा सामान्य विधि हो । ग्राफिकल निरीक्षण विधिबाट आउटलायर्सका बारेमा पूर्ण विश्वस्त हुनका लागि चाहिँ अरू विधिहरू प्रयोग गर्नुपर्दछ । कतिपय मानिसहरूले इन्टरक्वार्टायल रेन्ज (Interquartile Range) प्रयोग गरेर सो दायराभन्दा बाहिर पर्ने आँकडाहरूलाई आउटलायर्स मान्ने गरेको पनि पाइन्छ । कम्प्युटरमा नर्मल सम्भाव्य बक्र, बक्सप्लट जस्ता ग्राफिकल विधिबाट पनि आउटलायर्स पहिचान गरिन्छ । अध्ययनको उद्देश्य तथा नतिजाको संवेदनशीलताको आधारमा अध्येताहरूले यस बारेमा निर्णय गर्न उचित हुन्छ । जस्तै: तालिका ६.९ मा चितवन जिल्लाको अन्य स्थानीय तहमा भन्दा भरतपुर महानगरपालिकामा कडा निमोनियाका बिरामीको संख्या असामान्य धेरै देखिन्छ । चित्र नं. ६.४ को बक्सप्लटले पनि एउटा आँकडा मानलाई आउटलायर्सको रूपमा देखाएको छ । यसका पछाडि धेरै कारण हुन सक्दछन् । जस्तो कि, यहाँ ठुला अस्पताल तथा अनुभवी डाक्टर भएका कारण अन्य स्थानबाट प्रेषण गरिएका बिरामी पनि हुन सक्दछन् । जनसंख्या धेरै भएका कारण यही स्थानका बिरामी पनि हुन सक्दछन् । तर जे भए पनि स्थानीय तहअनुसार तथ्याङ्क विश्लेषण गर्ने बेलामा यस्ता असामान्य अङ्कहरूलाई विशेष ध्यान दिनुपर्ने हुन्छ । तालिका ६.९ को तथ्याङ्कको हकमा भरतपुर महानगरपालिकाबाहेक अन्य स्थानीय तहहरूको छुट्टै अध्ययन गर्न सकिन्छ । यद्यपि भरतपुरबाहेक अन्यलाई हेर्दा फेरि मादी नगरपालिका आउटलायर्स हुने सम्भावना पनि छ ।

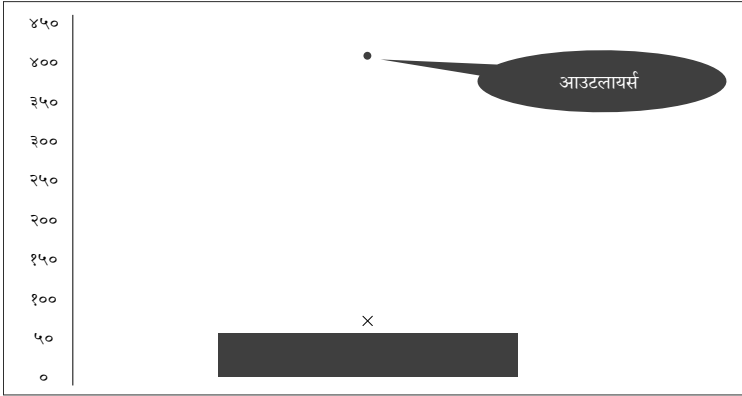
तालिका ६.९: चितवन जिल्लाको स्थानीय तहहरूमा
कडा निमोनियाका वार्षिक बिरामीको संख्या

स्थानीय तहको नाम	कडा निमोनिया बिरामीको संख्या
राप्ती नगरपालिका	१
कालिका नगरपालिका	०
इच्छाकामना गाउँपालिका	१
भरतपुर महानगरपालिका	४०८
रत्ननगर नगरपालिका	१
खैरहनी नगरपालिका	०

मादी नगरपालिका	५४
जम्मा	४६५

स्रोत: स्वास्थ्य सेवा विभाग, वार्षिक प्रतिवेदन, २०७६/७७¹³

चित्र नं. ६.४ चितवन जिल्लाको स्थानीय तहहरूमा कडा निमोनियाका वार्षिक बिरामीको संख्याको बक्सप्लट¹⁴



६.१० अभ्यासका लागि प्रश्नहरू

क) खाली ठाउँ भर्नुहोस् ।

- १) कुनै पनि अध्ययनका लागि सुरुमा संकलन गरेको तथ्याङ्कलाई भनिन्छ ।
- २) कुनै जनसंख्याको प्रतिनिधित्व हुने उप समूहलाई भनिन्छ ।
- ३),,, तथ्याङ्कका स्रोतहरू हुन् ।
- ४) तथ्याङ्कलाई विभिन्न उस्तै प्रकृतिका समूहमा विभाजन गर्ने कार्यलाई भनिन्छ ।
- ५) कुनै पनि तालिकाको लहर (Row) को शीर्षकलाई भनिन्छ ।

ख) तलका भनाइहरू ठिक वा बेठिक छुट्याउनुहोस् ।

- १) प्राथमिक तथ्याङ्कभन्दा माध्यमिक तथ्याङ्क संकलन गर्न कम खर्चिलो हुन्छ ।
- २) गणकहरूले तथ्याङ्क संकलन गर्न प्रयोग गर्ने प्रश्नावलीलाई सेड्युल भनिन्छ ।
- ३) तथ्याङ्क विश्लेषणमा आउटलायर्सहरूलाई ध्यान नदिए पनि हुन्छ ।

13 Department of Health Services. *Annual Reports*. Kathmandu, Nepal; Department of Health Services. Retrieved on November 21, 2021 from: <https://dohs.gov.np/annual-report-2076-77-2019-20/>

14 एक्सेल र स्टार्टामा बक्सप्लट तयार गर्ने विधिका बारेमा पाठ २२ र २३ मा चर्चा गरिएको छ ।

- ४) दुई वा दुईभन्दा बढी चर वा विशेषता दर्शाउने तालिकाहरूलाई सामान्य तालिका भनिन्छ ।
- ५) ठुला ठुला संख्या भएको तथ्याङ्कलाई बिग डाटा भनिन्छ ।
- ग) संक्षिप्त व्याख्या गर्नुहोस् ।
- १) नमुना
 - २) तथ्याङ्कीय एकाइ
 - ३) गणना
 - ४) बिग डाटा
 - ५) तथ्याङ्कको वर्गीकरण
- घ) प्राथमिक र माध्यमिक तथ्याङ्कबिचको फरक छुट्याउनुहोस् ।
- ङ) राष्ट्रिय जनगणना, २०७८ मा तथ्याङ्क संकलन गर्न तिनवटा विधि प्रयोग गरिएको थियो । ती विधिहरूका बारेमा संक्षेपमा चर्चा गर्नुहोस् ।

६.११ सन्दर्भ सामग्री

यस पाठ अन्तर्गतका विभिन्न विषयहरूका बारेमा थप जानकारीका लागि निम्नानुसारका सामग्री उपयोगी छन् ।

१. Fundamentals of Statistics by S.C. Gupta, Sixth Revised & Enlarged Edition, Himalaya Publishing House, 2009
२. Statistics and Probability, Concepts and Techniques by Hridaya B. Shrestha, Second Edition, Ekta Books Distributors Pvt. Ltd., 2008
३. United Nations. *UN Big Data*. Retrieved on November 23, 2021 from: <https://unstats.un.org/bigdata/>

६.१२ पदावलीहरू र तिनीहरूको अर्थ

१. अन्वेषक वा अनुसन्धानकर्ता (Investigator): कुनै निश्चित उद्देश्यका लागि तथ्याङ्क संकलन गर्ने प्रमुख व्यक्तिलाई अन्वेषक भनिन्छ ।
२. गणक (Enumerators): अन्वेषक वा अनुसन्धानकर्ताको तर्फबाट तथ्याङ्कको स्रोतबाट विभिन्न औजार तथा विधिको प्रयोग गरेर तथ्याङ्क संकलन गर्ने व्यक्तिलाई गणक भनिन्छ ।
३. उत्तरदाता (Respondents): तथ्याङ्क संकलन गर्ने गणकहरूलाई सत्य तथ्य विवरण उपलब्ध गराउने व्यक्तिलाई उत्तरदाता भनिन्छ ।

४. तथ्याङ्कीय एकाइ (Statistical Unit): कुनै पनि तथ्याङ्कीय अनुसन्धानमा मापन वका गणनाहरूसँग सम्बन्धित स्पष्ट रूपले परिभाषित वा पहिचान गर्न सकिने वस्तु वा वस्तुहरूको समूहलाई सांख्यिकीय एकाइ भनिन्छ । अर्थात् तथ्याङ्कीय एकाइ एक त्यस्तो संस्था हो जसको बारेमा सूचना खोजिएको हुन्छ र जसको लागि तथ्याङ्कहरू संकलन गरिन्छ । तथ्याङ्कीय एकाइलाई गणनाको एकाइ (Unit of Enumeration), तथ्याङ्क विश्लेषणको एकाइ (Unit of Analysis), तथ्याङ्क व्याख्या गर्ने एकाइ (Unit of Interpretation) जस्ता एकाइहरूको प्रयोग गरिएको हुन्छ । जस्तै: व्यक्ति, परिवार, वडा, गाउँपालिका आदि ।
५. जनसंख्या (Population): अध्ययन गर्न खोजिएको सम्पूर्ण वस्तु वा वस्तुहरूको समूहलाई जनसंख्या भनिन्छ । जस्तै: नेपालमा सञ्चालित उद्योगहरूको समूह, कीर्तिपुर नगरपालिका भित्र बसोबास गर्ने महिलाहरूको समूह, नेपालमा अक्सर बसोबास गर्ने मानिसहरूको समूह आदि ।
६. नमुना (Sample): कुनै पनि जनसंख्याको समूहबाट सो जनसंख्यालाई प्रतिनिधित्व गर्ने उपसमूहलाई नमुना भनिन्छ ।



तथ्याङ्क र सूचकसम्बन्धी अवधारणा

७.१ परिचय

तथ्याङ्कसम्बन्धी आधारभूत जानकारीका सम्बन्धमा हामीले पाठ ४ मा छलफल गरिसकेका छौं। यस पाठमा हामी आँकडा (Data - डाटा) लाई सूचनाको कच्चा पदार्थको रूपमा कसरी प्रयोग गरिन्छ भन्ने विषयमा थप छलफल गर्नेछौं। डाटालाई प्रशोधन गरेर सूचनाको रूप दिनुभन्दा पहिला तिनीहरूको स्वरूप वा विशेषताका बारेमा जानकारी राख्नुपर्ने हुन्छ। कतिपय चरको विशेषताको मापन पूर्ण संख्याले मात्र मापन गर्न सकिन्छ भने कतिपय विशेषता पूर्ण संख्याका साथसाथै भिन्न संख्याहरूको पनि प्रयोग गर्नुपर्दछ। जस्तै: कुनै विद्यालयमा अध्ययनरत छात्र तथा छात्राहरूको संख्या पूर्ण संख्या मात्र हुन्छ। अर्थात् ३४१, ४६४ वा यस्तै संख्या हुन्छ। तर ३४१.४, ४६४.८ जना भन्ने हुँदैन। यदि तिनै विद्यार्थीको तौल मापन गर्ने हो भने पूर्ण संख्याका साथसाथै भिन्न संख्या समेत प्रयोग गर्नुपर्ने हुन्छ। कुनै विद्यार्थीको तौल ५६ के.जी., कसैको ५८.५ के.जी. हुन सक्दछ। डाटाका यस्तै विशेषताहरूलाई समेटेर यसका प्रकारहरूका बारेमा तल चर्चा गरिएको छ। डाटाहरू मुख्यतया तिन प्रकारका हुन्छन्।

क) व्यक्तिगत डाटा (Individual Series),

ख) खण्डित डाटा (Discrete Series), र

ग) अविच्छिन्न डाटा (Continuous Series)

७.१.१ व्यक्तिगत डाटा

कुनै पनि अध्ययनका लागि छनोट भएका तथ्याङ्कीय एकाइहरूको विशेषता मापन गरी छुट्टाछुट्टै संख्याले प्रतिनिधित्व गरेर जस्ताको त्यस्तै राखिएको आँकडाहरूको समूहलाई व्यक्तिगत डाटा भनिन्छ। सामान्यतया आँकडामा थोरै संख्यामा अङ्कहरू छन् भने त्यस्तो आँकडामा कुनै विभाजनको आवश्यकता पर्दैन। जुनसुकै तथ्याङ्कीय विधिहरू (औसत गणना गर्ने, विस्तार गणना गर्ने, आदि) सिधै प्रयोग गर्न सकिन्छ। तर ठुलो संख्यामा अङ्कहरू भएको व्यक्तिगत श्रेणीलाई तथ्याङ्क विश्लेषणका लागि व्यवस्थापन गर्न असजिलो हुन्छ। कम्प्युटरका माध्यमबाट यस्ता व्यक्तिगत श्रेणीका आँकडाहरूलाई आवृत्ति वितरण तयार गरेर तालिकीकरण गर्ने गरिन्छ। उदाहरणका लागि नेपालको जीवनस्तर सर्वेक्षणमा छनोटमा परेका परिवारहरूको परिवार सदस्य संख्या छुट्टाछुट्टै संकलन गरिएको हुन्छ। ती सबै अङ्कहरू व्यक्तिगत श्रेणीका आँकडा हुन् र पछि तथ्याङ्क विश्लेषण गर्ने बेला तिनीहरूलाई वर्गीकरण गरेर विभिन्न आवृत्ति वितरण गरी तालिकीकरण गरिन्छ।

तालिका ७.१ मा ५,९८८ परिवारको परिवार संख्या आवृत्ति वितरण तालिकामा प्रस्तुत गरिएको छ। कम्प्युटरमा ती सबै परिवारको छुट्टाछुट्टै आँकडाहरू प्रविष्टि भएको हुन्छ र ती आँकडालाई व्यक्तिगत श्रेणी भनिन्छ। यहाँ ५,९८८ परिवारको छुट्टाछुट्टै आँकडाहरू प्रस्तुत गर्न स्थान अभावका कारण आवृत्ति वितरण गरी देखाइएको हो।

तालिका ७.१: नेपाल जीवनस्तर सर्वेक्षण २०१०/११ मा छनोटमा परेका परिवारहरूको परिवार सदस्य संख्याको आवृत्ति वितरण

परिवार सदस्य संख्या	जम्मा परिवार संख्या
१	२०७
२	६४६
३	९४५
४	१,२५४
५	१,०८६
६	७९८
७	४५०
८	२५७
९	१३५
१०	८१
११	४९
१२	२५
१३	१७
१४	८
१५	१२
१६	४
१७	४
१८	२
१९	५
२०	३
जम्मा	५,९८८

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

७.१.२ खण्डित डाटा

कुनै पनि वस्तुको वा तथ्याङ्कीय एकाइको विशेषता मापन गर्दा प्रयोग हुने अङ्कहरू पूर्ण संख्या (Whole Number) मात्र हुन्छन् भने त्यसरी बन्ने डाटाको समूहलाई खण्डित श्रेणी भनिन्छ । तालिका ७.१ मा दिइएको उदाहरणमा छनोट भएका परिवारहरूको परिवार सदस्य संख्याअनुसार समूह देखाइएको छ जसलाई खण्डित श्रेणी भनिन्छ । किनभने परिवारको सदस्य संख्या १ वा २ वा ३ वा ४ वा ५ वा यस्तै पूर्ण संख्या हुन्छ, ४.५, ५.३ भन्न हुँदैन । तालिका ७.१ जस्तै अन्य खण्डित श्रेणीको पनि आवृत्ति वितरणको माध्यमबाट सरलीकरण गरी अध्ययनलाई सहज बनाइन्छ ।

७.१.३ अविच्छिन्न डाटा

कुनै पनि वस्तुको वा तथ्याङ्कीय एकाइको विशेषता मापन गर्दा प्रयोग हुने अङ्कहरू पूर्ण संख्या मात्र नभएर जुनसुकै भिन्नात्मक संख्या पनि हुन्छन् भने त्यसरी बन्ने डाटाको समूहलाई अविच्छिन्न श्रेणी भनिन्छ । उदाहरणका लागि परिवारको वार्षिक खर्च रकम पूर्ण संख्याका साथसाथै भिन्न संख्या पनि हुन सक्दछ । यस्तो अवस्थामा उक्त आँकडालाई अन्तराल समूहमा विभाजन गरी आवृत्ति वितरण तालिका तयार गरेर हेर्न सकिन्छ । तालिका ७.२ मा प्रस्तुत गरिएको तथ्याङ्क अविच्छिन्न श्रेणी हो । यसमा कुनै पनि अन्तराल समूहमा पर्ने उपभोग रकम पूर्ण संख्या वा भिन्न संख्या जे पनि हुन सक्दछन् । जस्तै: ३८२,६२२.८, ५१०,७३३.१ आदि क्रमशः ३००,००० देखि ४००,००० को समूहमा र ४००,००० र सोभन्दा माथिको अन्तराल समूहमा परेका छन् ।

तालिका ७.२: नेपाल जीवनस्तर सर्वेक्षण २०१०/११ मा छनोटमा परेका परिवारहरूको प्रति परिवार वार्षिक उपभोग आवृत्ति वितरण

प्रति परिवार वार्षिक उपभोग (रु)	परिवार संख्या
१००,००० भन्दा कम	५,५३०
१००,००० देखि २००,०००	३७५
२००,००० देखि ३००,०००	५८
३००,००० देखि ४००,०००	१९
४००,००० र सोभन्दा माथि	६

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

आँकडाहरूको विशेषता अध्ययन गरिसकेपछि तिनीहरूलाई उपयुक्त गणितीय वा तथ्याङ्कीय विधिबाट प्रशोधन तथा विश्लेषण गरेर सूचक उत्पादन गरिन्छ । सूचकहरूले विभिन्न सूचना बोकेका हुन्छन् । सूचकहरूले खासगरी “Wh” प्रश्नहरूको जवाफ दिन्छन् ।

कसले, कहाँ, कसरी, किन जस्ता प्रश्नको उत्तर सूचकको प्रयोग गरेर बताउन सकिन्छ। उदाहरणका लागि नेपालको वि.सं. २०६८ मा कोरा मृत्युदर ७.३ (प्रति एकहजार जनसंख्यामा) रहेको छ। यहाँ, कोरा मृत्युदर एउटा सूचक हो। यो सूचकले नेपालमा प्रति एकहजार जनसंख्यामा वर्षेनी ७ जना भन्दा मानिसहरू मरिरहेका छन् भन्ने बुझाउँछ। त्यसैगरी, कुल गार्हस्थ्य उत्पादन वृद्धिदर अर्को एउटा आर्थिक सूचक हो। आर्थिक वर्ष २०७७/७८ मा नेपालको कुल गार्हस्थ्य उत्पादन आधार मूल्यमा वृद्धिदर ३.९८ प्रतिशत रहेको छ। यसले नेपालको कुल गार्हस्थ्य उत्पादन आर्थिक वर्ष २०७६/७७ को तुलनामा आर्थिक वर्ष २०७७/७८ मा ३.९८ प्रतिशतले बढी छ भन्ने बुझाउँछ। देशहरूले आर्थिक वृद्धि गर्ने लक्ष्य लिएका हुन्छन्। लक्ष्यअनुरूप प्रगति हुन सक्यो वा सकेन भनेर अनुगमन तथा मूल्याङ्कनका लागि यो सूचक प्रयोग गरिन्छ। त्यस्तै सूचकहरूले आर्थिक, सामाजिक, राजनैतिक, वातावरणीय लगायत विभिन्न अवस्थालाई प्रतिबिम्बित गर्नुका साथै राष्ट्रिय तथा अन्तर्राष्ट्रिय विकास लक्ष्यहरू मापन गर्न महत्त्वपूर्ण औजारका रूपमा सहयोग गरिरहेका हुन्छन्। यस पाठमा हामी यिनै सूचकहरूका बारेमा थप चर्चा गर्नेछौं।

७.२ सूचकका प्रकार (Types of Indicators)

कुनै पनि प्रणाली र प्रक्रियामा प्रयोग हुने हिसाबले मुख्यतया निम्नानुसार चार प्रकारका सूचक हुन्छन्।

७.२.१ लागत सूचक (List Indicator)

कुनै योजना वा कार्यक्रममा लाग्ने समय, पैसा, श्रम आदि सूचित गर्ने सूचकहरूलाई लागत सूचक भनिन्छ। जस्तै: कुनै गाउँपालिकाको कुल वार्षिक बजेटको कति प्रतिशत बजेट पूर्वाधार विकासमा छुट्ट्याइएको छ, भन्ने मापन गर्ने सूचक लागत सूचक हो। त्यस्तै गरी, काठमाण्डौं महानगरपालिकाको फोहर व्यवस्थापनमा खटिएका कर्मचारी संख्या पनि लागत सूचक हो। आर्थिक वर्ष २०७७/७८ का लागि नेपाल सरकारले शिक्षा क्षेत्रमा कुल बजेटको ११.६४ प्रतिशत बजेट छुट्ट्याएको छ, यो शिक्षा क्षेत्रको लागत सूचक हो। यस्ता सूचकहरू खासगरी परियोजनाहरू लागतप्रभावी भए भएनन् जाँच गर्नका लागि प्रयोग हुन्छन्।

७.२.२ प्रक्रियागत सूचक (Process Indicator)

कार्यक्रम सञ्चालन गर्दागर्दैको अवस्थामा कार्यप्रगतिको स्तर मापन गर्न प्रयोग गरिने सूचकहरू प्रक्रियागत सूचकहरू हुन्। जस्तै: कोभिड-१९ विरुद्धको खोप प्रदान गर्न आवश्यक स्वास्थ्यकर्मीहरूमध्ये कति प्रतिशतलाई तालिम दिइयो, जनगणनामा कति प्रतिशत परिवारले आफ्नो विवरण टिपाए, आदि। राष्ट्रिय जनगणना २०७८ मा मंसिर ५ गतेसम्ममा ८० प्रतिशत गणना सम्पन्न भयो भन्नु प्रक्रियागत सूचक हो। यस्ता

सूचकहरूले खासगरी कार्यक्रमको सूक्ष्म अनुगमन गर्न मद्दत गर्दछन्। निर्धारित लक्ष्यअनुरूप प्रगति हासिल भइरहेको छ वा थप हस्तक्षेपकारी भूमिका आवश्यक पर्दछ भन्ने कुराको निर्णय गर्न यस्ता सूचकको प्रयोग गरिन्छ।

७.२.३ प्रतिफलमापक सूचक (Output Indicator)

कुनै पनि कार्यक्रमको निर्धारित कार्य अवधि समापन भएपछि प्राप्त भएको नतिजा मापन गर्ने सूचकहरू नतिजामापक सूचक हुन्। जस्तै: एक वर्षभरिमा तयार भएको सडकहरूको कुल लम्बाइ, जनगणनामा कति परिवारको तथ्याङ्क संकलन गरियो, कति प्रतिशत मानिसले सही तरिकाले हात धुने बानीको विकास गरे, आदि। यसैगरी १५ वर्षमाथिका जनसंख्या मध्ये मदिंराको हानिकारक प्रयोग गर्ने जनसंख्या आर्थिक वर्ष २०७५/७६ को अन्तसम्म २ प्रतिशत थिए भन्नु प्रतिफलमापक सूचकको एउटा उदाहरण हो। यस्ता सूचकहरू लक्ष्यअनुरूप प्रगति हासिल हुन सक्यो वा सकेन भनेर थाहा पाउन प्रयोग गरिन्छ। तिनै तहका सरकारहरूले वार्षिक बजेटमार्फत् विभिन्न कार्यक्रमहरूको लक्ष्य निर्धारण गरेका हुन्छन्। आर्थिक वर्षको समाप्तिसँगै निर्धारित लक्ष्य हासिल गर्न सकियो वा सकिएन भनेर नतिजामापक सूचकहरूको प्रयोग गरी समीक्षा गरिन्छ।

७.२.४ प्रभाव वा असरमापक सूचक (Outcome Indicator)

विभिन्न निकायहरूले कार्यान्वयन गरेका कार्यक्रमको धेरै वर्षसम्म प्रभाव वा असर देखा पर्दछ। प्रभाव अलि लामो अवधिमा देखिने प्रतिफल हो भने प्रभावको तुलनामा असर केही छोटो अविधमा प्राप्त हुने प्रतिफल हो। जस्तै: मानिसहरूको हात धुने बानीको विकास भएसँगै कोभिड १९ लगायत अन्य सरूवा रोगहरूको प्रकोप नियन्त्रण हुन थाल्दछ। खुला दिसामुक्त क्षेत्र घोषणा भएपछि भाडापखाला लगायतका सरूवा रोगहरूको संक्रमण नियन्त्रण हुँदै जान्छ। कृषिमा बैकले लगानी गरेसँगै परिवारको गरिबी न्यूनीकरण हुँदै जान्छ। त्यसैगरी, आर्थिक वर्ष २०७५/७६ मा नेपालको मानव विकास सूचकाङ्क ०.५७९ थियो भन्नु प्रभावमापक सूचकको एउटा उदाहरण हो। अर्थात्, यस्ता कार्यक्रमबाट केही समयपछि असर देखिन्छ भने अलि लामो समयपछि प्रभाव देखिन थाल्दछ। तसर्थ, यस्ता विषयहरू मापन गर्ने सूचकलाई प्रभाव वा असरमापन सूचक भनिन्छ। जस्तै: कुल प्रजनन दर, मातृमृत्यु अनुपात, गरिबीको रेखामुनि रहेका परिवारको प्रतिशत, आदि। यस्ता सूचकहरूले कुनै पनि कार्यक्रमको दीर्घकालीन लक्ष्य (Goal) पूरा हुन सक्यो वा सकेन भनेर समीक्षा गर्ने आधार प्रदान गर्दछन्।

संघीय व्यवस्थामा तिनै तहका सरकारले समृद्धिका लागि विभिन्न समयसीमाका योजनाहरू बनाएका हुन्छन्। ती योजनाहरूको सोच, दृष्टि, उद्देश्य, गन्तव्य तथा लक्ष्य हुन्छन्। यस किसिमका लक्ष्यहरू हासिल गर्नका लागि विभिन्न कार्यक्रमहरू सञ्चालन

गरिएको हुन्छ । यस्ता कार्यक्रमहरूको निर्धारित लक्ष्य प्रगति मापन गर्नका लागि सूचकहरूको प्रयोग गरिन्छ । उदाहरणका लागि नेपालको पन्ध्रौँ योजनाको अनुसूची २ मा उल्लेख भएको नतिजा खाका हेर्न सकिन्छ ।¹⁴

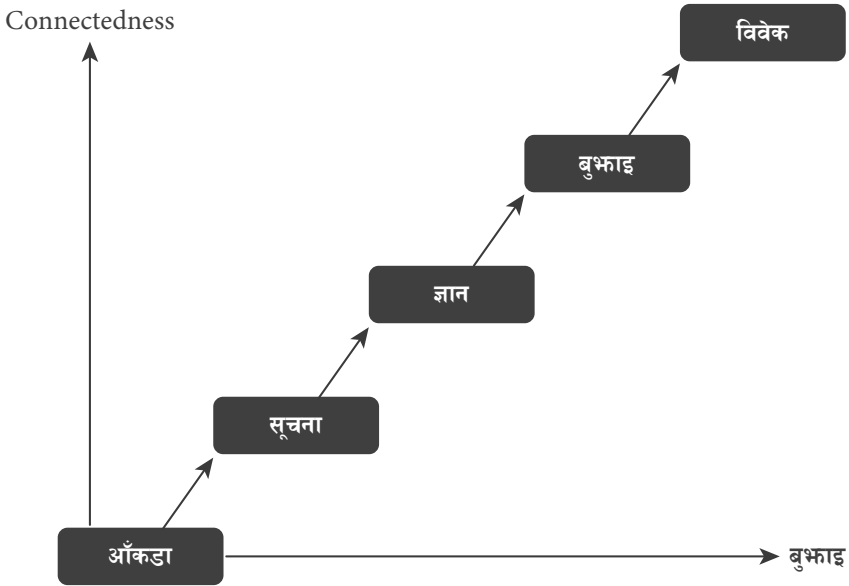
७.३ सूचनाको कच्चा पदार्थका रूपमा आँकडा

सूचनाको प्रमुख कच्चा पदार्थ भनेको आँकडा हो । आँकडाहरूलाई जबसम्म प्रशोधन र विश्लेषण गरिँदैन तबसम्म तिनीहरू अर्थपूर्ण हुँदैनन् । तसर्थ, आँकडाहरूलाई जोड, घटाउ, गुणा, भाग तथा अन्य गणितीय वा तथ्याङ्कीय क्रियाको प्रयोग गरेर अर्थपूर्ण सूचना व्यक्त गर्न सक्ने संख्या वा संकेतलाई सूचक भनिन्छ । तथ्याङ्क संकलनदेखि विश्लेषण सम्मका कार्यहरूको मुख्य उद्देश्य नै अध्ययनसँग सम्बन्धित क्षेत्रको बारेमा आवश्यक सूचना उपलब्ध गराउनु हो । अधिकांश सूचकहरू संख्यात्मक रूपमा व्यक्त गरिन्छ । सूचकहरू संख्या, दर, प्रतिशत, अनुपात आदिका रूपमा व्यक्त गरिन्छन् ।

सिद्धान्तकार प्राध्यापक रसेल एकफका अनुसार चित्र नम्बर ७.१ मा देखाइए जस्तै आँकडा, सूचना, ज्ञान, बुझाइ तथा विवेक रहेको आधारमा मानिसको मस्तिष्कलाई मुख्यतया ५ भागमा विभाजन गर्न सकिन्छ । आँकडा मात्र भएको मस्तिष्कमा बुझाइ शून्य हुन्छ र त्यसले कुनै अर्थ राख्दैन । जब आँकडाहरूलाई प्रशोधन गरेर सूचना बनाइन्छ र सूचना भएको मस्तिष्कलाई ज्ञानको भण्डार मानिन्छ । ज्ञानले बुझाइको स्तर कायम गर्दछ र यसको प्रयोगले मात्र विवेकको परिचय दिन्छ । तसर्थ, आँकडा सूचनाको कच्चा पदार्थ हो र यसैबाट सूचना उत्पादन हुन्छ । सूचनाले ज्ञानको तह वृद्धि गर्दछ र बुझाइको स्तरोन्नति गर्दछ । अन्ततः, ज्ञान र बुझाइको सही प्रयोगले नै विवेकको परिचय दिन्छ ।

14 National Planning Commission. (2019). *The Fifteenth Plan (FY 2019/20 – 2023/24)*. Kathmandu, Nepal; National Planning Commission. Retrieved on November 22, 2021 from: https://npc.gov.np/images/category/15th_plan_English_Version.pdf

चित्र नं. ७.१: आँकडा, सूचना, ज्ञान तथा विवेक बिचको सम्बन्ध



स्रोत: जिन बेलिङ्गर, दुर्वल क्यास्ट्रो, एन्थोनी मिल्स (२००४)¹⁵

७.४ सूचकको महत्त्व तथा प्रयोग

सूचक एउटा त्यस्तो औजार हो, जसले जुनसुकै कार्यक्रमको प्रगति अवस्था अनुगमन गर्नका लागि सहयोग गर्दछ। स्थानीय तहदेखि राष्ट्रिय तथा अन्तर्राष्ट्रिय तहसम्मका विकास लक्ष्य अनुगमन गर्नका लागि सूचककै प्रयोग गरिएको हुन्छ। कुनै संस्था विशेषले पनि आफूले सञ्चालन गरेका कार्यक्रमहरूको अनुगमन सूचककै सहायताले गरिरहेका हुन्छन्। उदाहरणका लागि, एउटा विद्यालयको विद्यार्थी भर्नादर, विद्यार्थीले विद्यालय छोड्ने दर, दलित विद्यार्थीको प्रतिशत, छात्राहरूको प्रतिशतजस्ता सूचकले सम्बन्धित विद्यालयको उद्देश्य तथा लक्ष्य मापन गर्न सहयोग गरिरहेका हुन्छन्। त्यसैगरी कुनै व्यावसायिक प्रतिष्ठानले दैनिक उत्पादन, दैनिक बिक्री, दैनिक खर्च जस्ता विषयमा सूचक निर्माण गरी अनुगमन गरिरहेका हुन्छन्। नेपालको सन्दर्भमा हेर्ने हो भने पन्ध्रौँ आवधिक योजनामा नतिजा खाका तयार गरी विभिन्न सूचकहरूको प्रयोगमार्फत् सम्बन्धित कार्यक्रमको अनुगमन गरिएको पाइन्छ। अन्तर्राष्ट्रियस्तरमा हेर्दा दिगो विकासका १७ वटा लक्ष्य, १६९ वटा गन्तव्य अनुगमनका लागि २३१ अद्वितीय सूचक र दोहोरिएको सहित २४७ सूचक

15 Gene Bellinger, Durval Castro, Anthony Mills. *Data, Information, Knowledge, and Wisdom. The Way of Systems*. Retrieved on November 21, 2021 from: <http://www.systems-thinking.org/dikw/dikw.htm>

प्रयोग भएको छ । दिगो विकास लक्ष्य अनुगमनकै लागि नेपालले ४७९ वटा सूचक निर्माण गरेको छ । मानवअधिकार, सुशासन, मानव विकास जस्ता विषयहरूको अवस्था अनुगमन गर्नका लागि पनि सूचककै प्रयोग हुन्छ । तसर्थ, संक्षेपमा भन्नुपर्दा सूचकहरू कुनै पनि संस्था, स्थानीय सरकार, प्रदेश सरकार, संघीय सरकारदेखि अन्तर्राष्ट्रिय स्तरसम्म नीति निर्माण, योजना तर्जुमा तथा विकास लक्ष्य अनुगमन गर्नका लागि प्रयोग भइरहेका छन् । साथै, यिनीहरूले अनुगमन तथा मूल्याङ्कनको भरपर्दो औजारका रूपमा ठुलो महत्त्व बोकेका छन् ।

चित्र नं. ७.२: सूचकको प्रयोग



कुनै पनि संस्थाका लागि तथ्याङ्कको प्रयोग निर्णय प्रक्रियामा प्रयोग हुन्छ भने व्यक्ति विशेषका लागि तथ्याङ्क निर्णय गर्नका साथसाथै मनोरञ्जन वा आत्मसन्तुष्टिका लागि पनि प्रयोग हुन्छ । व्यक्ति सधैं उत्सुक हुन्छ र त्यो उत्सुकता मेट्नका लागि सूचना आवश्यक पर्दछ । जस्तै: “आज कोभिड-१९ का कति जना नयाँ बिरामी भेटिए” भन्ने सूचना व्यक्तिका लागि उत्सुकता मेट्न प्रयोग हुन्छ भने संस्थाका लागि नीति निर्माण, योजना तर्जुमा वा निर्णय गर्न प्रयोग हुन्छ । संक्षिप्त रूपमा सूचकको प्रयोगका आयामहरूलाई चित्र नं. ७.२ मा प्रस्तुत गरिएको छ ।

७.५ असल सूचकमा हुनुपर्ने गुणहरू

सूचकहरूका महत्त्व र प्रयोगका बारेमा हामीले माथि चर्चा गरिसकेका छौं । यस्ता महत्त्वपूर्ण औजारहरूले निश्चित गुणहरू धारणा गरेकै हुनुपर्दछ । मापन गर्ने औजारमा गुणस्तर कायम हुन सकेमा मात्र मापन गर्न खोजिएको वस्तु वा विशेषता सही तरिकाले मापन हुन्छन् । तसर्थ, असल सूचकमा निम्नानुसारका गुणहरू हुनु अनिवार्य हुन्छ ।

क) सूचकहरू मापन गर्न खोजिएको उद्देश्यानुरूप हुनुपर्दछ ।

ख) सूचकहरू मापन गर्न सकिने खालको हुनुपर्दछ ।

ग) सूचकहरू प्रयोगकर्ताको पहुँचयोग्य हुनुपर्दछ ।

घ) सूचकहरू विषयवस्तु सान्दर्भिक हुनुपर्दछ ।

ङ) सूचकहरू निर्धारित समय अवधिका लागि निर्माण गरिएको हुनुपर्दछ ।

च) सूचकहरू स्पष्ट, नबाभिने र वैज्ञानिक तवरले निर्माण गरिएको हुनुपर्दछ ।

छ) सूचकहरू विश्वसनीय, सर्वमान्य, संवेदनशील, प्रतिनिधिमूलक हुनुपर्दछ ।

ज) सूचकहरू बुझ्न सकिने खालको हुनुपर्दछ ।

झ) सूचकहरू नैतिक धरातलमा बसेर निर्माण गरेको हुनुपर्दछ ।

७.६ तथ्याङ्क विश्लेषण र सूचक निर्माण

संकलित तथ्याङ्कको विधिपूर्वक प्रशोधन तथा विश्लेषण गरेर सूचक निर्माण गरिन्छ भन्ने बारेमा हामीले माथि संक्षेपमा चर्चा गरिसकेका छौं । यसरी तथ्याङ्क विश्लेषण गर्दा सूचकका केही अवयवहरूका बारेमा पूर्ण रूपमा ध्यान दिनु जरूरी हुन्छ । जस्तै: के मापन गरिँदैछ, मापनको तह के हो, कार्यक्रमको आधार तह के हो, कस्तो परिवर्तनको अपेक्षा गरिएको छ, अर्थात् उद्देश्य एवं लक्ष्य के रहेको छ, परिवर्तनको गुणस्तर कस्तो हुने, लक्षित जनसंख्या के हो, समयसीमा के हो, आदि जस्ता विषयहरूमा पूर्ण जानकारी भएपछि मात्र तथ्याङ्कको प्रशोधन एवं विश्लेषण गरेर सूचक निर्माण गर्नुपर्दछ । यसका अतिरिक्त सूचनाको कच्चा पदार्थको रूपमा प्रयोग गरिने तथ्याङ्कका बारेमा पनि जानकारी राख्नुपर्दछ । कस्तो प्रकारको तथ्याङ्क प्रयोग गरिँदैछ, तिनीहरूको स्रोत तथा संकलन विधि के हो, गुणस्तर कस्तो छ र तथ्याङ्कलाई नै व्याख्या गर्ने मेटाडाटा उपलब्ध छ, छैन जस्ता कुराहरू सदैव ख्याल गर्नुपर्दछ ।

सामान्यतया सूचकहरू संख्या, अनुपात, दर, प्रतिशतका रूपमा प्रस्तुत गरिएको हुन्छ । जे जस्तो रूपमा प्रस्तुत गरिएको भए तापनि तिनीहरूको एकाइ, गणना गर्ने विधिका बारेमा पूर्ण जानकारी राख्नुपर्दछ । विभिन्न सूचकहरू निर्माण गर्ने विधि तथा प्रयोगका बारेमा हामी पाठ ८ र ९ मा थप चर्चा गर्नेछौं । यस पाठमा हामी तथ्याङ्क विश्लेषणसम्बन्धी आधारभूत विषयमा छलफल गर्दैछौं ।

तथ्याङ्क विश्लेषण सूचना चक्रको एउटा महत्त्वपूर्ण अवयव हो । तथ्याङ्क प्रशोधनपछिको चरण भनेको तथ्याङ्कको विश्लेषण हो (चित्र नं. ६.२) । यो तथ्याङ्क व्यवस्थापनका कार्यहरूमध्येको महत्त्वपूर्ण कार्य पनि हो । कच्चा रूपमा संकलित आँकडाहरूलाई अर्थपूर्ण सूचनामा परिणत गर्ने कार्य तथ्याङ्क विश्लेषणअन्तर्गत पर्दछ । यसले तथ्याङ्कलाई एकत्रित गरी तालिकीकरण गर्नमा समेत सहयोग गर्दछ । तथ्याङ्क विश्लेषणको सहयोगले मात्र कार्यक्रम अनुसन्धानसम्बन्धी सोधिने प्रश्नहरूको जवाफ प्रदान गर्न सकिन्छ । संक्षेपमा भन्नुपर्दा, तथ्याङ्क विश्लेषण भनेको तथ्याङ्कीय प्रविधि र तर्कलाई प्रयोग गरी सङ्कलन गरिएको तथ्याङ्कलाई व्यवस्थित रूपमा व्याख्या गर्न, तुलना गर्न, वर्गीकरण गर्न र निष्कर्षमा पुऱ्याउनको लागि प्रयोग गरिने प्रक्रिया हो । तथ्याङ्क विश्लेषण सामान्यतया वर्णनात्मक र अनुमानात्मक गरी दुई प्रकारका हुन्छन् । केन्द्रीय प्रवृत्तिको मापन, विचलनको मापनजस्ता कार्यहरू वर्णनात्मक प्रकृतिका विश्लेषण हुन् भने परिकल्पना परीक्षण, प्रतिगमन विश्लेषण, विश्वसनीय अन्तराल निर्माण, भेरियन्स विश्लेषण, मोडलिङ आदि अनुमानात्मक विश्लेषणका उदाहरण हुन् । वर्णनात्मक विश्लेषणको तुलनामा अनुमानात्मक विश्लेषणका लागि बढी प्राविधिक ज्ञान र सीपको आवश्यकता पर्दछ । वर्तमान समयमा तथ्याङ्क विश्लेषणका लागि कम्प्युटरको प्रयोग लगभग अनिवार्य जस्तै भइसकेको छ । तसर्थ, यस पुस्तकको पाठ २२ र २३ हामी कम्प्युटरका केही सफ्टवेयरको प्रयोग सम्बन्धमा चर्चा गर्नेछौं ।

७.७ तथ्याङ्क विश्लेषणका उद्देश्य

तथ्याङ्क विश्लेषण गर्नुका प्रमुख उद्देश्य भनेको विषयवस्तुसँग सम्बन्धित “WHAT” प्रश्नहरूको उत्तर प्रदान गर्नु हो । तथ्याङ्क विश्लेषणले “WHY” प्रश्नको उत्तर भने दिन सक्दैन । कुनै पनि विषय वा वस्तुको विशेषताको वर्णन गर्न तथ्याङ्क विश्लेषण गरिन्छ भने विवेकपूर्ण तथ्याङ्कको व्याख्याले मात्र कारण र प्रभावका बारेमा बताउन सकिन्छ । उदाहरणका लागि कुनै गाउँपालिकाले एक वर्ष भित्रमा शतप्रतिशत जनसंख्यालाई साक्षर बनाउने अभियानका साथ परियोजना सञ्चालन गरेको थियो । परियोजनाको अवधि सकिएपछि, परियोजनाको उद्देश्य प्राप्त भयो वा भएन भनेर थाहा पाउन तथ्याङ्क संकलन तथा संकलित तथ्याङ्कको विश्लेषण गर्नुपर्दछ । विश्लेषण गरिएका तथ्याङ्कको आधारमा सञ्चालित कार्यक्रमहरूको लक्ष्य र प्रगति तुलना गरेर मात्र उद्देश्य पूरा भए नभएको यकिन गर्न सकिन्छ । तर यदि लक्ष्यअनुरूप प्रगति हासिल हुन सकेन भने सोको कारण उक्त तथ्याङ्क विश्लेषणको नतिजाले मात्र पत्ता लगाउन वा बताउन सक्दैन । त्यसका लागि विवेकपूर्ण तथ्याङ्क व्याख्याको आवश्यकता हुन्छ । कतिपय अवस्थामा तथ्याङ्क व्याख्या गर्न थप तथ्याङ्क विश्लेषणको आवश्यकता पनि पर्न सक्दछ ।

उदाहरण

विश्लेषण: श्रावण महिनामा भन्दा भाद्र महिनामा एक लाख रूपैयाँ बराबरको बढी सामान बिक्री भयो ।

व्याख्या: भाद्रमा तीज लगायतका चाडपर्वहरूका कारण बिक्री बढेको हो ।

७.८ सूचक सन्दर्भ पाना

तथ्याङ्क विश्लेषणबाट प्राप्त हुने प्रत्येक सूचकहरूको विस्तृत जानकारी सहितको अभिलेख गर्ने औजारलाई सूचक सन्दर्भ पाना भनिन्छ । यस पानाले सूचकको आधारभूत जानकारी, यसको व्याख्या, यसका लागि आवश्यक पर्ने तथ्याङ्क संकलन, विश्लेषणको योजना, तथ्याङ्कको गुणस्तर, आधार वर्षको मान तथा लक्ष्य सम्बन्धमा जानकारी उपलब्ध गराउँदछ । सूचकको वैज्ञानिक तवरले प्रयोग गर्नका लागि सूचक सन्दर्भ पाना तयार गरिएको हुनुपर्दछ । सूचक सन्दर्भ पानाको नमुना तालिका ७.३ मा दिइएको छ ।

तालिका ७.३: सूचक सन्दर्भ पानाको नमुना

सूचक सन्दर्भ पाना
उद्देश्य तथा प्रतिफल:
सूचक:
सूचकसम्बन्धी संक्षिप्त विवरण:
सूचकको परिभाषा:
सूचक मापनको एकाइ:
सूचक निकाल्ने विधि: अंश: हर:
व्यवस्थापकीय उपयोग:
तथ्याङ्क संकलनको योजना:
तथ्याङ्क संकलनको विधि र स्रोत:
तथ्याङ्क संकलनको समय तथा पटक:
तथ्याङ्क संकलनको लागत र बजेट स्रोत:
जिम्मेवार निकाय:
तथ्याङ्क भण्डारण गर्ने स्थान:
तथ्याङ्क विश्लेषणको योजना (कार्यतालिका, जिम्मेवारी):
तथ्याङ्क विश्लेषणको विधि:

तथ्याङ्कको प्रस्तुति:
तथ्याङ्कको गुणस्तर नियन्त्रण योजना:
तथ्याङ्कका सीमा (भएमा मात्र):
तथ्याङ्कका सीमा सम्बोधन गर्न गरिएका प्रयास:
कार्यसम्पादन तथ्याङ्क तालिका:
आधार वर्ष तथा लक्ष्य र तिनीहरूको छनोटको औचित्य:
टिप्पणी:

७.९ अभ्यासका लागि प्रश्नहरू

क) तथ्याङ्क केका लागि प्रयोग गरिन्छ ?

- १) योजना तर्जुमा
- २) नीति निर्माण
- ३) अध्ययन अनुसन्धान
- ४) माथिका सबै

ख) तथ्याङ्कका प्रयोगकर्ताहरू किन सजग हुनुपर्छ ?

- १) तथ्याङ्कलाई बुझ्न
- २) तथ्याङ्कको गलत प्रयोगबाट हुनसक्ने जोखिमबाट मुक्त हुन
- ३) तथ्याङ्कका सीमाहरूबारे जानकारी प्राप्त गर्न
- ४) माथिका सबै

ग) सार्क राष्ट्रहरूको विगत ५ वर्षको कुल गार्हस्थ्य उत्पादनका आँकडा कस्तो डाटा हो ?

- १) खण्डित डाटा
- २) बिग डाटा
- ३) अविच्छिन्न डाटा
- ४) मेटाडाटा

घ) निम्नलिखित विषयमा टिप्पणीसहित व्याख्या गर्नुहोस् ।

- १) कुनै सडकमा कुनै वर्षमा गणना गरिएको सडक दुर्घटनाबाट भएको मृत्यु विवरण यसप्रकार छ: सडक बिच मा मृत्यु: १२४ जना,
सडकपेटीमा मृत्यु: ४०४८ जना ।
उल्लिखित आँकडाको आधारमा निम्नानुसारको निचोड निकाल्न सकिन्छ कि सकिँदैन, किन ?

“सडक बिचमा थोरै मृत्यु भएका कारण सडक पेटीबाटभन्दा बिचमा हिँडनु बढी सुरक्षित देखिन्छ”।

ड) तलको तालिका अध्ययन गरी सोधिएको प्रश्नको जवाफ दिनुहोस् ।

कक्षा ७ मा अध्ययनरत दुई विद्यार्थीको नाम	प्रथम त्रैमासिक परीक्षाको प्राप्ताङ्क (प्रतिशत)	दोस्रो त्रैमासिक परीक्षाको प्राप्ताङ्क (प्रतिशत)	अन्तिम परीक्षाको प्राप्ताङ्क (प्रतिशत)	औसत वार्षिक प्राप्ताङ्क (प्रतिशत)
सुजन खतिवडा	८०	६०	४०	६०
सजना दर्लामी	५५	६०	६५	६०

“औसत वार्षिक प्राप्ताङ्क प्रतिशत सूचकका आधारमा सुजन र सजनाको अध्ययन प्रगति समान छ ।” यो भनाइ कतिको सही छ, यसले तथ्याङ्क विश्लेषणमा सूचक प्रयोग गर्दा के कुरामा ख्याल गर्नुपर्दछ भन्ने कुरा इङ्कित गर्न खोजेको छ, व्याख्या गर्नुहोस् । (उत्तरका लागि संकेत: चित्र नं. ७.१ अनुसार व्याख्या गर्नुहोस्)

च) जोडा मिलाउनुहोस् ।

आँकडा

सूचक

"WHAT" प्रश्न

" WHY" प्रश्न

परिवारको सदस्य संख्या

परिवारको वार्षिक आमदानी

तथ्याङ्कको व्याख्या

खण्डित श्रेणी

अविच्छिन्न श्रेणी

सूचनाको कच्चा पदार्थ

निर्णय प्रक्रिया

तथ्याङ्कको विश्लेषण

छ) फरक छुट्याउनुहोस् ।

१) तथ्याङ्कको विश्लेषण र तथ्याङ्कको व्याख्या

२) खण्डित श्रेणी र अविच्छिन्न श्रेणी

३) आँकडा र सूचना

४) प्रतिफल मापक सूचक र असर मापक सूचक

७.१० सन्दर्भ सामग्री

तथ्याङ्क र सूचकसम्बन्धी अवधारणाका बारेमा विस्तृत अध्ययन गर्नका लागि निम्न बमोजिमको सामग्री उपयोगी छ ।

१. World Bank (1996). Performance Monitoring Indicators: A handbook for task managers. Retrieved on November 25, 2021 from: <https://www.measureevaluation.org/resources/training/capacity-building-resources/basic-me-concepts-portuguese/indicators.pdf>

७.११ शब्दावली

अनुगमन (Monitoring): “अनुगमन” भन्नाले योजना कार्यान्वयनमा झोत तथा साधनको प्रयोग उचित ढङ्गले भए नभएको तथा अपेक्षाकृत नतिजा हासिल भए, नभएको सम्बन्धमा निरन्तर वा आवधिक रूपमा निगरानी, निरीक्षण, अवलोकन, सूचना सङ्कलन, विश्लेषण तथा सुधार गर्ने कार्यलाई बुझाउँछ ।

मूल्याङ्कन (Evaluation): “मूल्याङ्कन” भन्नाले कार्यान्वयन चरणमा रहेका वा सम्पन्न भएका योजनाको सान्दर्भिकता, सामञ्जस्यता, प्रभावकारिता, कार्यदक्षता, प्रभाव तथा दिगोपनाका सम्बन्धमा उद्देश्यपूर्ण र व्यवस्थित तरिकाले गरिने मूल्याङ्कनलाई बुझाउँछ । यो शब्दले योजना सुरू गर्नु अघिको अवस्थामा सोको औचित्य र लाभ-लागतका आधारमा गरिने पूर्व मूल्याङ्कन समेतलाई जनाउँछ ।



तथ्याङ्क प्रस्तुतिकरण

८.१ परिचय

सूचना चक्रको एउटा महत्वपूर्ण अवयवको रूपमा तथ्याङ्क प्रस्तुतिकरणलाई लिइन्छ । तथ्याङ्कको प्रस्तुतिकरण भनेको तथ्याङ्कको संक्षेपीकरण हो । यसले तथ्याङ्क विश्लेषण र तथ्याङ्कको व्याख्याको बिचमा पुलको काम गर्दछ । अर्थात्, तथ्याङ्कको विश्लेषण गरिसकेपछि व्याख्या गर्नका लागि तथ्याङ्क प्रस्तुतिकरण गरिन्छ । प्रस्तुतिकरणका विभिन्न तरिकामध्ये तालिकीकरणका सम्बन्धमा यस अगाडिका पाठहरूमा वर्णन गरिसकिएको छ । यस पाठमा खासगरी डायग्राम र ग्राफका माध्यमबाट तथ्याङ्क प्रस्तुत गर्ने विधिका बारेमा छलफल गर्नेछौं ।

८.२ तथ्याङ्क प्रस्तुतिको महत्त्व

प्रस्तुतिकरणले ठुलो आयातनको तथ्याङ्कलाई संक्षेपमा देखाउन सहयोग गर्दछ । यसले ठुला ठुला संख्याको जटिलतालाई पनि सरलताका साथ बुझ्न मद्दत गर्दछ । चरको प्रवृत्ति (Trend) र चरहरूबिचको सम्बन्ध सरलताका साथ प्रस्तुत गर्दछ जसले गर्दा विभिन्न आयामबाट तुलनात्मक अध्ययन गर्न सहज हुन्छ । तथ्याङ्कको सरल प्रस्तुतिकरणले सामान्य व्यक्तिले पनि त्यसभित्र अन्तरनिहित सूचना बोध गर्न सक्दछन् । यसले दर्शक वा पाठकलाई दृश्यात्मक छाप प्रदान गर्ने हुनाले ग्रहण गर्ने क्षमताको विकास गरिदिन्छ । धेरै सूचनाहरूलाई पनि एकैसाथ प्रस्तुत गर्न सक्ने विशेषताले गर्दा नै तथ्याङ्ककर्मी तथा सरोकारवालाहरू बिचमा तथ्याङ्कको प्रस्तुतिकरण निकै लोकप्रिय छ । हामी जो कोहीलाई कुनै सामग्री पढ्दा त्यसमा अन्तर्निहित सूचना वा तथ्याङ्क छुट्टै तालिका वा चित्रात्मक विधिद्वारा प्रस्तुत गरिएको भएमा निकै रुचि लाग्दछ । अर्थात्, प्रस्तुतिकरणका माध्यमबाट तथ्याङ्कलाई आकर्षक बनाउन सकिन्छ ।

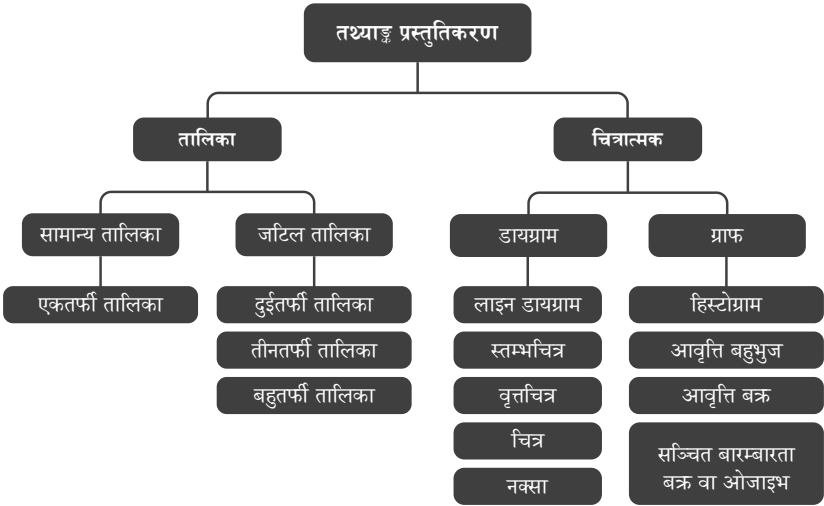
तथ्याङ्कको प्रस्तुतिकरण तथ्याङ्कका सरोकारवालाहरूका लागि विभिन्न दृष्टिकोणबाट महत्वपूर्ण हुन्छ । विद्यार्थीहरूले अध्ययनको दायरा फराकिलो बनाउन विभिन्न सूचनाहरूको जानकारी लिनुपर्ने हुन्छ । तथ्याङ्कको प्रस्तुतिकरण, विश्लेषण र व्याख्या गर्ने एउटा कला हो । यसले विद्यार्थीको बौद्धिक तथा तार्किक क्षमता अभिवृद्धि गर्दछ । विभिन्न तरिकाले प्रस्तुत गरिएका तथ्याङ्कलाई बुझ्न र आफ्नै शैक्षिक क्रियाकलापमा तथ्याङ्कको प्रयोग गर्न विद्यार्थीहरूले तथ्याङ्क प्रस्तुतिकरणका बारेमा जानकारी राख्नु निकै महत्वपूर्ण मानिन्छ । साथै, भविष्यमा अपनाउने व्यावसायिक वा पेशागत जीवनका लागि पनि विद्यार्थी जीवनदेखि नै तथ्याङ्क प्रस्तुत गरेर तार्किक क्षमता वृद्धि गर्नुपर्ने हुन्छ । विकासका

साभेदारहरूले आफ्ना कार्यक्रम तथा तिनका प्रगतिसम्बन्धमा तथ्याङ्कको सहयोगमा सरोकारवालाहरूलाई जानकारी गराउँदा होस् वा विकास योजनामा सहकार्य गर्नका लागि होस् तथ्याङ्कको प्रभावकारी प्रस्तुतिकरण हुनु जरूरी ठानिन्छ। संघ, प्रदेश तथा स्थानीय तहका सरकारहरूले सार्वजनिक चासोको विषयमा जनताको सुसूचित हुने अधिकारको सम्मान गर्न सम्बन्धित विषयका सूचनाहरू सार्वजनिक पहुँचमा हुने गरी सरल भाषामा प्रस्तुत गर्नुपर्दछ। सरकारका विभिन्न अङ्गहरूले प्राप्त गर्ने बजेट र खर्च तथा अन्य भौतिक प्रगति दर्शाउनका लागि पनि तथ्याङ्कको सरल तरिकाले प्रस्तुत गरी प्रसार गर्नुपर्दछ। तसर्थ, सूचना प्राप्त गर्न तथा प्रसार गर्नका लागि तथ्याङ्कको प्रस्तुति सरोकारवाला सबैलाई निकै महत्त्वपूर्ण हुन्छ।

८.३ तथ्याङ्क प्रस्तुतिका प्रकार

आँकडाको प्रकृति हेरेर अथवा आवश्यकता अनुसार सामान्यतया आँकडालाई दुई तरिकाबाट प्रस्तुत गरिन्छ: तालिका प्रस्तुतिकरण (Tabular Presentation) र चित्रात्मक वा ग्राफिकल वा रेखाचित्र प्रस्तुतिकरण (Diagrammatic or Graphical Presentation)। चित्र नं. ८.१ मा तथ्याङ्कको प्रस्तुतिका प्रकारलाई संक्षेपमा उदाहरणसहित प्रस्तुत गरिएको छ। तालिकाहरू मूलतः सामान्य र जटिल गरी दुई प्रकारका तालिका हुन्छन्। यस सम्बन्धमा हामीले पाठ ६ मा छलफल गरिसकेको हुनाले यस पाठमा सामान्य पुनरावलोकन मात्र गर्नेछौं। त्यस्तै, तथ्याङ्कको चित्रात्मक प्रस्तुतिकरणमा पनि धेरै प्रकारका चित्र तथा ग्राफहरूको प्रयोग हुन्छ। तर यस पाठमा हामीले धेरैजसो प्रयोग हुने केही डायग्राम र ग्राफिकल प्रस्तुतिका बारेमा मात्र छलफल गर्नेछौं। यस्ता चित्रात्मक प्रस्तुति बनाउनका लागि कम्प्युटरको प्रयोग गर्ने सम्बन्धमा पाठ २२ र २३ मा चर्चा गर्नेछौं। अविस्मरणीय कुरा के छ भने, तथ्याङ्कको प्रकृति हेरेर कुन प्रकारको प्रस्तुतिकरण विधि अपनाउने भन्ने बारेमा प्रस्तुतकर्ताले पूर्ण रूपमा जानकारी राखेको हुनुपर्दछ।

चित्र ८.१: तथ्याङ्क प्रस्तुतिका प्रकार



८.४ तथ्याङ्क प्रस्तुतिकरण गर्दा ध्यान दिनुपर्ने कुराहरू

तथ्याङ्कलाई तालिका वा चित्रात्मक जुनसुकै विधिबाट प्रस्तुत गर्दा केही महत्त्वपूर्ण कुराहरूलाई सदैव मध्यनजर गर्नुपर्दछ । पहिलो कुरा तथ्याङ्कको प्रकृति हेरेर प्रस्तुतिको विधि छनोट गर्न सक्नुपर्दछ । उदाहरणका लागि नेपालको जनसंख्या वृद्धिको समय शृङ्खलालाई लाइन ग्राफ (Line Graph) को प्रयोग गर्नु निकै उत्तम हुन्छ । दुई वा दुईभन्दा बढी समूहहरूको विशेषता तुलनात्मक अध्ययन गर्नका लागि स्तम्भ चित्र वा वृत्तचित्र उत्तम हुन्छ । भौगोलिक सूचनाहरूलाई नक्साको प्रयोग उत्तम हुन्छ । तथ्याङ्क प्रस्तुतिका लागि जुनसुकै विधि अपनाउँदा पनि निम्न अनुसारका कुराहरूमा विशेष ध्यान दिनुपर्दछ ।

क) तथ्याङ्कको प्रकृति हेरी उपयुक्त प्रस्तुतिको छनोट गर्नुपर्दछ ।

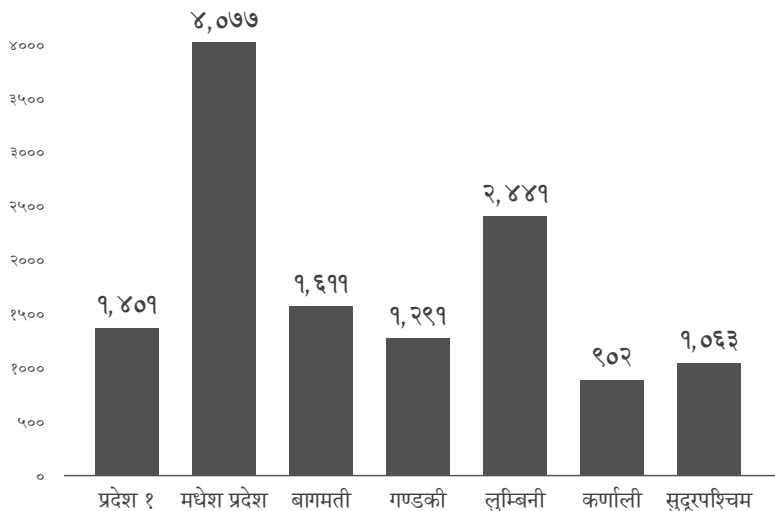
ख) तालिका वा चित्रात्मक, जुनसुकै प्रस्तुति पनि अति सामान्य, सफा, सबैले बुझ्ने खालको र स्वयं वर्णित हुनसक्ने किसिमको हुनुपर्दछ । प्रस्तुति “चित्र आफै बोल्दछ” भन्ने भनाइप्रति समर्पित हुनुपर्दछ । अर्थात्, प्रस्तुति सरल र स्पष्ट हुनुपर्दछ । आवश्यकता भन्दा बढी रङ्ग, अक्षर, सूचना वा तथ्याङ्कको प्रयोग र अन्य यस्तै कार्यले प्रस्तुतिलाई अनावश्यक भद्दा बनाउँछ ।

ग) अत्यधिक धेरै सूचना राखेर एउटै प्रस्तुति बनाउनु हुँदैन । धेरै सूचना राखेर एउटै प्रस्तुति बनाउनाले पाठक वा दर्शकलाई भ्रम सिर्जना गर्दछ र प्रस्तुति आकर्षक नभई भ्रन्धट लाग्ने खालको हुन्छ ।

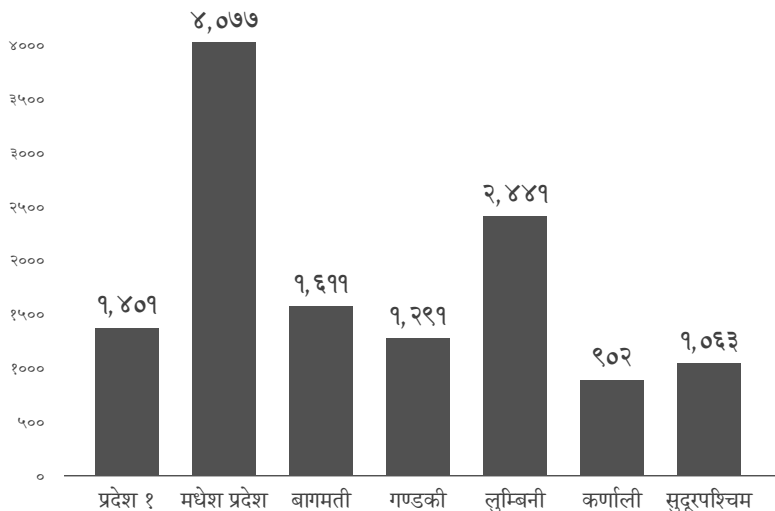
घ) प्रस्तुतिको लागि उपयुक्त शीर्षक दिनुपर्दछ । शीर्षक स्पष्ट, संक्षिप्त, स्वयं वर्णनात्मक, अलमल नहुने खालको हुनुपर्दछ । चित्र नं. ८.२ र ८.३ एउटै तथ्याङ्क प्रस्तुत गर्ने स्तम्भ

चित्र हुन् । तर चित्र नं. ८.२ को शीर्षक निकै छोटो र अस्पष्ट छ । चित्र नं. ८.३ मा प्रयोग गरिएको शीर्षक संक्षिप्त, स्वयं वर्णनात्मक, अलमल नहुने खालको छ ।

चित्र नं ८.२: अपराध



चित्र नं. ८.३: महिला, बालबालिका तथा जेष्ठ नागरिकविरुद्धको अपराध तथ्याङ्क



स्रोत: अपराध अनुसन्धान विभाग, २०७६/७७¹⁶

16 Crime Investigation Department (CID). *Data on Crime against Women, Children and Senior Citizen*. Retrieved on November 27, 2021 from: <https://cid.nepalpolice.gov.np/cid-wings/women-children-and-senior-citizen-service-directorate/>

तालिका ८.१: विगत ११ वर्षमा भएको महिला, बालबालिका तथा जेष्ठ नागरिकविरुद्धको अपराध तथा

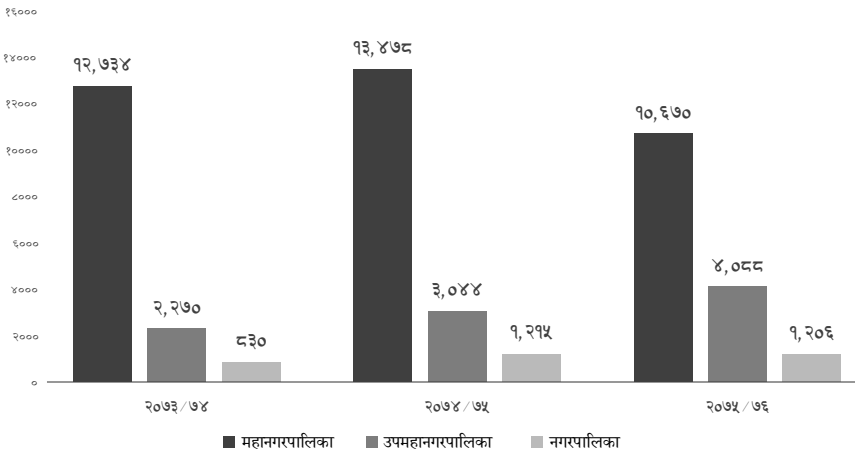
वर्ष	बलात्कार	बलात्कार प्रयास	गर्भपतन	बहुविवाह	बालविवाह	घरेलु हिंसा	बोक्सीको आरोप	छुवाछुत	बाल यौन दुर्व्यवहार
०६६/०६७	३७६	१०१	८	१४६	७	९८३	४		
०६७/०६८	४८१	१५१	१२	१९७	२	१,३५५	३९		
०६८/०६९	५५५	१५६	१३	२४९	१२	२,२५०	३५		
०६९/०७०	६७७	२४५	२८	३५०	१९	१,८००	२८		
०७०/०७१	९१२	४१४	१८	४२१	१५	६,८३५	३९	१४	
०७१/०७२	९८१	५६२	१७	५१८	२३	८,२३८	४३	१०	
०७२/०७३	१,०८९	४५२	२२	४६३	२०	९,३९८	२८	१९	
०७३/०७४	१,१३१	५३६	२२	४६४	२६	११,६२९	२४	१७	
०७४/०७५	१,४८०	७२७	१८	६०२	५९	१२,२२५	४८	१८	
०७५/०७६	२,२३०	७८६	२७	१,००१	८६	१४,७७४	४६	४३	२११
०७६/०७७	२,१४४	६८७	२९	७३४	६४	११,७३८	३४	३०	२३२

स्रोत: अपराध अनुसन्धान विभाग, २०७६/७७

नोट: यो तथ्याङ्क नेपाल प्रहरीमा आएका उजुरीका आधारमा संकलन गरिएको हो । समाजमा योभन्दा बढी अपराध घटेका हुन सक्छन् ।

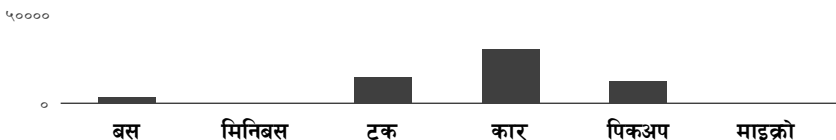
- ड) प्रत्येक प्रस्तुतिमा प्रयोग भएका तथ्याङ्कको स्रोत उल्लेख गर्नुपर्दछ। चित्र नं. ८.३ मा स्रोत खुलाइएको छ तर ८.२ मा स्रोत खुलाइएको छैन। तसर्थ, चित्र ८.२ अपुरो छ।
- च) आवश्यकताअनुसार नोट लेख्नुपर्दछ। प्रस्तुत गरिएका विवरण, तथ्याङ्क वा चित्रका बारेमा थप स्पष्ट पार्नुपर्ने भएमा शीर्षक नोट (Headnote), पाद नोट (Footnote) लेख्नुपर्दछ। तालिका ८.१ मा पाद नोट लेखिएको छ।
- छ) चित्रात्मक प्रस्तुतिमा आवश्यकताअनुसार अनुक्रमणिका (Index) लेख्नुपर्दछ। चित्रमा प्रयोग गरिएका रङ तथा संकेत वा अन्य यस्तै कुराहरूलाई थप स्पष्ट बनाउन अनुक्रमणिका उपलब्ध गराउनुपर्दछ। चित्र नं. ८.४ मा प्रत्येक स्तम्भमा प्रयोग भएको रङले बुझाउने अर्थ भल्कने गरी अनुक्रमणिका उपलब्ध गराइएको छ।

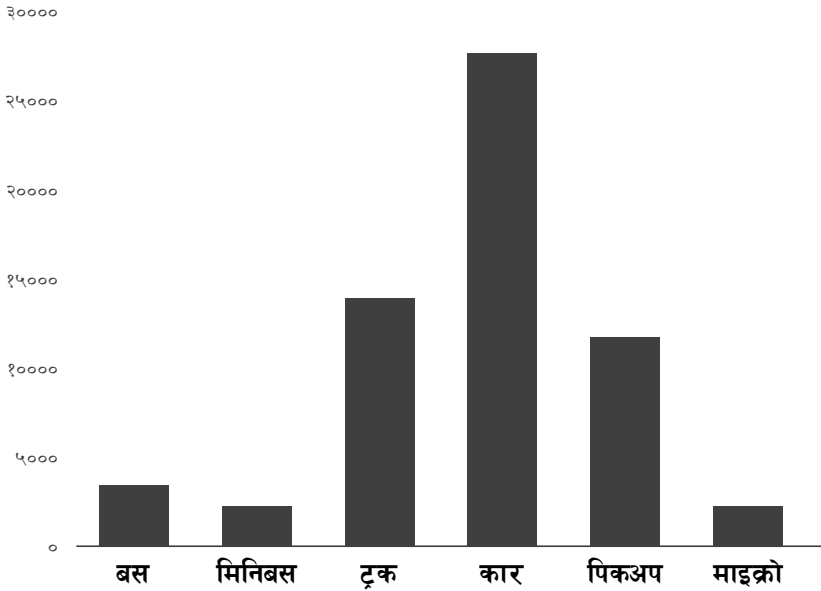
चित्र नं. ८.४: नगरपालिकाहरूद्वारा सङ्कलन गरिएको जैविक फोहोरको औसत वार्षिक मात्रा



- ज) चित्रात्मक प्रस्तुतिको लागि उपयुक्त स्केलको चयन गर्नुपर्दछ। यसले गर्दा चित्र र सूचनाबिचको तालमेल मिलेर प्रस्तुत गर्न खोजिएको सन्देश यथार्थ रूपमा प्रदान हुन्छ। उदाहरणका लागि चित्र नं. ८.५ (क) र चित्र नं. ८.५ (ख) हेरौं। चित्र ८.५ (क) मा प्रस्तुत गरिएको स्तम्भहरूको उचाइ लगभग उस्तै उस्तै देखिनाले तुलनात्मक अध्ययन प्रभावकारी हुँदैन। यस चित्रमा प्रयोग गरेको स्केल दर्शकको ध्यान आकर्षण गर्ने खालको छैन। यही स्तम्भ चित्रलाई चित्र नं. ८.५ (ख) मा सही तरिकाले प्रस्तुत गरिएको छ।

चित्र नं. ८.५ (क): आ.व. २०७४/७५ मा दर्ता भएका सवारी साधन



चित्र नं. ८.५ (ख): आ.व. २०७४/७५ मा दर्ता भएका सवारी साधन¹⁷

हामीसँग भएको आँकडाको विभिन्न सारांश वा विशेषताहरू तालिकामा प्रस्तुत गर्नुलाई नै तालिका प्रस्तुतिकरण भनिन्छ। विभिन्न समूहहरूलाई अङ्कको माध्यमद्वारा तुलना गर्न तालिकाको प्रयोग गरिन्छ। कुनै चरको एउटा मात्र विशेषता जनाउने तालिकालाई एक आयाम तालिका भनिन्छ भने दुई वा दुईभन्दा बढी चरहरूको विशेषता एकसाथ प्रस्तुत गर्ने तालिकालाई जटिल तालिका भनिन्छ। तालिका ८.२ एक आयामिक तालिका वा सामान्य तालिका हो। जटिल तालिका दुई आयामिक, तिन आयामिक वा बहुआयामिक हुन्छन्। उदाहरणको लागि नेपालको ७ वटा प्रदेशहरूको जनसंख्या विभिन्न वर्षहरूमा कति कति छ भनेर तालिका ८.३ मा प्रस्तुत गरिएको छ। यसमा प्रदेश एउटा चर भयो भने वर्ष अर्को चर भयो, तसर्थ, यो तालिका दुई आयामिक तालिका हो। यस तालिकाबाट कुनै प्रदेशको कुन वर्षमा कति जनसंख्या थियो भनेर तुलना गर्न सकिन्छ साथै कुनै एउटा समयमा कुन कुन प्रदेशको जनसंख्या कति थियो भनेर तुलना गर्न पनि सकिन्छ। यही तालिकामा प्रदेशअनुसार वर्तमान जनसंख्या मात्र प्रस्तुत गरेको भए एक आयामिक तालिका हुन्थ्यो।

17 Department of Transport Management. *Details of Registered Vehicles*. Retrieved on November 27, 2021 from: <https://www.dotm.gov.np/MainData/OldStatistics>

तालिका ८.२: विभिन्न स्थानको वायुको गुणमापक सूचकाङ्क

स्थान	वायुको गुणमापक सूचकाङ्क
रारा	६
भरतपुर	४३
भक्तपुर	५४
हेटौडा	१०२
जनकपुर	१५४
सिमरा	२१६

स्रोत: वायु गुणस्तर निगरानी। नोट: यो तथ्याङ्क २०७८ मंसिर ८ गते बिहान ६:५० बजेको हो।¹⁸

तालिका ८.३ प्रदेशअनुसार जनसांख्यिक विवरण (संख्या लाखमा)

प्रदेश	१९७१	१९८१	१९९१	२००१	२०११
प्रदेश १	२१.८३	२९.५५	३५.२०	४२.०२	४५.३५
मधेश प्रदेश	२२.०१	२८.८३	३६.०५	४६.०५	५४.०४
बागमती प्रदेश	२२.७९	२७.८१	३५.०५	४५.७०	५५.२९
गण्डकी प्रदेश	१३.६९	१७.४६	२०.१९	२३.८२	२४.७८
लुम्बिनी प्रदेश	१८.२०	२४.४४	३१.१३	३९.३५	४४.७४
कर्णाली प्रदेश	७.४५	८.९४	१०.४९	१२.६७	१५.२२
सुदूरपश्चिम प्रदेश	९.५८	१३.२०	१६.७९	२१.९१	२५.५३

स्रोत: केन्द्रीय तथ्याङ्क विभाग, विभिन्न वर्षका राष्ट्रिय जनगणनाहरू

८.५ चित्रात्मक प्रस्तुतिकरण

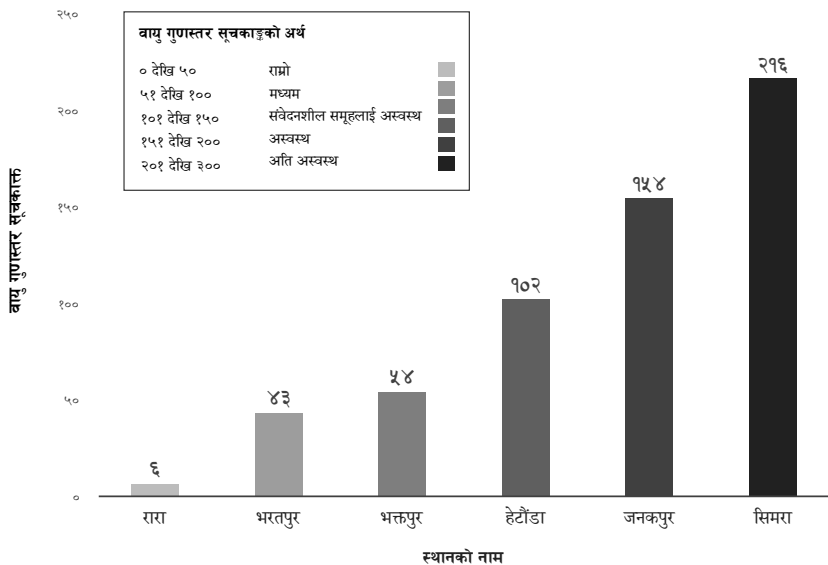
कुनै पनि वस्तु वा विशेषताको परिमाणलाई विभिन्न चित्रको सहायताले प्रस्तुत गर्ने विधिलाई चित्रात्मक प्रस्तुतिकरण भनिन्छ। चित्रात्मक प्रस्तुति मुख्यतया डायग्राम र ग्राफ गरी दुई प्रकारका हुन्छन्। यस खण्डमा हामी डायग्राम प्रस्तुतिका लागि स्तम्भचित्र (Bar Diagram), वृत्तचित्र (Pie-Chart) र नक्सा (Map or Cartograms) तथा ग्राफिकल प्रस्तुतिका लागि हिस्टोग्राम (Histogram), समय श्रृङ्खलाको ग्राफ (Graphs of Time Series or Histograms) का बारेमा अध्ययन गर्नेछौं।

८.५.१ स्तम्भचित्र

विभिन्न वर्ग, समूह वा श्रेणी (Category) मा विभक्त विशेषताहरूलाई तुलनात्मक अध्ययन गर्नका लागि स्तम्भचित्र प्रयोग गरिन्छ। दुई वा सोभन्दा बढी समूहहरूलाई कुनै विशेषताको आधारमा तुलना गर्दा स्तम्भचित्र (बार ग्राफ) प्रयोग गरिन्छ। स्तम्भचित्र

बनाउँदा तेर्सो अक्ष (X-axis) मा समूहको नाम र ठाडो अक्ष (Y-axis) मा तुलना गर्नुपर्ने अङ्क राखिन्छ । स्तम्भचित्रमा स्तम्भको उचाइले सम्बन्धित समूहको आवृत्ति वा तुलना गर्नुपर्ने अङ्कलाई प्रतिनिधित्व गर्दछ ।

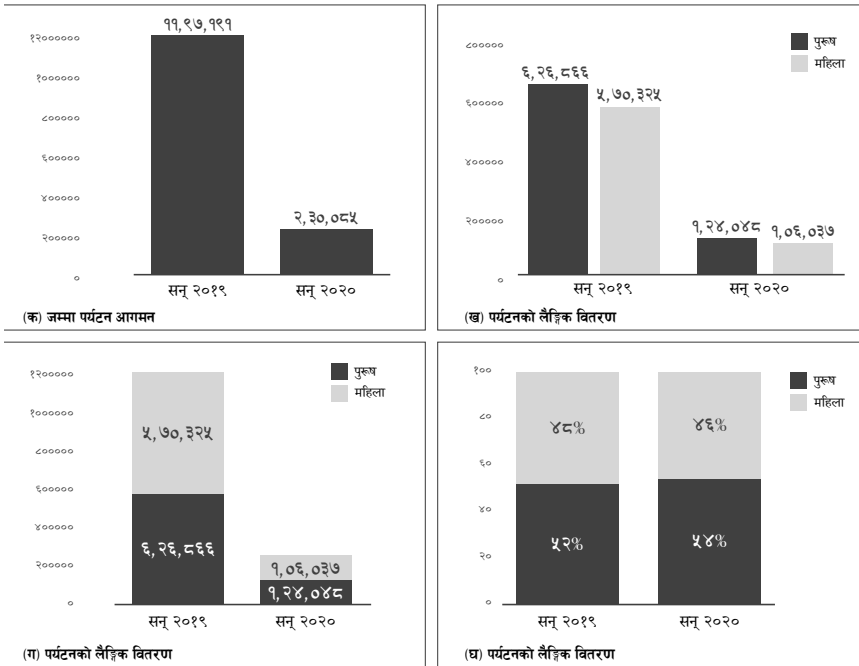
चित्र नं. ८.६: विभिन्न स्थानको वायुको गुणस्तर



स्रोत: वायु गुणस्तर निगरानी । नोट: यो तथ्याङ्क २०७८ मंसिर ८ गते बिहान ६:५० बजे

माथि चित्र नं. ८.६ मा प्रस्तुत गरिएको स्तम्भचित्रले नेपालका विभिन्न स्थानको वायु गुणस्तर अवस्था प्रतिनिधित्व गरेको छ । हरेक घण्टामा मापन गरिएको सूचकाङ्कलाई स्तम्भचित्रमा प्रस्तुत गर्दा ती स्थानहरूको वायु गुणस्तरको अवस्था थाहा पाइनुका साथै एक आपसमा तुलनात्मक अध्ययन गर्न समेत सहयोग पुऱ्याएको छ । उदाहरणका लागि मिति २०७८ साल मंसिर ८ गते बिहान ६:५० बजे रारा र भरतपुरको वायुको गुणस्तर राम्रो रहेको छ । यद्यपि राराको तुलनामा भरतपुरको वायु गुणस्तर कमसल देखिन्छ । सोही समयमा जनकपुरको वायु अस्वस्थ देखिन्छ भने सिमराको वायु निकै अस्वस्थ देखिन्छ । तुलनात्मक रूपमा हेटौँडाभन्दा जनकपुर बढी प्रदूषित छ । यसरी स्तम्भचित्रको सहायताले सरल र सहज तरिकाले तथ्याङ्क प्रस्तुति, विश्लेषण तथा व्याख्या गर्न सकिन्छ ।

चित्र नं. ८.७: नेपालमा पर्यटन आगमनको लैङ्गिक अवस्था



स्रोत: संस्कृति पर्यटन तथा नागरिक उड्डयन मन्त्रालय, नेपाल पर्यटन तथ्याङ्क, २०२०^{१९}

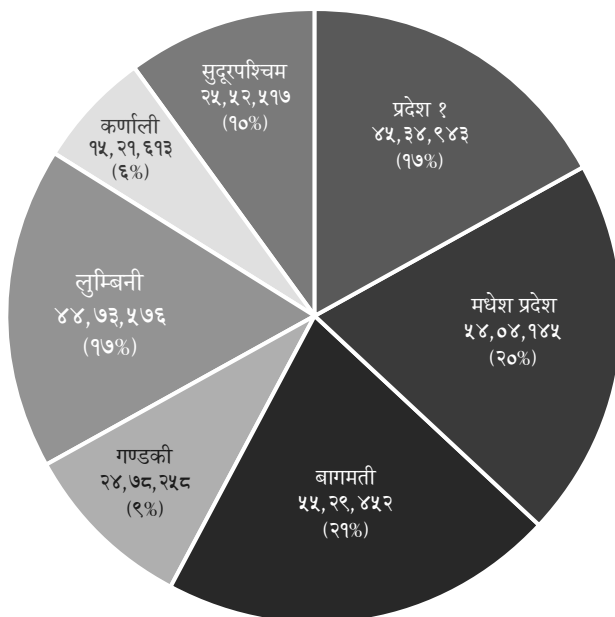
स्तम्भचित्र विभिन्न प्रकारका हुन्छन् । केही प्रमुख स्तम्भचित्रहरूमा सामान्य स्तम्भचित्र (Simple Bar Diagram), बहुस्तम्भचित्र (Multiple Bar Diagram), उप-विभाजित स्तम्भचित्र (Sub-divided Bar Diagram), प्रतिशत स्तम्भचित्र (Percentage Bar Diagram), आदि पर्दछन् र तिनीहरूलाई क्रमशः चित्र नं. ८.७ को प्यानल (क), (ख), (ग), (घ) मा देखाइएको छ ।

८.५.२ वृत्तचित्र (Pie chart)

कुनै आँकडाको कुल जम्मामा विभिन्न समूहहरूको कति कति हिस्सा छ भनेर तुलना गर्नुपर्दा वृत्तचित्रको प्रयोग गरिन्छ । जस्तै: नेपालको कुल जनसंख्यामा प्रदेश अनुसारको हिस्सा कति कति छ भनेर हेर्दा चित्र नं. ८.८ बमोजिमको वृत्तचित्रको प्रयोग गर्न सकिन्छ ।

19 Ministry of Culture, Tourism and Civil Aviation. (2021). *Nepal Tourism Statistics 2020*. Kathmandu, Nepal: Ministry of Culture, Tourism and Civil Aviation. Retrieved on November 24, 2021 from: https://www.tourism.gov.np/files/publication_files/316.pdf

चित्र नं. ८.८: वि.सं. २०६८ को प्रदेशगत जनसंख्या विवरण



स्रोत: केन्द्रीय तथ्याङ्क विभाग, राष्ट्रिय जनगणना, २०६८

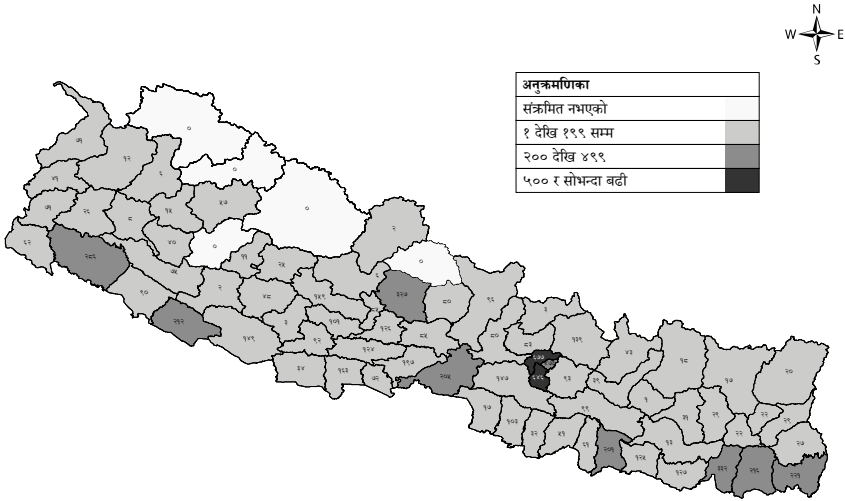
यसबाट नेपालको कुल जनसंख्यामा सबैभन्दा ठुलो हिस्सा (२१%) बागमती प्रदेशको देखिन्छ। त्यस्तै, प्रदेश नं. २ को २०% हिस्सा, प्रदेश नं. १ को १७% हिस्सा देखिन्छ र सबैभन्दा कम कर्णाली प्रदेशको ६% हिस्सा देखिन्छ। यसरी वृत्तचित्रका माध्यमबाट कुनै वस्तु वा विशेषताको कुल अंशमा विभिन्न समूहहरूको हिस्सा कति कति छ भनेर तुलनात्मक अध्ययन गर्न उपयुक्त मानिन्छ। तर एउटा मात्र समूहका लागि वृत्तचित्र प्रयोग गरिँदैन भने १० भन्दा बढी समूहहरूको लागि पनि वृत्तचित्र उपयुक्त मानिँदैन।

८.५.३ नक्सा

भौगोलिक सूचनाहरूलाई नक्साका माध्यमबाट प्रस्तुत गर्दा निकै प्रभावकारी हुन्छ। नक्साको प्रयोगले विभिन्न भौगोलिक एकाइहरूको सम्बन्धित सूचनाका आधारमा तुलनात्मक अध्ययन गर्न सजिलो हुन्छ। उदाहरणका लागि जिल्लाअनुसार मिति २०७८ मंसिर ७ मा कोभिड १९ को संक्रमण अवस्था चित्र नं. ८.९ मा प्रस्तुत गरिएको छ। यस नक्साले नेपालको कुन जिल्लामा उक्त दिन संक्रमित संख्या उच्च रहेको र कुन जिल्लामा संक्रमित नै नभएको भनी एकैछिनमा हेर्न, बुझ्न र बताउन सकिन्छ। जस्तै: चित्र ८.९ बाट मिति २०७८ मंसिर ७ गते काठमाण्डौँ उपत्यकामा संक्रमण उच्च रहेको र हुम्ला, डोल्पा, जाजरकोट, मनाङ, र ओखलढुङ्गामा कुनै पनि संक्रमित नरहेको भनेर

सजिलै पत्ता लगाउन सकिन्छ । यही तथ्याङ्क अनुच्छेद वा तालिकामा प्रस्तुत गरिएको भए पाठक वा दर्शकलाई बुझ्न, सम्झन र बुझाउन निकै समय लाग्ने थियो । संघ, प्रदेश तथा स्थानीय तहका नीतिनिर्माता एवं सरोकारवालाहरूलाई यसरी भौगोलिक खण्डीकृत तथ्याङ्क (Disaggregated Data) लाई नक्साको माध्यमबाट प्रस्तुत गर्न सकियो भने यसले कार्यकौशलतामा मात्र सुधार नगरी सही नीति तय गर्नमा पनि उल्लेख्य भूमिका निर्वाह गर्दछ ।

चित्र नं. ८.९: कोभिड-१९ संक्रमण अवस्था, २०७८ मंसिर ७ गते



स्रोत: स्वास्थ्य तथा जनसंख्या मन्त्रालय, स्थिति प्रतिवेदनहरू²⁰

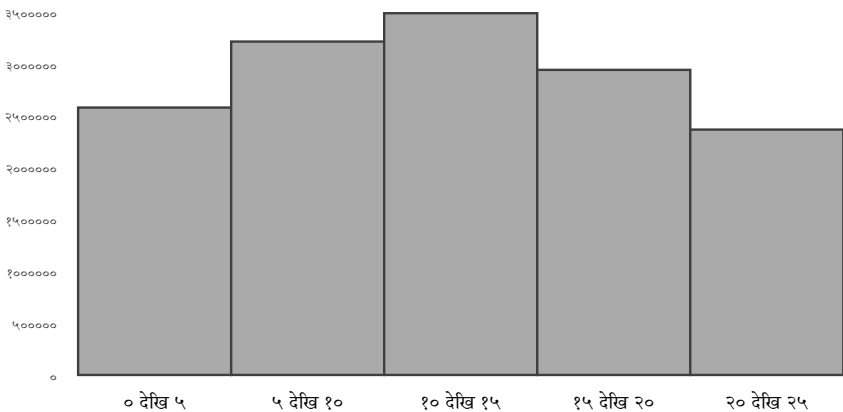
आजकल राष्ट्रिय तथा अन्तर्राष्ट्रिय सरकारी तथा गैर सरकारी निकायहरूले अधिकांश भौगोलिक सूचनाहरूलाई नक्साको प्रयोग गरी प्रस्तुत गर्ने गरेको पाइन्छ । केन्द्रीय तथ्याङ्क विभागले पनि विभिन्न क्षेत्रका भौगोलिक सूचनाहरूलाई नक्साको माध्यमबाट प्रस्तुत गरी प्रसार गरिरहेको छ । विभागले प्रकाशन गरेका नक्सा प्रयोग गर्नका लागि विभागको वेबसाइट <https://cbs.gov.np/gis-maps/> हेर्न सकिन्छ । यस्ता नक्साहरू भौगोलिक सूचना प्रणाली (Geographic Information System- GIS) सफ्टवेयरका माध्यमबाट तयार गरिन्छन् ।

20 Ministry of Health and Population. COVID-19: Situation Reports. Kathmandu, Nepal; Ministry of Health and Population. Retrieved on November 24, 2021 from: <https://covid19.mohp.gov.np/situation-report>

८.५.४ हिस्टोग्राम (Histogram)

अन्तराल समूहमा आवृत्ति वितरण गरिएको तथ्याङ्कको ग्राफिकल प्रस्तुति गर्ने उत्तम उपाय भनेको हिस्टोग्राम हो । जब अन्तराल समूहको आवृत्ति वितरणमा तथ्याङ्क प्रस्तुत गरिन्छ तब व्यक्तिगत आवृत्ति हराएर जान्छ । तर अन्तराल समूहको आवृत्तिका आधारमा ग्राफ बनाउन सकिन्छ । तसर्थ, लम्बाइ तथा चौडाइ अर्थात् क्षेत्रफलको अर्थ राख्ने स्तम्भचित्रहरू हिस्टोग्राम हुन् । हिस्टोग्रामको ग्राफमा ठाडो अक्षले आवृत्ति प्रतिनिधित्व गरेको हुन्छ । हिस्टोग्राममा प्रयोग हुन आयातको उचाइ सम्बन्धित अन्तरालसँग समानुपातिक हुन्छ भने चौडाइले अन्तराल प्रतिनिधित्व गरेको हुन्छ । चित्र ८.१० मा वि.सं. २०६८ को जनगणना बमोजिम २५ वर्ष मुनिका जनसंख्याको ५ वर्षे उमेर समूहअनुसार जनसंख्या वितरण प्रस्तुत गरिएको छ । यसले उमेर समूहअनुसार जनसंख्या तुलना गर्न सहजीकरण गर्दछ । जस्तै: चित्र ८.१० बमोजिम २५ वर्ष मुनिका जनसंख्यामा १० देखि १५ वर्ष उमेर समूहका जनसंख्या अन्य उमेर समूहभन्दा धेरै छन् भने २० देखि २५ वर्ष उमेर समूहका जनसंख्या अन्य उमेर समूहभन्दा थोरै छन् ।

चित्र नं. ८.१०: पाँच वर्ष उमेर समूहअनुसार २५ वर्ष भित्रका जनसंख्या वितरण



स्रोत: केन्द्रीय तथ्याङ्क विभाग, राष्ट्रिय जनगणना, २०६८²¹

८.५.५ समय शृङ्खलाको ग्राफ

कुनै अङ्क समयअनुसार अविच्छिन्न रूपमा कसरी परिवर्तन भएको छ भनेर हेर्न रेखाचित्रको प्रयोग गरिन्छ । रेखाचित्र बनाउँदा X-axis समय र Y-axis मा तुलना

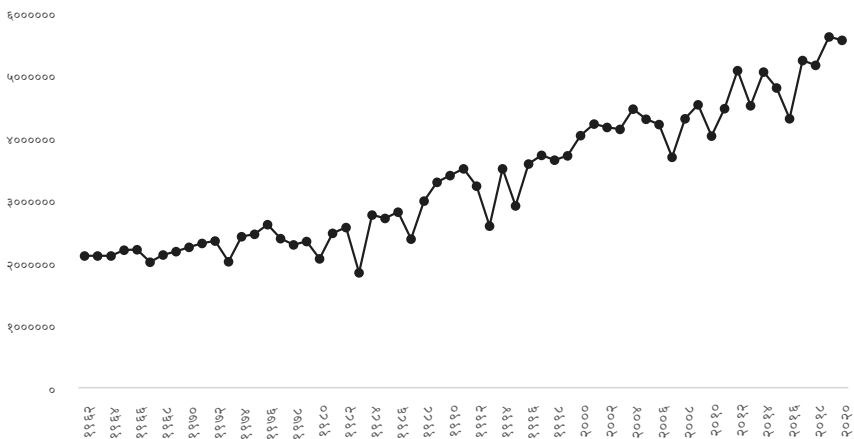
21 National Planning Commission and Central Bureau of Statistics.

National Data Profile. Kathmandu, Nepal; National Planning Commission and Central Bureau of Statistics. Retrieved on November 24, 2021 from:

<http://nationaldata.gov.np/>

गर्नुपर्ने अङ्क राखेर उक्त अङ्कहरूलाई समयअनुसार बिन्दुहरूबाट प्रस्तुत गरिन्छ। उक्त बिन्दुहरूलाई रेखाले जोडेेर समय श्रृङ्खलाको ग्राफ बनाइन्छ। जस्तै: नेपालमा सन् १९६२ देखि २०२० सम्म वार्षिक रूपमा कति कति धान उत्पादन भएको थियो भनेर चित्र ८.११ मा प्रस्तुत गरिएको छ। यसबाट नेपालमा धानको उत्पादन कुन वर्षमा कति भयो भनेर हेर्नुका साथ साथै कुन वर्षमा बढ्यो वा कुन वर्षमा घट्यो भनेर तुलना समेत गर्न सकिन्छ। यसका अतिरिक्त समग्रमा धानको उत्पादन बढ्दो वा घट्दो प्रवृत्ति (Trend) मा छ भनेर हेर्न सकिन्छ। चित्र नं. ८.११ अनुसार नेपालमा धानको उत्पादन सन् १९६२ देखि २०२० सम्म आइपुग्दा बिचका वर्षहरूमा केही घटबढ भए तापनि बढ्दो प्रवृत्तिमा रहेको छ।

चित्र नं. ८.११: नेपालमा धानको उत्पादन (मे. टन)



स्रोत: कृषि तथा पशुपन्छी विकास मन्त्रालय, Statistical Information on Nepalese Agriculture²²

८.६ डायग्राम र ग्राफको प्रयोगबाट तथ्याङ्कको विश्लेषण तथा व्याख्या

डायग्राम र ग्राफको सहायताले तथ्याङ्कको विश्लेषण गर्न सजिलो हुने कुरा हामीले माथिका खण्डहरूमा उल्लेख गरेका उदाहरणहरूबाट स्पष्ट भइसकेको छ। तथ्याङ्कको प्रस्तुतिकरणले वास्तविक संसारको चित्र देखिन्छ र के देखियो भन्ने कुरा व्यक्त गर्न सकिन्छ। धेरै सूचनाहरूलाई संक्षेपीकरण गरेर एकैसाथ प्रस्तुत गर्दा आकर्षक हुनुका साथै

22 Ministry of Agriculture and Livestock Development. *Statistical Information on Nepalese Agriculture*. Kathmandu, Nepal; Ministry of Agriculture and Livestock Development. Retrieved on November 24, 2021 from: <https://www.moald.gov.np/publication/Agriculture%20Statistics>

सम्भन सजिलो पनि हुन्छ । तसर्थ, तथ्याङ्क प्रयोग गर्ने अध्येता, सञ्चार जगतलगायत नीतिनिर्माताहरूले तथ्याङ्कको प्रस्तुतिकरणलाई निकै महत्त्व दिएका हुन्छन् । हामीले हेर्ने दैनिक पत्रपत्रिकामा पनि प्रायजसो कुनै न कुनै सूचनाहरूलाई चित्रात्मक प्रस्तुति गरेर तथ्याङ्क विश्लेषण गरेको पाइन्छ । वैज्ञानिक अनुसन्धानात्मक लेखहरूमा पनि तथ्याङ्कको विभिन्न विधि प्रयोग गरेर विश्लेषण एवं व्याख्या गरेको पाइन्छ । सार्वजनिक निकायहरूले आफ्ना प्रगति प्रतिवेदनहरूमा तथ्याङ्कका विभिन्न प्रस्तुतिकरण विधि प्रयोग गरेर विश्लेषण एवं व्याख्या गरिरहेका हुन्छन् । तथ्याङ्कको असल प्रस्तुतिले सुशासनको अभिन्न अवयव पारदर्शिता कायम गर्न मद्दत गर्दछ ।

८.७ अभ्यासका लागि प्रश्नहरू

क) तथ्याङ्कअनुसार प्रस्तुतिकरणको कुन विधि उत्तम होला ?

१) नेपालको विगत १० वर्षको कुल गार्हस्थ्य उत्पादन वृद्धिदर

२) पाँच वर्ष उमेर समूहअनुसार विद्यार्थीहरूको वितरण

३) कुनै परिवारले खाना, कपडा, औषधी, मनोरञ्जन र अन्य शीर्षकमा गरेको मासिक खर्च

४) कुनै गाउँपालिकाको वडाअनुसार निरपेक्ष गरिब परिवारको संख्या

५) विगत तिन वर्षमा भारत, चीन, अमेरिका र जापानले नेपाललाई दिएको अनुदान रकम रू.

ख) तलका भनाइहरू ठिक वा बेठिक कारणसहित बताउनुहोस् ।

१) तथ्याङ्कको प्रस्तुतिकरण नगरिकन पनि तथ्याङ्कको विश्लेषण र व्याख्या गर्न सकिन्छ ।

२) स्तम्भचित्रको प्रयोगले तुलनात्मक अध्ययन गर्न सजिलो हुन्छ ।

३) वृत्तचित्र ग्राफको उदाहरण हो ।

४) नीतिनिर्माताहरूका लागि तथ्याङ्क प्रस्तुतिकरणले खासै महत्त्व राख्दैन ।

ग) फरक छुट्याउनुहोस् ।

१) डायग्राम र ग्राफ

२) स्तम्भचित्र र हिस्टोग्राम

३) तथ्याङ्क विश्लेषण र तथ्याङ्कको व्याख्या

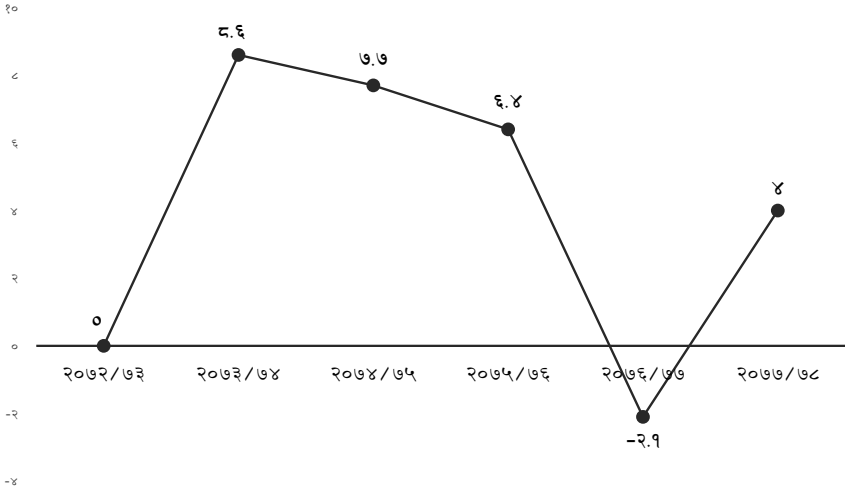
घ) तल दिइएको तथ्याङ्कलाई उपयुक्त विधिबाट प्रस्तुति गर्नुहोस् । (नोट: यो तालिकामा दिइएको तथ्याङ्कको स्रोत केन्द्रीय तथ्याङ्क विभाग हो । यो तथ्याङ्क केन्द्रीय तथ्याङ्क विभागले प्रकाशन गर्ने कुल गार्हस्थ्य उत्पादनसम्बन्धी तथ्याङ्कीय तालिकाबाट लिइएको हो । उक्त तथ्याङ्क <https://cbs.gov.np/national-accounts-of-nepal-2020-21/> मा उपलब्ध छ ।)

तालिका ८.४: औद्योगिक क्षेत्रको समूह कुल गार्हस्थ्य उत्पादनमा योगदान

औद्योगिक क्षेत्रको समूह	कुल गार्हस्थ्य उत्पादनमा योगदान (%)
प्राथमिक क्षेत्र	२६.४
द्वितीय क्षेत्र	१२.५
सेवा क्षेत्र	६१.१

ड) तलको चित्र अध्ययन गरेर सोधिएका प्रश्नहरूको जवाफ दिनुहोस् ।

चित्र नं. ८.१२: कुल गार्हस्थ्य उत्पादन वार्षिक वृद्धिदर



स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपालको राष्ट्रिय लेखा अनुमान, २०७७/७८

- १) यो चित्र कुन प्रकारको चित्र हो ?
 - २) चित्रमा के देखाइएको छ ?
 - ३) कुन आ.व.को वृद्धिदर नकारात्मक छ ?
 - ४) सबैभन्दा बढी वृद्धिदर कुन आ.व. देखिन्छ ?
 - ५) चित्रमा प्रस्तुत गरिएको तथ्याङ्क विश्लेषण गरी एक अनुच्छेद लेख्नुहोस् ।
- च) तलको तालिका अध्ययन गरेर सोधिएका प्रश्नको जवाफ दिनुहोस् ।

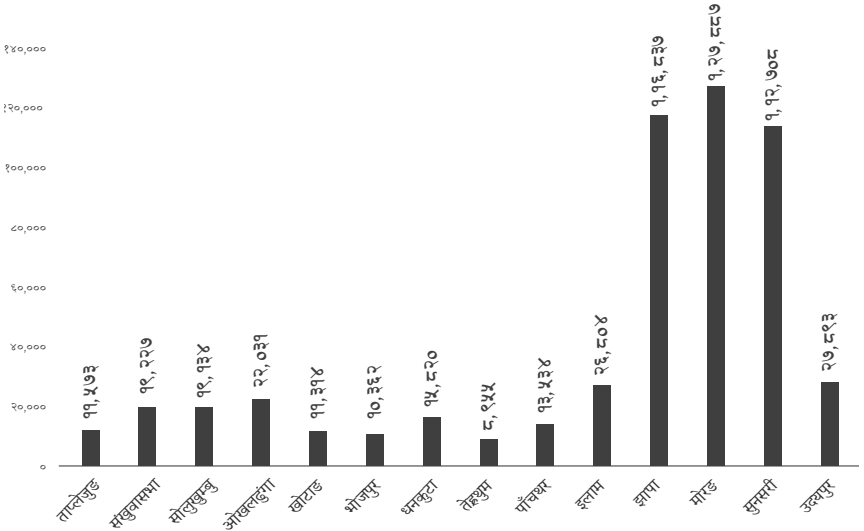
तालिका ८.४: विगत १० वर्षमा पेट्रोलियम पदार्थको आयात परिमाण

आ.व.	पेट्रोल (किलो लिटर)	डिजेल	मट्टितेल	हवाई इन्धन	हलुका डिजेल	फर्तान्स	ग्याँस (मे.टन.)
२०७६/७७	५,१२,१२८	१४,७३,५३६	१८,९२४	१,३७,४२४	०	०	४,४९,०६३
२०७५/७६	५,६६,८२७	१७,१४,९१७	२५,००४	२,००,१०८	०	०	४,२९,६०९
२०७४/७५	४,८८,६७५	१५,८८,८६९	२२,३३७	१,९७,२२०	०	०	३,७०,५६०
२०७३/७४	४,०७,२७०	१३,१९,८७३	१९,६०७	१,६४,८३६	०	३६	३,१२,९२८
२०७२/७३	२,४०,३८६	७,८५,६८५	१४,१९४	८३,८१९	०	७७	२,१४,१९४
२०७१/७२	२,८७,४७३	९,२१,७१४	१९,६५३	१,४१,४०४	०	८८३	२,५८,२९९
२०७०/७१	२,५३,३८१	८,०८,५६७	१८,४०९	१,२५,६७८	०	२,१७२	२,३२,६६०
२०६९/७०	२,२३,०८७	७,२१,२०३	२४,०६५	१,१५,८९६	२६०	२,४५६	२,०७,०३८
२०६८/६९	२,०२,४६७	६,५३,५६०	४१,६०९	१,०९,९०४	०	४४०	१,८१,४११
२०६७/६८	१,८८,०८२	६,५२,७६४	४३,३९९	९९,९९०	२२८	१,४३४	१,५९,२८६

स्रोत: नेपाल आयात निगम लिमिटेड^{२३}

- १) यो तालिकामा के प्रस्तुत गरिएको छ ?
 - २) पेट्रोल, डिजेल र मट्टितेल सबैभन्दा बढी आयात भएको वर्षहरू कुन कुन हुन् ?
 - ३) ग्याँसको आयात सबैभन्दा बढी र सबैभन्दा कम कुन कुन वर्षमा भएको छ ?
 - ४) पेट्रोल र डिजेलको विगत १० वर्षको आयात परिमाणलाई ग्राफमा प्रस्तुत गर्नुहोस् ।
 - ५) आ.व. २०७६/७७ मा पेट्रोल, डिजेल, मट्टितेल, र हवाई इन्धनको आयात परिमाणलाई वृत्तचित्रमा प्रस्तुत गर्नुहोस् ।
- छ) तलको चित्र हेरेर निम्नानुसारका प्रश्नको जवाफ दिनुहोस् ।

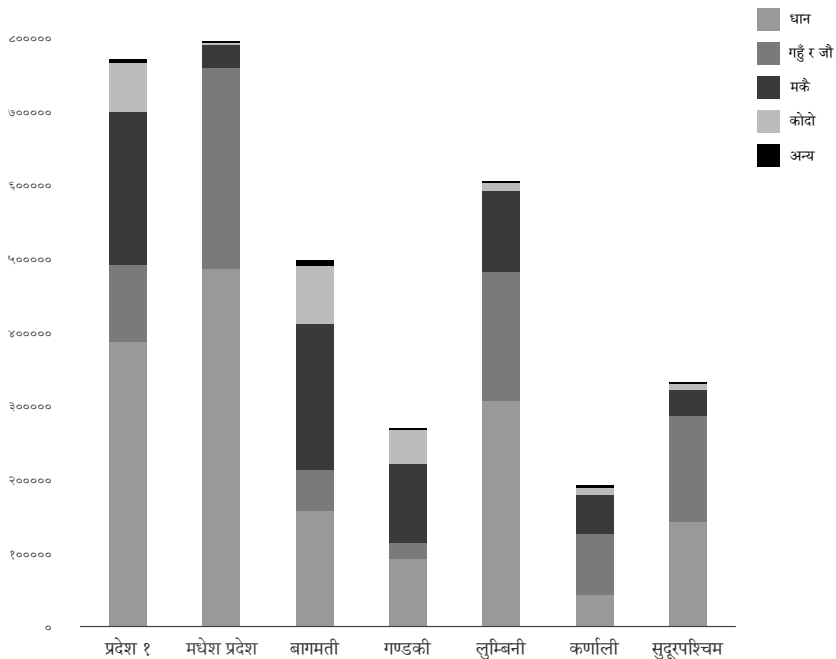
चित्र नं. ८.१३: प्रदेश १ अन्तर्गतका जिल्लाहरूमा आर्थिक क्रियाकलापमा संलग्न जनसंख्या विवरण



स्रोत: केन्द्रीय तथ्याङ्क विभाग, आर्थिक गणना, २०१८

- १) यो चित्र कुन प्रकारको चित्र हो ?
 - २) यो चित्रले के देखाउन खोजेको छ ?
 - ३) प्रदेश १ मा आर्थिक क्रियाकलापमा संलग्न सबैभन्दा बढी र सबैभन्दा कम जनसंख्या भएको जिल्ला कुन कुन हुन् ?
 - ४) यो तथ्याङ्कलाई तालिकामा प्रस्तुत गर्नुहोस् ।
 - ५) तालिका र यो चित्र तुलना गरेर तथ्याङ्क विश्लेषणका लागि कुन बढी सजिलो अनुभव भयो, छलफल गर्नुहोस् ।
- ज) तलको तालिका अध्ययन गरेर सोधिएका प्रश्नको जवाफ दिनुहोस् ।

चित्र नं. ८.१४: प्रदेश र अन्नअनुसार अन्न खेती गरिने जग्गाको क्षेत्रफल (हेक्टर)



स्रोत: केन्द्रीय तथ्याङ्क विभाग, राष्ट्रिय कृषिगणना, २०६८^{२४}

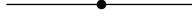
- १) यो कस्तो प्रकारको चित्र हो ?
 - २) यस चित्रमा केको तथ्याङ्क प्रस्तुत गरिएको छ ?
 - ३) सबैभन्दा बढी धान खेती गर्ने प्रदेश कुन हो ?
 - ४) कोदोको सबैभन्दा बढी खेती गर्ने प्रदेश कुन हो ?
 - ५) मकैको सबैभन्दा कम खेती गर्ने प्रदेश कुन हो ?
 - ६) यो चित्रमा प्रस्तुत गरिएको तथ्याङ्क विश्लेषण गरी एक अनुच्छेद लेख्नुहोस् ।
- भ) तथ्याङ्क प्रस्तुतिकरणको महत्त्व माथि प्रकाश पादौ प्रस्तुतिकरणमा ध्यान दिनुपर्ने मुख्य मुख्य कुराहरू उल्लेख गर्नुहोस् ।

24 Central Bureau of Statistics. (2011). *Province Summary of Agriculture Census 2011*. Kathmandu, Nepal; Central Bureau of Statistics. Retrieved on November 28, 2021 from: <https://cbs.gov.np/province-summary-of-agriculture-census-2011/>

द.द सन्दर्भ सामग्री

तथ्याङ्क प्रस्तुतिकरणका धेरै विधिहरूका बारेमा विस्तृत जानकारीका लागि निम्न सामग्री उपयोगी छ ।

१. Fundamentals of Statistics by S.C. Gupta, Sixth Revised & Enlarged Edition, Himalaya Publishing House, 2009



तथ्याङ्क विश्लेषणका लागि
आधारभूत गणितीय सूचक

९.१ परिचय

सूचकहरू प्रायः शुद्ध संख्या, दर, अनुपात, समानुपात, प्रतिशत वा प्रतिशत बिन्दुका रूपमा प्रयोग भएका हुन्छन् । जस्तै: नेपालको कुल जनसंख्या, कोरा जन्मदर, लिङ्ग अनुपात, कोभिड-१९ विरूद्ध खोप पाएका जनसंख्याको प्रतिशत आदि । तथ्याङ्क विश्लेषणका लागि साधारण गणितीय क्रियाहरू जस्तै: जोड, घटाउ, गुणा तथा भागको प्रयोग गरी गुणस्तरीय कच्चा पदार्थ (तथ्याङ्क वा आँकडा) बाट दर, अनुपात, समानुपात तथा प्रतिशतजस्ता सूचकहरू निर्माण हुन्छन् । यस पाठमा हामीले यिनै सूचकहरूको गणना कसरी गर्ने र तिनीहरूलाई कसरी व्याख्या गर्ने वा प्रयोग गर्ने भन्ने सन्दर्भमा छलफल गर्दछौं । कतिपय प्रयोगकर्ताहरूले यी सूचकहरूको प्रयोग चाहिँ गर्ने तर सही तरिकाले प्रयोग नगर्ने गरेको पनि देखिन्छ । उदाहरणका लागि मातृमृत्यु अनुपात र मातृमृत्यु दर भन्नुमा ठुलो फरक हुन्छ । कतिपय मानिसले यस्ता कुरामा ध्यान दिएको पाइँदैन । सामान्यतया मातृमृत्यु (Maternal Mortality) मापन गर्नका लागि अनुपात प्रयोग गरिन्छ । दर र अनुपात बिचको फरक के हो भन्ने विषयमा पनि कतिपय मानिसहरूलाई थाहा नहुन सक्दछ । सूचक प्रयोग गर्दा तिनीहरूको अर्थ र तिनीहरूले देखाउन खोजेको विषयलाई यथार्थ रूपमा व्यक्त गरियो भने मात्र तथ्याङ्क प्रयोगको सान्दर्भिकता पुष्टि हुन्छ । यस्तै रोचक विषयहरूमा यस पाठमा चर्चा गरिएको छ ।

९.२ संख्याको सामान्य ज्ञान

तथ्याङ्क विश्लेषण गरी सूचक निर्माण गर्न तथा प्रयोग गर्नका लागि संख्याको सामान्य ज्ञान हुन जरूरी हुन्छ । संख्याहरूलाई मुख्यतया निम्नानुसारका समूहमा विभाजन गर्न सकिन्छ ।

क) प्राकृतिक संख्या (Natural Numbers): सामान्य दैनिक जीवनमा प्रयोग हुने गन्तीका संख्याहरूलाई प्राकृतिक संख्या भनिन्छ । जस्तै: १, २, ३, अनन्तसम्म ।

ख) पूर्ण संख्या (Whole Numbers): प्राकृतिक संख्या र शून्य समेतको समूहलाई पूर्ण संख्या भनिन्छ । जस्तै: ०, १, २, ३, अनन्तसम्म ।

ग) पूर्णाङ्क (Integers): पूर्ण संख्या र नकारात्मक संख्याहरूको समूहलाई पूर्णाङ्क भनिन्छ । जस्तै: अनन्तदेखि, -३, -२, -१, ०, १, २, ३, अनन्तसम्म ।

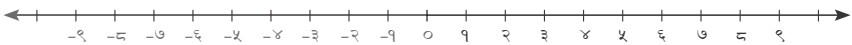
घ) भिन्नमा व्यक्त गर्न सकिने संख्या (Rational Numbers): पूर्णाङ्कमा भिन्नात्मक संख्या सहितको समूहलाई भिन्नमा व्यक्त गर्न सकिने संख्या भनिन्छ । जस्तै: ८, ११, ०.३,

७/९, आदि । यहाँ, पूर्ण संख्या ८, ११ जस्ता संख्याहरूलाई पनि भिन्नमा व्यक्त गर्न सकिन्छ । जस्तै: $८ = ८/१$ । यी संख्याहरू दशमलवमा व्यक्त गर्दा कि दोहोरिने प्रकृतिका हुन्छन् कि टुङ्गिने प्रकृतिका हुन्छन् । जस्तै: $१०/३ = ३.३३३३३३...$ (दोहोरिने), $२२/७ = ३.१४२८५७१४२८५७१४३ ..$ (दोहोरिने), $४५/२ = २२.५$ (टुङ्गिने), आदि ।

ड) भिन्नमा व्यक्त गर्न नसकिने संख्या (Irrational Numbers): भिन्नको रूपमा व्यक्त गर्न नसकिने संख्याहरूलाई भिन्नमा व्यक्त गर्न नसकिने संख्या भनिन्छ । जस्तै: $\sqrt{२}$, $\sqrt{३}$, $\sqrt{७}$ आदि । यी संख्याहरू भिन्नको रूपमा व्यक्त गर्न पनि सकिँदैन र दशमलवको प्रयोग गर्दा नदोहोरिने वा नटुङ्गिने खालका हुन्छन् । जस्तै: $\sqrt{२} = १.४१४२१३५६२३७३०९५$ (दोहोरिएको पनि छैन र टुङ्गिएको पनि छैन ।)

च) वास्तविक संख्या (Real Number): वास्तविक संख्या भनेको भिन्नमा व्यक्त गर्न सकिने र नसकिने संख्याहरूको समूह जसलाई संख्याको रेखामा प्रस्तुत गरिन्छ । जस्तै: चित्र नं. ९.१ मा वास्तविक संख्या रेखा देखाइएको छ । जहाँ वाणको संकेतले यो रेखा जति पनि (अनन्त) लम्ब्याउन सकिन्छ भन्ने बुझाउँछ भने रेखाका हरेक बिन्दु प्रतिनिधित्व गर्ने भिन्नमा व्यक्त गर्न सकिने र नसकिने संख्याहरू हुन्छन् । अर्थात्, संख्या ३ र ४ को बिचमा अगणित संख्याहरू छन् ।

चित्र नं. ९.१: वास्तविक संख्या रेखा



छ) काल्पनिक संख्या (Imaginary Number): काल्पनिक संख्या त्यस्ता संख्या हुन् जसलाई गणितीय क्रियाहरूले गणना गर्न सम्भव हुँदैन । जस्तै: $\sqrt{-१}$, $\sqrt{-४}$ आदि । यी संख्याहरूलाई संकेत अंग्रेजी वर्ण वा ग्रिक संकेतहरूले जनाइन्छ । जस्तै: $\sqrt{-१} = i$, वृत्तको परिधि र व्यासको अनुपात = π आदि ।

ज) जटिल संख्या (Complex Number): वास्तविक संख्या र काल्पनिक संख्याको मिश्रणबाट जटिल संख्या बन्दछन् । जस्तै: $२ + ३i$, $(\pi - ०.६)$, आदि ।

९.३ भिन्नको सामान्य ज्ञान

विभिन्न विशेषताहरू भएका समूह वा मिश्रणमा कुनै एक विशेषताको हिस्सा कति छ भनेर जनाउने गणितीय विधिलाई भिन्न (Fraction) भनिन्छ । अर्को शब्दमा भन्दाखेरि कुनै सिङ्गो वस्तुलाई बराबर दुई वा दुईभन्दा बढी भागमा विभाजन गरी केही भागलाई जम्मा भागको कति छ भनेर जनाउने संख्या भिन्न संख्या हो । गणितीय भाषामा भिन्न भनेको

$a \div b$ को रूपमा लेख्न सकिने संख्या हो, जहाँ 'a' ले पूर्ण संख्या र 'b' ले प्राकृतिक संख्या प्रतिनिधित्व गरेका छन् । जस्तै: नेपालको कुल जनसंख्या २,६४,९४,५०४ मध्ये महिला १,३६,४५,४६३ र पुरुष १,२८,४९,०४१ रहेका छन् । यो आँकडालाई भिन्नमा यसरी व्यक्त गर्न सकिन्छ: महिलाको हिस्सा = $(१,३६,४५,४६३) / (२,६४,९४,५०४)$, पुरुषको हिस्सा = $(१,२८,४९,०४१) / (२,६४,९४,५०४)$ । त्यस्तै, एउटा सिङ्गो स्याउलाई ४ भागमा बराबर विभाजन गरेर सो मध्ये ३ भाग जनाउनका लागि $३/४$ प्रयोग गरिन्छ । दर, अनुपात, प्रतिशत जस्ता गणितीय सूचकहरू निकाल्नको लागि भिन्नको प्रयोग हुन्छ । तसर्थ, भिन्नका बारेमा केही महत्त्वपूर्ण जानकारी लिइराखौं ।

क) भिन्नका अंश (Numerator) र हर (Denominator) गरेर दुई भाग हुन्छन् । भिन्न $३/४$ मा ३ अंश हो र ४ हर हो । गणितीय क्रियाका लागि अंशलाई हरले भाग गरेर भिन्नलाई सरल वा दशमलवमा व्यक्त गरिन्छ ।

उदाहरण: $१६/२० = ४/५ = ०.८$, यो उदाहरणमा दुईवटा भिन्नहरू $१६/२०$ र $४/५$ बराबर छन् र यस्ता बराबर भिन्नहरूलाई समानुपात भनिन्छ । यहाँ २० ले १६ लाई भाग गर्दा र ५ ले ४ लाई भाग गर्दा ०.८ नै आउँछ । यस उदाहरणबाट अर्को महत्त्वपूर्ण कुरा पनि जान्नुपर्ने हुन्छ । कुनै पनि भिन्नलाई साभ्ना संख्याले अंश र हरमा एकैसाथ भाग वा गुणा गर्न पाइन्छ । अर्थात्, $१६/२०$ को अंश र हर दुवैलाई ४ ले भाग गरेर $४/५$ प्राप्त भएको हो । अर्कोतर्फ $४/५$ को अंश र हर दुवैलाई ४ ले गुणा गरेर $१६/२०$ प्राप्त भएको हो ।

ख) दुई वा दुईभन्दा बढी भिन्नहरू तुलना गर्नुपर्दा सबैलाई समान हर बनाउनु पर्दछ । साथै भिन्नहरूको जोड, घटाउ गर्न पनि समान हर आवश्यक हुन्छ । जस्तै: ०, १, २, ३, अनन्तसम्म ।

उदाहरण: $४/५$ र $१९/२०$ मा कुन ठुलो भन्नका लागि दुवैलाई समान हर बनाऔं ।
 $४/५ = (४ \times ४) / (५ \times ४) = १६/२०$
 $१६/२०$ र $१९/२०$ मा $१९/२०$ ठुलो भयो ।

ग) कुनै पनि भिन्नमा अंशभन्दा हर ठुलो संख्या छ भने Proper Fraction भनिन्छ र अंशभन्दा हर सानो छ भने Improper Fraction भनिन्छ । अंश र हर बराबर भएको भिन्न पूर्ण संख्या हो ।

उदाहरण:

$४/५ = \text{Proper Fraction}$

$९/७ = \text{Improper Fraction}$

$१३/१३ = १$ पूर्ण संख्या

घ) ठुला ठुला संख्या प्रयोग भएका भिन्नहरू पढ्न बुझ्न त्यति रूचि लाग्दैन । तसर्थ, भिन्नहरूलाई प्रतिशतमा व्यक्त गरिन्छ । जस्तै माथिको नेपालको जनसंख्यामा महिलाको हिस्सा प्रतिनिधित्व गर्ने भिन्न साङ्ख्यिकी ठुलो संख्याले बनेको छ, यसलाई प्रतिशतमा व्यक्त गर्ने हो भने महिलाको ५१.५% र पुरुषको ४८.५% हुन आउँछ ।

९.४ अनुपात, समानुपात, दर तथा प्रतिशतको व्याख्या र प्रयोग

९.४.१ अनुपात (Ratio)

दुईवटा संख्याको तुलना गर्दा अनुपात प्रयोग गरिन्छ । जस्तै: लिङ्ग अनुपात, डाक्टर र जनसंख्याको अनुपात आदि । राष्ट्रिय जनगणना, २०६८ बमोजिम नेपालको लिङ्ग अनुपात ९४% छ । यसको अर्थ, नेपालमा प्रति १०० महिलाको जनसंख्या हुँदा पुरुषको जनसंख्या ९४ जना छन् भन्ने हुन्छ । त्यस्तै, सन् २०१८ को तथ्याङ्क अनुसार प्रति १००० जनसंख्याका लागि ०.७४९ डाक्टर उपलब्ध छन् । अर्थात् १००० जनसंख्याका लागि मुश्किलले एक जना डाक्टर उपलब्ध छन् ।²⁵ यहाँ, डाक्टरको संख्या भिन्नात्मक भएको हुनाले यसलाई समान अनुपात प्रयोग गरी प्रति दश हजार जनसंख्यामा ७ जना डाक्टर उपलब्ध छन् भनेर भन्न सकिन्छ । यसलाई अनुपातमा व्यक्त गर्दा डाक्टर र जनसंख्याको अनुपात ७:१०,००० हुन्छ । अनुपातलाई a to b , $a \div b$, $a:b$ भनेर जनाउने प्रचलन छ । उदाहरण: कुनै विद्यालयमा २६ जना शिक्षक र ४०३ जना विद्यार्थी छन् भने विद्यार्थी शिक्षक अनुपात कति होला ?

विद्यार्थी शिक्षक अनुपात निकाल्दा अंशमा विद्यार्थीको संख्या र हरमा शिक्षकको संख्या राख्नुपर्दछ ।

अर्थात्, विद्यार्थी शिक्षक अनुपात $403/26 = 31/2 = 31:2$ हुन्छ । यसले ३१ जना विद्यार्थीका लागि २ जना शिक्षक उपलब्ध छन् भन्ने बुझाउँछ । यही उदाहरणलाई शिक्षक विद्यार्थी अनुपात भन्ने हो भने $2:31$ हुन्छ र यसले पनि ३१ जना विद्यार्थीका लागि २ जना शिक्षक उपलब्ध छन् भन्ने नै बुझाउँछ । अनुपातलाई व्याख्या गर्ने बेला त्यसको भाव र सार नबिग्रने गरी गर्नुपर्दछ ।

अर्को उदाहरण हेरौं,

केन्द्रीय तथ्याङ्क विभाग, नेपाल बहुसूचक सर्वेक्षण, २०७५/७६ को अनुसार नेपालमा आर्थिक वर्ष २०७५/७६ मा जम्माजम्मी १३,३२९ छोरा र १२,३७९ छोरी जन्मिएका

25 The World Bank. *Physicians per 1000 people in Nepal*. Retrieved on November 25, 2021 from: <https://data.worldbank.org/indicator/SH.MED.PHYS.ZS?locations=NP>

थिए, जसमध्ये १२,६३४ छोरा र ११,८२८ छोरी जीवित थिए भने ६९४ छोरा र ५५१ छोरीको मृत्यु भइसकेको थियो।²⁶ यस आँकडाबाट विभिन्न अनुपातहरू निकालेर हेर्दा, जन्मदाको लिङ्ग अनुपात १.०८ अर्थात् १०८:१००, जीवित बच्चाको लिङ्ग अनुपात १.०७ अर्थात् १०७:१०० र मृतक बच्चाको लिङ्ग अनुपात १.२६ अर्थात् १२६:१०० देखिन्छ। यी अनुपातहरूले के बताइरहेका छन् भने नेपालमा १०० छोरी जन्मिँदै गर्दा १०८ जना छोरा जन्मिरहेका छन् र १०० छोरीको मृत्यु भइरहँदा १२६ छोराको मृत्यु भइरहेको छ। जीवितमध्ये १०० छोरीहरू र १०७ छोरा रहेका छन्। यी सूचनाहरू लैङ्गिक संवेदनशीलता अध्ययन गर्न र सोअनुसार रणनीति तय गर्न निकै महत्त्वपूर्ण मानिन्छन्। उदाहरणका लागि १०० छोरी जन्मिँदै गर्दा १०८ छोरा जन्मनु भनेको प्राकृतिक वितरण हो वा होइन, सोचनीय प्रश्न हुनसक्दछ।²⁷ छोरा र छोरीको जन्म बराबर संख्यामा वा केहीमात्र घटबढ किन भएन ? यस्तो अवस्थामा यदि लिङ्ग पहिचान गरेर गर्भपतन गराइएको हो भने त्यसलाई रोक्न थप रणनीति आवश्यक पर्नसक्दछ। त्यस्तै मृत्यु हुने अनुपातमा पनि छोराको संख्या किन बढी भयो ? यसले कतै प्राकृतिक असन्तुलन त निम्त्याउँदै छैन ? यस्ता गम्भीर प्रश्नहरू आउन सक्दछन्। तसर्थ, तथ्याङ्क व्यवस्थापनमा अनुपातको प्रयोग र महत्त्व निकै ठुलो छ।

अनुपातलाई प्रतिशत, प्रति हजार, प्रति दश हजार, प्रति लाख आदिमा व्यक्त गर्ने चलन छ। उदाहरणका लागि मातृमृत्यु अनुपातलाई हेरौं। सन् २०१६ मा नेपालको मातृमृत्यु अनुपात प्रति एक लाख जीवित जन्ममा २३९ थियो। यसको अर्थ, नेपालमा एक वर्षमा एक लाख जीवित जन्म भइरहँदा गर्भ वा गर्भसँग सम्बन्धित समस्याका कारण २३९ जना आमाहरूले ज्यान गुमाउँछन्। अर्थात्, यो अनुपात प्रति लाखमा व्यक्त भएको छ। यो सूचना निम्न बमोजिमको मातृमृत्यु अनुपातको सूत्र प्रयोग गरी निकालिएको हो।

मातृमृत्यु अनुपात = (निश्चित अवधिमा निश्चित स्थानमा भएको मातृमृत्यु संख्या) / (सोही अवधिमा र स्थानमा जीवित जन्मिएका बच्चाहरूको कुल संख्या) × १०००००

26 Central Bureau of Statistics. 2020. *Multiple Indicator Cluster Survey 2019*. Kathmandu, Nepal. Central Bureau of Statistics. Retrieved on 28 November, 2021 from: https://cbs.gov.np/wp-content/uploads/2020/12/Nepal_MICS6-2019_Fi-nal-Online-Version-15-Dec-2020-v2.pdf

27 विभिन्न अध्ययनहरूले जन्मदाको लिङ्ग अनुपातका बारेमा चर्चा गरेका छन्। उदाहरणका लागि यो लिङ्गमा हेर्न सकिन्छ: <https://ourworldindata.org/gender-ratio> (retrieved on 28 November 2021)

अनुपातका सम्बन्धमा विशेष ख्याल गर्नुपर्ने कुरा के छ भने यसमा प्रयोग हुने भिन्नको अंश र हरमा कुनै पनि साभ्भा भाग वा विशेषता हुँदैन । दुई फरक वस्तु वा विशेषताको तुलना गर्न अनुपात प्रयोग गरिन्छ । जस्तै: लिङ्ग अनुपातमा अंशमा पुरूषको संख्या र हरमा महिलाको संख्या प्रयोग भएको हुन्छ, जहाँ कुनै साभ्भा संख्या हुँदैन । त्यसैगरी मातृमृत्यु अनुपातमा पनि अंशमा मातृमृत्यु संख्या र हरमा जीवित जन्म संख्या हुन्छ, जहाँ कुनै साभ्भा संख्या हुँदैन ।

१.४.२ समानुपात (Proportion)

अंशमा भएका सबै विशेषता वा भाग वा वस्तु हरमा पनि समावेश छ भने त्यस्तो भिन्नलाई समानुपात भनिन्छ । यसलाई सिङ्गो समूहको कुनै भागको तुलना गर्न प्रयोग गरिन्छ । समानुपातलाई प्रायः प्रतिशतमा उल्लेख गर्ने गरिन्छ । उदाहरणका लागि नेपालको कुल जनसंख्यामा महिला ५१.५% रहेका छन् । यसको अर्थ, कुल जनसंख्यामा महिलाको हिस्सा ५१.५% छ भन्ने हुन्छ । यहाँ सिङ्गो समूह कुल जनसंख्या हो र यसको कुनै एक भाग महिलाको जनसंख्या हो । अनुपात र समानुपातको फरक हेर्नका लागि तलका सूत्रहरू हेरौं ।

अनुपात = (पुरूषको संख्या) / (महिलाको संख्या)

समानुपात = (महिलाको जनसंख्या) / (कुल जनसंख्या)

माथिका दुई सूत्रहरूमध्ये अनुपातको सूत्रमा हरमा महिलाको जनसंख्या मात्र छ, तर कुनै पनि पुरूषको जनसंख्या समावेश छैन । अर्थात् अंशमा भएको विशेषता हरमा छैन । दुई फरक विशेषताको तुलना गरिएको छ । अर्कोतर्फ समानुपातको सूत्रमा अंशमा महिलाको जनसंख्या छ र यो जनसंख्या हरमा भएको कुल जनसंख्यामा पनि समावेश छ ।

१.४.३ दर (Rate)

निश्चित समयमा अवधिमा कुनै घटना घट्ने आवृत्ति दर्साउने सूचक नै दर हो । दरलाई पनि भिन्नको सहायताले गणना गरिन्छ । अंशमा निश्चित अवधिमा घटेका जम्मा घटनाहरूको संख्या र हरमा सो घटना प्रत्यक्ष प्रभावित हुनसक्ने सम्भावना भएका कुल जनसंख्या राखेर दर निकालिन्छ । उदाहरणका लागि अंशमा कुनै स्थानको कुनै निश्चित अवधिभित्र भएको मातृमृत्यु संख्या र हरमा सोही स्थान र समयमा मातृमृत्यु हुनसक्ने सम्भावना भएका गर्भवती वा सुत्केरी आमाहरूको कुल जनसंख्या राखेर हिसाब गर्ने हो भने मातृमृत्यु दर आउँछ । यसरी मातृमृत्यु अनुपात र मातृमृत्यु दरमा ठुलो अन्तर छ । दरलाई पनि प्रतिशतमा, प्रति हजारमा, प्रति दश हजारमा, प्रति लाखमा व्यक्त गर्ने प्रचलन छ । कतिमा व्यक्त गर्ने भन्ने सन्दर्भमा सम्बन्धित घटना घट्ने आवृत्ति कतिको धेरै छ भन्ने कुराले निर्धारण गर्दछ । जस्तै: जन्मदर वा मृत्युदर प्रतिशतमा हेरियो भने सानो संख्या आउँछ । त्योभन्दा

प्रति हजारमा हेरियो भने व्याख्या गर्न र बुझाउन सजिलो हुने संख्या आउँछ। उदाहरणका लागि नेपालको राष्ट्रिय जनगणना, २०६८ बमोजिम नेपालको कोरा मृत्युदर प्रति हजार जनसंख्यामा ७.३ रहेको छ। दरले एक किसिमको सम्भाव्यता (Probability) पनि बताउँछ। जस्तै: कोरा मृत्युदर ७.३ भन्दै गर्दा, सो अवधिमा १००० मानिसमा ७.३ जना मर्ने सम्भावना छ भन्न पनि सकिन्छ।

समय अन्तरालसँगै विभिन्न कुराहरूको घटबढ हुने क्रम चलिरहेको हुन्छ। यस्तो अवस्थामा कुनै कुरा कुन दरमा बढेको वा घटेको छ भनी सोको मात्रा जनाउनका लागि वृद्धिदर (Growth Rate) निकाल्ने गरिन्छ। वृद्धिदर निकाल्ने सरल सूत्र यसप्रकार छ:

$$\text{वृद्धिदर} = \frac{(\text{जम्मा वृद्धि भएको संख्या})}{(\text{वृद्धि हुन सुरु भएको बेलाको संख्या} \times \text{समयावधि})} \times १००$$

उदाहरण: वि.सं. २०७६/७७ को नेपालको आधार मूल्यको कुल गार्हस्थ्य उत्पादन (स्थिर मूल्यमा) = २०,६४,५९९ दश लाख

वि.सं. २०७७/७८ को नेपालको आधार मूल्यको कुल गार्हस्थ्य उत्पादन (स्थिर मूल्यमा) = २१,४६,८२४ दश लाख

एक वर्षको अवधिमा वृद्धि भएको कुल गार्हस्थ्य उत्पादन = ८२,२२४ दश लाख

समयावधि = १ वर्ष

$$\text{आर्थिक वृद्धिदर} = \frac{(८२,२२४)}{(२०,६४,५९९ \times १)} \times १०० = ३.९८\%$$

यो सूचकलाई प्रस्तुत गर्दा, नेपालको आर्थिक वर्ष २०७७/७८ मा आर्थिक वृद्धिदर ३.९८% रहेको छ भन्नुपर्दछ। अर्थात्, आर्थिक वर्ष २०७६/७७ को तुलनामा आर्थिक वर्ष २०७७/७८ मा कुल गार्हस्थ्य उत्पादनमा ३.९८% ले वृद्धि भएको छ।

१.४.४ प्रतिशत (Percentage)

विभिन्न भिन्नहरूलाई तुलना गर्न सहज गराउने उद्देश्यले समान भाग अर्थात् १०० भागमा परिवर्तन गरिने अङ्कलाई प्रतिशत भनिन्छ। कुनै पनि भिन्नलाई प्रतिशतमा बदल्नका लागि १०० ले गुणा गरी % चिन्ह दिनुपर्दछ। खासगरी दर, अनुपात तथा समानुपातहरूलाई प्रतिशतमा व्यक्त गरिन्छ। जस्तै: कोभिड १९ को संक्रमण दर २% भनेको एक सय जनसंख्यामा २ जनालाई संक्रमण भएको भन्ने बुझिन्छ। प्रतिशतको प्रयोगले कुल जम्मा फरक-फरक हुँदा पनि अङ्कहरूलाई तुलना गर्न सजिलो हुन्छ। जस्तै: कुनै दुईवटा समितिहरूमा क्रमशः ५० र ८० जना सदस्यहरू छन्, जसमा पहिलो समितिमा २५ र दोस्रो समितिमा ३० जना महिला छन्। यी दुई समितिमा महिला सहभागिता तुलना गर्दा

अङ्गलाई मात्रै तुलना गर्दा गाह्रो पर्दछ । भिन्नमा तुलना गर्दा पनि गाह्रो नै पर्न सक्दछ । तर दुवैमा महिलाको प्रतिशत निकाल्यो भने बुझ्न सहज हुन्छ ।

पहिलो समितिमा महिला सहभागिताको प्रतिशत = $२५/५० \times १००\% = ५०\%$

दोस्रो समितिमा महिला सहभागिताको प्रतिशत = $३०/८० \times १००\% = ३७.५\%$

यसरी, संख्यामा हेर्दा पहिलोमा भन्दा दोस्रो समितिमा महिला बढी देखिए पनि महिलाको सहभागिता पहिलो समूहमा बढी देखिन्छ ।

नेपालको प्रतिनिधि सभामा कुल २७५ सांसदमध्ये ९० जना महिला सांसद रहेका छन् । प्रतिशतका हिसाबले यो ३२.७३% हो ।^{२८} विभिन्न समितिहरूमा महिला प्रतिनिधित्वको अवस्था अध्ययन गर्नका लागि तालिका ९.१ हेरौं ।

तालिका ९.१: प्रतिनिधि सभाका केही समितिहरूमा माननीयहरूको लैङ्गिक प्रतिनिधित्व

समिति	जम्मा	पुरुष	महिला	महिला (%)
अर्थ समिति	२५	१९	६	२४.००%
अन्तर्राष्ट्रिय सम्बन्ध समिति	२३	१२	११	४८.००%
कानून, न्याय तथा मानव अधिकार समिति	२१	१५	६	२९.००%
कृषि, सहकारी तथा प्राकृतिक स्रोत समिति	२१	१५	६	२९.००%
महिला तथा सामाजिक समिति	२४	७	१७	७१.००%
शिक्षा तथा स्वास्थ्य समिति	२४	१४	१०	४२.००%

स्रोत: प्रतिनिधि सभा, नेपाल^{२९}

तालिका ९.१ मा प्रस्तुत गरिएका समितिहरू मध्ये सबैभन्दा बढी महिला सहभागिता भएको समिति महिला तथा सामाजिक समिति हो । अर्थ समिति तथा कानून, न्याय तथा मानव अधिकार समिति र कृषि, सहकारी तथा प्राकृतिक स्रोत समितिमा महिलाको संख्या बराबर भए तापनि अर्थ समितिमा अरू दुई समितिको भन्दा महिला प्रतिनिधित्व प्रतिशतका हिसाबले कम देखिन्छ ।

२८ House of Representatives, Nepal. (2075 B.S.). *Women's Participation in Parliament (2015 B.S – 2074 B.S)*. Kathmandu, Nepal; House of Representatives, Nepal. Retrieved on November 28, 2021 from: <https://hr.parliament.gov.np/uploads/attachments/lnoonrxme6omsulm.pdf>

२९ House of Representatives, Nepal. Retrieved on November 28, 2021 from: <https://hr.parliament.gov.np/#>

१.४.५ प्रतिशत बिन्दु (Percentage Point)

कुनै दुईवटा प्रतिशत मानहरूको भिन्नतालाई प्रतिशत बिन्दु भनिन्छ। माथिको उदाहरणमा अर्थ समिति र कानुन, न्याय तथा मानव अधिकार समितिमा महिला सहभागिताको प्रतिशतको भिन्नता ५% छ, यसैलाई प्रतिशत बिन्दु भनिन्छ। तर कानुन, न्याय तथा मानव अधिकार समितिको तुलनामा अर्थ समितिमा महिला सहभागिता ५% ले बढी छ भनेर बुझ्नुचाहिँ गलत हुन्छ। यसलाई कानुन, न्याय तथा मानव अधिकार समितिको महिला सहभागिताको हिस्साभन्दा अर्थ समितिको महिला सहभागिताको हिस्सा ५ प्रतिशत बिन्दुले बढी छ भनेर भन्नुपर्दछ। अर्को उदाहरण हेरौं, आ.व. २०७७/७८ को नेपालको कुल गार्हस्थ्य उत्पादनमा शिक्षा र स्वास्थ्य क्षेत्रको योगदान क्रमशः ८.०५% र १.७३% रहेको छ।^{३०} तसर्थ, कुल गार्हस्थ्य योगदानमा स्वास्थ्य क्षेत्रको भन्दा कृषि क्षेत्रको योगदान ६.३१ प्रतिशत बिन्दुले बढी छ भनेर भन्न सकिन्छ।

अनुपात, समानुपात, दर, प्रतिशत आदिको प्रयोग कार्यक्रमको लागत, प्रतिफल, असर तथा प्रभाव मापन गर्न वा अनुगमन तथा मूल्याङ्कन गर्न प्रयोग हुन्छ। यसका अतिरिक्त घरायसी गणित तथा अन्य विविध क्षेत्रमा पनि यिनीहरूको प्रयोग हुन्छ। जस्तै: प्रकाशले वार्षिक परीक्षामा ८९% अङ्क हासिल गरे, बाबुको उमेर र छोरीको उमेरको अनुपात २:१ छ आदि।

१.५ दर, अनुपात, समानुपात, प्रतिशत, प्रतिशत बिन्दु प्रयोग गर्दा ध्यान दिनुपर्ने कुराहरू कुनै पनि सूचक प्रयोग गर्नुपूर्व ती सूचक निर्माणको सम्पूर्ण प्रक्रियाको बारेमा पूर्ण जानकारी राख्नुपर्ने हुन्छ। अन्यथा, सूचकले इङ्कित गरेको भन्दा भिन्न तरिकाले व्याख्या गरिने सम्भावना हुन्छ। तसर्थ, दर, अनुपात, समानुपात, प्रतिशत, प्रतिशत बिन्दु प्रयोग गर्दा निम्न कुराहरूमा ध्यान दिनुपर्दछ।

- दर र अनुपात फरक-फरक अवधारणा हुन्।
- प्रतिशत र प्रतिशत बिन्दु फरक-फरक अवधारणा हुन्।
- कुनै पनि सूचक कुन एकाइमा मापन गरिएको छ, सोही अनुसार व्याख्या गर्नुपर्दछ।
- प्रति हजार, प्रति दश हजार, प्रति लाखमा मापन गरिएका दर, अनुपात वा समानुपातलाई प्रतिशत भन्नु हुँदैन।

30 Central Bureau of Statistics. (2021). *National Accounts of Nepal 2020/21*. Kathmandu, Nepal; Central Bureau of Statistics. Retrieved on November 28, 2021 from: <https://cbs.gov.np/national-accounts-of-nepal-2020-21/>

१.६ अभ्यासका लागि प्रश्नहरू

क) खाली ठाउँ भर्नुहोस् ।

- १) पूर्ण संख्या भनेको सहितको प्राकृतिक संख्या हो ।
- २) दुई संख्याहरूको तुलना गर्न को प्रयोग गरिन्छ ।
- ३) जन्मदर, मृत्युदर जस्ता दरहरू प्राय प्रति मा व्यक्त गरिन्छ ।
- ४) सय भागको केही भागलाई भनिन्छ ।
- ५) अंशभन्दा हर ठुलो भएको भिन्नलाई भनिन्छ ।

ख) ठिक र बेठिक छुट्याउनुहोस् ।

- १) $\sqrt{16}$ भिन्नमा व्यक्त गर्न नसकिने संख्या हो ।
- २) \angle पनि एउटा भिन्न संख्या हो ।
- ३) $i + 5$ एउटा जटिल संख्या हो ।
- ४) दर र अनुपात एउटै कुरा हुन् ।
- ५) भिन्नलाई प्रतिशतमा लैजान १०० ले भाग गर्नुपर्दछ ।

ग) कुनै गाउँपालिकाको कुल १५,००० जनसंख्यामा १४,५०० साक्षर रहेछन् भने कति प्रतिशत मानिस निरक्षर रहेछन् ?

घ) कुनै एउटा नगरपालिकामा एक वर्ष मुनिका बालबालिकाको जनसंख्या २४० रहेको छ । उक्त नगरपालिकामा एक वर्षभरि जन्मएका शिशुमध्ये एक वर्ष उमेर पनि नपुग्दै ३० जनाको मृत्यु भएको रहेछ भने उक्त नगरपालिकाको शिशु मृत्युदर कति रहेछ ? मानौं, मृत्यु हुने ३० जनामध्ये १२ जना केटा र १८ जना केटी रहेछन् र २४० जनामध्ये ११० जना केटा र १३० जना केटीको जनसंख्या रहेछ भने केटाको मृत्युदर र केटीको मृत्युदरमा कुनचाहिँ बढी छ, निकाल्नुहोस् ।

ङ) कुनै गाउँपालिकामा १८,००० जनसंख्या रहेछन् । त्यहाँ २५ जना नर्स कार्यरत छन् भने नर्स र जनसंख्याको अनुपात निकालेर व्याख्या गर्नुहोस् ।

च) अनुपात, दर र प्रतिशत प्रयोग भएका $5/5$ वटा सूचकहरूको उदाहरण दिनुहोस् । ती सूचकहरूको मान निकाल्न प्रयोग हुने गणितीय सूत्रसमेत लेख्नुहोस् ।

छ) तालिका १.२ मा वि.सं. २०६८ को नेपालको पाँच वर्षे उमेर समूहअनुसार जनसंख्याको वितरण दिइएको छ । उक्त तालिकाको आँकडाको आधारमा सोधिएका प्रश्नको जवाफ दिनुहोस् ।

तालिका १.२: पाँच वर्षे उमेर समूहअनुसार जनसंख्या वितरण

उमेर समूह	पुरुष	महिला	जम्मा
९५+	४,२०१	७,००७	११,२०८
९०-९४	८,९४०	११,३९५	२०,३३५
८५-८९	२५,८१०	२६,७१६	५२,५२६
८०-८४	६२,७८७	६५,९९०	१,२८,७७७
७५-७९	१,१७,३५८	१,१७,७७७	२,३५,१३५
७०-७४	१,९९,६१०	१,९५,५४३	३,९५,१५३
६५-६९	२,७७,७८२	२,७६,६६७	५,५४,४४९
६०-६४	३,६८,४५१	३,८८,३७६	७,५६,८२७
५५-५९	४,१२,८९२	४,०५,३७१	८,१८,२६३
५०-५४	५,०५,८६४	४,९९,६१२	१०,०५,४७६
४५-४९	५,७५,१०१	५,९७,८५८	११,७२,९५९
४०-४४	६,६०,२९०	७,२५,८३१	१३,८६,१२१
३५-३९	७,४०,२००	८,६४,११९	१६,०४,३१९
३०-३४	७,७०,५७७	९,६४,७२८	१७,३५,३०५
२५-२९	९,१७,२४३	११,६२,१११	२०,७९,३५४
२०-२४	१०,४३,९८१	१३,१४,०९०	२३,५८,०७१
१५-१९	१४,४३,१९१	१४,८८,७८९	२९,३१,९८०
१०-१४	१७,६४,६३०	१७,१०,७९४	३४,७५,४२४
५-९	१६,३५,१७६	१५,६९,६८३	३२,०४,८५९
०-४	१३,१४,९५७	१२,५३,००६	२५,६७,९६३
जम्मा	१,२८,४९,०४१	१,३६,४५,४६३	२,६४,९४,५०४

स्रोत: केन्द्रीय तथ्याङ्क विभाग, राष्ट्रिय जनगणना, २०६८

- १) कुल जनसंख्याको लिङ्ग अनुपात निकाल्नुहोस् ।
- २) ५ वर्ष मुनिको उमेर समूहको लिङ्ग अनुपात निकाल्नुहोस् ।
- ३) कुल जनसंख्यामा महिला र पुरुष कति कति प्रतिशत रहेका छन्, निकाल्नुहोस् ।
- ४) ९५ वा सोभन्दा बढी उमेर समूहमा महिलाको प्रतिशत कति छ ?
- ५) पाँच वा सोभन्दा बढी उमेरका साक्षर पुरुष ८६,६६,७११ जना र साक्षर महिला ७१,१२,०३१ जना छन् भने कुल साक्षरता दर, महिला साक्षरता दर र पुरुष साक्षरता दर प्रतिशतमा निकाल्नुहोस् ।

९.७ सन्दर्भ सामग्री

यस पाठका विषयवस्तुको विस्तृत अध्ययनका लागि विद्यालयस्तरको कुनै एउटा गणित पुस्तक अध्ययन गर्न सकिन्छ ।

९.८ शब्दावली

अनन्त संख्या (Infinite Number): असाध्यै ठुलो संख्यालाई अनन्त संख्या भनिन्छ । उदाहरणका लागि आकाशमा कतिबटा ताराहरू छन् भनेर गणना गर्ने हो भने जति ठुलो संख्याले पनि पुग्दैन । यस्तो अवस्थामा हामीले अनन्त संख्याको सहायता लिनुपर्दछ । यसलाई ग्रीक लेटर “Infinity (∞)” ले जनाइन्छ । त्यस्तै प्राकृतिक संख्याको सबैभन्दा ठुलो संख्या कुन हो भनेर सोधियो भने कसैले पनि उत्तर बताउन सक्दैन । यस्तो अवस्थामा पनि अनन्त संख्याको सहायता लिनुपर्दछ ।



सूचकाङ्क र यसको प्रयोग

१०.१ परिचय

अधिल्ला पाठहरूमा तथ्याङ्क विश्लेषणका लागि प्रयोग हुने विभिन्न गणितीय सूचकहरूका बारेमा अध्ययन गरिसकेका छौं । यस पाठमा हामी तथ्याङ्क विश्लेषणमा वर्णनात्मक (Descriptive) औजारका रूपमा प्रयोग हुने सूचकाङ्क (Index Number) र यसको प्रयोगका बारेमा छलफल गर्नेछौं । सूचकाङ्क भनेको एक अवधिदेखि अर्को अवधिसम्म कुनै पनि मान वा मूल्यको सापेक्षिक परिवर्तन (Relative Change of Price) मापन गर्ने तथ्याङ्कीय औजार हो । समय परिवर्तनसँगै कुनै पनि वस्तु वा विशेषताको मान वा मूल्यमा पनि परिवर्तन भइरहेको हुन्छ । यस्ता परिवर्तनहरूले देश वा समाजको आर्थिक, सामाजिक पक्षलाई प्रभाव पारिरहेको हुन्छ । तसर्थ, मूल्य वा मानको परिवर्तन सापेक्ष रूपमा मापन गरेर नीतिनिर्माता तथा योजनाविद्हरूलाई नीति पृष्ठपोषण गर्नका लागि सूचकाङ्कहरूको ठुलो योगदान रहन्छ । यसका अतिरिक्त कतिपय सूचकाङ्कहरू कुनै पनि वस्तुको स्थिर मूल्य वा चालु मूल्य मापनका लागि पनि प्रयोग हुन्छन् । आम मानिसको क्रयशक्तिका बारेमा पनि सूचकाङ्कले बताउन सहयोग गर्दछन् । मूल्य सूचकाङ्क (Price Index), मान सूचकाङ्क (Value Index) यसका उदाहरणहरू हुन् । मूल्य सूचकाङ्कका उदाहरणमा उपभोक्ता मूल्य सूचकाङ्क (Consumer Price Index), उत्पादकको मूल्य सूचकाङ्क (Producer's Price Index), आदि हुन् भने मान सूचकाङ्कका उदाहरणमा औद्योगिक उत्पादन सूचकाङ्क (Index of Industrial Production), कृषि उत्पादन सूचकाङ्क (Index of Agricultural Production), आदि रहेका छन् ।

सूचकाङ्कहरू खासगरी आर्थिक क्रियाकलापमा आउने सापेक्षिक परिवर्तन मापन गर्नका लागि प्रयोग हुन्छन् । आर्थिक क्रियाकलापका उपज वस्तु वा सेवाको उत्पादन परिमाण, मूल्य तथा मानमा परिवर्तन भइरहेको हुन्छ । कुनै खास अवधि, स्थान वा परिस्थितिसँग तुलना गरेर यस्ता उत्पादित वस्तु वा सेवाको मात्रा, मूल्य वा मानमा आउने सापेक्षिक परिवर्तनहरूको मापन गर्ने काम सूचकाङ्कले गर्दछ । सूचकाङ्कको सहायताले कुनै पनि वस्तु वा सेवाको मूल्य, परिमाण वा मानको दुई फरक समय वा स्थानबिच हुने तुलनात्मक परिवर्तनहरूको बारेमा जानकारी प्राप्त गर्न सकिन्छ । संक्षेपमा भन्नुपर्दा, आधार अवधि, स्थान वा परिस्थितिको तुलनामा कुनै पनि वस्तु वा सेवाको मूल्य, मात्रा वा मानमा सापेक्षिक परिवर्तन व्यक्त गर्ने संख्यालाई सूचकाङ्क भनिन्छ । अर्को शब्दमा भन्दा, सूचकाङ्क भनेका विशेष प्रकारका अनुपात, दर, प्रतिशत हुन् जसले दुई वा बढी अवस्थाहरूमा भिन्न तर सम्बन्धित चरहरूको समूहको सामान्य स्तरको परिमाण प्रतिबिम्बित गर्दछन् ।

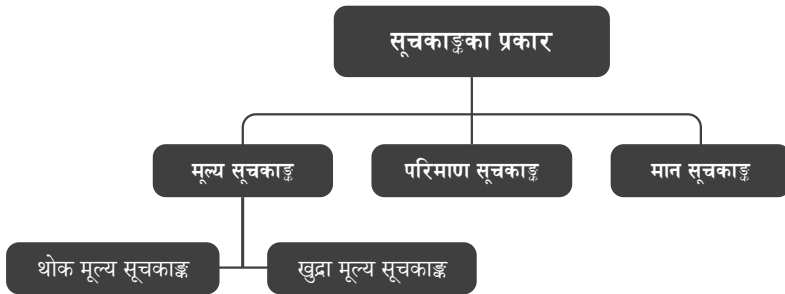
उदाहरणका लागि आ.व. २०७१/७२ का लागि खाद्य पदार्थको मूल्यलाई १०० मान्ने हो भने आ.व. २०७७/७८ मा यसको मूल्य बढेर १३५.९७ पुगेको छ।^{३१} यस उदाहरणमा आ.व. २०७१/७२ लाई आधार वर्षको रूपमा लिइएको छ भने यही वर्षको तुलनामा आ.व. २०७७/७८ को मूल्य सूचकाङ्क प्रस्तुत गरिएको छ। यसलाई अर्को तरिकाले पनि व्याख्या गर्न सकिन्छ: आ.व. २०७१/७२ मा रू. १०० पर्ने खाद्य पदार्थ आ.व. २०७७/७८ मा रू. १३५.९७ पर्दछ।

१०.२ सूचकाङ्कका प्रकार र नेपालमा निर्माण गरिने सूचकाङ्कहरू

१०.२.१ सूचकाङ्कका प्रकार

सूचकाङ्कलाई मुख्य रूपमा तिन भागमा विभाजन गर्न सकिन्छ जसलाई चित्र नं. १०.१ मा प्रस्तुत गरिएको छ।^{३२}

चित्र नं. १०.१: सूचकाङ्कका प्रकार



क) मूल्य सूचकाङ्क

मूल्य सूचकाङ्कले वस्तु तथा सेवाको मूल्यमा आउने परिवर्तनलाई मापन गर्दछ। यसलाई थोक तथा खुद्रा मूल्य सूचकाङ्क गरी दुई भागमा विभाजन गर्न सकिन्छ। थोक मूल्य सूचकाङ्कले देशको साधारण मूल्यस्तर प्रतिबिम्बित गरेको हुन्छ। खुद्रा मूल्य सूचकाङ्कले विभिन्न वस्तुहरूको खुद्रा मूल्यमा भएको परिवर्तन मापन गर्दछ। खुद्रा मूल्य सूचकाङ्कको विशेष प्रकारको सूचकाङ्कका रूपमा उपभोक्ता मूल्य सूचकाङ्कलाई लिइन्छ जसलाई

31 Nepal Rastra Bank. *Current Macroeconomic and Financial Situation*. Kathmandu, Nepal; Nepal Rastra Bank. Retrieved on November 26, 2021 from: <https://www.nrb.org.np/category/current-macroeconomic-situation/page/2/?department=red>

32 यहाँ उल्लेख गरिएका बाहेक विभिन्न किसिमका समाष्टित सूचकाङ्क (Composite Index) हरू प्रचलनमा रहेका छन्। उदाहरणका लागि मानव विकास सूचकाङ्क, लैङ्गिकसँग सम्बन्धित सूचकाङ्क, मानव गरिबीसम्बन्धी सूचकाङ्क, ह्यापिनेस सूचकाङ्क, सुशासन सम्बन्धी सूचकाङ्क, आदि। यी सूचकाङ्कहरूका बारेमा यस पाठमा थप चर्चा गरिएको छैन।

जीविकोपार्जन सूचकाङ्क पनि भनिन्छ । यसले उपभोग्य वस्तुहरूको मूल्यको सापेक्षिक परिवर्तन मापन गर्दछ र जीविकोपार्जनका लागि मानिसको क्रयशक्तिको सामर्थ्य मापन गर्न सहयोग गर्दछ ।

उदाहरण: नेपाल राष्ट्र बैंकले प्रकाशन गरेको तथ्याङ्क बमोजिम:³³

राष्ट्रिय थोक मूल्य सूचकाङ्क (National Wholesale Price Index): आ.व. २०७४/७५ लाई आधार मानेर हेर्दा आ.व. २०७७/७८ मा खाद्य पदार्थको थोक मूल्य सूचकाङ्क १३१.९१ थियो, त्यसैको अघिल्लो वर्ष आ.व. २०७६/७७ मा उक्त सूचकाङ्क ११९.६१ थियो । यस हिसाबले २०७७/७८ मा खाद्य पदार्थको थोक मूल्य औसतमा १०.२८% ले बढेको देखिन्छ ।

राष्ट्रिय उपभोक्ता मूल्य सूचकाङ्क (National Consumer Price Index): आ.व. २०७१/७२ लाई आधार मानेर हेर्दा आ.व. २०७७/७८ मा खाद्य पदार्थको उपभोक्ता मूल्य सूचकाङ्क १३५.९७ थियो, त्यसैको अघिल्लो वर्ष आ.व. २०७६/७७ मा उक्त सूचकाङ्क १२९.४९ थियो । यस हिसाबले २०७७/७८ मा खाद्य पदार्थको खुद्रा मूल्य औसतमा ५% ले बढेको देखिन्छ ।

ख) परिमाण सूचकाङ्क (Quantity Index)

उत्पादन, उपभोग वा वितरण गरिने वस्तुहरूको मात्रामा भएको सापेक्षिक परिवर्तन मापन गर्ने औजारलाई परिमाण सूचकाङ्क भनिन्छ । जस्तै: कृषि उत्पादन सूचकाङ्क, औद्योगिक उत्पादन सूचकाङ्क आदि । यस्ता सूचकाङ्क कुनै पनि अर्थतन्त्रको भौतिक प्रतिफलको स्तर अध्ययन गर्नका लागि निकै उपयोगी हुन्छन् ।

उदाहरण: केन्द्रीय तथ्याङ्क विभागको त्रैमासिक औद्योगिक उत्पादन सूचकाङ्क बमोजिम:

२०७१/७२ लाई आधार मानेर हेर्दा आ.व. २०७५/७६ मा औद्योगिक उत्पादन सूचकाङ्क १३३.३ थियो, त्यसैको अघिल्लो वर्ष आ.व. २०७४/७५ मा उक्त सूचकाङ्क १२५.८९ थियो । यस हिसाबले २०७५/७६ मा औद्योगिक उत्पादन औसतमा ५.८८% ले बढेको देखिन्छ ।

33 Nepal Rastra Bank. *Current Macroeconomic and Financial Situation*. Kathmandu, Nepal; Nepal Rastra Bank. Retrieved on November 26, 2021 from: <https://www.nrb.org.np/category/current-macroeconomic-situation/page/2/?department=red>

ग) मान सूचकाङ्क (Value Index)

मूल्य र परिमाणको परिवर्तन संयुक्त रूपमा अध्ययन गर्न प्रयोग गरिने सूचकाङ्कलाई मान सूचकाङ्क भनिन्छ। यी सूचकाङ्कहरू धेरै प्रयोगमा छैनन् तथापि राष्ट्रिय लेखा अनुमानमा आन्तरिक प्रयोजनका लागि निकाल्ने गरिन्छ।

१०.२.२ नेपालमा निर्माण गरिने सूचकाङ्कहरू

नेपालमा सूचकाङ्कको नियमित निर्माण तथा प्रकाशन गर्ने प्रमुख निकायहरूमा नेपाल राष्ट्र बैंक, केन्द्रीय तथ्याङ्क विभाग र नेपाल स्टक एक्सचेन्ज हुन्। यी निकायहरूले निर्माण तथा प्रकाशन गर्ने प्रमुख सूचकाङ्कहरू यसप्रकार छन्।

क) नेपाल राष्ट्र बैंक: राष्ट्र बैंकले राष्ट्रिय उपभोक्ता मूल्य सूचकाङ्क, काठमाण्डौं उपत्यका उपभोक्ता मूल्य सूचकाङ्क, पहाडी उपभोक्ता मूल्य सूचकाङ्क, तराई उपभोक्ता मूल्य सूचकाङ्क, हिमाली मूल्य सूचकाङ्क, राष्ट्रिय थोक मूल्य सूचकाङ्क, राष्ट्रिय तलब ज्यालादर सूचकाङ्कहरू मासिक, त्रैमासिक र वार्षिक रूपमा नियमित प्रकाशन गर्दै आइरहेको छ। यस्ता सूचकाङ्कहरू बैंकको वेबसाइटमा उपलब्ध छन्।³⁴

ख) केन्द्रीय तथ्याङ्क विभाग: नेपालमा सबैभन्दा पहिला सूचकाङ्क निर्माण गर्ने निकाय केन्द्रीय तथ्याङ्क विभाग हो। आवधिक योजनाको सुरुवातसँगै नेपालमा केही सूचकाङ्कहरू निर्माण गर्न थालिएको थियो। हाल विभागले उत्पादनसँग सम्बन्धित केही सूचकाङ्क र उत्पादनको परिमाणसँग सम्बन्धित केही सूचकाङ्क निर्माण गर्दै आइरहेको छ। विभागले निर्माण क्षेत्रको लागत मूल्य सूचकाङ्क, औद्योगिक उत्पादन सूचकाङ्क, औद्योगिक उत्पादकको मूल्य सूचकाङ्क, कृषि उत्पादकको मूल्य सूचकाङ्क त्रैमासिक रूपमा उत्पादन तथा प्रकाशन गरिरहेको छ। यी सूचकाङ्कहरू विभागको वेबसाइटमा उपलब्ध छन्।³⁵

ग) नेपाल स्टक एक्सचेन्ज: नेपाल स्टक एक्सचेन्ज (नेप्से) ले धितोपत्र बजार मूल्यसम्बन्धी विभिन्न सूचकाङ्कहरू निर्माण गरी प्रकाशन गर्दै आइरहेको छ। यी सूचकाङ्कहरू नेप्सेको वेबसाइटमा उपलब्ध छन्।³⁶

34 Nepal Ratsra Bank. *Economic Bulletin*. Kathmandu, Nepal; Nepal Rastra Bank. Retrieved on November 26, 2021 from: <https://www.nrb.org.np/category/economic-bulletin/?department=red>

35 Central Bureau of Statistics. *Prices and Indices*. Kathmandu, Nepal; Central Bureau of Statistics. Retrieved on November 26, 2021 from: <https://cbs.gov.np/prices-indices/>

36 Nepal Stock Exchange Ltd (NEPSE). (2021). *NEPSE Index from Date 2021/05/25 to 2021/11/25*. Kathmandu, Nepal; Nepal Stock Exchange Ltd (NEPSE). Retrieved on November 26, 2021 from: <http://www.nepalstock.com/indices>

१०.३ सूचकाङ्क निर्माणका तरिका तथा सूचकाङ्क निर्माणमा आइपने समस्या

सूचकाङ्क निर्माणको सबैभन्दा पहिलो चरणो काम भनेको सूचकाङ्क निर्माण गरिने समूहमा पर्ने वस्तु वा सेवाहरूको छनोट गर्नु हो । उदाहरणका लागि तरकारीको मूल्य सूचकाङ्क निकाल्नुपर्दा काउली, बन्दा, सिमी, आलु, लौका, फर्सी, जस्ता कैयौँ तरकारीहरूमध्ये कुन कुन प्रतिनिधिमूलक तरकारी छनोट गर्ने भन्ने बारेमा निर्णय गरी एउटा बास्केट तयार पार्नुपर्दछ । यसरी छनोट गर्दा सम्बन्धित समूहलाई उपयुक्त तरिकाले प्रतिनिधित्व गर्ने तरकारीहरू छनोट गर्नुपर्दछ । त्यसपछि छनोटमा परेका तरकारीहरूको आधार समयका लागि र सूचकाङ्क निकाल्ने समयका लागि प्रति एकाइ मूल्य संकलन गर्नुपर्दछ । मानौँ, तरकारीको मूल्य सूचकाङ्क निकाल्नका लागि बकुल्ला, बन्दा, काउली, र आलु छनोटमा परेका रहेछन् र तिनीहरूको मूल्य संकलन गर्दा तालिका १०.१ बमोजिम पाइयो ।

तालिका १०.१: तरकारीको मूल्य

तरकारी	एकाइ	मूल्य	
		२०७२	२०७८
बरेला	के.जी.	४५	१२५
बन्दा	के.जी.	४३	२५
काउली	के.जी.	४५	६५
आलु रातो	के.जी.	३८	५८
जम्मा		१७१	२७३

नोट: यो तथ्याङ्क कालिमाटी फलफूल तथा तरकारी विकास समितिको वेबसाइटबाट संकलन गरिएको हो ।^{३७}

तरकारीको भारविनाको मूल्य सूचकाङ्क निकाल्नका लागि निम्न बमोजिमको सूत्र प्रयोग गरिन्छ ।

$$P_{01} = \frac{\sum P_1}{\sum P_0} \times 100 \quad (१)$$

जहाँ, P_{01} भनेको मूल्य सूचकाङ्क, P_1 भनेको दिइएको वर्षको वस्तु वा सेवाको मूल्य र P_0 भनेको आधार वर्षमा सोही वस्तु वा सेवाको मूल्य बुझ्नुपर्दछ । उदाहरणका लागि P_{01} वि.सं. २०७२ को तुलनामा वि.सं. २०७८ को तरकारीको मूल्य सूचकाङ्क प्रतिनिधित्व

37 Kalimati Fruits & Vegetable Market Development Board. (2021). *Daily Price Information of Fruits and Vegetables*. Kathmandu, Nepal; Kalimati Fruits & Vegetables Market Development. Retrieved on November 26, 2021: <https://kalimatimarket.gov.np/price>

गरेको छ । P_1 भनेको वि.सं. २०७८ को तरकारीको मूल्य र P_0 भनेको वि.सं. २०७२ को तरकारीको मूल्य हो । Σ ले मूल्यहरूको योग लिने काम गर्दछ । अब तालिका १०.२ बाट सूत्र (१) मा मानहरू प्रतिस्थापन गर्दा,

$$P_{01} = \frac{२७३}{१७१} \times 100 = १५९.६४ \quad (२)$$

यसप्रकारले भारविनाको मूल्य सूचकाङ्क निकाल्ने गरिन्छ । यो उदाहरण अभ्यासका लागि मात्र प्रस्तुत गरिएको छ । वास्तविक मूल्य सूचकाङ्क निकाल्नका लागि तरकारीहरूको छनोट पूर्ण रूपमा गर्नुपर्दछ र मूल्य पनि विभिन्न स्थानबाट संकलन गर्नुपर्दछ । तर यस उदाहरणमा तरकारीहरूको छनोटका कार्य र विभिन्न स्थानबाट मूल्य संकलन कार्य गरिएको छैन । यी विषयहरू अति प्राविधिक छन् र यस पुस्तकको दायराभन्दा बाहिरको कार्य हुन् ।

माथि तालिका १०.२ मा प्रस्तुत गरिएका तरकारीहरूको अलग अलग मूल्य सूचकाङ्क वा सापेक्ष मूल्य (Price Relative) निकाल्न पनि सकिन्छ । जसका लागि निम्न सूत्र प्रयोग गरिन्छ ।

$$P_{01} = \frac{p_1}{p_0} \times 100 \quad (३)$$

उदाहरणका लागि वि.सं. २०७२ को तुलनामा वि.सं. २०७८ को काउलीको सापेक्ष मूल्य: $P_{01} = \frac{p_1}{p_0} \times 100 = \frac{६५}{४५} \times 100 = १४४.४४$ हुन्छ ।

यसबाट हामीले वि.सं. २०७२ को तुलनामा वि.सं. २०७८ मा काउलीको मूल्य १४४.४४% छ वा ४४.४४% ले वृद्धि भएको छ भनेर भन्न सक्दौं ।

सूत्र नं. (१) मा यदि मूल्यको स्थानमा परिमाण (Quantity) प्रयोग गरियो भने परिमाण सूचकाङ्क निकाल्न सकिन्छ । अर्थात्,

$$Q_{01} = \frac{\Sigma q_1}{\Sigma q_0} \times 100 \quad (४)$$

जहाँ, Q_{01} ले आधार समयको तुलनामा सूचकाङ्क निकाल्ने समयको तरकारीको परिमाण सूचकाङ्क प्रतिनिधित्व गरेको छ । q_1 भनेको सूचकाङ्क निकाल्ने समयको तरकारीको परिमाण र q_0 भनेको आधार समयको तरकारीको परिमाण हो । Σ ले परिमाणहरूको योग लिने काम गर्दछ ।

माथि समीकरण (२) मा निकालिएको मूल्य सूचकाङ्कले छनोटमा परेका सबै तरकारीहरूलाई बराबर भार दिएको छ । वास्तविक संसारमा त्यस्तो हुँदैन । बरेलाको प्रयोग र आलुको प्रयोग समान नहुन सक्छ । तसर्थ, नमुना छनोटमा परेका वस्तु वा

सेवाहरूलाई फरक-फरक भार दिनुपर्दछ । त्यस्तो भार सहितको मूल्य सूचकाङ्कले मात्र वास्तविक सूचना प्रदान गर्दछ । भार सहितको मूल्य सूचकाङ्क निकाल्नको लागि निम्न सूत्र प्रयोग गरिन्छ ।

$$P_{01} = \frac{\sum wP_1}{\sum wP_0} \times 100 \quad (५)$$

जहाँ, w ले सम्बन्धित वस्तु वा सेवाको भार प्रतिनिधित्व गर्दछ । भारत मूल्य सूचकाङ्कका लागि लास्पियरको, पास्चेको, फिसरको, तरिकाहरू बढी प्रचलनमा छन् । यी विद्वानहरूले भार केलाई मान्ने भन्नेमा आ-आफ्नो तर्कसहित सूत्र प्रस्तुत गरेका छन् ।

क) लास्पियरको मूल्य सूचकाङ्क:

$$P_{01} = \frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times 100 \quad (६)$$

अर्थात्, लास्पियरले आधार समयको सम्बन्धित वस्तुको परिमाण Q_0 लाई भार मानेका छन् ।

ख) पास्चेको मूल्य सूचकाङ्क:

$$P_{01} = \frac{\sum P_1 Q_1}{\sum P_0 Q_1} \times 100 \quad (७)$$

अर्थात्, पास्चेले सूचकाङ्क निकाल्ने समयको सम्बन्धित वस्तुको परिमाण Q_1 लाई भार मानेका छन् ।

ग) पास्चेको मूल्य सूचकाङ्क:

फिसरले लास्पियर र पास्चेको मूल्य सूचकाङ्कको ज्यामितीय औसतलाई मूल्य सूचकाङ्क मानेका छन् । यसलाई आदर्श मूल्य सूचकाङ्क पनि भनिन्छ किनभने यो ज्यामितीय औसतमा आधारित छ र आधार अवधि तथा तोकिएको अवधि दुवैको मूल्य तथा परिमाण लिइएको हुन्छ । सूत्रमा,

$$\text{फिसरको मूल्य सूचकाङ्क} = (\text{लास्पियरको मूल्य सूचकाङ्क} \times \text{पास्चेको मूल्य सूचकाङ्क})^{\frac{1}{2}} \quad (८)$$

सूचकाङ्क निर्माणमा देहायबमोजिमका समस्याहरू पर्न सक्दछन् ।

- वस्तु तथा सेवाको मापन एकाइमा हुने विविधताले सूचकाङ्क निर्माणमा समस्या पैदा गर्दछन् ।
- प्रतिनिधिमूलक वस्तु तथा सेवा छनोट गर्न कठिन हुन्छ ।
- आधार समय छनोट गर्दा कुनै उतारचढाव नभएको समय छान्नुपर्दछ । व्यवहारतः यो कार्य कठिन हुन्छ ।

- आधार समयमा बास्केटमा परेका वस्तु वा समय तोकिएको समयमा प्राप्त नहुन सक्छन् वा गुणस्तरमा परिवर्तन हुन सक्दछ । यसो भएमा सो वस्तु वा सेवालाई अर्को उस्तै प्रकृतिको वस्तु वा सेवाले प्रतिस्थापन गर्नुपर्ने हुन्छ, जुन कठिन कार्य हो ।
- बास्केटमा परेका वस्तुहरूको मूल्य विभिन्न स्थानबाट प्रतिनिधित्व हुनेगरी संकलन गर्नु चुनौतीपूर्ण कार्य हो ।

१०.४ सूचकाङ्कको व्याख्या एवं प्रयोग

सूचकाङ्कलाई अर्थतन्त्रको ब्यारोमिटर पनि भन्ने गरिन्छ । सूचकाङ्कहरूलाई विभिन्न अवधिहरूको तुलनात्मक अध्ययन गरेर अर्थतन्त्रको कार्यसम्पादन र अन्य आर्थिक चरहरूको मापन गर्न प्रयोग गरिन्छ । सूचकाङ्कको प्रयोग अर्थतन्त्रका अतिरिक्त व्यवसाय एवं व्यवस्थापनमा पनि गरिन्छ । सारांशमा भन्दा सूचकाङ्कको प्रयोग:

- अर्थतन्त्रमा हुने मूल्यवृद्धि वा संकुचन मापन गर्न,
- मुद्रास्फीति र संकुचनको लक्ष्य निर्धारण गर्न,
- ब्याजदर निर्धारण गर्न,
- उत्पादन, उपभोग, लगानी, व्यापार, श्रमशक्तिसँग सम्बन्धित मूल्य, परिमाण वा मानको चलायमान प्रवृत्तिको पहिचान गर्न,
- समय शृङ्खला तथ्याङ्कको प्रवृत्ति मापन गर्न,
- नीति निर्माण (विशेष रूपमा मौद्रिक नीति) तथा निर्णय गर्न,
- मुद्राको क्रयशक्ति मापन गर्न,
- स्थिर मूल्यमा मापन गरिएका वस्तु वा सेवालाई चालु मूल्यमा रूपान्तरण गर्न र चालु मूल्यमा मापन गरिएका वस्तु वा सेवालाई स्थिर मूल्यमा रूपान्तरण गर्न - Inflate and Deflate, especially in National Accounts,
- तलब, ज्याला, भाडा, कानुनी सम्झौता, सामाजिक सुरक्षा भुक्तानी, आदिको निर्धारण गर्न,
- अन्तर्राष्ट्रिय संघ संस्था (IMF, ILO, ADB, WB, Eurostat, UNESCAP, etc.) हरूले अन्तर्राष्ट्रियस्तरको तुलना गर्न र अनुगमन गर्न ।

माथि उल्लिखित कतिपय बुँदाहरू स्थानीय तह तथा प्रदेश सरकारका लागि पनि महत्त्वपूर्ण हुन्छन् । मूल्यवृद्धि, ज्यालादर, भाडादर, कानुनी सम्झौता (जस्तै: ठेक्कापट्टामा वस्तु वा सेवाको मूल्यवृद्धिको अवस्था हेरेर निर्णय गर्न सकिन्छ), सामाजिक सुरक्षालगायत अन्य क्षेत्रमा पनि स्थानीय तथा प्रदेश सरकारले मूल्य सूचकाङ्कको प्रयोग गर्न सक्दछन् ।

सूचकाङ्क जहिले पनि आधार समयावधिको तुलनामा अन्य समयावधिको मूल्य, परिमाण वा मान कति छ भनेर हेर्नको लागि प्रयोग गरिन्छ । तसर्थ, यसको विश्लेषण वा व्याख्या

गर्दा आधार समयका लागि कुनै वस्तु वा सेवाको मूल्य, परिमाण वा मानलाई १०० मान्नुपर्दछ र यसको तुलनामा अन्य समयको लागि मूल्य, परिमाण वा मान कति हुन्छ भनेर उल्लेख गर्नुपर्दछ। उदाहरणका लागि नेपाल राष्ट्र बैंकले प्रकाशन गरेको समग्र उपभोक्ता सूचकाङ्क, तरकारी, लत्ताकपडा, स्वास्थ्य र शिक्षाको उपभोक्ता मूल्य सूचकाङ्क तालिका हेरौं।

तालिका १०.२: राष्ट्रिय उपभोक्ता मूल्य सूचकाङ्क

क्र. सं.	आ.व.	उपभोक्ता मूल्य सूचकाङ्क (आधार वर्ष: २०१४/१५)				
		समग्र	तरकारी	लत्ताकपडा	स्वास्थ्य	शिक्षा
१	२००६/०७	४९.८	३०.६	४६.६	६६.३	५३.६
२	२००७/०८	५३.२	३२.९	४८.१	७०	५६.२
३	२००८/०९	५९.९	३६.६	५२.२	७३.५	६०.९
४	२००९/१०	६५.६	४४.१	५६.१	७६	६७.८
५	२०१०/११	७१.९	५९.६	६३.६	७९.१	७१.५
६	२०११/१२	७७.८	७४	७३	८२.७	७८.५
७	२०१२/१३	८५.५	७८.३	८१.९	८८.३	८८
८	२०१३/१४	९३.३	९४.३	९१	९४.८	९४.७
९	२०१४/१५	१००	१००	१००	१००	१००
१०	२०१५/१६	१०९.९	११०.३	११४.२	१०२.६	११०.१
११	२०१६/१७	११४.८	१०७.५	१२४.७	१०५.३	१२०.९
१२	२०१७/१८	११९.७	१२१.१	१३२.१	१०७.६	१३०.६
१३	२०१८/१९	१२५.१	११६.३	१४१	११०.१	१३७.१
१४	२०१९/२०	१३२.८	१४३	१४८.७	११४.९	१४४.४

स्रोत: नेपाल राष्ट्र बैंक, त्रैमासिक आर्थिक बुलेटिन, मध्य-अप्रिल २०२१^{३८}

तालिका १०.२ मा आ.व. २०१४/१५ को लागि सबै मूल्य सूचकाङ्कहरू १०० छन्। तसर्थ, यो वर्षलाई आधार वर्ष मानिएको छ। समग्र उपभोक्ता मूल्य २०१४/१५ मा १०० मान्दा आ.व. २०१९/२० मा आइपुग्दा बढेर १३२.८ पुगेको छ। आ.व. २०१९/२० मा तुलना गरिएका वस्तु र सेवामध्ये सबैभन्दा बढी मूल्य सूचकाङ्क लत्ताकपडाको छ भने

38 Nepal Rastra Bank. *Economic Bulletin*. Kathmandu, Nepal; Nepal Rastra Bank. Retrieved on November 26, 2021 from: <https://www.nrb.org.np/category/economic-bulletin/?department=red>

सबैभन्दा कम मूल्य सूचकाङ्क स्वास्थ्य क्षेत्रको देखिन्छ । आ.व. २०१८/१९ र आ.व. २०१९/२० को मूल्य सूचकाङ्क अध्ययन गरेर एक वर्षमा भएको मूल्यवृद्धि अनुमान गर्न सकिन्छ । उक्त एक वर्षको अवधिमा समग्रमा ६.२% मूल्यवृद्धि भएको देखिन्छ भने सबै भन्दा बढी तरकारीमा (२२.९६%) र सबैभन्दा कम स्वास्थ्यमा (४.४%) मूल्यवृद्धि देखिन्छ । सोही अवधिमा लत्ताकपडा र शिक्षाको मूल्य वृद्धिदर क्रमशः ५.५% र ५.३% रहेको देखिन्छ ।

नोट: मूल्य वृद्धिदर (Inflation Rate) निकाल्नका लागि पाठ ९ मा अध्ययन गरे बमोजिमको वृद्धिदरको सूत्र प्रयोग गरिएको छ ।

जस्तै: आ.व. २०१८/१९ को तुलनामा आ.व. २०१९/२० मा तरकारीको मूल्य वृद्धिदर

$$= \frac{(१४३.० - ११६.३)}{११६.३} \times १००\% = २२.९६\% \text{ रहेको छ ।}$$

१०.५ सूचकाङ्क प्रयोग गर्दा ध्यान दिनुपर्ने कुराहरू

सूचकाङ्क प्रयोग गर्दा निम्नानुसारका विषयमा ध्यान दिनुपर्दछ ।

- सूचकाङ्कको आधार समय कुन हो भन्ने बारेमा जानकारी हुनुपर्दछ । आधार समय दिन, महिना, त्रैमासिक, अर्धवार्षिक, वार्षिकमध्ये कुनै एउटा हुनसक्दछ । केको सूचकाङ्क हो भन्ने बारेमा जानकारी हुनुपर्दछ ।
- सूचकाङ्क निर्माण गर्दा अंश र हरमा भएको मूल्य वा परिमाण वा मानको एकाइ एकले अर्कोलाई भाग गरी हराएर जाने हुनाले सूचकाङ्कको एकाइ हुँदैन तर प्रतिशतमा व्यक्त गर्न भने सकिन्छ ।
- सूचकाङ्क निकालिएको विधिकाबारेमा पूर्ण जानकारी हुनुपर्दछ ।
- सूचकाङ्कहरू अनुमान हुन् र सोहीअनुसार प्रयोग गर्नुपर्दछ ।

१०.६ आधार समय परिवर्तन र सूचकाङ्कको स्प्लाइसिड तथा डिफ्लेटिड

कुनै पनि सूचकाङ्कको आधार समय एउटै हुन सक्दैन । लामो समयअवधिको मूल्यसँग तुलना गर्दा सूचकाङ्कले खासै अर्थ नराख्न सक्दछ । मानिसको रहनसहनमा भएको परिवर्तन, वस्तु तथा सेवाको गुणस्तरमा भएको परिवर्तन, मानिसको जीवनस्तरमा भएको परिवर्तनजस्ता विषयहरूले वस्तु तथा सेवाको मूल्य, मात्रा तथा मानमा ठुलै फड्को मारेको हुन्छ । कतिपय वस्तुहरू लोप भएर जान सक्दछन् । तसर्थ, सूचकाङ्कको आधार समयलाई समय समयमा अनुकूलता मिलाएर परिवर्तन गर्दै लैजानु पर्दछ र परिवर्तित आधार समयबमोजिम पुरानो आधार वर्षका सूचकाङ्कहरूलाई पनि नयाँ श्रेणीमा प्रस्तुत गर्नुपर्दछ ।

कुनै एउटा आधार समयमा प्रस्तुत गरेको सूचकाङ्कको समय शृङ्खलालाई आधार समय

परिवर्तन गरेर बनेको नयाँ समय श्रृङ्खलासँग समायोजन गरी एउटै आधार समयको पुरानो श्रेणी र नयाँ श्रेणीको संयुक्त श्रेणी बनाउने कार्यलाई स्प्लाइसिड भनिन्छ। स्प्लाइसिड प्रक्रियाबाट धेरै पुराना सूचकाङ्कहरू पनि नयाँ आधार समयको तुलनामा प्रस्तुत गरिने हुँदा तुलनात्मक अध्ययनका लागि उपयोगी हुन्छन्।

बढेको (फुलेको) मूल्य, मान, वा मात्रालाई समायोजन वा संकुचन गरी मूल्य, मान वा मात्रामा भएको वृद्धिको असर हटाउने प्रक्रियालाई डिफ्लेटिड भनिन्छ। यस्तो कार्य वस्तु वा सेवाको वास्तविक मूल्य पत्ता लगाउन गरिन्छ। उदाहरणका लागि आ.व. २०७७/७८ मा कृषि क्षेत्रको गार्हस्थ्य उत्पादन चालु मूल्यमा १२,७९,४४,३०,००,००० रूपैयाँ रहेको छ। कृषि क्षेत्रको यो उत्पादनलाई अन्य आ.व. हरूको उत्पादनसँग तुलना गर्नुपर्‍यो भने मूल्यवृद्धिको असर हटाउनुपर्दछ। मूल्यवृद्धिको असर हटाउनका लागि मूल्य सूचकाङ्कको प्रयोग गरी उक्त उत्पादनलाई डिफ्लेट गरिन्छ। डिफ्लेट गर्नका लागि कुल गार्हस्थ्य उत्पादन (जिडिपी) डिफ्लेटर प्रयोग गरिन्छ। जिडिपी डिफ्लेटर पनि एक किसिमको मूल्य सूचकाङ्क नै हो। आ.व. २०७७/७८ का लागि कृषि क्षेत्रको जिडिपी डिफ्लेटर १४९.६२ छ। आ.व. २०६७/६८ लाई आधार मानेर मूल्य वृद्धिको असर हटाउँदा आ.व. २०७७/७८ को कृषि क्षेत्रको गार्हस्थ्य उत्पादन स्थिर मूल्यमा,

$$\frac{१२,७९,४४,३०,००,०००}{१४९.६२} = ८,५५,१०,००,००० \text{ रूपैयाँ हुन्छ।}$$

यहाँ स्मरणीय कुरा के छ भने, चालु मूल्यलाई स्थिर मूल्यमा लैजान डिफ्लेटर (मूल्य सूचकाङ्क) ले भाग गर्नुपर्दछ। साथै, डिफ्लेट गर्न यहाँ स्मरणीय कुरा के छ भने, चालु मूल्यलाई स्थिर मूल्यमा लैजान डिफ्लेटर (मूल्य सूचकाङ्क) ले भाग गर्नुपर्दछ। साथै, डिफ्लेट गर्न प्रयोग गरिने मूल्य सूचकाङ्कको आधार समय र गार्हस्थ्य उत्पादनको आधार समय एउटै हुनुपर्दछ।

१०.७ जीविकोपार्जन सूचकाङ्क: निर्माण र प्रयोग

कुनै पनि समाजमा विभिन्न वर्गका मानिसहरू बसोबास गरेका हुन्छन्। कसैको आर्थिक अवस्था सबल हुन्छ भने कसैको कमजोर हुन्छ। जीविकोपार्जनका लागि सबैले एकै किसिमले वस्तु वा सेवा खरिद गरेर उपभोग गर्दैनन्। उदाहरणका लागि सम्पन्न मानिसले खाने चामल र सामान्य मानिसले खाने चामल फरक हुन्छ। धनी मानिसले विलासिताका सामान प्रयोग गर्दछन् भने सामान्य मानिसले जीविकोपार्जनका लागि आवश्यक सामान मात्र प्रयोग गर्दछन्। बास्मती चामलको मूल्य बढेर गरिब मानिसलाई खासै असर गर्दैन भने कोदोको मूल्य बढेर धनी मानिसलाई असर नगर्ला। तसर्थ, निश्चित समूहका मानिसहरूका लागि जीविकोपार्जन गर्न कति मूल्य आवश्यक पर्दछ भनेर अनुमान गर्नका

लागि वर्ग, स्थान वा अन्य यस्तै विशेषताका आधारमा मानिसहरूलाई समूहीकृत गरेर तोकिएको वर्ग वा समूहले उपभोग गर्ने वस्तु वा सेवाको प्रतिनिधिमूलक बास्केट तयार गरेर निकालिने खुद्रा मूल्य सूचकाङ्कलाई जीविकोपार्जन सूचकाङ्क (Cost of Living Index Number) भनिन्छ। जीविकोपार्जन सूचकाङ्क लास्पियरको मूल्य सूचकाङ्क सूत्र प्रयोग गरेर निकाल्न सकिन्छ। अर्थात्,

$$\text{जीविकोपार्जन सूचकाङ्क } (P_{01}) = \frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times 100 \quad (९)$$

जहाँ, P_1 भनेको बास्केटका वस्तुहरूको दिइएको वर्षको मूल्य हो, P_0 भनेको बास्केटका वस्तुहरूको आधार वर्षको मूल्य हो र Q_0 भनेको बास्केटका वस्तुहरूको आधार वर्षको उपभोग भएको परिमाण हो। सूत्र (९) लाई अर्को तरिकाले पनि प्रस्तुत गर्न सकिन्छ, जस्तै:

$$\text{जीविकोपार्जनको सूचकाङ्क} = \frac{\text{दिइएको वर्षको जम्मा खर्च}}{\text{आधार वर्षको जम्मा खर्च}} \times १०० \quad (१०)$$

जीविकोपार्जन सूचकाङ्क मुद्राको क्रयशक्ति मापन गर्न र वास्तविक तलब ज्याला अनुमान गर्नका लागि व्यापकरूपमा प्रयोग गरिन्छ। यसका अतिरिक्त वास्तविक आय पत्ता लगाउन, मूल्यसम्बन्धी नीति तय गर्न, ज्याला तथा अन्य आर्थिक नीति तय गर्नका लागि पनि यसको प्रयोग गरिन्छ।

$$\text{मुद्राको क्रयशक्ति} = \frac{१}{\text{जीविकोपार्जन सूचकाङ्क}} \quad (११)$$

$$\text{वास्तविक तलब ज्याला} = \frac{\text{तलब ज्याला}}{\text{जीविकोपार्जन सूचकाङ्क}} \times १०० \quad (१२)$$

वास्तविक तलब वा ज्यालाको ठुलो महत्त्व हुन्छ। सरकारले वा अन्य प्रतिष्ठानहरूले आफ्ना कर्मचारीहरूलाई तलब ज्याला तोक्न पर्दा वा बढाउन घटाउन पर्दा यसको प्रयोग गर्दछन्। महङ्गी भत्ता, अन्य भत्ता, बोनस वा यस्तै सुविधा प्रदान गर्दा पनि यसको प्रयोग हुन्छ। वास्तविक ज्यालाको अर्थ बुझ्नको लागि एउटा उदाहरण हेरौं। कुनै कम्पनीमा काम गर्ने कामदारहरूको तलब यस वर्ष १० प्रतिशतले बढेको छ। यसै वर्ष उक्त कामदारहरूको जीविकोपार्जन मूल्य सूचकाङ्कले प्रतिबिम्बित गरेबमोजिम बजारको मूल्यवृद्धि पनि १० प्रतिशत नै बढेको छ भने, उक्त कामदारहरूको वास्तविक आमदानी बढ्ला कि नबढ्ला? यस उदाहरणमा बढेको तलबले बजारको बढेको सामान खरिद गर्न खर्च हुने हुनाले कामदारको आमदानी जस्ताको तस्तै हुने भयो। यहाँनेर रोचक कुरा के छ भने, तलब वृद्धि भएकोमा कामदारहरू खुसी हुन्छन् र उत्पादकत्व बढाउँछन् तर प्रतिष्ठानका मालिकहरूले

भने कामदारको तलब बढाउँदा पनि उत्पादनको मूल्य वृद्धिसँगै अभ बढी नाफा कमाइरहेका हुन्छन् । उता कामदारको भने वास्तविक आमदानीमा कुनै असर परेको हुँदैन । त्यसैले, अर्थतन्त्रलाई चलायमान बनाई राख्न ठिक्कको सीमाभित्र मुद्रास्फीति आवश्यक हुन्छ भनिन्छ ।

१०.८ सूचकाङ्कका सीमाहरू

देशको आर्थिक तथा व्यावसायिक क्रियाकलाप अध्ययनमा सूचकाङ्कको ठुलो महत्त्व भए तापनि यसका केही सीमाहरू छन् । तसर्थ, सूचकाङ्क प्रयोगमा चनाखो हुनुपर्दछ । सूचकाङ्कका मुख्य मुख्य सीमाहरू यसप्रकार छन् ।

- सूचकाङ्क नमुना तथ्याङ्कमा आधारित भएर गणना गरिने भएकोले वास्तविक परिवर्तनलाई प्रतिबिम्बित नगर्न सक्दछ ।
- वस्तु तथा सेवाको छनोट गरी बास्केट तयार गर्दा विभिन्न प्रकारका भ्रमांशहरू (Errors) समावेश हुन सक्दछन् ।
- आधार समय छनोट गर्दा साधारण किसिमको समय छनोट नहुन सक्दछ ।
- मूल्य, मात्रा तथा मानसम्बन्धी तथ्याङ्क संकलनमा भ्रमांशहरू समावेश हुन सक्दछन् ।
- वस्तुहरूको औसत निकाल्न प्रयोग गरिने केन्द्रीय मापनको छनोट सही नहुन सक्दछ ।
- समयसँगै मानिसहरूको स्वाद, पहिरन, चलन आदिमा आउने परिवर्तनले आधार समयका परिस्थितिसँग तोकिएको समयको विशेषता मेल नखान सक्दछ । यस्तो अवस्थामा सूचकाङ्कले सापेक्ष परिवर्तनको सही मापन नगर्न सक्दछ ।
- सूचकाङ्क निकाल्ने सूत्रहरूमा कुनै न कुनै किसिमका पूर्वाग्रह समावेश भएको हुन्छ ।
- आधार समय छनोट गर्दा साधारण समय छनोट गर्ने सन्दर्भमा फेरि पनि व्यक्तिगत पूर्वाग्रह समावेश हुने सम्भावना रहन्छ ।

सूचकाङ्कका यस्ता किसिमका सीमाहरू भए तापनि यसले अर्थतन्त्र तथा व्यवसायका लागि पुऱ्याउने योगदान निकै महत्त्वपूर्ण छ । तसर्थ, वैज्ञानिक सोचबुझका साथ यसको प्रयोग गर्नुपर्दछ ।

१०.९ अभ्यासका लागि प्रश्नहरू

क) तलका पदावलीहरूको संक्षेपमा परिभाषा दिनुहोस् ।

१) सूचकाङ्क

२) सूचकाङ्क निर्माणमा प्रयोग हुने शब्द: बास्केट

३) आधार वर्ष

४) सापेक्ष मूल्य (Price Relative)

५) सूचकाङ्क निर्माणमा प्रयोग हुने शब्द: भार (Weight)

ख) तलका भनाइहरूलाई कारणसहित ठिक वा बेठिक छुट्याउनुहोस् ।

- १) सूचकाङ्कहरू अर्थतन्त्रका ब्यारोमिटर हुन् ।
 - २) जीविकोपार्जन सूचकाङ्क एक प्रकारको परिमाण सूचकाङ्क हो ।
 - ३) वास्तविक ज्याला पत्ता लगाउन थोक सूचकाङ्क प्रयोग गरिन्छ ।
 - ४) पास्चेको सूचकाङ्कमा आधार वर्षको परिमाणलाई भारको रूपमा प्रयोग गरिन्छ ।
 - ५) सूचकाङ्क निर्माणमा तथ्याङ्कको गुणस्तरको कुनै असर हुँदैन ।
- ग) जोडा मिलाउनुहोस् ।

लास्पियरको मूल्य सूचकाङ्क $P_{01} = \frac{\sum P_1 Q_1}{\sum P_0 Q_1} \times 100$

सापेक्ष मूल्य $P_{01} = \frac{\sum P_1 Q_0}{\sum P_0 Q_0} \times 100$

वास्तविक तलब ज्याला $P_{01} = \frac{p_1}{p_0} \times 100$

पास्चेको मूल्य सूचकाङ्क $\frac{\text{तलब ज्याला}}{\text{जीविकोपार्जन सूचकाङ्क}} \times 100$

मुद्राको क्रयशक्ति $\frac{1}{\text{जीविकोपार्जन सूचकाङ्क}}$

घ) सूचकाङ्कको महत्त्वका बारेमा प्रकाश पार्नुहोस् ।

ङ) सूचकाङ्कको निर्माण गर्दा भारको प्रयोग किन गरिन्छ, उदाहरणसहित व्याख्या गर्नुहोस् ।

च) तलको तालिकामा विभिन्न औद्योगिक क्षेत्रहरूको आ.व. २०७६/७७ र २०७७/७८ को चालु मूल्यमा मूल्य अभिवृद्धि र डिफ्लेटर दिइएको छ । उक्त मूल्य अभिवृद्धिलाई औद्योगिक क्षेत्रअनुसार डिफ्लेट गरी आ.व. २०६६/६७ को स्थिर मूल्यमा निकाल्नुहोस् । साथै, औद्योगिक क्षेत्रअनुसार २०७७/७८ को मूल्य अभिवृद्धि वृद्धिदर पनि निकाल्नुहोस् ।

तालिका १०.३: औद्योगिक क्षेत्रअनुसार मूल्य अभिवृद्धि तथा जिडिपी डिफ्लेटर

औद्योगिक क्षेत्र	औद्योगिक क्षेत्रअनुसार मूल्य अभिवृद्धि (चालु मूल्यमा)		जिडिपी डिफ्लेटर (आधार वर्ष २०६७/६८)	
	२०७६/७७	२०७७/७८	२०७६/७७	२०७७/७८
कृषि, वन तथा मत्स्य	९,०५,३४४	९,६४,४२९	१४४.१६	१४९.६२
खानी तथा उत्खनन्	२०,४४३	२१,५३१	१३८.१६	१३५.३७

औद्योगिक उत्पादन	१,७३,८६१	१,८९,६५७	१५२.८६	१६०.५६
विद्युत्, ग्याँस, वाफ तथा वातानुकूलित आपूर्ति	४२,५४५	४५,७८२	१४३.४५	१४३.२७
पानी आपूर्ति, ढल निकास, फोहोर व्यवस्थापन र सुधार गतिविधिहरू	२०,१२४	२०,४८१	१२७.०२	१२७.२३
निर्माण	२,०४,८६२	२,११,९९४	१४१.११	१३८.३३
थोक तथा खुद्रा व्यापार, मोटरवाहन तथा मोटरसाइकल मर्मत	५,१९,५६५	५,८५,०१३	१६५.१९	१७६.६८
यातायात तथा गोदाम	१,७९,०११	२,०२,६६२	१६१.६	१७२.४१
खाना तथा बाससम्बन्धी सेवा	५०,२६१	५८,१७३	२०७.९५	२१६.४५
सूचना तथा सञ्चार	७७,१६६	८१,००५	९६.६	९९.९५
वित्तीय तथा बिमासम्बन्धी क्रियाकलाप	२,३५,२५४	२,५६,९८३	१९९.३३	२०५.७६
घरजग्गा कारोबार तथा भाडासम्बन्धी क्रियाकलाप	३,१७,०११	३,५२,०६८	१८४.०४	१९९.१४
व्यावसायिक, वैज्ञानिक र प्राविधिक गतिविधिहरू	३५,२४१	३९,७२९	१८१.५२	२००
प्रशासनिक तथा सहयोगी सेवा गतिविधिहरू	२७,१९२	२८,४०५	१७७.७९	१८१.७८
सार्वजनिक प्रशासन र रक्षा; अनिवार्य सामाजिक सुरक्षा	२,७६,११०	२,८७,४६८	२७१.७७	२७३.४
शिक्षा	२,८८,४५९	३,००,३५८	२३२.८१	२३३.९९
मानव स्वास्थ्य तथा सामाजिक कार्य	६०,३२८	६४,६४३	२१९.३६	२२०.६५
कला, मनोरञ्जन, अन्य सेवा गतिविधिहरू र रोजगारदाताको रूपमा घरपरिवारका गतिविधिहरू, अविभाज्य वस्तु र सेवा उत्पादन गतिविधिहरू	२१,३६४	२२,८९२	२१२.९७	२२१.३६

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपालको राष्ट्रिय लेखा, २०७७/७८^{३९}

39 Central Bureau of Statistics. (2021). *National Accounts of Nepal 2020/21*. Kathmandu, Nepal; Central Bureau of Statistics. Retrieved on November 27, 2021 from: <https://cbs.gov.np/national-accounts-of-nepal-2020-21/>

१०.१० सन्दर्भ सामग्री

सूचकाङ्कसम्बन्धी विस्तृत जानकारीका लागि निम्नानुसारका सामग्री उपयोगी छन् ।

१. विष्ट, विकास (२०६६) । मूल्य सूचकाङ्क । केन्द्रीय तथ्याङ्क विभाग, नेपालमा आधिकारिक तथ्याङ्क प्रणाली तथा विधि (पृष्ठ २४४-२४९) । काठमाण्डौं, नेपाल: केन्द्रीय तथ्याङ्क विभाग ।
२. Statistical Techniques in Business & Economics by Douglas A. Lind, William G. Marchal Samuel A. Wathen, Seventeenth Edition, McGraw-Hill Education, 2018
३. Fundamentals of Statistics by S.C. Gupta, Sixth Revised & Enlarged Edition, Himalaya Publishing House, 2009



भाग ३

वर्णनात्मक र अनुमानात्मक तथ्याङ्क

केन्द्रीय प्रवृत्ति मापन

११.१ परिचय

यस अधिका पाठहरूमा हामीले मूलतः तथ्याङ्कको परिचयदेखि यसको संकलन र प्रस्तुतिकरणका सम्बन्धमा चर्चा गर्छौं। यसका अतिरिक्त तथ्याङ्क व्यवस्थापनका लागि आवश्यक केही गणितीय सूचकहरूका बारेमा पनि जानकारी हासिल गर्छौं। अब यस पाठमा हामीले तथ्याङ्कका थप विशेषताहरू अध्ययन गर्नका लागि तथ्याङ्कीय विधिहरूको प्रयोग गरेर केन्द्रीय प्रवृत्तिको मापन गर्ने सम्बन्धमा चर्चा गर्नेछौं। केन्द्रीय प्रवृत्ति तथ्याङ्कको औसत मानसँग सम्बन्धित विषय हो। तथ्याङ्कको औसत मान भन्नाले एउटा त्यस्तो एकल मान वा आँकडा हो, जसले तोकिएको तथ्याङ्क समूहको प्रतिनिधित्व गर्दछ। मध्यक (Mean), मध्यिका (Median), बहुलक (Mode) केन्द्रीय प्रवृत्तिका धेरै प्रयोगमा आउने उदाहरण हुन्। अर्को शब्दमा भन्नुपर्दा, कुनै पनि आँकडालाई समग्रमा परिभाषित गर्न उक्त आँकडा समग्रमा कुन अङ्कको वरिपरि बढी रहेको छ वा केन्द्रित छ भनेर हेर्न केन्द्रीय प्रवृत्तिको मापन गरिन्छ। तथ्याङ्कको प्रकृति र विश्लेषणको उद्देश्य अनुसार मध्यक, मध्यिका वा बहुलकमध्ये कुनै एक वा दुई वा सबैको प्रयोग गर्न सकिन्छ। उदाहरणका लागि मिति २०७८ मंसिर १० देखि १६ गतेसम्म संकलित राजस्वको तथ्याङ्क क्रमशः २०७२०, ३१८७०, २३१७७, २१७३१, ४१५२४, १७१, १२१६९ लाख रूपैयाँ पाइयो।¹ एक हप्ताको यस तथ्याङ्कबाट औसत एक दिनमा कति राजस्व संकलन हुने रहेछ भनेर हेर्दा २१६२३ लाख रूपैयाँ पाइयो। यसबाट सरकारले राजस्व संकलनमा राखेको दैनिक लक्ष्य प्राप्त हुँदै छ वा छैन भनेर अनुमान लगाउन सकिन्छ। त्यस्तै, कुनै गाउँपालिकाले मानौं विद्यालयस्तरको शिक्षालाई निःशुल्क बनाउने नीति निर्माण गर्न खोजिरहेको छ भने उक्त गाउँपालिकाले यससम्बन्धी खर्चको अनुमान गर्न औसतमा एक जना विद्यार्थीलाई कति खर्च व्यहोर्नु पर्दछ भनेर निकाल्नुपर्ने हुन्छ। त्यसका लागि प्रत्येक कक्षामा अध्ययनरत विद्यार्थीको मासिक शुल्क संकलन गरेर उक्त संकलित तथ्याङ्कको केन्द्रीय मान पत्ता लगाउनु पर्ने हुन्छ। यस्तो प्रकृतिको तथ्याङ्क र उद्देश्यका लागि मध्यक सही मापन हुन्छ। त्यस्तै, मानौं, कुनै विद्यालयले कुनै कक्षामा अध्ययनरत विद्यार्थीको गुणस्तर सुधार गर्न दुईवटा एकरूपका समूह बनाएर अतिरिक्त कक्षा सञ्चालन गर्ने योजना बनाउँदै छ। एउटा

1 Financial Comptroller General Office. (2021). *Daily Receipts and Payment Status*. Kathmandu, Nepal; Financial Comptroller General Office. Retrieved on December 3, 2021 from: https://www.fcgo.gov.np/daily-budgetary-analysis;-date_type=bs&eng_search_date=&np_search_date

समूहमा तुलनात्मक रूपमा कमजोर र अर्कोमा तुलनात्मक रूपमा केही सबल विद्यार्थीको समूह बनाउनका लागि उनीहरूले परीक्षामा प्राप्त गरेको अंकको मध्यिका निकालेर प्रयोग गर्न सकिन्छ । पसलेहरूले बहुसंख्यक मानिसलाई मिल्ने जुत्ताको साइज पत्ता लगाउन बहुलक प्रयोग गर्न सक्दछन् ।

११.२ असल केन्द्रीय प्रवृत्तिका गुणहरू

एउटा मानका आधारमा समग्र तथ्याङ्कको विशेषता सम्बन्धमा अवधारणा बनाउनु पर्ने हुनाले केन्द्रीय प्रवृत्ति आदर्श हुनुपर्दछ । आदर्श केन्द्रीय प्रवृत्तिका प्रमुख गुणहरू यसप्रकार छन् ।

- केन्द्रीय प्रवृत्ति स्पष्ट रूपमा परिभाषित भएको हुनुपर्दछ ।
- यसलाई बुझ्न र पत्ता लगाउन सजिलो हुनुपर्दछ ।
- सकेसम्म तथ्याङ्कका सबै आँकडाहरू प्रयोग गरी निकालेको हुनुपर्दछ ।
- यसलाई अन्य थप गणितीय हिसाबहरूमा प्रयोग गर्न सकिने खालको हुनुपर्दछ ।
- नमुना छनोटको फेरबदलले कम प्रभावित हुने खालको हुनुपर्दछ । अर्थात् एउटै जनसंख्याबाट जतिपटक नमुना लिए तापनि सामान्यतया औसतहरू उस्तै उस्तै आउनुपर्दछ ।
- तथ्याङ्कमा भएका अति कम वा अति बढी परिमाण बोकेका आँकडाहरूबाट कम प्रभावित हुने खालको हुनुपर्दछ ।

माथि सूचीमा उल्लेख गरे बमोजिम सबै गुणहरू कुनै पनि औसतमा हुँदैन । तर अधिकांश गुण धारण गरेका औसतलाई बढी प्रयोग गरिन्छ ।

११.३ केन्द्रीय प्रवृत्ति मापन

यस खण्डमा हामीले मध्यक, मध्यिका र बहुलक कसरी निकाल्ने भन्ने बारेमा अध्ययन एवं अभ्यास गर्दछौं ।

११.३.१ मध्यक (Mean)

कुनै आँकडाको अध्ययन गर्दा प्रत्येक एकाइको वा एकाइमानको विश्लेषण गर्नु मनासिब र सम्भव हुँदैन । यस्तो अवस्थामा ती एकाइहरूको केन्द्रीय मान (औसत) निकालेर त्यस आँकडाको समष्टिगत विशेषताका बारेमा अर्थ निकाल्नु पर्ने वा धारणा बनाउनु पर्ने हुन्छ । कुनै आँकडामा छरिएर रहेका अङ्कहरूको केन्द्रीय प्रवृत्ति वा चरित्र वा मान के हो भनेर बुझाउने तथ्याङ्कशास्त्रीय साधनलाई मध्यक भनिन्छ । यो अवस्था र विषय प्रकृति हेरी अङ्कगणितीय वा ज्यामितीय (गुणोत्तर) वा हार्मोनिक औसतका रूपमा निकाल्ने र प्रयोग गर्ने गरिन्छ ।

क) अङ्कगणितीय मध्यक (Arithmetic Mean)

यो औसतहरू मध्येको सबैभन्दा सरल र बढी प्रयोग हुने प्रकारको मध्यक हो । आँकडामा भएका सबै अङ्कहरूलाई जोडेर तिनीहरूको जम्मा संख्याले भाग गरेपछि यो मध्यक प्राप्त हुन्छ ।

उदाहरणः

तालिका ११.१: सन् २०१९ मा नेपाल भित्रिएका पर्यटकको मासिक संख्या विवरण

महिना	पर्यटक संख्या
जनवरी	८१,२७३
फेब्रुअरी	१,०२,४२३
मार्च	१,२७,३५१
अप्रिल	१,०९,३९९
मे	७८,३२९
जुन	७४,८८३
जुलाई	७०,९१६
अगस्ट	९४,७४९
सेप्टेम्बर	९२,६०४
अक्टोबर	१,३४,०९६
नोभेम्बर	१,३०,३०२
डिसेम्बर	१,००,८६६

स्रोत: नेपाल पर्यटन बोर्ड

माथिको तालिकाबाट सन् २०१९ मा नेपालमा मासिक औसत कति पर्यटक भित्रिए भनेर थाह पाउनका लागि अंकगणितीय मध्यक निकालेर हेर्न सकिन्छ । त्यसका लागि बाह्रै महिनाका सबै पर्यटकका संख्यालाई जोडेर आएको जोडफललाई १२ ले भाग गर्नुपर्दछ । अर्थात्,

$$\frac{८१,२७३ + १,०२,४२३ + \dots + १,३०,३०२ + १,००,८६६}{१२}$$

$$= \frac{११,९७,१९१}{१२} = ९९,७६६$$

सन् २०१९ मा औसतमा मासिक ९९,७६६ पर्यटक नेपाल भित्रिएका थिए ।

औसतहरू निकाल्नका लागि आँकडाहरू कसरी उपलब्ध छन् भन्ने कुरा प्रमुख हुन्छन् । माथि उल्लेख गरिएको तरिका व्यक्तिगत श्रेणीको लागि हो । खण्डित र अविच्छिन्न श्रेणीका तथ्याङ्कको मध्यक निकाल्नका लागि फरक-फरक विधि अनुसरण गर्नुपर्ने हुन्छ । यी तिन किसिमका तथ्याङ्कको अंकगणितीय मध्यक निकाल्ने विधि यसप्रकार छ ।

व्यक्तिगत श्रेणी (Individual Series)

कुनै पनि व्यक्तिगत श्रेणीमा भएको आँकडा अर्थात् सम्पूर्ण आँकडालाई जस्ताको त्यस्तै राखिएको (कुनै बारम्बारता वितरणमा नराखिएको) अवस्थामा अङ्कगणितीय औसत निकाल्ने प्रचलित सूत्र यसप्रकार छ:

$$\bar{X} = \quad (9)$$

जहाँ, $\sum X$ ले सबै अङ्कहरूको कुल जम्मालाई जनाउँछ र N ले जम्मा कतिवटा अङ्क समावेश गरिएको छ भन्ने जनाउँछ ।

खण्डित श्रेणी (Discrete Series)

खण्डित श्रेणीका आँकडाहरूको आवृत्ति वितरण गरी राखिएको तथ्याङ्कको अंकगणितीय मध्यक निकाल्न तलको सूत्रको प्रयोग गर्न सकिन्छ ।

$$\bar{X} = \frac{\sum fX}{\sum f} \quad (2)$$

जहाँ, f भनेको सम्बन्धित आँकडाको आवृत्ति (बारम्बारता) हो, $\sum fX$ भनेको बारम्बारता र सम्बन्धित आँकडाको गुणनफलहरूको योग हो भने $\sum f$ भनेको बारम्बारताको योगफल हो ।

उदाहरण:

तालिका ११.२: नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ मा छनोटमा परेका परिवारको परिवार सदस्य संख्या

क्र.सं.	परिवार सदस्य संख्या (x)	परिवार संख्या (f)	$f x$
१	१	२०७	२०७
२	२	६४६	१,२९२
३	३	९४५	२,८३५
४	४	१,२५४	५,०१६
५	५	१,०८६	५,४३०

६	६	७९८	४,७८८
७	७	४५०	३,१५०
८	८	२५७	२,०५६
९	९	१३५	१,२१५
१०	१०	८१	८१०
११	११	४९	५३९
१२	१२	२५	३००
१३	१३	१७	२२१
१४	१४	८	११२
१५	१५	१२	१८०
१६	१६	४	६४
१७	१७	४	६८
१८	१८	२	३६
१९	१९	५	९५
२०	२०	३	६०
जम्मा		५,९८८	२८,४७४

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

तालिका ११.२ खण्डित श्रेणीको आवृत्ति वितरण गरी राखिएको तथ्याङ्कको उदाहरण हो। यो तालिकाबाट हामीले अंकगणितीय मध्यक अथवा औसत परिवार सदस्य संख्या निकाल्न सक्दछौं।

$$\text{अंकगणितीय मध्यक } (\bar{X}) = \frac{\sum fX}{\sum f} = \frac{२८४७४}{५९८८} = ४.८ \approx ५$$

तसर्थ, छोटोमा परेका परिवारहरूको औसत परिवार सदस्य संख्या ४.८ अर्थात्, एउटा परिवारमा लगभग ५ जना सदस्य रहेछन्।

अविच्छिन्न श्रेणी (Continuous Series)

अविच्छिन्न श्रेणीका समूहीकृत तथ्याङ्कको अंकगणितीय मध्यक निकाल्न निम्नानुसारको सूत्र प्रयोग गरिन्छ।

$$\bar{X} = \frac{\sum fX}{\sum f}$$

जहाँ, m भनेको दिइएको समूहको मध्यमान हो।

उदाहरण:

तालिका ११.३: नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ मा छनोटमा परेका परिवारको परिवार सदस्यहरूको उमेर समूहअनुसार आवृत्ति वितरण

उमेर समूह	संख्या (f)	मध्यमान (m)	fm
० देखि १० वर्ष	६,३१०	५	३१,५५०
१० देखि २० वर्ष	७,००९	१५	१,०५,१३५
२० देखि ३० वर्ष	४,२४१	२५	१,०६,०२५
३० देखि ४० वर्ष	३,४७८	३५	१,२१,७३०
४० देखि ५० वर्ष	२,७४५	४५	१,२३,५२५
५० देखि ६० वर्ष	२,३५६	५५	१,२९,५८०
६० देखि ७० वर्ष	१,५१९	६५	९८,७३५
७० देखि ८० वर्ष	७६८	७५	५७,६००
८० देखि ९० वर्ष	२१७	८५	१८,४४५
९० देखि १०० वर्ष	२७	९५	२,५६५
जम्मा	२८,६७०		७,९४,८९०

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

तालिका ११.३ मा दिइएको तथ्याङ्क अविच्छिन्न श्रेणीको उदाहरण हो। यसबाट छनोटमा परेका परिवारका सदस्यहरूको अंकगणितीय मध्यक अथवा औसत उमेर निकाल्न सकिन्छ। त्यसका लागि प्रत्येक उमेर समूहको मध्यमान निकाल्नु पर्दछ। मध्यमान निकाल्नका लागि समूहको दुईवटा सीमाका संख्याहरूको मध्यक निकाल्नु पर्दछ। जस्तै, समूह ० देखि $\frac{0+10}{2} = 5$, १० देखि २० का लागि $\frac{10+20}{2} = 15$,.....।

र मध्यमानको गुणा गरी गुणनफलहरूको योगफल निकाल्नु पर्दछ। त्यसपछि, अंकगणितीय मध्यक अथवा औसत उमेर सूत्रको प्रयोग गरेर निकाल्न सकिन्छ।

$$\text{अर्थात्, औसत उमेर } (\bar{X}) = \frac{\sum fm}{\sum f} = \frac{794890}{28670} = 27.73$$

तसर्थ, छनोटमा परेका परिवारका सदस्यहरूको औसत उमेर २७.७३ वर्ष रहेको देखिन्छ।

ख) ज्यामितीय (गुणोत्तर) मध्यक (Geometric Mean)

ज्यामितीय औसत पनि अङ्कगणितीय औसत जस्तै कुनै आँकडालाई प्रतिनिधित्व गर्ने गरी केन्द्रीय प्रवृत्ति हेर्ने एउटा तथ्याङ्कीय विधि हो। विभिन्न प्रकारका तथ्याङ्कहरूको

ज्यामितीय मध्यक निकाल्ने विधि यसप्रकार छः

व्यक्तिगत श्रेणी

ज्यामितीय मध्यक निकाल्नको लागि दिइएका आँकडाहरूको गुणनफल निकालेर जतिवटा अङ्कहरूको गुणनफल निकालिएको हो त्यति नै अङ्कको मूलांक (Root) निकाल्नु पर्दछ ।

$$\bar{X}_{geom} = \sqrt[N]{\prod_{i=1}^N x_i} = \sqrt[N]{X_1 \cdot X_2 \dots \dots \dots X_N}$$

जहाँ \bar{X}_{geom} भनेको ज्यामितीय मध्यक हो, $\sqrt[N]{\dots}$ ले N औँ मूलांक जनाउँछ भने $\prod_{i=1}^N \dots$ ले N औँ पदहरूको गुणनफल लिने भन्ने बुझाउँछ ।

उदाहरणः नेपालको आ.व. २०६८/६९ देखि २०७५/७६ सम्मको कुल गार्हस्थ्य उत्पादन वृद्धिदर तालिका ११.४ मा दिइएको छ । यस आँकडाबाट अवधिको औसत आर्थिक वृद्धिदर निकाल्नुहोस् ।

तालिका ११.४: वार्षिक आर्थिक वृद्धिदर (स्थिर मूल्यमा)

२०६८/ ६९	२०६९/ ७०	२०७०/ ७१	२०७१/ ७२	२०७२/ ७३	२०७३/ ७४	२०७४/ ७५	२०७५/ ७६
४.९५	३.०७	५.७४	३.५१	०.००३	८.५९	७.३७	६.३९

स्रोतः केन्द्रीय तथ्याङ्क विभाग, नेपालको राष्ट्रिय लेखा

वृद्धिदरको औसत निकाल्न ज्यामितीय मध्यक प्रयोग गरिन्छ ।

$$\bar{X}_{geom} = \sqrt[N]{\prod_{i=1}^N x_i} = \sqrt[8]{4.95 \times 3.07 \times 5.74 \times 3.51 \times 0.003 \times 8.59 \times 7.37 \times 6.39}$$

यो गणितीय क्रियाकलाप गर्न निकै फन्फटिलो छ । तसर्थ, हामीले क्याल्कुलेटर वा कम्प्युटरको प्रयोग गरेर औसत आर्थिक वृद्धिदर २.०५ भन्न सकिन्छ ।

टिप्पणीः यस उदाहरणबाट मुख्यतया दुईवटा कुराहरू विशेष ध्यान दिनुपर्ने हुन्छः आँकडाको कुनै मान शून्य हुँदा वा कुनै मान नकारात्मक हुँदा यो मध्यक प्रयोग गर्न सकिँदैन ।

खण्डित श्रेणी र अविच्छिन्न श्रेणी

खण्डित श्रेणी र अविच्छिन्न श्रेणीको ज्यामितीय मध्यक निकाल्न निकै कठिन हुन्छ ।

क्याल्कुलेटर वा कम्प्युटरको प्रयोग गरेर मात्र यी मध्यकहरू निकाल्न सकिन्छ । यद्यपि यहाँ, पाठकवर्गको जानकारीका लागि सूत्रहरू मात्र उल्लेख गरिएको छ ।

$$\text{खण्डित श्रेणी} = \text{Antilog} \left[\frac{1}{N} \sum_{i=1}^N \log X_i \right]$$

$$\text{अविच्छिन्न श्रेणी} = \text{Antilog} \left[\frac{1}{N} \sum_{i=1}^N \log M_i \right]$$

ग) लयबद्ध मध्यक (Harmonic Mean)

लयबद्ध औसत पनि अङ्कगणितीय र ज्यामितीय औसत जस्तै कुनै तथ्याङ्कलाई प्रतिनिधित्व गर्ने गरी केन्द्रीय प्रवृत्ति हेर्ने एउटा तथ्याङ्कीय विधि हो । अधिल्ला दुई औसतभन्दा यो विरलै प्रयोगमा ल्याइन्छ । खासमा यो अङ्कगणितीय औसतको विलोम औसत भने पनि हुन्छ । यो मध्यक निकाल्नका लागि पहिला सबै आँकडाहरूको रेसिप्रोकल निकालेर तिनीहरूको अंकगणितीय मध्यक निकाल्नु पर्छ र पुनः त्यसको रेसिप्रोकल निकाल्नु पर्दछ । यहाँ, यो मध्यक निकाल्ने सूत्रमात्र प्रस्तुत गरिएको छ ।

व्यक्तिगत श्रेणी

$$HM = \frac{1}{\frac{1}{N} \left[\frac{1}{X_1} + \frac{1}{X_2} + \dots + \frac{1}{X_N} \right]} + \frac{N}{\sum_{i=1}^N \frac{1}{X_i}}$$

खण्डित श्रेणी

$$HM = \frac{1}{\frac{1}{N} \left[\frac{f_1}{X_1} + \frac{f_2}{X_2} + \dots + \frac{f_N}{X_N} \right]} + \frac{N}{\sum_{i=1}^N \frac{f_i}{X_i}}$$

अविच्छिन्न श्रेणी

$$HM = \frac{1}{\frac{1}{N} \left[\frac{f_1}{M_1} + \frac{f_2}{M_2} + \dots + \frac{f_N}{M_N} \right]} + \frac{N}{\sum_{i=1}^N \frac{f_i}{M_i}}$$

जहाँ, HM भनेको हार्मोनिक मध्यक हो, f भनेको बारम्बारता (आवृत्ति) हो र M भनेको वर्ग समूहको मध्यमान हो ।

११.३.२ मध्यिका (Median)

मध्यिका भन्नेबित्तिकै बिचको भन्ने बुझिन्छ । मध्यिकाले दिइएका आँकडाहरूलाई बराबर दुई भागमा विभाजन गरिदिन्छ । कुनै अङ्क वा वस्तुहरूलाई सानोदेखि ठुलो क्रममा राख्दा बिचमा पर्ने अङ्क वा वस्तुलाई मध्यिका भनिन्छ । मध्यिका भन्दा-साना अङ्क वा वस्तु ५०% र सो भन्दा-ठुलो अङ्क वा वस्तु ५०% को सङ्ख्यामा छन् भन्ने बुझिन्छ । दिइएको तथ्याङ्कलाई अथवा बारम्बारताको वितरणलाई सानोदेखि ठुलोसम्म वा

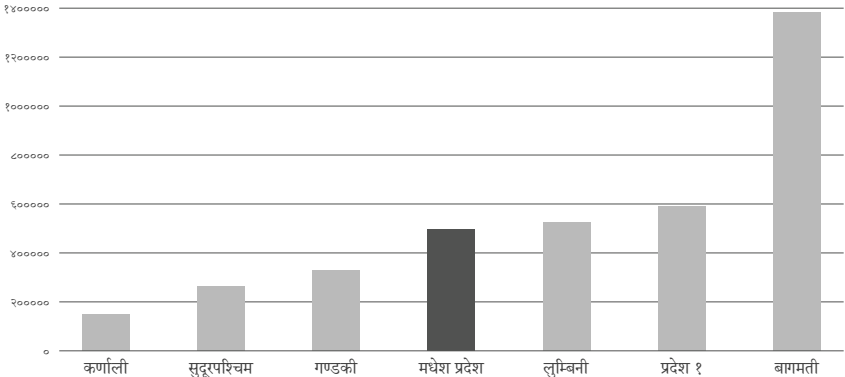
तुलोदेखि सानोसम्मको क्रम मिलाएर राखेपछि ठिक मध्य भागमा जुन मान पर्दछ त्यो नै मध्यिका हो । उदाहरणको लागि सातवटा प्रदेशको आ.व. २०७७/७८ को वार्षिक चालु मूल्यमा कुल मूल्य अभिवृद्धि (आधार मूल्यमा) तालिका ११.५ मा देखाइएको छ । जसअनुसार मधेश प्रदेशको मूल्य अभिवृद्धि ५००७५ ले उक्त आँकडाहरूलाई बराबर दुई भागमा विभाजन गरेका कारण ५००७५ नै मध्यिका हो । चित्रमा मध्यिका कस्तो देखिन्छ त ? यहाँ, हेरौं । चित्र ११.१ मा देखाइए अनुसार सातवटा प्रदेशहरूको मूल्य अभिवृद्धिको मध्यिका मधेश प्रदेश को मूल्य अभिवृद्धिले प्रतिनिधित्व गरेको छ । यसले के बुझाउँछ भने मधेश प्रदेश भन्दा-कम मूल्य अभिवृद्धि भएका र बढी मूल्य अभिवृद्धि भएका प्रदेशहरू बराबर छन् अर्थात्, यहाँ तिन तिनवटा छन् ।

तालिका ११.५: प्रदेशको आ.व. २०७७/७८ को वार्षिक चालु मूल्यमा
कुल मूल्य अभिवृद्धि (आधार मूल्यमा)

प्रदेश	मूल्य अभिवृद्धि
कर्णाली प्रदेश	१,५१,८१६
सुदूरपश्चिम प्रदेश	२,६१,१६८
गण्डकी प्रदेश	३,२६,२४८
मधेश प्रदेश	५,००,५७५
लुम्बिनी प्रदेश	५,२६,३१२
प्रदेश १	५,९०,३४०
बागमती प्रदेश	१३,७६,८१४

स्रोत: केन्द्रीय तथ्याङ्क विभाग, प्रदेशको कुल गार्हस्थ्य उत्पादन, २०७७/७८

चित्र नं. ११.१: प्रदेशको आ.व. २०७७/७८ को वार्षिक चालु मूल्यमा
कुल मूल्य अभिवृद्धि (आधार मूल्यमा)



विभिन्न प्रकारका तथ्याङ्कका लागि मध्यिका निकाल्ने विधि र सूत्र यसप्रकार छ ।

व्यक्तिगत श्रेणी

कुनै पनि व्यक्तिगत श्रेणीमा भएको आँकडा अर्थात् सम्पूर्ण आँकडालाई जस्ताको त्यस्तै राखिएको (कुनै बारम्बारता वितरणमा नराखिएको) अवस्थामा आँकडालाई बढ्दो वा घट्दो क्रममा राख्ने, त्यसपछि जम्मा आँकडाको संख्यामा १ जोडेर २ ले भाग गर्ने, जति आउँछ त्यति औँ स्थानको अङ्क नै मध्यिका हुन्छ । यसका चरणहरू उदाहरणसहित हेरौं । तालिका ११.६ मा दिइएको तथ्याङ्क व्यक्तिगत श्रेणीको तथ्याङ्क हो ।

१. दिइएको आँकडाहरूलाई सानोदेखि ठुलो क्रममा राखौं ।

२६१७७०, ४०९८७९, ४१३३००, ६५८७७६, ६७२९२७, ६९७२९३, ७१७१४८

२. यहाँ, जम्मा आँकडाहरूको संख्या ७ छ । अब $\frac{(७+१)}{२} = ४$ निकालौं ।

३. चौथो स्थानको आँकडा अर्थात्, ६५८७७६ नै मध्यिका हो ।

तालिका ११.६: प्रदेशअनुसार कृषक परिवार संख्या

प्रदेश	कृषक परिवार संख्या
प्रदेश १	७,१७,१४८
मधेश प्रदेश	६,७२,९२७
बागमती	६,५८,७७६
गण्डकी	४,१३,३००
लुम्बिनी	६,९७,२९३
कर्णाली	२,६१,७७०
सुदूरपश्चिम	४,०९,८७९

स्रोत: केन्द्रीय तथ्याङ्क विभाग, राष्ट्रिय कृषि गणना, २०६८

टिप्पणी: यदि दिइएका आँकडाहरूको संख्या जोर भएमा खोजेको संख्या भेटिँदैन । यस्तो अवस्थामा $((N+1))/2$ को मान दशमलवमा आउने गर्दछ, त्यस्तो बेलामा अगाडिको अङ्क र पछाडिको अङ्कको औसत अर्थात् बिचको अङ्कलाई लिने गरिन्छ । उदाहरणको लागि ६ जना व्यक्तिहरूको उचाइ क्रमशः १८०, १५०, १७०, १८०, १६०, १७५ सेन्टिमिटर छ भने मध्यिका निकाल्दा,

पहिला सानोदेखि ठुलो क्रममा मिलाऊं । यहाँ, जम्मा संख्या $(N) = ६$ छ ।

१५०, १६०, १७०, १७५, १८०, १८०

$$\text{अव, } \frac{(N+1)}{2} = \frac{(६+१)}{२} = ३.५ \text{ हुन्छ।}$$

यस्तो अवास्थामा ३.५ औं व्यक्तिको उचाइ भन्ने हुँदैन। त्यसैले तेस्रो र चौथो स्थानमा पर्ने व्यक्तिहरूको औसत वा बिचको अङ्क नै मध्यिका हुन्छ। अर्थात्,

पहिलो	दोस्रो	तेस्रो	मध्यिका (३.५ औं)	चौथो	पाँचौं	छैटौं
१५०	१६०	१७०	$(१७०+१७५)/२=१७२.५$	१७५	१८०	१८०

खण्डित श्रेणी

आवृत्ति वितरणमा रहेका तथ्याङ्कको मध्यिका निकाल्न निम्नानुसार गर्नुपर्दछ।

उदाहरण:

तालिका ११.७: नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ मा छनोटमा परेका परिवारको परिवार सदस्य संख्या

दिइएको तथ्याङ्क (आवृत्ति वितरण)		पहिलो चरण
परिवार सदस्य संख्या (आँकडा)	जम्मा परिवार (बारम्बारता)	भन्दा-कम सञ्चित बारम्बारता निकाल्ने
१	२०७	२०७
२	६४६	८५३
३	९४५	१,७९८
४	१,२५४	३,०५२
५	१,०८६	४,१३८
६	७९८	४,९३६
७	४५०	५,३८६
८	२५७	५,६४३
९	१३५	५,७७८
१०	८१	५,८५९
११	४९	५,९०८
१२	२५	५,९३३
१३	१७	५,९५०
१४	८	५,९५८
१५	१२	५,९७०

१६	४	५,९७४
१७	४	५,९७८
१८	२	५,९८०
१९	५	५,९८५
२०	३	५,९८८
N = ५,९८८		

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

यस तथ्याङ्कबाट मध्यिका निकाल्नुहोस् ।

१. दिइएको आवृत्ति वितरणलाई क्रम मिलाएर राख्ने । (तालिका ११.७ मा क्रमबद्धता सानोबाट ठुलो मिलेको छ ।)
२. माथिदेखि तलसम्म बारम्बारताहरूलाई जोडेर भन्दा-कम सञ्चित बारम्बारता (Less Than Cumulative Frequency) निकाल्ने ।
३. $\frac{N}{2} = \frac{५९८८}{२} = २९९४$ निकाल्ने
४. भन्दा-कम सञ्चित बारम्बारतामा २९९४ भन्दा नजिकको तर ठुलो संख्या खोज्ने । त्यस्तो संख्या ३०५२ फेला पऱ्यो ।
५. यही ३०५२ सँग सम्बन्धित परिवार सदस्य संख्या अर्थात् ४ नै मध्यिका हो । यसबाट के देखिन्छ भने कुल छनोट भएका ५,९८८ परिवारमध्ये परिवार सदस्य संख्या ४ भन्दा कम र ४ भन्दा बढी हुने परिवार संख्या बराबर छन् । यही तथ्याङ्कको अंकगणितीय मध्यक ४.८ आएको थियो ।

अविच्छिन्न श्रेणी

समूहीकृत आवृत्ति वितरणमा उपलब्ध तथ्याङ्कबाट मध्यिका निकाल्न निम्नानुसार गर्नुपर्दछ ।

उदाहरण: तलको तालिका ११.८ बाट मध्यिका उमेर निकाल्नुहोस् ।

१. दिइएको आवृत्ति वितरणलाई क्रम मिलाएर राख्ने । (तालिका ११.८ मा क्रमबद्धता सानोबाट ठुलो मिलेको छ ।)
२. माथिदेखि तलसम्म बारम्बारताहरूलाई जोडेर भन्दा-कम सञ्चित बारम्बारता (Less Than Cumulative Frequency) निकाल्ने ।

$$३. \frac{N}{2} = \frac{२८६७०}{२} = १४३३५ \text{ निकाल्ने ।}$$

४. भन्दा-कम सञ्चित बारम्बारतामा १४३३५ भन्दा नजिकको तर ठुलो संख्या खोज्ने । त्यस्तो संख्या १७,५६० फेला पत्थो ।

५. यही १७,५६० सँग सम्बन्धित उमेर समूह अर्थात् २० देखि ३० वर्ष नै मध्यिका समूह हो । यही समूहको कुनै संख्या मध्यिका हुन्छ । उक्त मध्यिका पत्ता लगाउन निम्न सूत्र प्रयोग गर्नुपर्दछ ।

$$\text{मध्यिका (Md)} = l + \frac{\frac{N}{2} - cf}{f} \times h,$$

जहाँ, l, f, cf, h भनेको क्रमशः मध्यिका पर्ने समूहको तल्लो मान, मध्यिका पर्ने समूहको बारम्बारता, मध्यिका पर्ने समूहभन्दा अधिल्लो समूहको भन्दा-साना सञ्चित बारम्बारता, र मध्यिका पर्ने समूहको अन्तराल हो ।

तलको उदाहरणमा $l=२०$, $f=४२४१$, $cf=१३३१९$, $h=१०$ छ । यी मानहरू सूत्र (१०) मा राखौं ।

$$\text{मध्यिका (Md)} = L + \frac{\frac{N}{2} - cf}{f} \times h = २० + \frac{१४३३५ - १३३१९}{४२४१} \times १० = २२.४ \text{ वर्ष}$$

तसर्थ, मध्यिका २२.४ वर्ष भयो ।

तालिका ११.८: नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ मा छनोटमा परेकापरिवारको परिवार सदस्यहरूको उमेर समूहअनुसार आवृत्ति वितरण

उमेर समूह	संख्या (f)	भन्दा-कम सञ्चित बारम्बारता निकाल्ने
० देखि १० वर्ष	६,३१०	६,३१०
१० देखि २० वर्ष	७,००९	१३,३१९
२० देखि ३० वर्ष	४,२४१	१७,५६०
३० देखि ४० वर्ष	३,४७८	२१,०३८
४० देखि ५० वर्ष	२,७४५	२३,७८३
५० देखि ६० वर्ष	२,३५६	२६,१३९
६० देखि ७० वर्ष	१,५१९	२७,६५८
७० देखि ८० वर्ष	७६८	२८,४२६
८० देखि ९० वर्ष	२१७	२८,६४३

१० देखि १०० वर्ष	२७	२८,६७०
जम्मा	N = २८,६७०	

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

११.३.३ बहुलक (Mode)

कुनै पनि तथ्याङ्कमा सबैभन्दा बढी दोहोरिएको मान नै बहुलक हो। सामान्यतया, बहुलकको वरिपरि अन्य आँकडाहरू घनारूपमा वितरित भएका हुन्छन्। व्यवसायमा बहुलक निकै महत्त्वपूर्ण मानिन्छ। उमेरअनुसार वितरण गरिएको नेपालको जनसंख्यामा सबैभन्दा धेरै जनसंख्या १० वर्ष उमेर समूहमा (८,११,७४८ जना) रहेको छ। तसर्थ, उमेर अनुसारको जनसंख्याका लागि १० वर्ष उमेर बहुलक हुन्छ।^२ कुनै कपडा उत्पादन गर्ने कम्पनीले कुन कुन आकारका कपडा कति उत्पादन गर्ने भन्ने विषय बहुलकसँग सम्बन्धित छ। धेरै साना र धेरै ठुला आकारका कपडाभन्दा मध्यम आकारका कपडा बढी बिक्री हुनसक्दछ। यस्तो मध्यम आकारको पहिचान बहुलकले गर्दछ। बहुलकले सबैभन्दा बढी माग गरिने आकार बताउँछ। कतिपय तथ्याङ्कको बहुलक निकाल्न साह्रै सजिलो हुन्छ। साधारण वितरण (*Normal Distribution*) मा रहेका तथ्याङ्कहरूको दोहोरिएको पटक हेरेर वा बारम्बारता हेरेर बहुलक पत्ता लगाउन सकिन्छ।^३ साधारण वितरणमा नरहेका तथ्याङ्कहरूको बहुलक निकाल्न अलि बढी प्राविधिक भएको हुनाले यहाँ चर्चा गरिएको छैन। साधारण वितरणमा रहेका तथ्याङ्कहरूको बहुलक निकाल्ने उदाहरणसहित निम्नानुसार वर्णन गरिएको छ।

व्यक्तिगत श्रेणी

व्यक्तिगत श्रेणीका आँकडाहरू दोहोरिएको पटक हेरेर सबैभन्दा बढी दोहोरिएको आँकडालाई बहुलक भन्नुपर्दछ। अर्थात्, बहुलक भनेको दिइएको आँकडामा सबैभन्दा बढी दोहोरिएको अङ्कमान हो। अर्को शब्दमा उच्चतम बारम्बारता भएको अङ्कमान नै बहुलक हो। यो पनि आँकडाको केन्द्रीय प्रवृत्तिको मापन गर्ने एक विधि हो। कहिलेकाहीँ कुनै दुई आँकडा सबैभन्दा बढी तर बराबर पटक दोहोरिएको पनि हुन सक्दछन्। यस्तो अवस्थामा द्वि-बहुलक (*Bimodal*) तथ्याङ्क भनिन्छ।

उदाहरण: कुनै पाँच जना व्यक्तिहरूको उचाइ क्रमशः १८०, १५०, १७०, १७०, १६० सेन्टिमिटर छ भने यो आँकडामा १७० सबैभन्दा धेरै (दुई) पटक दोहोरिएको छ।

२ National Planning Commission and Central Bureau of Statistics. *Population by Age Group*. National Planning Commission and Central Bureau of Statistics. Retrieved on December 03, 2021 from: <http://nationaldata.gov.np/>

३ साधारण वितरणसम्बन्धी थप जानकारी पाठ १५ मा दिइएको छ।

त्यसैले १७० बहुलक हो ।

अर्को उदाहरण हेरौं । नेपाल जीवनस्तर सर्वेक्षण २०१०/११ मा छनोटमा परेका परिवार मध्ये घरभाडा लगाउने ४८९ परिवार रहेछन् । तीमध्ये मासिक भाडा ६००० बुझ्ने सबै भन्दा धेरै ३४ परिवार रहेछन् । यस अवस्थामा बहुलक ६००० हुन्छ ।

खण्डित श्रेणी

आवृत्ति वितरणमा रहेको खण्डित श्रेणीको सबैभन्दा-ठुलो बारम्बारता हेरेर बहुलक पत्ता लगाउन सकिन्छ ।

उदाहरण:

तालिका ११.९: नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ मा छनोटमा परेका परिवारको परिवार सदस्य संख्या

दिइएको तथ्याङ्क (आवृत्ति वितरण)	
परिवार सदस्य संख्या (आँकडा)	जम्मा परिवार (बारम्बारता)
१	२०७
२	६४६
३	९४५
४	१,२५४
५	१,०८६
६	७९८
७	४५०
८	२५७
९	१३५
१०	८१
११	४९
१२	२५
१३	१७
१४	८
१५	१२
१६	४
१७	४
१८	२

१९	५
२०	३

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

यो तालिकामा सबैभन्दा-ठुलो बारम्बारता १२५४ हो र यससँग सम्बन्धित परिवार सदस्य संख्या ४ हो। तसर्थ, यो तथ्याङ्कको बहुलक ४ हो।

अविच्छिन्न श्रेणी

अविच्छिन्न श्रेणीमा प्रस्तुत गरिएका तथ्याङ्कको बहुलक निकाल्न सबैभन्दा पहिला ठुलो बारम्बारताको आधारमा बहुलक रहने समूह निकाल्नु पर्दछ। त्यसपछि बहुलक निकाल्न निम्न सूत्र प्रयोग गर्नुपर्दछ।

$$\text{बहुलक } (M_0) = l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times h$$

जहाँ, l , f_1 , f_0 , f_2 , र h भनेको क्रमशः बहुलक समूहको सानो संख्या वा तल्लो मान, बहुलक समूहको बारम्बारता, बहुलक समूहभन्दा अगाडिको समूहको बारम्बारता, बहुलक समूहभन्दा पछाडिको समूहको बारम्बारता र बहुलक समूहको अन्तराल हो।

उदाहरण: तलको तालिकाबाट बहुलक निकाल्नुहोस्।

तालिका ११.१०: नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ मा छनोटमा परेका परिवारको परिवार सदस्यहरूको उमेर समूहअनुसार आवृत्ति वितरण

उमेर समूह	संख्या (f)
० देखि १० वर्ष	६,३१०
१० देखि २० वर्ष	७,००९
२० देखि ३० वर्ष	४,२४१
३० देखि ४० वर्ष	३,४७८
४० देखि ५० वर्ष	२,७४५
५० देखि ६० वर्ष	२,३५६
६० देखि ७० वर्ष	१,५१९
७० देखि ८० वर्ष	७६८
८० देखि ९० वर्ष	२१७
९० देखि १०० वर्ष	२७

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

यहाँ, सबैभन्दा-ठुलो बारम्बारता ७००९ भएको हुनाले १० देखि २० वर्ष उमेर समूह नै बहुलक उमेर समूह हो ।

तसर्थ,

$l=१०$, $f_1=७००९$, $f_0=६३१०$, $f_2=४२४१$, र $h=१०$ हुने भयो । अब यी मानहरूलाई सूत्र (११) मा प्रतिस्थापन गर्दा,

$$M_0 = १० + \frac{७००९ - ६३१०}{२(७००९) - ६३१० - ४२४१} \times १०$$

$$M_0 = १० + \frac{६९९}{३४६७} \times १०$$

$$M_0 = १० + २.०२$$

तसर्थ, बहुलक = १२.०२ वर्ष हुन्छ । अर्थात्, छनोट भएका परिवारका सदस्यहरूमध्ये सबै भन्दा बढी १२ वर्षका थिए ।

११.४ मध्यक, मध्यिका र बहुलक बिचको सम्बन्ध

१) अंकगणितीय मध्यक (AM), ज्यामितीय मध्यक (GM) र हार्मोनिक मध्यक (HM) बिचको सम्बन्ध:

$$AM \geq GM \geq HM \quad (१२)$$

२) अंकगणितीय मध्यक (Mean), मध्यिका (Median) र बहुलक (Mode) बिचको सम्बन्ध:

क) सिमेट्रिकल (Symmetrical) वितरणका लागि,^४

$$\text{Mean} = \text{Mode} = \text{Median} \quad (१३)$$

ख) स्क्युड (Skewed or Non-Symmetrical) वितरणका लागि,

$$\text{Mode} = \text{Mean} - 3(\text{Mean} - \text{Median}) \quad (१४)$$

११.५ केन्द्रीय प्रवृत्तिको गुण, अवगुण तथा प्रयोगका क्षेत्र

केन्द्रीय प्रवृत्ति मापनका लागि हामीले अंकगणितीय मध्यक, ज्यामितीय मध्यक, हार्मोनिक मध्यक, मध्यिका, र बहुलक विभिन्न किसिमका तथ्याङ्कबाट कसरी निकाल्न सकिन्छ भन्ने बारेमा जानकारी प्राप्त गरिसक्यौं । यी सबै औसतहरू सबै प्रकारका तथ्याङ्कलाई उपयोगी नहुन पनि सक्दछन् । यिनीहरूको आ-आफ्नै सुन्दर पक्ष र सीमाहरू छन् । तसर्थ, यिनीहरूको प्रयोग गर्नुपूर्व यिनीहरूको गुण, अवगुण र प्रयोग गर्न सकिने विषय वा क्षेत्रका बारेमा जानकारी लिनुपर्दछ । तलको तालिका ११.८ मा यिनै केन्द्रीय

प्रवृत्तिका प्रमुख विशेषताहरू प्रस्तुत गरिएको छ ।

तालिका ११.११: केन्द्रीय प्रवृत्ति मापनका प्रमुख विशेषताहरू

क्र.सं.	औसत	गुण	अवगुण	प्रयोगका क्षेत्र
१	अंकगणितीय मध्यक	<ul style="list-style-type: none"> स्पष्ट रूपमा परिभाषित छ, निकाल्न र बुझ्न सजिलो छ, तथ्याङ्कका सबै आँकडामा आधारित छ, अन्य गणितीय क्रियाका लागि पनि उपयोगी छ, नमुना छनोटको फेरबदलले कम प्रभावित हुन्छ । 	<ul style="list-style-type: none"> तथ्याङ्कमा भएको अति ठुलो संख्याको आँकडा वा अति सानो आँकडाले निकै प्रभावित हुन्छ, खुला मान भएका समूहीकृत (Open End Class) तथ्याङ्कका लागि यो मध्यक निकाल्न सकिँदैन, आँकडा हेरेर वा ग्राफको सहायताले यसलाई निकाल्न सकिँदैन, गुणमापक विशेषताहरूलाई यो मध्यक प्रयोग गर्न मिल्दैन जस्तै: इमान्दारिता, सुन्दरता आदिको अंकगणितीय मध्यक हुँदैन, एउटा मात्र आँकडाको अनुपस्थितिमा पनि यो मध्यक निकाल्न सकिँदैन, एकरूपता नभएको श्रेणीका (Heterogeneous) लागि यो मध्यक उपयुक्त मानिँदैन, एकातर्फ अति नै भुकेको (Extremely Skewed) तथ्याङ्क वितरणमा यो मध्यक उपयुक्त मानिँदैन, खाण्डित आँकडाको मध्यक दशमलवमा आउन सक्दछ । 	<ul style="list-style-type: none"> अनुपात तहमा मापन गरिएका अधिकांश तथ्याङ्कमा यसको प्रयोग हुन्छ । जस्तै: औसत मासिक विप्रेषण, औसत उमेर आदि ।

4 सिमेट्रिकल भनेको सिङ्गो चिजलाई कुनै लाइनले उस्तै आकारका दुई भागमा विभाजन गर्न सक्ने भन्ने बुझाउँछ । जस्तै: मानिसको टाउकोदेखि एउटा सिधा रेखा ठाडो तान्ने हो भने उस्तै बराबर दुई भागमा विभाजन हुन्छ: कान एउटा एउटा, आँखा एउटा एउटा, मुख आधाआधा तर उस्तै, हात एउटा एउटा, । तथ्याङ्कको वितरणमा पनि सिमेट्रिकल हुनका लागि आँकडाहरू यसरी छरिनु पर्यो कि कुनै काल्पनिक रेखाले तिनीहरूको वितरणलाई बराबर र उस्तै दुई भागमा विभाजन गरोस् । चित्रको सहायताले हामी यस्तो वितरण साधारण वितरण वा नर्मल डिस्ट्रिब्युशनमा देख्दछौं, जुन हामी पाठ १५ मा अध्ययन गर्नेछौं ।

२ ज्यामितीय मध्यक	<ul style="list-style-type: none"> • स्पष्ट रूपमा परिभाषित छ, • तथ्याङ्कका सबै आँकडामा आधारित छ, • अन्य गणितीय क्रियाका लागि पनि उपयोगी छ, • तथ्याङ्कमा भएको अति ठुलो संख्याको आँकडा वा अति सानो आँकडाले कम प्रभावित हुन्छ, • नमुना छनोटको फेरबदलले कम प्रभावित हुन्छ । 	<ul style="list-style-type: none"> • निकाल्न र बुझ्न तुलनात्मक रूपले गाह्रो छ, • एउटा मात्र आँकडाको अनुपस्थितिमा पनि यो मध्यक निकाल्न सकिँदैन, • कुनै एउटा आँकडा शून्य वा नकारात्मक छ भने यो मध्यक निकाल्न सकिँदैन, • खुला मान भएका समूहीकृत (Open End Class) तथ्याङ्कका लागि यो मध्यक निकाल्न सकिँदैन, • आँकडा हेरेर वा ग्राफको सहायताले यसलाई यकिनका साथ भन्न वा निकाल्न सकिँदैन । 	<ul style="list-style-type: none"> • अनुपात, दर तथा प्रतिशतहरूको मध्यक निकाल्न यो मध्यकको प्रयोग गरिन्छ, • सूचकाङ्क निर्माणमा या मध्यक प्रयोग गरिन्छ, साना संख्यालाई बढी र ठुला संख्यालाई कम महत्त्व दिनुपर्ने अवस्थामा यो मध्यक प्रयोग गर्न सकिन्छ । जस्तै: जनसंख्या वृद्धिदरको औसत, शिक्षक विद्यार्थी अनुपातको औसत, प्राप्ताङ्क प्रतिशतको औसत आदि ।
३ हार्मोनिक मध्यक	<ul style="list-style-type: none"> • स्पष्ट रूपमा परिभाषित छ, • तथ्याङ्कका सबै आँकडामा आधारित छ, • अन्य गणितीय क्रियाका लागि पनि उपयोगी छ, • तथ्याङ्कमा भएको अति ठुलो संख्याको आँकडा वा अति सानो आँकडाले कम प्रभावित हुन्छ, • नमुना छनोटको फेरबदलले कम प्रभावित हुन्छ । 	<ul style="list-style-type: none"> • निकाल्न र बुझ्न तुलनात्मक रूपले गाह्रो छ, • एउटा मात्र आँकडाको अनुपस्थितिमा पनि यो मध्यक निकाल्न सकिँदैन, • कुनै एउटा आँकडा शून्य छ भने यो मध्यक निकाल्न सकिँदैन, • खुला मान भएका समूहीकृत (Open End Class) तथ्याङ्कका लागि यो मध्यक निकाल्न सकिँदैन, • आँकडा हेरेर वा ग्राफको सहायताले यसलाई यकिनका साथ भन्न वा निकाल्न सकिँदैन । 	<ul style="list-style-type: none"> • समय चरको रूपमा रहने विशेष प्रकारका अनुपात वा दर जस्तै: वेग, गति आदिको औसत निकाल्न उपयुक्त हुन्छ ।

४ मध्यिका	<ul style="list-style-type: none"> • स्पष्ट रूपमा परिभाषित छ, • निकाल्न र बुझ्न निकै सजिलो छ, • तथ्याङ्कमा भएको अति ठुलो संख्याको आँकडा वा अति सानो आँकडाले प्रभावित हुँदैन, • खुला मान भएका समूहीकृत (Open End Class) तथ्याङ्कका लागि पनि यो मध्यक निकाल्न सकिन्छ, • आँकडा हेरेर वा ग्राफको सहायताले यसलाई कहिले काहीँ भन्न वा निकाल्न सकिन्छ । 	<ul style="list-style-type: none"> • तथ्याङ्कका सबै आँकडामा आधारित छैन, • अन्य गणितीय क्रियाका लागि खासै उपयोगी छैन, • जोर संख्याका आँकडाहरूको मध्यका फेरि अंकगणितीय मध्यकमा निर्भर गर्दछ, • नमुना छनोटको फेरबदलले तुलनात्मक रूपमा बढी प्रभावित हुन्छ । 	<ul style="list-style-type: none"> • गुणमापक विशेषताहरूलाई यो मध्यक प्रयोग गर्न सकिन्छ । जस्तै: इमान्दारी, सुन्दरता आदिको मध्यिका निकालेर अध्ययन गर्न सकिन्छ । • क्रमबद्ध तहमा मापन गरिएका तथ्याङ्कको लागि मध्यिका उपयोगी हुन्छ ।
५ बहुलक	<ul style="list-style-type: none"> • निकाल्न र बुझ्न निकै सजिलो छ, • तथ्याङ्कमा भएको अति ठुलो संख्याको आँकडा वा अति सानो आँकडाले प्रभावित हुँदैन, • खुला मान भएका समूहीकृत (Open End Class) तथ्याङ्कका लागि पनि यो मध्यक निकाल्न सकिन्छ, 	<ul style="list-style-type: none"> • स्पष्ट रूपमा परिभाषित छैन, • तथ्याङ्कका सबै आँकडामा आधारित छैन, • अन्य गणितीय क्रियाका लागि खासै उपयोगी छैन, • अन्य गणितीय क्रियाका लागि खासै उपयोगी छैन, • साधारण वितरण नभएका तथ्याङ्कको बहुलक निकाल्न भन्झटिलो हुन्छ, • नमुना छनोटको फेरबदलले तुलनात्मक रूपमा बढी प्रभावित हुन्छ । 	<ul style="list-style-type: none"> • बजार व्यवस्थापनमा सबैभन्दा लोकप्रिय वस्तु, सेवा सामान पत्ता लगाउन बहुलकको प्रयोग गरिन्छ । तसर्थ, यसका लगाउन बहुलकको प्रयोग गरिन्छ । तसर्थ, यसका प्रयोग बजार, व्यापार, व्यवसाय तथा प्रतिष्ठानहरूमा बढी हुन्छ ।

• आँकडा हेरेर
वा ग्राफको
सहायताले
यसलाई कहिले
काहीं भन्न
वा निकाल्न
सकिन्छ ।

जस्तै: बढी बिक्री
हुने जुत्ताको
आकार, बढी
प्रयोग हुने पत्रिका,
बढी हेरिने टि.भी.
च्यानल आदि ।

११.६ केन्द्रीय प्रवृत्ति मापनको छनोट

प्रायजसो तथ्याङ्क विश्लेषण गर्दा केन्द्रीय प्रवृत्तिको प्रयोग गर्ने गरिन्छ । केन्द्रीय प्रवृत्तिले समग्र आँकडालाई प्रतिनिधित्व गर्ने हुनाले यसलाई हेरेर नीति, नियम तथा योजना बनाउने गरिन्छ । तथ्याङ्क विश्लेषणमा अन्य विधिहरू प्रयोग गर्दा समेत यसको प्रयोग हुने गर्दछ । यसरी वृहत् रूपमा प्रयोगको दायरा भएको केन्द्रीय प्रवृत्ति मापनको छनोट गर्दा भने विशेष ध्यान दिनुपर्दछ । केन्द्रीय प्रवृत्तिले वास्तवमा आँकडाहरूको केन्द्रीकृत मान पत्ता लगाउन मद्दत गर्दछ । तथ्याङ्कको प्रकृति र अध्ययनको विषय वा क्षेत्रअनुसार सही केन्द्रीय प्रवृत्तिको चयन गर्न सकिएन भने अपेक्षित उद्देश्य प्राप्त नहुन सक्दछ । माथि तालिका ११.११ मा प्रमुख केन्द्रीय प्रवृत्तिका प्रयोगका क्षेत्रका बारेमा उल्लेख गरिएको छ । तसर्थ, ती क्षेत्रअनुसार अनुकूल हुने केन्द्रीय प्रवृत्तिको छनोट गर्नुपर्दछ । अन्यथा, कहिलेकाहीं गलत केन्द्रीय प्रवृत्तिको प्रयोग गरियो भने अर्थहीन नतिजामा पनि पुग्न सकिन्छ । जस्तै: महिला र पुरुषको आँकडाको अंकगणितीय मध्यक निकालियो भने, त्यसलाई के भन्ने ?

आँकडामा एकदमै धेरै विचलन भएको बेलामा अथवा एकदमै सानो वा एकदमै ठुलो अङ्क भएर अङ्कगणितीय मध्यक प्रयोग गर्न नमिल्ने अवस्थामा ज्यामितीय मध्यकको प्रयोग गर्नुपर्ने हुन्छ । समय चर भएका अनुपात वा दरका लागि हार्मोनिक मध्यक उपयुक्त हुन्छ भने अन्य दर, अनुपात तथा प्रतिशतहरूलाई ज्यामितीय मध्यक उपयुक्त हुन्छ । त्यसैगरी गुणसँग सम्बन्धित तथ्याङ्कको विश्लेषणमा मध्यिका प्रयोग गर्न सकिन्छ भने बजार व्यवस्थापनसँग सम्बन्धित लोकप्रिय वस्तु वा सेवा पत्ता लगाउन बहुलक प्रयोग गर्न सकिन्छ । तथ्याङ्कको कुनै आँकडा नभएको अवस्थामा, कुनै आँकडा शून्य भएको अवस्थामा, कुनै आँकडा नकारात्मक भएको अवस्थामा कुन केन्द्रीय प्रवृत्ति छनोट गर्ने भन्ने बारेमा विशेष ध्यान दिनुपर्दछ । केन्द्रीय प्रवृत्तिका गुण तथा अवगुणहरूको अध्ययनबाट कुन औसत कतिबेला छनोट गर्ने भन्ने बारेमा लगभग जानकारी प्राप्त भइसकेको छ । यद्यपि कुशल तथ्याङ्कीय विश्लेषणका लागि केन्द्रीय प्रवृत्ति छनोट निम्न कुराहरूमा निर्भर गर्दछ:

१) तथ्याङ्कको प्रकृति र उपलब्धता,

- २) तथ्याङ्क विश्लेषणमा संलग्न गरिएका चरहरूको प्रकृति,
- ३) अध्ययनको उद्देश्य,
- ४) तथ्याङ्क वर्गीकरणको प्रणाली,
- ५) थप तथ्याङ्कीय गणनाको लागि आवश्यक केन्द्रीय प्रवृत्ति हो वा होइन ।

असल केन्द्रीय प्रवृत्तिका गुणहरू, प्रत्येक केन्द्रीय प्रवृत्तिका गुण, अवगुण तथा प्रयोगका क्षेत्र, तथा छनोटका आधार समेतको अध्ययनबाट अंकगणितीय मध्यक अरूभन्दा उत्तम पाइन्छ र यसैको प्रयोग धेरै भएको पाइन्छ । तसर्थ, यसलाई सर्वोत्कृष्ट औसत भनिन्छ ।

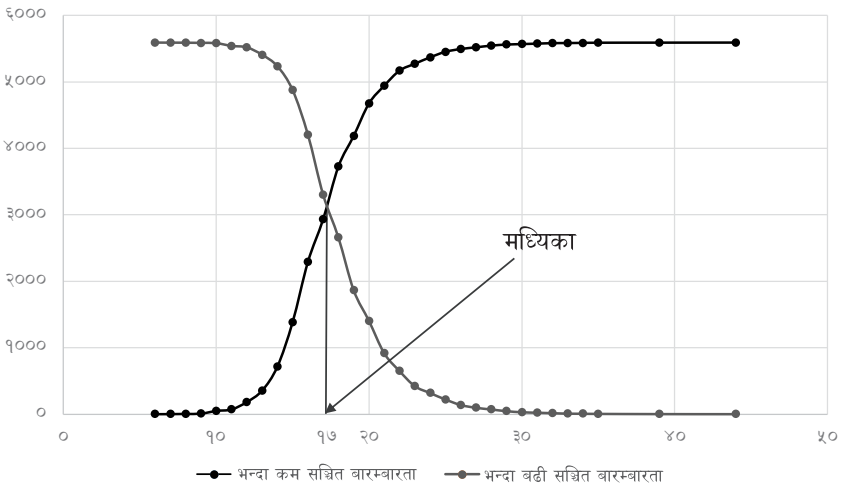
११.७ मध्यिका र बहुलकको ग्राफिकल मापन

केन्द्रीय प्रवृत्ति मापनका विभिन्न विधिहरूमध्ये मध्यिका र बहुलक पत्ता लगाउन तथ्याङ्कीय वा गणितीय गणनाका अतिरिक्त ग्राफको प्रयोग पनि गर्न सकिन्छ । मध्यिका एक किसिमको स्थानविशेष मान (Positional Value) हो र बहुलक सबैभन्दा बढी दोहोरिने मान हो । तसर्थ, यिनीहरूलाई ग्राफको सहायताले पत्ता लगाउनु सकिन्छ ।

मध्यिका

सञ्चित बारम्बारता बक्ररेखा (Cumulative Frequency Curve or Ogive) को सहायताले मध्यिका पहिचान गर्न सकिन्छ । भन्दा-साना ओजाइभ र भन्दा-ठुलो ओजाइभ बक्ररेखा खिचेर तिनीहरू भेट भएको बिन्दुबाट नर्मल रेखा X अक्षमा खिचियो भने, उक्त नर्मल रेखाले X अक्षमा छुने बिन्दु नै मध्यिका हुन्छ ।

चित्र नं. ११.२: पहिलो पटक विवाह गर्दाको उमेर अनुसारको महिलाको संख्याको भन्दा-सानो र भन्दा-ठुलो ओजाइभ बक्ररेखाहरू



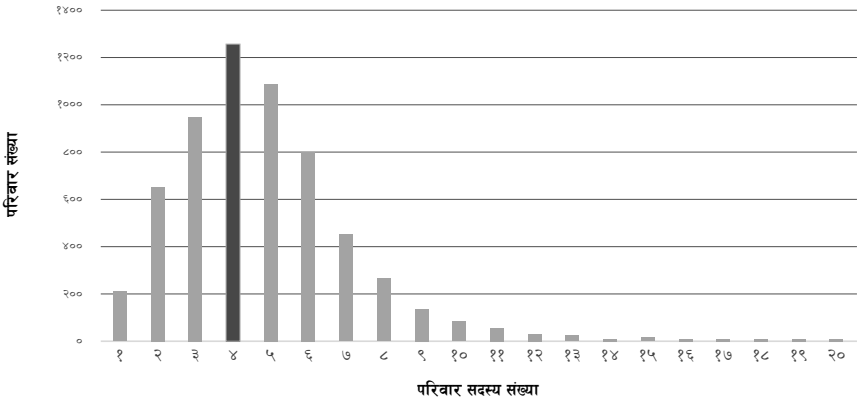
माथि चित्र ११.२ मा बिन्दु १७ ले मध्यिका प्रतिनिधित्व गरेको छ ।

नोट: ओजाइभ बक्ररेखाहरू खिचनका लागि पहिला “भन्दा-कम सञ्चित बारम्बारता (Less Than Cumulative Frequency)” र “भन्दा बढी सञ्चित बारम्बारता (More Than Cumulative Frequency)” निकाल्नु पर्दछ । बारम्बारतालाई माथिबाट तलसम्म जोड्दै गयो भने “भन्दा-कम सञ्चित बारम्बारता” निस्कन्छ भने तलबाट माथिसम्म जोड्दै गयो भने “भन्दा-बढी सञ्चित बारम्बारता” निस्कन्छ ।

बहुलक

बहुलक पत्ता लगाउनका लागि बारम्बारताको साधारण स्तम्भचित्र प्रयोग गर्न सकिन्छ ।

चित्र नं. ११.३: परिवारको आकार (बारम्बारताको स्तम्भचित्र)



यसैगरी अविच्छिन्न श्रेणीका लागि पनि ओजाइभ बक्ररेखा र हिस्टोग्राममार्फत् क्रमशः मध्यिका र बहुलक निकाल्न सकिन्छ । यस पुस्तकको दायराभित्र यसलाई समावेश गरिएको छैन ।

११.८ औसतका सीमा

बहुउपयोगी हुँदाहुँदै पनि औसतका केही सीमाहरू छन् । केही प्रमुख सीमाहरू निम्नानुसार रहेका छन् ।

१) औसत एउटा संख्याले सम्पूर्ण तथ्याङ्कको बारेमा प्रतिनिधित्व गर्नुपर्ने भएकोले यसको प्रयोग गर्दा निकै ध्यान दिनुपर्ने हुन्छ । अन्यथा अर्थको अनर्थ लाग्न सक्दछ । उदाहरणका लागि एउटा खोला तर्ने पुरानो कथा सान्दर्भिक हुन्छ । पाँच जना सदस्य भएको एउटा परिवार ४ फिट गहिरो खोला तर्नका लागि सदस्यहरूको औसत उचाइ निकाल्छन् । एक जना ३ फिट, एक जना ४ फिट, एकजना ५ फिट र बाँकी २ जना ६ फिट उचाइ भएका सदस्यहरूले औसत उचाइ ४.८ फिट निकाले र खोला तर्न मिल्ने निष्कर्ष पनि

निकाले । तर बिडम्बना ३ फिट र ४ फिट हुने सदस्यले खोला पार गर्न सकेनन् । यस कथाबाट के स्पष्ट हुन्छ भने औसतको व्याख्या र प्रयोग पर्याप्त सोच विचारविना गरियो भने यो खतरनाक पनि हुन सक्दछ ।

२) औसतको छनोट सही हुन सकेन भने गलत निर्णयमा पुग्न सकिन्छ ।

३) औसतले पूरै तथ्याङ्क वितरणको चित्र प्रस्तुत गर्न सक्दैन ।

४) औसतले खण्डित तथ्याङ्कको दशमलवमा मान दिन सक्दछ, जुन हास्यास्पद हुन्छ । जस्तै औसत परिवार आकार ४.६ भन्नुको तात्विक अर्थ हुँदैन । अर्थात् ४.६ जना भनेर के बुझ्ने, अलमल हुन्छ ।

११.९ विभाजक गर्ने मान

सानोदेखि ठुलो वा ठुलोदेखि सानोसम्म क्रमबद्ध रूपमा मिलाएर राखिएका तथ्याङ्क श्रेणीलाई निश्चित संख्याका बराबर भागमा विभाजन गर्ने मानहरूलाई विभाजन गर्ने मान (Partition Values) भनिन्छ । माथि छलफल गरिएको मध्यिका पनि एक प्रकारको विभाजन गर्ने मान हो किनभने यसले दिइएको तथ्याङ्कको श्रेणीलाई बराबर दुई भागमा विभाजन गर्दछ । त्यसैगरी दिइएको श्रेणीलाई बराबर चार भागमा विभाजन गर्ने मानलाई चतुर्थांश (Quartiles), बराबर दश भागमा विभाजन गर्ने मानलाई दशमक (Deciles) र बराबर सय भागमा विभाजन गर्ने मानलाई प्रतिशतक (Percentiles) भनिन्छ । यस खण्डमा हामीले यिनै विभाजन गर्ने मानहरूका बारेमा संक्षिप्त जानकारी हासिल गर्नेछौं ।

११.९.१ चतुर्थांश (Quartile)

सानोदेखि ठुलो वा ठुलोदेखि सानोसम्म क्रमबद्ध रूपमा मिलाएर राखिएका तथ्याङ्क श्रेणीलाई बराबर चार भागमा विभाजन गर्ने मानहरूलाई चतुर्थांश भनिन्छ । यसरी विभाजित प्रत्येक भागमा २५% आँकडाहरू हुन्छन् । चतुर्थांशका तिनवटा मानहरू हुन्छन् । तिनवटा चतुर्थांशहरूलाई Q1, Q2 र Q3 ले जनाइन्छ । जसमध्ये Q2 भनेको मध्यिका हो ।

तालिका ११.१२: विभिन्न स्थानको वायु प्रदूषण सूचकाङ्क

विराटनगर	रत्नपार्क	जनकपुर	कीर्तिपुर	शंखपार्क	धनगढी	सिमरा	धनकुटा
२३५	१७२	१६४	१५२	१५१	९७	७८	५९
२५%		२५%		२५%		२५%	
Q1			Q2			Q3	

स्रोत: वायु गुणस्तर निगरानी, २०७८ मंसिर २९ गते बिहान ५ बजे^५

मध्यिका निकाल्दा जसरी नै आँकडाको प्रकृति अनुसार चतुर्थांश निकाल्ने तरिका फरक हुन्छ।

व्यक्तिगत श्रेणी (Individual Series)

कुनै पनि व्यक्तिगत श्रेणीमा भएको आँकडा अर्थात् सम्पूर्ण आँकडालाई जस्ताको त्यस्तै राखिएको (कुनै बारम्बारता वितरणमा नराखिइकन) अवस्थामा आँकडालाई बढ्दो वा घट्दो क्रममा राख्ने, त्यसपछि तलका सूत्रहरू प्रयोग गर्न सकिन्छ।

पहिलो चतुर्थांश $(Q1) = \left(\frac{n+1}{4}\right)$ औँ स्थानको मान

माथि तालिका ११.१२ को लागि $(Q1) = \left(\frac{5+1}{4}\right)$ औँ स्थानको मान
 $= २.२५$ औँ स्थानको मान

यसका लागि हामीले दोस्रो र तेस्रो स्थानको मानको मध्यक लिनुपर्ने हुन्छ। अर्थात्,

$$Q1 = \left(\frac{१७२+१६४}{२}\right) = १६८$$

दोस्रो चतुर्थांश $(Q2) = 2\left(\frac{n+1}{4}\right) = \frac{n+1}{२}$ औँ स्थानको मान अर्थात् मध्यिका

$$= \left(\frac{5+1}{२}\right) \text{ औँ स्थानको मान}$$

$$= ४.५ \text{ औँ स्थानको मान}$$

$$Q2 = \left(\frac{१५२+१५१}{२}\right) = १५१.५$$

तेस्रो चतुर्थांश $(Q3) = ३\left(\frac{n+1}{4}\right)$ औँ स्थानको मान

$$= ३\frac{5+1}{४} \text{ औँ स्थानको मान}$$

$$= ६.७५ \text{ औँ स्थानको मान}$$

5 Department of Environemnt. Air Quality Monitoring. Retrieved on December 15, 2021 from: <https://pollution.gov.np/portal/#/home; k=89ij52>

$$Q3 = \left(\frac{97+98}{2} \right) = 97.5$$

तसर्थ, यहाँ १६८, १५१.५, ८७.५ ले दिइएको आँकडालाई बराबर ४ भागमा विभाजन गरेका छन् ।

खण्डित श्रेणी (Discrete Series)

खण्डित श्रेणीका लागि चतुर्थांश निकाल्दा निम्नानुसार गर्नुपर्दछ । यसका लागि एउटा उदाहरणसहित व्याख्या गरिएको छ ।

उदाहरण: नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ को आधारमा परिवार सदस्य संख्याको तथ्याङ्क तालिका ११.१३ मा दिइएको छ ।

तालिका ११.१३: नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ मा छनोटमा परेका परिवारको परिवार सदस्य संख्या

परिवार सदस्य संख्या (आँकडा)	जम्मा परिवार (बारम्बारता)	भन्दासानो सञ्चित बारम्बारता
१	२०७	२०७
२	६४६	८५३
३	९४५	१,७९८
४	१,२५४	३,०५२
५	१,०८६	४,१३८
६	७९८	४,९३६
७	४५०	५,३८६
८	२५७	५,६४३
९	१३५	५,७७८
१०	८१	५,८५९
११	४९	५,९०८
१२	२५	५,९३३
१३	१७	५,९५०
१४	८	५,९५८
१५	१२	५,९७०
१६	४	५,९७४
१७	४	५,९७८

१८	२	५,९८०
१९	५	५,९८५
२०	३	५,९८८

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

क) बारम्बारताको तालिकामा माथिदेखि तलसम्म बारम्बारतालाई जोड्दै सञ्चित बारम्बारता (Less Than Cumulative Frequency) निकाल्ने । यो कार्य तालिका ११.१३ को अन्तिम महलमा गरिएको छ ।

ख) पहिलो चतुर्थांश (Q_1) = $\left(\frac{n}{4}\right) = \left(\frac{५,९८८}{४}\right) = १,४९७$ भन्दा नजिकको ठुलो मान सञ्चित बारम्बारतामा हेर्ने । त्यो भनेको १,७९८ हो । फेरि १,७९८ सँग सम्बन्धित परिवार सदस्य संख्या मान तालिकाबाट हेर्ने । त्यो भनेको ३ हो र यो नै पहिलो चतुर्थांश हो ।

ग) दोस्रो चतुर्थांश (Q_2) = $२\left(\frac{n}{४}\right) = २\left(\frac{५,९८८}{४}\right) =$ भन्दा नजिकको ठुलो मान सञ्चित बारम्बारतामा हेर्ने ।

त्यो भनेको ३,०५२ हो । फेरि ३,०५२ सँग सम्बन्धित परिवार सदस्य संख्या मान तालिकाबाट हेर्ने । त्यो भनेको ४ हो र यो नै दोस्रो चतुर्थांश हो । यो मान दिइएको तथ्याङ्कको मध्यिका पनि हो (तालिका ११.७ मा मध्यिका निकालिएको विधि पुनः हेर्नुहोस्) ।

घ) तेस्रो चतुर्थांश (Q_3) = $३\left(\frac{n}{४}\right) = ३\left(\frac{५,९८८}{४}\right) =$ भन्दा नजिकको ठुलो मान सञ्चित बारम्बारतामा हेर्ने ।

त्यो भनेको ४,९३६ हो । फेरि ४,९३६ सँग सम्बन्धित परिवार सदस्य संख्या मान तालिकाबाट हेर्ने । त्यो भनेको ६ हो र यो नै तेस्रो चतुर्थांश हो ।

ङ) तसर्थ, माथिको आँकडाको चतुर्थांशहरू ३, ४ र ६ हुन् ।

अविच्छिन्न श्रेणी (Continuous Series)

अविच्छिन्न श्रेणीका लागि चतुर्थांश निकाल्दा निम्नानुसार गर्नुपर्दछ । यसका लागि एउटा उदाहरणसहित व्याख्या गरिएको छ ।

उदाहरण: नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ को आधारमा परिवार सदस्य संख्याको तथ्याङ्क तालिका ११.१३ मा दिइएको छ ।

तालिका ११.१४: नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ मा छनोटमा परेका परिवारको परिवार सदस्यहरूको उमेर समूहअनुसार आवृत्ति वितरण

उमेर समूह	संख्या (f)	भन्दा-कम सञ्चित बारम्बारता निकाल
० देखि १० वर्ष	६,३१०	६,३१०
१० देखि २० वर्ष	७,००९	१३,३१९
२० देखि ३० वर्ष	४,२४१	१७,५६०
३० देखि ४० वर्ष	३,४७८	२१,०३८
४० देखि ५० वर्ष	२,७४५	२३,७८३
५० देखि ६० वर्ष	२,३५६	२६,१३९
६० देखि ७० वर्ष	१,५१९	२७,६५८
७० देखि ८० वर्ष	७६८	२८,४२६
८० देखि ९० वर्ष	२१७	२८,६४३
९० देखि १०० वर्ष	२७	२८,६७०
जम्मा	N = २८,६७०	

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

क) बारम्बारताको तालिकामा माथिदेखि तलसम्म बारम्बारतालाई जोड्दै सञ्चित बारम्बारता (Less Than Cumulative Frequency) निकाल्ने । यो कार्य तालिका ११.१४ को अन्तिम महलमा गरिएको छ ।

ख) पहिलो चतुर्थांश (Q1) पर्ने समूह पत्ता लगाउने

ग) दोस्रो चतुर्थांश (Q2) = $\left(\frac{n}{2}\right) = \left(\frac{२८,६७०}{२}\right) = १४,३३५$ हुन्छ । अब, १४,३३५.५ भन्दा नजिकको ठुलो मान सञ्चित बारम्बारतामा हेर्ने । यो मान १३,३१९ हो र यससँग सम्बन्धित समूह १० देखि २० वर्ष हो । तसर्थ, पहिलो चतुर्थांश पर्ने समूहको तल्लो मान (L) = १०, पहिलो चतुर्थांश पर्ने समूहको बारम्बारता (f) = ७,००९, पहिलो चतुर्थांश पर्ने समूहभन्दा अघिल्लो समूहको Cf (Cf) = ६,३१०, पहिलो चतुर्थांश पर्ने समूहको अन्तराल (h) = १०

$$\text{पहिलो चतुर्थांश (Q1)} = L + \frac{\frac{n}{4} - Cf}{f} \times h$$

$$= 90 + \frac{\frac{25,670}{4} - 6390}{7009} \times 90 \quad (१५)$$

$$Q1 = 11.22$$

घ) दोस्रो चतुर्थांश (Q2) पर्ने समूह पत्ता लगाउने

$$\text{ड) दोस्रो चतुर्थांश पर्ने समूह पत्ता लगाउन } 2\left(\frac{n}{4}\right) = \left(\frac{25,670}{2}\right) = 12,835 \text{ हुन्छ।}$$

अब, १४,३३५ भन्दा नजिकको ठुलो मान सञ्चित बारम्बारतामा हेर्ने। यो मान १७,५६० हो र यससँग सम्बन्धित समूह २० देखि ३० वर्ष हो। तसर्थ, दोस्रो चतुर्थांश पर्ने समूहको तल्लो मान (L) = २०, दोस्रो चतुर्थांश पर्ने समूहको बारम्बारता (f) = ४,२४१, दोस्रो चतुर्थांश पर्ने समूहभन्दा अघिल्लो समूहको Cf (Cf) = १३,३१९, दोस्रो चतुर्थांश पर्ने समूहको अन्तराल (h) = १०

$$\text{दोस्रो चतुर्थांश (Q2)} = L + \frac{\frac{n}{2} - Cf}{f} \times h$$

$$= 20 + \frac{2 \frac{25,670}{4} - 13,319}{4,241} \times 10 \quad (१६)$$

च) Q2 = २२.३९ र यो मान भनेको मध्यिका पनि हो। माथि मध्यिका निकाल्ने उदाहरण तालिका ११.८ मा प्रस्तुत गरिएको तथ्याङ्क र यहाँ प्रस्तुत गरिएको तथ्याङ्क एउटै भएकोले तुलना गरेर हेर्न सकिन्छ।

छ) तेस्रो चतुर्थांश (Q3) पर्ने समूह पत्ता लगाउने

$$\text{ज) तेस्रो चतुर्थांश पर्ने समूह पत्ता लगाउन } 3\left(\frac{n}{4}\right) = 3\left(\frac{25,670}{4}\right) = 23,502 \text{ हुन्छ।}$$

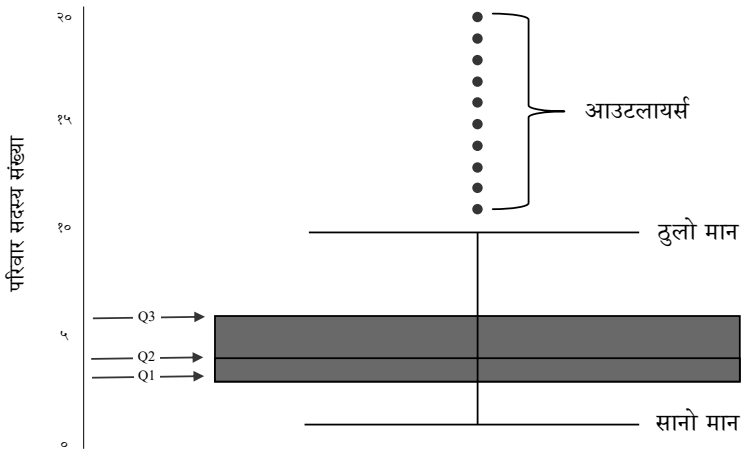
अब, २१,५०२ भन्दा नजिकको ठुलो मान सञ्चित बारम्बारतामा हेर्ने। यो मान २३,७८३ हो र यससँग सम्बन्धित समूह ४० देखि ५० वर्ष हो। तसर्थ, दोस्रो चतुर्थांश पर्ने समूहको तल्लो मान (L) = ४०, तेस्रो चतुर्थांश पर्ने समूहको बारम्बारता (f) = २७४५, तेस्रो चतुर्थांश पर्ने समूहभन्दा अघिल्लो समूहको Cf (Cf) = २१,०३८, तेस्रो चतुर्थांश पर्ने समूहको अन्तराल (h) = १०

$$\begin{aligned}
 \text{तेस्रो चतुर्थांश (Q3)} &= L + \frac{\frac{3n}{4} - Cf}{f} \times h \\
 &= 80 + \frac{3 \frac{28,690}{4} - 29,038}{2,985} \times 90 \\
 Q3 &= 81.69
 \end{aligned} \tag{१७}$$

उक्त आँकडाको चतुर्थांशहरूको मान ११.२२, २२.३९, ४१.६९ भयो ।

चतुर्थांशहरूलाई *Box Plot* को माध्यमबाट हेर्न सकिन्छ ।

चित्र नं. ११.४: परिवार सदस्य संख्याको बक्स प्लट



स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

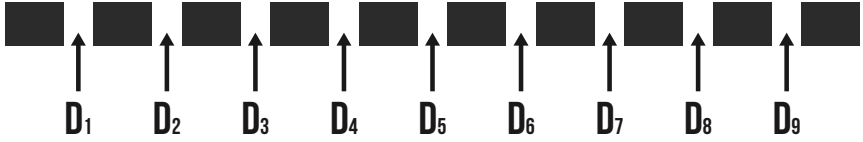
माथिको चित्र नं. ११.४ मा *STATA* को प्रयोग गरी नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ मा छानिएका परिवारको सदस्य संख्याहरूको बक्स प्लट बनाइएको छ । सफ्टवेयरले आउटलायर्स छाडेर अन्य मानको बक्स प्लट तयार गरेको छ । जसमा सबै भन्दा तुलो मान, सबैभन्दा सानो मान, पहिलो चतुर्थांश, दोस्रो चतुर्थांश, तेस्रो चतुर्थांश देखाइएको छ ।

११.९.२ दशमक (Decile)

सानोदेखि तुलो वा तुलोदेखि सानोसम्म क्रमबद्ध रूपमा मिलाएर राखिएका तथ्याङ्क श्रेणीलाई बराबर दश भागमा विभाजन गर्ने मानहरूलाई दशमक भनिन्छ । यसरी विभाजित प्रत्येक भागमा १०% आँकडाहरू हुन्छन् । दशमकका नौ वटा मानहरू हुन्छन् । नौ वटा चतुर्थांशहरूलाई {D1, D2, D3, D4, D5, D6, D7, D8, र D9} ले जनाइन्छ । जसमध्ये D5 भनेको मध्यिका हो ।

दशमक निकाल्ने विधि चतुर्थांश निकाल्ने विधिजस्तै हुन्छ । तसर्थ, यहाँ दशांश निकाल्ने सूत्रमात्र उल्लेख गरिएको छ ।

चित्र नं. ११.५: दशमकको नमुना



व्यक्तिगत श्रेणीको लागि

i^{th} दशमक (D_i) = $i \frac{n+1}{10}$ औँ मान, i भनेको १, २, ३, , ८, ९ हो । तसर्थ,

$$(D_1) = 1 \times \frac{n+1}{10} = \frac{n+1}{10} \text{ औँ मान,}$$

$$(D_2) = 2 \frac{n+1}{10} \text{ औँ मान,}$$

.....

$$(D_9) = 9 \frac{n+1}{10}$$

खण्डित श्रेणीको लागि

i^{th} दशमक प्रतिनिधित्व गर्ने आँकडा निकाल्नुपर्ने हुन्छ । त्यसका लागि सञ्चित बारम्बारताबाट $i \frac{n}{10}$ भन्दा नजिकको मान हेरेर त्यसको सम्बन्धित आँकडा i^{th} दशमक हुन्छ ।

अविच्छिन्न श्रेणीको लागि

i^{th} दशमक पर्ने समूह निकाल्नुपर्ने हुन्छ । त्यसका लागि सञ्चित बारम्बारताबाट $i \frac{n}{10}$ भन्दा नजिकको मान हेरेर त्यसको सम्बन्धित आँकडा i^{th} दशमक हुन्छ । i^{th} दशमकसँग सम्बन्धित समूह पत्ता लगाइसकेपछि निम्न सूत्रको प्रयोग गर्नुपर्दछ ।

$$i^{\text{th}} \text{ दशमक } (D_i) = L + \frac{i \frac{n}{10} - cf}{f} h, \tag{१८}$$

जहाँ, L भनेको i^{th} दशमक पर्ने समूहको तल्लो मान हो, f भनेको i^{th} दशमक पर्ने समूहको बारम्बारता हो, cf भनेको i^{th} दशमक पर्ने समूहभन्दा अघिल्लो समूहको सञ्चित बारम्बारता हो र h भनेको i^{th} दशमक पर्ने समूहको अन्तराल हो ।

११.९.३ प्रतिशतक (Percentile)

सानोदेखि ठुलो वा ठुलोदेखि सानोसम्म क्रमबद्ध रूपमा मिलाएर राखिएका तथ्याङ्क श्रेणीलाई बराबर सय भागमा विभाजन गर्ने मानहरूलाई प्रतिशतक भनिन्छ । यसरी विभाजित प्रत्येक भागमा १% आँकडाहरू हुन्छन् । प्रतिशतकका उनान्सय वटा मानहरू हुन्छन् । उनान्सय वटा प्रतिशतकहरूलाई P1, P2, P98, र P99 ले जनाइन्छ । जसमध्ये P50 भनेको मध्यिका हो ।

प्रतिशतक निकाल्ने विधि पनि चतुर्थांश निकाल्ने विधि जस्तै हुन्छ । तसर्थ, यहाँ प्रतिशतक निकाल्ने सूत्र मात्र उल्लेख गरिएको छ ।

i^{th} प्रतिशतक (Pi) = $i \frac{n+1}{100}$ आँ मान, i भनेको १, २, ३,, ९८, ९९ हो । तसर्थ,

$$(P1) = 1 \times \frac{n+1}{100} = \frac{n+1}{100} \text{ आँ मान,}$$

$$(P2) = 2 \frac{n+1}{100} \text{ आँ मान,}$$

.....

$$(P99) = 99 \frac{n+1}{100} \text{ आँ मान हुन्छ ।}$$

खण्डित श्रेणीको लागि

i^{th} प्रतिशतक प्रतिनिधित्व गर्ने आँकडा निकाल्नुपर्ने हुन्छ । त्यसका लागि सञ्चित बारम्बारताबाट $i \frac{n}{10}$ भन्दा नजिकको मान हेरेर त्यसको सम्बन्धित आँकडा i^{th} प्रतिशतक हुन्छ ।

अविच्छिन्न श्रेणीको लागि

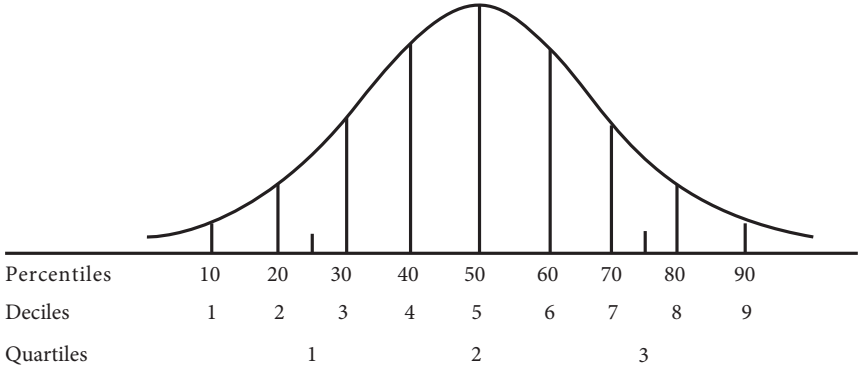
i^{th} प्रतिशतक पर्ने समूह निकाल्नुपर्ने हुन्छ । त्यसका लागि सञ्चित बारम्बारताबाट $i \frac{n}{10}$ भन्दा नजिकको मान हेरेर त्यसको सम्बन्धित आँकडा i^{th} प्रतिशतक हुन्छ । i^{th} प्रतिशतकसँग सम्बन्धित समूह पत्ता लगाइसकेपछि निम्न सूत्रको प्रयोग गर्नुपर्दछ ।

$$i^{\text{th}} \text{ दशमक (Di)} = L + \frac{i \frac{n}{10} - cf}{f} h, \quad (१९)$$

जहाँ, L भनेको i^{th} प्रतिशतक पर्ने समूहको तल्लो मान हो, f भनेको i^{th} प्रतिशतक पर्ने समूहको बारम्बारता हो, cf भनेको i^{th} प्रतिशतक पर्ने समूहभन्दा अधिल्लो समूहको सञ्चित बारम्बारता हो र h भनेको i^{th} प्रतिशतक पर्ने समूहको अन्तराल हो ।

चतुर्थांश, दशमक र प्रतिशतकको सम्बन्ध निम्न चित्र नं. ११.६ बाट हेर्न सकिन्छ ।

चित्र नं. ११.६: चतुर्थांश, दशमक र प्रतिशतकको सम्बन्ध



११.९.४ विभाजन गर्ने मानहरूको प्रयोग तथा महत्त्व

माथि उल्लेख गरिएका चतुर्थांश, दशमक, प्रतिशतकबाहेक अरू पनि विभाजन गर्ने मान हुन्छन् । जस्तै: कुनै तथ्याङ्कलाई बराबर पाँच भागमा विभाजन गर्ने मानलाई पञ्चमक (Quintiles) भनिन्छ । क्विनटाइल्स निकाल्नका लागि पनि चतुर्थांश, दशमक, प्रतिशतक निकाल्ने विधिकै अनुसरण गर्न सकिन्छ । यस्ता विभाजन गर्ने मानहरूले दिइएको तथ्याङ्कका मानहरूलाई निश्चित संख्यामा बराबर भागमा विभाजन गरेका हुन्छन् । यसबाट तथ्याङ्कलाई समूहीकृत गरेर अध्ययन गर्न सजिलो हुन्छ । उदाहरणका लागि नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ को प्रतिवेदनमा विभिन्न विभाजन गर्ने मान प्रयोग गरेर तथ्याङ्क विश्लेषण एवं व्याख्या गरिएको छ । तीमध्ये दशमक प्रयोग गरेको एउटा तालिका निम्नानुसार रहेको छ । यस तालिका ११.१५ बाट नेपालमा धनीवर्गले गर्ने खर्च र गरिबवर्गले गर्ने खर्चका बारेमा तुलनात्मक रूपले अध्ययन गर्न सकिन्छ । जस्तै सबैभन्दा धनी १०% परिवारले गरेको उपभोग खर्च सबैभन्दा गरिब १० % परिवारले गरेको उपभोग खर्चभन्दा ९ गुणा बढी देखिएको छ (तालिका ११.१५) ।

तालिका ११.१५: चालु मूल्यमा दशमकअनुसार प्रतिव्यक्ति उपभोग खर्च (रु)

दशमक वर्ग	औसत उपभोग खर्च (रु)	उपभोगमा हिस्सा
सबभन्दा गरिब (न्यूनतम)	११,०९३	३.२
दोस्रो	१५,२४१	४.४
तेस्रो	१७,८९३	५.१
चौथो	२०,७४४	६
पाँचौँ	२४,२३८	७
छैटौँ	२८,२७०	८.१

सातौँ	३३,१९६	९.५
आठौँ	४०,७३०	११.७
नवौँ	५४,२०३	१५.६
सबभन्दा धनी (अधिकतम)	१,०२,७७२	२९.५
नेपाल	३४,८२९	१००

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

नोट: साह्रै थोरै संख्याका आँकडाहरूको विभाजकले त्यति महत्त्व राख्दैन। जस्तै नेपालका सात प्रदेशहरूको जनसंख्याको प्रतिशतक वा दशमक मानले खासै महत्त्वपूर्ण भूमिका निर्वाह गर्दैन। यस उदाहरणमा जम्मा सातवटा मानहरू हुन्छन् जसलाई सय भागमा विभाजन गर्नुको खासै औचित्य हुँदैन।

११.१० अभ्यासका लागि प्रश्नहरू

क) तलका पदावलीहरूलाई छोटकरीमा व्याख्या गर्नुहोस्।

- १) अंकगणितीय मध्यक
- २) मध्यिका
- ३) बहुलक

ख) तलका तथ्याङ्कहरूको लागि उपयुक्त औसत छान्नुहोस्।

- १) एक वर्षको मासिक ब्याजदरहरूको औसत (अंकगणितीय मध्यक, ज्यामितीय मध्यक, हार्मोनिक मध्यक, मध्यिका, बहुलक)
- २) विद्यार्थीले परीक्षामा प्राप्त गरेको प्रतिशतको औसत (अंकगणितीय मध्यक, ज्यामितीय मध्यक, हार्मोनिक मध्यक, मध्यिका, बहुलक)
- ३) केन्द्रीय तथ्याङ्क विभागका कर्मचारीहरूको मासिक औसत तलब (अंकगणितीय मध्यक, ज्यामितीय मध्यक, हार्मोनिक मध्यक, मध्यिका, बहुलक)
- ४) मन्त्रालयहरूको विद्युत महसुलको औसत मासिक खर्च (अंकगणितीय मध्यक, ज्यामितीय मध्यक, हार्मोनिक मध्यक, मध्यिका, बहुलक)
- ५) सम्पत्तिका आधारमा देशका परिवारलाई बराबर दुई भाग लगाउने औसत (अंकगणितीय मध्यक, ज्यामितीय मध्यक, हार्मोनिक मध्यक, मध्यिका, बहुलक)

ग) आ.व. २०६८/६९ देखि २०७७/७८ सम्म शिक्षा क्षेत्रको मूल्य अभिवृद्धि वृद्धिदर क्रमशः ५.६१%, ५.८२%, ४.९५%, ५.४८%, ७.१५%, ७.२१%, ५.८३%, ५.९८%, ३.२०%, ३.६०% रहेको छ। यस आँकडाको आधारमा उपयुक्त औसत छनोट गरी विगत दश वर्षको औसत वृद्धिदर निकाल्नुहोस्।

घ) तालिका ११.९ मा नेपालका महिला र पुरूषले पहिलो पटक विवाह गर्दाको उमेर दिइएको छ। यस तालिकाबाट महिलाका लागि, पुरूषका लागि र जम्मा जनसंख्याका लागि अंकगणितीय मध्यक, मध्यिका र बहुलक छुट्टाछुट्टै निकालेर तुलना गर्नुहोस्। साथै प्राप्त नतिजाका बारेमा तथ्याङ्क विश्लेषण गरी एक अनुच्छेद लेख्नुहोस्।

तालिका ११.१६: पहिलो पटक विवाह गर्दाको उमेर

उमेर समूह	महिला	पुरूष
० देखि १० वर्ष	१,१५,१५०	२२,८६५
१० देखि १५ वर्ष	११,०१,८८५	२,६१,२३२
१५ देखि २० वर्ष	४३,४०,४९७	२१,७६,७८४
२० देखि २५ वर्ष	१५,५१,३६४	२२,६६,९३२
२५ देखि ३० वर्ष	२,६६,५०५	८,५९,९२९
३० देखि ३५ वर्ष	५४,०६४	२,१९,६५१
३५ देखि ४० वर्ष	१५,५२७	५३,१२२
४० देखि ४५ वर्ष	५,५१०	१५,५०३
४५ देखि ५० वर्ष	२,००१	४,८८९
५० देखि ५५ वर्ष	१,६०६	३,४८०
जम्मा	७४,५४,१०९	५८,८४,३८७

स्रोत: केन्द्रीय तथ्याङ्क विभाग, राष्ट्रिय जनगणना, २०६८^६

नोट: केन्द्रीय तथ्याङ्क विभागको तथ्याङ्कमा सुरूको उमेर समूह र अन्तिम उमेर समूह खुला अन्त भएको समूह छन्। यहाँ अभ्यास गर्ने प्रयोजनका लागि बन्द अन्त भएको समूह बनाइएको छ।

११.११ सन्दर्भ सामग्री

केन्द्रीय प्रवृत्तिको मापन विस्तृतरूपमा अध्ययन गर्नका लागि तलको पुस्तक उपयोगी छ।

१. Fundamentals of Statistics by S.C. Gupta, Sixth Revised & Enlarged Edition, Himalaya Publishing House, 2009

6 नोट: केन्द्रीय तथ्याङ्क विभागको तथ्याङ्कमा सुरूको उमेर समूह र अन्तिम उमेर समूह खुला अन्त भएको समूह छन्। यहाँ अभ्यास गर्ने प्रयोजनका लागि बन्द अन्त भएको समूह बनाइएको छ।

विचलनको मापन

१२.१ परिचय र अर्थ

तथ्याङ्क वितरणमा केन्द्रीय मानका बारेमा औसतहरूले बताउँछन् । अर्थात्, आँकडाहरू कहाँनेर बढी केन्द्रित छन् भन्ने बारेमा केन्द्रीय प्रवृत्तिको अध्ययनबाट प्रष्ट हुन्छ तर औसतले मात्र तथ्याङ्कक वितरणको सम्पूर्ण विशेषता बताउन सक्दैनन् । यो केवल सम्पूर्ण आँकडाको सामान्य प्रतिनिधि पात्र मात्र हो किनभने एउटै औसत भएका दुई वा दुईभन्दा बढी तथ्याङ्क वितरणहरू हुन सक्दछन् । जस्तै: दुई जना विद्यार्थीले तिनवटा विषयमा क्रमशः ४०, ५०, ६० र ३०, ५०, ७० प्राप्त गरेका रहेछन् भने दुवै जनाको औसत ५० नै हुन्छ । तर दुई जनाले प्राप्त गरेको प्राप्ताङ्कहरूमा निकै भिन्नता छ । तसर्थ, तथ्याङ्कका विशेषताहरूलाई थप स्पष्ट पार्न, तथ्याङ्क विश्लेषणमा केन्द्रीय प्रवृत्तिलाई सहयोग गर्ने अर्को मापनको रूपमा विचलनको मापनलाई लिइन्छ । तथ्याङ्क विश्लेषणमा तथ्याङ्क वितरणमा हुने विचलनको अध्ययन निकै महत्त्वपूर्ण मानिन्छ । विचलनको अर्थ छरिनु, फरक हुनु वा भिन्नता हुनु वा समानता नहुनु हो । कुनै पनि तथ्याङ्कमा आँकडाहरू कतिको छरिएर रहेका छन् भनेर हेर्ने विधिलाई विचलनको मापन भनिन्छ । विचलनको मापनले आँकडाहरूको समानता वा विषमताको बारेमा बताउन सक्दछ । अर्को शब्दमा, आँकडाहरू एकापसमा कतिको भिन्न छन् वा छरिएका छन् भनेर विचलनको मापनद्वारा अध्ययन गरिन्छ । यस पाठमा हामी विचलन मापनका विभिन्न विधिहरू मध्ये धेरै प्रयोगमा आउने विस्तार, मध्य विचलन, स्तरीय भिन्नता र भेरियन्सका बारेमा चर्चा गर्ने छौं । यसका अतिरिक्त विचलन मापनका लागि उपयोगी रेखा चित्रात्मक अध्ययनका बारेमा पनि चर्चा गर्नेछौं ।

१२.२ असल विचलन मापनका गुणहरू

तथ्याङ्क विश्लेषणमा केही प्रतिनिधि मापनले सम्पूर्ण आँकडाको विशेषताका बारेमा धारणा बनाउनुपर्ने हुनाले त्यस्ता मापनहरूले सकेसम्म राम्रा गुणहरू बहन गरेको हुनुपर्दछ । असल केन्द्रीय प्रवृत्तिका गुणहरूका बारेमा हामीले पाठ ११ मा चर्चा गरिसकेका छौं । असल केन्द्रीय प्रवृत्तिका गुण जस्तै असल विचलनका केही प्रमुख गुणहरू निम्नानुसार छन् ।

- क) यो स्पष्ट रूपमा परिभाषित भएको हुनुपर्दछ,
- ख) दिइएको आँकडाहरूबाट सजिलै निकाल्न र बुझ्नका लागि सजिलो हुनुपर्दछ,
- ग) सबै आँकडाहरूमा आधारित भएर निकालिएको हुनुपर्दछ,

- घ) थप गणितीय क्रियाहरूमा उपयोगी हुनुपर्दछ,
 ङ) नमुना छनोटको फेरबदलले कमभन्दा कम प्रभावित हुनुपर्दछ,
 च) अत्यन्त ठुला वा साना आँकडाले प्रभावित पार्नु हुँदैन ।

सबै विचलनका मापनले यी सबै गुण बोकेका भने हुँदैनन् । विचलनको छनोट गर्दा सकेसम्म धेरै गुणहरू धारण गरेको विचलनको मापनलाई छान्नुपर्दछ ।

१२.३ विचलनको निरपेक्ष र सापेक्ष मापन

निरपेक्ष विचलनको मापन भनेको आँकडाको मौलिक एकाइमा प्रस्तुत गरिएका विचलनहरू हुन् । जस्तै: परिवारको मासिक खर्च रूपैयाँमा प्रस्तुत गरिएको छ भने उक्त तथ्याङ्कको विचलन मापन पनि रूपैयाँमा नै प्रस्तुत गरियो भने त्यसलाई निरपेक्ष विचलनको मापन भनिन्छ । तर प्रायजसो निरपेक्ष विचलनको मापन दुई वा दुईभन्दा बढी तथ्याङ्क वितरणको तुलनात्मक अध्ययनका लागि उपयुक्त मानिँदैन । निरपेक्ष विचलनले आँकडाको मौलिक एकाइमा विचलन प्रस्तुत गर्दछ जसलाई तुलना गर्न गाह्रो हुन्छ । जस्तै: के.जी.मा व्यक्त गरिएका तौलका तथ्याङ्कको विचलन के.जी. मा हुन्छ र फिटमा व्यक्त गरिएका उचाइका तथ्याङ्कको विचलन फिटमा हुन्छ, यी दुई के.जी. र फिटको तुलना गर्न सकिँदैन । तसर्थ, तुलनात्मक अध्ययनका लागि विचलनको सापेक्ष मापन प्रयोग गरिन्छ । विचलनको सापेक्ष मापनमा अनुपात वा प्रतिशतको प्रयोग गरी आँकडाको मौलिक एकाइलाई हटाइन्छ । यस पाठमा हामीले अध्ययन गरिने विचलनहरूको निरपेक्ष जस्तै: विस्तार, मध्य विचलन, स्तरीय भिन्नता, भेरियन्स र विचलनको सापेक्ष मापन जस्तै: विस्तारको गुणाङ्क, मध्य विचलनको गुणाङ्क, स्तरीय भिन्नताको गुणाङ्क आदि दुवै मापनको चर्चा गर्नेछौं ।

१२.४ विचलनको विभिन्न मापन

विचलनको मापनका लागि संख्यात्मक तथा रेखा चित्रात्मक विधिहरूको अवलम्बन गर्न सकिन्छ । संख्यात्मक मापनका लागि विस्तार (Range), मध्य विचलन (Mean Deviation), स्तरीय भिन्नता (Standard Deviation), र भेरियन्स (Variance) का बारेमा निरपेक्ष तथा सापेक्ष मापनका बारेमा चर्चा गर्नेछौं । साथै, रेखा चित्रात्मक विधिका बारेमा स्क्याटर डायग्राम (Scatter Diagram) र लरेन्ज बक्र (Lorenz Curve) का बारेमा अध्ययन गर्नेछौं ।

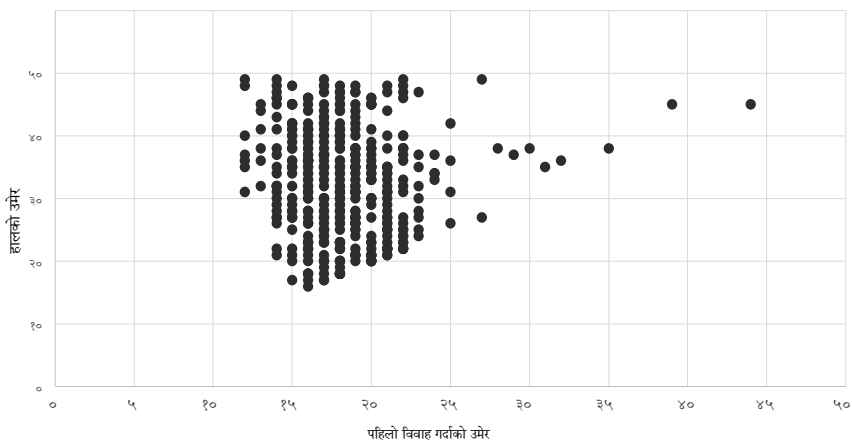
१२.४.१ स्क्याटर डायग्राम

कुनै पनि तथ्याङ्कको विश्लेषण गर्नु अगाडि उक्त तथ्याङ्कको सामान्य विशेषताका बारेमा जनाकारी लिनको लागि स्क्याटर डायग्राम हेर्नु उपयुक्त मानिन्छ । यस चित्रले दिइएको तथ्याङ्कका बारेमा केही जानकारी प्रदान गर्नुका साथै तथ्याङ्कको सामान्य

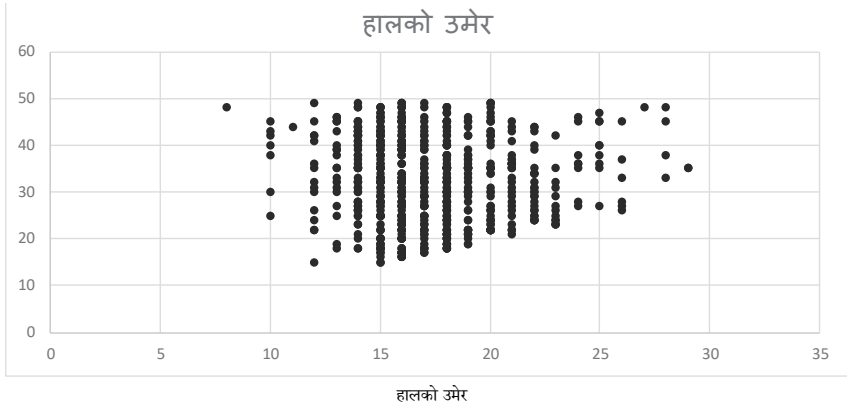
प्रकृतिका बारेमा जानकारी दिन्छ । स्क्याटर डायग्रामले विचलनको कच्चा धारणा बनाउनसमेत सहयोग गर्दछ । उदाहरणका लागि चित्र नं. १२.१ मा नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ मा छनोटमा परेका पहाडी र तराई सहरका १५ देखि ४९ वर्ष उमेर समूहका कम्तीमा एकपटक विवाह गरेका महिलाहरूको पहिलोपटक विवाह गर्दाको उमेरको स्क्याटर डायग्राममा प्रस्तुत गरिएको छ । यस चित्रबाट सामान्य हेराइको आधारमा कुन उमेरमा धेरै महिलाले पहिलोपटक विवाह गर्दछन् भन्ने कुराका साथै पहिलोपटक विवाह गर्दाको उमेर कति छरिएको छ भन्ने कुरा पनि देखिन्छ । साथै, सानो उमेरमा विवाह भएको र ठुलो उमेरमा विवाह भएको पनि प्रष्टै देखिन्छ । यसका अतिरिक्त पहिलो विवाह गर्दाको उमेर पहाडी सहर र पहाडी तराईमध्ये कुनमा बढी छरिएको छ भनेर चित्रका थोप्लाहरू कति हदसम्म छरपष्ट छन् भनेर हेर्न सकिन्छ । तसर्थ, यस्तो चित्र विचलनको तुलनात्मक अध्ययनका लागि पनि उपयोगी हुन्छ । यद्यपि कतिपय अवस्थामा चित्रकै अध्ययनले पूर्ण यकिनका साथ विचलन तुलना गर्न गाह्रो पनि हुन्छ ।

चित्र नं. १२.१: नेपाली महिलाहरूको पहिलोपटक विवाह गर्दाको उमेर

क) पहाडी सहर



(ख) तराई सहर



स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

१२.४.२ विस्तार (Range)

विचलन मापनको सबैभन्दा सजिलो तर कच्चा मापनको रूपमा विस्तारलाई लिइन्छ। यसले आँकडाहरूको फैलावटको सामान्य जानकारी दिन्छ। उदाहरणका लागि नेपाली महिलाहरूको पहिलो विवाह गर्दाको उमेरको विस्तार २१ वर्ष भनेर भन्दा खासै अर्थपूर्ण धारणा बनाउन सकिँदैन। तथापि सामान्य फैलावट हेर्नका लागि अन्य सूचनाहरूका साथमा यसको प्रयोग गरिन्छ। जस्तै नेपाली महिलाहरूको पहिलो विवाह गर्दाको उमेर ८ वर्षदेखि २९ वर्षसम्म रहेको र विस्तार २१ वर्ष हो भन्दा केही अर्थ दिन्छ। तथ्याङ्कमा भएको सबैभन्दा सानो र सबैभन्दा ठुलो मानको फरकलाई नै विस्तार भनिन्छ। सूत्रमा उल्लेख गर्दा,

$$\text{विस्तार} = \text{सबैभन्दा ठुलो आँकडा} - \text{सबैभन्दा सानो आँकडा} \quad (१)$$

सूत्र (१) मा उल्लेख गरेको विस्तार निरपेक्ष मापन हो भने यसको सापेक्ष मापनका लागि विस्तारको गुणाङ्क देहायको सूत्र प्रयोग गरेर निकालिन्छ।

$$\frac{\text{सबैभन्दा ठुलो आँकडा} - \text{सबैभन्दा सानो आँकडा}}{\text{सबैभन्दा ठुलो आँकडा} + \text{सबैभन्दा सानो आँकडा}} \quad (२)$$

विस्तार निकाल्नको लागि आँकडाको ठुलो र सानो मानमात्र आवश्यक पर्ने हुनाले आवृत्ति वितरणमा रहेका खण्डित श्रेणी वा अविच्छिन्न श्रेणी जे भए पनि एउटै विधि प्रयोग हुन्छ।

उदाहरण:

तालिका १२.१ (क): नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ मा छनोट भएका परिवारका सदस्यहरूको उमेर वितरण (खण्डित श्रेणी)

उमेर	संख्या
०	५४९
१	५३१
२	५९०
.....
९६	२
९७	०
९८	०
९९	५
जम्मा	२८,६७०

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

माथिको तालिका १२.१ (क) बाट विस्तार गणना गर्दा, $९९ - ० = ९९$ हुन्छ।

तालिका १२.१ (ख): नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ मा छनोट भएका परिवारका सदस्यहरूको उमेर वितरण (अविच्छिन्न श्रेणी)

उमेर समूह	संख्या
० देखि १०	७,१२९
१० देखि २०	६,७२५
२० देखि ३०	४,२३५
३० देखि ४०	३,४५३
४० देखि ५०	२,६७६
५० देखि ६०	२,२६०
६० देखि ७०	१,३५९
७० देखि ८०	६४७
८० देखि ९०	१६९
९० देखि ९९	१७

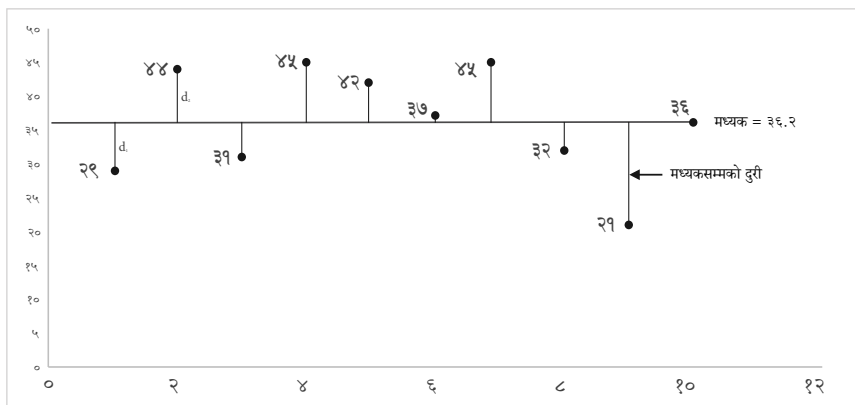
स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

माथिको तालिका १२.१ (ख) अविच्छिन्न आवृत्ति वितरणको उदाहरण हो । यस तथ्याङ्कको विस्तार $९९ - ० = ९९$ हुन्छ । अर्थात्, मानिसहरूको उमेर ० देखि ९९ वर्षसम्म भएकोले उमेर ९९ वर्षमा फैलिएको छ ।

१२.४.३ मध्य विचलन (Mean Deviation)

मध्य विचलनले तथ्याङ्क वितरणमा आँकडाहरू केन्द्रीय प्रवृत्तिबाट कति टाढासम्म छरिएका छन् भन्ने बारेमा सामान्य जानकारी दिन्छ । अर्थात्, औसतबाट प्रत्येक आँकडासम्मको दुरीको औसत मान नै मध्य विचलन हो । चित्र १२.२ मा देखाइए जस्तै प्रत्येक आँकडाबाट मध्यकसम्मका दुरी d_1, d_2, \dots, d_{10} को औसत लिएपछि मध्य विचलन आउँछ ।

चित्र १२.२: नेपाल जीवनस्तर सर्वेक्षणमा हिमाली कुनै एक गणना क्षेत्रको छानिएका १० जना महिलाको उमेर



यहाँ, औसतको रूपमा अंकगणितीय मध्यक, मध्यिका वा बहुलक भन्ने बुझिन्छ । सूत्रमा उल्लेख गर्दा,

व्यक्तिगत श्रेणीका लागि,

$$\text{मध्य विचलन} = \frac{1}{N} \sum_{i=1}^N |X_i - A|, \quad (३)$$

जहाँ, N भनेको आँकडाहरूको जम्मा संख्या हो, \sum योगफल लिने भन्ने बुझाउँछ, X_i ले आँकडाहरूको मान प्रतिनिधित्व गरेको छ, A भनेको मध्यक वा मध्यिका वा बहुलकमध्ये कुनै एक औसत भन्ने बुझाउँछ र "।" नकारात्मक संख्यालाई सकारात्मक बनाउन निर्देशन गर्दछ ।

तालिका १२.१ (ग): सन् २०१९ मा नेपाल भित्रिएका पर्यटकको मासिक संख्या विवरण

महिना	पर्यटक संख्या
जनवरी	८१,२७३
फेब्रुअरी	१,०२,४२३
मार्च	१,२७,३५१
अप्रिल	१,०९,३९९
मे	७८,३२९
जुन	७४,८८३
जुलाई	७०,९१६
अगस्ट	९४,७४९
सेप्टेम्बर	९२,६०४
अक्टुबर	१,३४,०९६
नोभेम्बर	१,३०,३०२
डिसेम्बर	१,००,८६६

स्रोत: नेपाल पर्यटन बोर्ड

माथिको तालिकाबाट सन् २०१९ मा नेपालमा मासिक औसत कति पर्यटक भित्रिए भनेर थाहा पाउनका लागि अंकगणितीय मध्यक निकालेर हेर्न सकिन्छ। त्यसका लागि बाह्रै महिनाका सबै पर्यटकका संख्यालाई जोडेर आएको जोडफललाई १२ ले भाग गर्नुपर्दछ। अर्थात्,

$$\frac{८१,२७३ + १,०२,४२३ + \dots + १,३०,३०२ + १,००,८६६}{१२}$$

$$= \frac{११,९७,१९१}{१२} = ९९,७६५.९२ \approx ९९,७६६$$

सन् २०१९ मा औसतमा मासिक ९९,७६६ पर्यटक नेपाल भित्रिएका थिए।

अब, प्रत्येक आँकडाबाट मध्यक घटाएर तालिकामा राखौं। त्यसपछि नकारात्मक मानहरूलाई सकारात्मक बनाएर योगफल निकालौं। अब, निकालिएका मानहरू निम्न तालिकाबमोजिम पाइन्छ।

तालिका १२.१ (घ): सन् २०१९ मा नेपाल भित्रिएका पर्यटकको मासिक संख्याको मध्य विचलन निकाल्ने तरिका

महिना	पर्यटक संख्या (X)	X - Mean	$ X_i - A $
जनवरी	८१,२७३	-१८,४९३	१८,४९३
फेब्रुअरी	१,०२,४२३	२,६५७	२,६५७
मार्च	१,२७,३५१	२७,५८५	२७,५८५
अप्रिल	१,०९,३९९	९,६३३	९,६३३
मे	७८,३२९	-२१,४३७	२१,४३७
जुन	७४,८८३	-२४,८८३	२४,८८३
जुलाई	७०,९१६	-२८,८५०	२८,८५०
अगस्ट	९४,७४९	-५,०१७	५,०१७
सेप्टेम्बर	९२,६०४	-७,१६२	७,१६२
अक्टुबर	१,३४,०९६	३४,३३०	३४,३३०
नोभेम्बर	१,३०,३०२	३०,५३६	३०,५३६
डिसेम्बर	१,००,८६६	११,००	१,१००
जम्मा	११,९७,१९१		२,११,६८३

$$\text{मध्य विचलन} = \frac{1}{N} \sum_{i=1}^N |X_i - A| = \frac{२,११,६८३}{१२} = १७,६४०$$

यसबाट के थाहा हुन्छ भने मासिक पर्यटकको संख्याहरू मध्यकबाट औसतमा १७,६४० सम्म फैलिएका छन् ।

खण्डित आवृत्ति वितरणका लागि,

$$\text{मध्य विचलन} = \frac{1}{N} \sum_{i=1}^N f |X_i - A|, \quad (४)$$

जहाँ, f भनेको बारम्बारता वा आवृत्ति हो ।

उदाहरण:

तालिका १२.१ (ड): नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ मा छनोटमा परेका परिवारको परिवार सदस्य संख्या

क्र.सं.	परिवार सदस्य संख्या (x)	परिवार संख्या (f)	f x
१	१	२०७	२०७
२	२	६४६	१,२९२
३	३	९४५	२,८३५
४	४	१,२५४	५,०१६
५	५	१,०८६	५,४३०
६	६	७९८	४,७८८
७	७	४५०	३,१५०
८	८	२५७	२,०५६
९	९	१३५	१,२१५
१०	१०	८१	८१०
११	११	४९	५३९
१२	१२	२५	३००
१३	१३	१७	२२१
१४	१४	८	११२
१५	१५	१२	१८०
१६	१६	४	६४
१७	१७	४	६८
१८	१८	२	३६
१९	१९	५	९५
२०	२०	३	६०
जम्मा		५,९८८	२८,४७४

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

तालिका १२.१ (ड) खण्डित श्रेणीको आवृत्ति वितरण गरी राखिएको तथ्याङ्कको उदाहरण हो। यो तालिकाबाट हामीले अंकगणितीय मध्यक अथवा औसत परिवार सदस्य संख्या निकाल्न सक्दछौं।

$$\text{अंकगणितीय मध्यक } (\bar{X}) = \frac{\sum fX}{\sum f} = \frac{२८,४७४}{५,९८८} = ४.८$$

अब, प्रत्येक मानबाट मध्यक घटाऔँ र नकारात्मक मानलाई सकारात्मक बनाऔँ । त्यसपछि सम्बन्धित बारम्बारताले गुणा गरेर योगफल निकालौँ । यी कार्यहरूको नतिजा तलको तालिकामा प्रस्तुत गरिएको छ ।

तालिका १२.१ (च): परिवार सदस्य संख्या वितरणको मध्य विचलन निकाल्ने तरिका

क्र.सं.	परिवार सदस्य संख्या (x)	परिवार संख्या (f)	f x	X - A abs	X-A abs	F(X-A) abs
१	१	२०७	२०७	-३.८	३.८	७७७.३
२	२	६४६	१,२९२	-२.८	२.८	१,७७९.८
३	३	९४५	२,८३५	-१.८	१.८	१,६५८.६
४	४	१,२५४	५,०१६	-०.८	०.८	९४७
५	५	१,०८६	५,४३०	०.२	०.२	२६५.९
६	६	७९८	४,७८८	१.२	१.२	९९३.४
७	७	४५०	३,१५०	२.२	२.२	१,०१०.२
८	८	२५७	२,०५६	३.२	३.२	८३३.९
९	९	१३५	१,२१५	४.२	४.२	५७३.१
१०	१०	८१	८१०	५.२	५.२	४२४.८
११	११	४९	५३९	६.२	६.२	३०६
१२	१२	२५	३००	७.२	७.२	१८१.१
१३	१३	१७	२२१	८.२	८.२	१४०.२
१४	१४	८	११२	९.२	९.२	७४
१५	१५	१२	१८०	१०.२	१०.२	१२२.९
१६	१६	४	६४	११.२	११.२	४५
१७	१७	४	६८	१२.२	१२.२	४९
१८	१८	२	३६	१३.२	१३.२	२६.५
१९	१९	५	९५	१४.२	१४.२	७१.२
२०	२०	३	६०	१५.२	१५.२	४५.७
जम्मा		५,९८८	२८,४७४			१,०३२५.६

अब,

$$\text{मध्य विचलन} = \frac{1}{N} \sum_{i=1}^N f|X_i - A| = \frac{१०३२५.६}{५९८८} = १.७२$$

यसबाट के थाहा हुन्छ भने परिवारको आकार मध्यकबाट औसतमा १.७२ सम्म छरिएका छन् ।

अविच्छिन्न आवृत्ति वितरणका लागि,

$$\text{मध्य विचलन} = \frac{1}{N} \sum_{i=1}^N f|M_i - A|, \quad (५)$$

जहाँ, M_i भनेको अन्तराल समूहको मध्यमान हो ।

मध्य विचलन निकाल्दा औसत र प्रत्येक आँकडाको फरक लिएर यदि सो फरक नकारात्मक आएमा सकारात्मक बनाएर प्राप्त हुने फरकहरूको आँकडाको मध्यक निकाल्नु पर्दछ । यदि नकारात्मक मानलाई सकारात्मक बनाइएन र औसतमा अंकगणितीय मध्यक छनोट गरियो भने मध्य विचलनको मान नै शून्य आउँछ । तसर्थ, माथिका सूत्रहरूमा सकारात्मक बनाउने गरिएको हो ।

उदाहरण:

माथि तालिका १२.१ को तथ्याङ्कबाट मध्य विचलन निकाल्नका लागि हामीले देहायबमोजिमको तालिका तयार गर्नुपर्दछ । यसरी तालिका तयार गर्नका लागि मध्य विचलनको सूत्र हेरेर के के आवश्यक पर्ने हो सोअनुसार गर्नुपर्दछ । यस उदाहरणमा औसतका लागि अंकगणितीय मध्यक छनोट गरिएको छ । तसर्थ, पहिला अंकगणितीय मध्यक निकाल्नुपर्दछ ।

तालिका १२.२: नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ मा छनोटमा परेका परिवारका सदस्यहरूको उमेर वितरणका मध्यमान निकाले तरिका

उमेर समूह	संख्या (f)	मध्यमान (M)	fM	Mi - A	Mi - A	f M _i - A
१	२	३	४	५	६	७
० देखि १०	७,१२९	५	३५,६४५	-२२	२२	१,५६,८३८
१० देखि २०	६,७२५	१५	१,००,८७५	-१२	१२	८०,७००
२० देखि ३०	४,२३५	२५	१,०५,८७५	-२	२	८,४७०
३० देखि ४०	३,४५३	३५	१,२०,८५५	८	८	२९,२२३
४० देखि ५०	२,६७६	४५	१,२०,४२०	१८	१८	४९,४०८
५० देखि ६०	२,२६०	५५	१,२४,३००	२८	२८	६४,३२७
६० देखि ७०	१,३५९	६५	८८,३३५	३८	३८	५२,२७१
७० देखि ८०	६४७	७५	४८,५२५	४८	४८	३१,३५६
८० देखि ९०	१६९	८५	१४,३६५	५८	५८	९,८८०
९० देखि १००	१७	९५	१,६१५	६८	६८	१,१६४
जम्मा	२८,६७०		७,६०,८१०			४,८३,६३७

तरिका :

मध्य विचलन निकालनका लागि हामीलाई औसतको आवश्यकता पर्दछ। यस उदाहरणमा अंकगणितीय मध्यक लिइएको हुनाले कोलम ३ मा प्रत्येक अन्तरालहरूको मध्यमान निकालिएको हो। अंकगणितीय मध्यकको सूत्रअनुसार,

$$\text{मध्यक} = \frac{1}{N} \sum fM = \frac{9}{27670} \times 960590 = 26.28 \text{ हुन्छ।}$$

अब, कोलम ५ मा प्रत्येक मध्यमान र मध्यकको फरक लिइएको छ। कोलम ६ मा नकारात्मक मानहरूलाई सकारात्मक बनाइएको छ। कोलम ७ मा बारम्बारता (f) र कोलम ६ का अंक गुणन गरी राखिएको छ।

$$\text{मध्य विचलन} = \frac{1}{N} \sum_{i=1}^N f|M_i - A| = \frac{9}{27670} \times 873637 = 28.17$$

तसर्थ, दिइएको तथ्याङ्कको मध्य विचलन २८.१७ रहेको पाइयो। यसको अर्थ, दिइएको तथ्याङ्कमा रहेका आँकडाहरू औसतमा अंकगणितीय मध्यकभन्दा २८.१७ ले फरक छन्।

मध्य विचलनको सापेक्ष मापन

$$\text{मध्य विचलनको गुणाङ्क} = \frac{\text{मध्य विचलन}}{\text{प्रयोग गरिएको औसत}} \quad (६)$$

जहाँ, प्रयोग गरिएको औसतको ठाउँमा मध्यक, मध्यिका वा बहुलक जे हो त्यही राख्ने। अर्थात्,

$$\text{मध्यकबाट निकालिएको मध्य विचलनको गुणाङ्क} = \frac{\text{मध्य विचलन}}{\text{मध्यक}},$$

$$\text{मध्यिकाबाट निकालिएको मध्य विचलनको} = \frac{\text{मध्य विचलन}}{\text{मध्यिका}}, \text{ र}$$

$$\text{बहुलकबाट निकालिएको मध्य विचलनको गुणाङ्क} = \frac{\text{मध्य विचलन}}{\text{बहुलक}}, \text{ हुन्छ।}$$

१२.४.४ स्तरीय भिन्नता (Standard Deviation)

मध्य विचलनमा औसत र आँकडाबिचको फरक लिएर नकारात्मक भएमा सकारात्मक बनाउने कार्यलाई कतिपय मानिसहरूले आलोचना गर्ने गर्दछन्। मध्य विचलनको यो सीमालाई समाधान गर्नका लागि स्तरीय भिन्नताको विकास गरिएको हो भन्ने गरिन्छ। स्तरीय भिन्नता मापन गर्नका लागि आँकडाहरू र अंकगणितीय मध्यकको फरकलाई वर्ग गर्ने अनि तिनीहरूको योगफल निकालेर जम्मा आँकडा संख्याले भाग गर्नुपर्दछ। यसरी

आएको भागफलको सकारात्मक वर्गमूल नै स्तरीय भिन्नता हो। सरल रूपमा बुझ्नुपर्दा कुनै तथ्याङ्कमा समावेश मानहरूको औसत मानबाट ती प्रत्येक मानहरू कति टाढा छन् वा छरिएका छन् भनेर निकालिने मान नै स्तरीय भिन्नता हो। यसलाई σ (Greek Letter, Pronounce as SIGMA) ले जनाइन्छ। अर्थात्,

व्यक्तिगत श्रेणीका लागि,

$$\text{स्तरीय भिन्नता } (\sigma) = \sqrt{\frac{\sum(X-\bar{X})^2}{N}}, \quad (७)$$

उदाहरण: सन् २०१९ मा नेपाल भित्रिएका पर्यटकको संख्या

तालिका १२.२(क): सन् २०१९ मा नेपाल भित्रिएका पर्यटकको संख्या

महिना	पर्यटक संख्या (X)
जनवरी	८१,२७३
फेब्रुअरी	१,०२,४२३
मार्च	१,२७,३५१
अप्रिल	१,०९,३९९
मे	७८,३२९
जुन	७४,८८३
जुलाई	७०,९१६
अगस्ट	९४,७४९
सेप्टेम्बर	९२,६०४
अक्टुबर	१,३४,०९६
नोभेम्बर	१,३०,३०२
डिसेम्बर	१,००,८६६
जम्मा	११,९७,१९१

$$\frac{८१,२७३ + १,०२,४२३ + \dots + १,३०,३०२ + १,००,८६६}{१२}$$

$$= \frac{११,९७,१९१}{१२} = ९९,७६५.९२ \approx ९९,७६६$$

सन् २०१९ मा औसतमा मासिक ९९,७६६ पर्यटक नेपाल भित्रिएका थिए।

अब, प्रत्येक आँकडाबाट मध्यक घटाएर तालिकामा राखौं। त्यसपछि ती मानहरूको वर्ग निकालौं। यसरी निकालिएका मानहरू निम्न तालिका बमोजिम पाइन्छ।

तालिका १२.२ (ख): सन् २०१९ मा नेपाल भित्रिएका पर्यटकको स्तरीय भिन्नता निकाल्ने तरिका

महिना	पर्यटक संख्या (X)	$X - \bar{x}$	$(X - \bar{x})^2$
जनवरी	८१,२७३	-१८,४९३	३४,१९,८७,९६७
फेब्रुअरी	१,०२,४२३	२,६५७	७०,६०,०९२
मार्च	१,२७,३५१	२७,५८५	७६,०९,३६,८२३
अप्रिल	१,०९,३९९	९,६३३	९,२७,९६,२९५
मे	७८,३२९	-२१,४३७	४५,९५,४१,३९६
जुन	७४,८८३	-२४,८८३	६१,९१,५९,५४२
जुलाई	७०,९१६	-२८,८५०	८३,२३,१७,६९२
अगस्ट	९४,७४९	-५,०१७	२,५१,६९,४५३
सेप्टेम्बर	९२,६०४	-७,१६२	५,१२,९३,०५०
अक्टुबर	१,३४,०९६	३४,३३०	१,१७,८५,५४,६२२
नोभेम्बर	१,३०,३०२	३०,५३६	९३,२४,५२,३८५
डिसेम्बर	१,००,८६६	१,१००	१२,१०,१८३
जम्मा	११,९७,१९१		५,३०,२४,७९,४९९

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

तालिका ११.२ (ग) खण्डित श्रेणीको आवृत्ति वितरण गरी राखिएको तथ्याङ्कको उदाहरण हो। यो तालिकाबाट हामीले अंकगणितीय मध्यक अथवा औसत परिवार सदस्य संख्या निकाल्न सक्दछौं।

$$\text{अंकगणितीय मध्यक } (\bar{X}) = \frac{\sum fX}{\sum f} = \frac{२८४७४}{५९८८} = ४.८$$

अब, प्रत्येक मानबाट मध्यक घटाऔं वर्ग लिऔं। त्यसपछि सम्बन्धित बारम्बारताले गुणा गरेर योगफल निकालौं। यी कार्यहरूको नतिजा तलको तालिकामा प्रस्तुत गरिएको छ।

तालिका १२.२(ग): नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ मा छनोटमा परेका परिवारको परिवार सदस्य संख्याको स्तरीय भिन्नता निकाल्ने तरिका

क्र.सं.	परिवार सदस्य संख्या (x)	परिवार संख्या (f)	f x	$X - \bar{x}$	$(X - \bar{x})^2$	$f(X - \bar{x})^2$
१	१	२०७	२०७	-३.८	१४.१	२९१९
२	२	६४६	१,२९२	-२.८	७.६	४,९०३.८

३	३	९४५	२,८३५	-१.८	३.१	२,९११.२
४	४	१,२५४	५,०१६	-०.८	०.६	७१५.१
५	५	१,०८६	५,४३०	०.२	०.१	६५.१
६	६	७९८	४,७८८	१.२	१.५	१,२३६.६
७	७	४५०	३,१५०	२.२	५	२,२६७.७
८	८	२५७	२,०५६	३.२	१०.५	२,७०५.९
९	९	१३५	१,२१५	४.२	१८	२,४३२.५
१०	१०	८१	८१०	५.२	२७.५	२,२२८.२
११	११	४९	५३९	६.२	३९	१,९१०.९
१२	१२	२५	३००	७.२	५२.५	१,३१२.२
१३	१३	१७	२२१	८.२	६८	१,१५५.६
१४	१४	८	११२	९.२	८५.५	६८३.७
१५	१५	१२	१८०	१०.२	१०५	१,२५९.५
१६	१६	४	६४	११.२	१२६.४	५०५.८
१७	१७	४	६८	१२.२	१४९.९	५९९.७
१८	१८	२	३६	१३.२	१७५.४	३५०.९
१९	१९	५	९५	१४.२	२०२.९	१,०१४.६
२०	२०	३	६०	१५.२	२३२.४	६९७.२
जम्मा		५,९८८	२८,४७४			३१,८७५.१

$$\text{स्तरीय भिन्नता } (\sigma) = \sqrt{\frac{\sum f(X - \bar{X})^2}{N}} = \sqrt{\frac{३१,८७५.१}{५,९८८}} = २.३$$

तसर्थ, परिवारको आकार मध्यकभन्दा औसतमा २.३ ले छरिएका छन् ।

अविच्छिन्न श्रेणीका लागि,

$$\text{स्तरीय भिन्नता } (\sigma) = \sqrt{\frac{\sum f(M - \bar{X})^2}{N}}, \quad (९)$$

जहाँ, X भनेको आँकडाको मान हो, \bar{X} भनेको अंकगणितीय मध्यक हो $[\bar{X} = \frac{\sum X}{N}]$, M भनेको समूह अन्तरालको मध्यमान हो, f भनेको बारम्बारता हो, N भनेको आँकडाको जम्मा संख्या हो । अविच्छिन्न श्रेणीको हकमा X को स्थानमा M अर्थात् समूह अन्तरालको मध्यमान राखिन्छ ।

तथ्याङ्कका विभिन्न सूत्रहरूको प्रयोग गरेर दिइएको तथ्याङ्कबाट आवश्यक मानहरू निकाल्नका लागि सूत्रलाई हेरेर क्रमशः हिसाब गर्दै जानुपर्दछ। उदाहरणको लागि सूत्र (७) बमोजिम व्यक्तिगत श्रेणीको स्तरीय भिन्नता निकाल्नु पर्‍यो भने,

- पहिला सबै आँकडालाई जोडेर कुल जम्मा निकाल्ने,
- कुल जम्मालाई आँकडाको जम्मा संख्याले भाग गरेर मध्यक निकाल्ने,
- सम्पूर्ण आँकडाबाट पालैपालो मध्यक घटाएर तालिकामा राख्ने,
- घटाएर आएका मानहरूलाई वर्ग गरी तालिकामा राख्ने,
- वर्ग गरेपछि आएका मानहरू सबै जोड्ने,
- आएको योगफललाई बारम्बारताको योगफलले भाग गरी वर्गमूल लिने,
- अर्थात्, माथिको सूत्र (७) मा राखेर स्तरीय भिन्नता निकाल्ने,
- यसैगरी सूत्र (८) अनुसार खण्डित आवृत्ति वितरणमा भएको तथ्याङ्कको स्तरीय भिन्नता निकाल्नु पर्‍यो भने,
- पहिला सबै आँकडालाई पालैपालो सम्बन्धित बारम्बारतासँग गुणा गरेर तालिकामा राख्ने,
- यसरी राखिएका गुणनफलहरूलाई जोड्ने र आएको योगफललाई बारम्बारताको योगफलले भाग गरेर मध्यक निकाल्ने,
- सम्पूर्ण आँकडाबाट पालैपालो मध्यक घटाएर तालिकामा राख्ने,
- घटाएर आएका मानहरूलाई वर्ग गरी तालिकामा राख्ने,
- यसरी आएका वर्गहरूलाई बारम्बारतासँग गुणन गरेर तालिकामा राख्ने,
- गुणा गरेर आएका गुणनफलहरूलाई जोड्ने र आएको योगफललाई बारम्बारताको योगफलले भाग गरी वर्गमूल लिने,
- अर्थात्, माथिको सूत्र (८) मा राखेर स्तरीय भिन्नता निकाल्ने।

सूत्र (९) बमोजिम अविच्छिन्न श्रेणीको स्तरीय भिन्नता निकाल्नका लागि उदाहरण सहित हेरौं। यसका लागि तालिका १२.१ (ख) को तथ्याङ्कको स्तरीय भिन्नता निकाल्ने अभ्यास गरौं।

तालिका १२.२ (ख): नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ मा छनोट भएका परिवारका सदस्यहरूको उमेर वितरण (अविच्छिन्न श्रेणी)

उमेर समूह	संख्या
० देखि १०	७,१२९
१० देखि २०	६,७२५
२० देखि ३०	४,२३५
३० देखि ४०	३,४५३

४० देखि ५०	२,६७६
५० देखि ६०	२,२६०
६० देखि ७०	१,३५९
७० देखि ८०	६४७
८० देखि ९०	१६९
९० देखि ९९	१७

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

- तलको तालिका १२.३ मा कोलम १ र २ दिइएको तथ्याङ्कको अविच्छिन्न श्रेणी हो ।
- अन्तराल समूहमा दिएको तथ्याङ्कको लागि मध्यमान निकाल्नुपर्ने भएकोले कोलम ३ निकालिएको हो ।
- कोलम ३ मा यसमा ० र १० को मध्यमान ५, १० र २० को मध्यमान १५, हुन्छ ।
- कोलम ४ मा मध्यमान र बारम्बारता गुणा गरिएको छ ।
- कोलम ४ मध्यक निकाल्नको लागि आवश्यक हुन्छ ।
- सूत्र $\bar{X} = \frac{\sum fM}{N}$ प्रयोग गरी मध्यकको मान २६.५४ निकालियो ।
- कोलम ५ मा प्रत्येक मध्यमानबाट मध्यक घटाइएको छ ।
- कोलम ६ मा कोलम ५ का संख्याको वर्ग लिइएको छ ।
- कोलम ७ मा कोलम ६ सँग बारम्बारता (कोलम २) गुणा गरिएको छ ।

तालिका १२.३: नेपाल जीवनस्तर सर्वेक्षण, तेस्रोमा छनोट भएका परिवारका सदस्यहरूको उमेर वितरणको स्तरीय भिन्नता निकाल्ने तरिका

उमेर समूह	संख्या (f)	मध्यमान (M)	fM	M- \bar{x}	(M- \bar{x}) ²	f(M- \bar{x}) ²
१	२	३	४	५	६	७
० देखि १०	७,१२९	५	३५,६४५	-२२	४६४	३३,०६,६७०
१० देखि २०	६,७२५	१५	१,००,८७५	-१२	१३३	८,९५,०८२
२० देखि ३०	४,२३५	२५	१,०५,८७५	-२	२	१०,००२
३० देखि ४०	३,४५३	३५	१,२०,८५५	८	७२	२,४७,३२४
४० देखि ५०	२,६७६	४५	१,२०,४२०	१८	३२१	९,१२,२२१
५० देखि ६०	२,२६०	५५	१,२४,३००	२८	८१०	१८,३०,९४८
६० देखि ७०	१,३५९	६५	८८,३३५	३८	१,४७९	२०,१०,५२९
७० देखि ८०	६४७	७५	४८,५२५	४८	२,३४९	१५,१९,५९७
८० देखि ९०	१६९	८५	१४,३६५	५८	३,४१८	५,७७,६३३
९० देखि १००	१७	९५	१,६१५	६८	४,६८७	७९,६८३
जम्मा	N=२८,६७०		७,६०,८१०			१,१३,८९,६८९

यसरी तालिका १२.३ तयार गरिसकेपछि सूत्र (९) मा सम्बन्धित मानहरू राखेर स्तरीय भिन्नता निकाल्न सकिन्छ। अर्थात्,

$$\text{स्तरीय भिन्नता } (\sigma) = \sqrt{\frac{\sum f(M-\bar{X})^2}{N}}, = \sqrt{\frac{११३८९६८९}{२८६७०}} = १९.९$$

तसर्थ, दिइएको उमेर समूह आवृत्ति वितरण तथ्याङ्कको स्तरीय भिन्नता १९.९ हुन्छ। यसले परिवारका सदस्यहरूको उमेरको औसत मानबाट उनीहरू प्रत्येकको उमेर औसतमा १९.९ ले टाढा छन् वा छरिएका छन् भनेर भन्न सकिन्छ।

१२.४.५ स्तरीय भिन्नताको मानाङ्क

स्तरीय भिन्नताको गुणाङ्क (*Coefficient of Standard Deviation*) पनि आँकडामा भएको सापेक्ष विचलन हेर्ने एक विधि हो। यसले मध्यकको तुलानामा स्तरीय भिन्नता कति भाग छ भनेर देखाउँछ। यसलाई प्रतिशतमा व्यक्त गरियो भने विचलनको गुणाङ्क (*Coefficient of Variation = CV*) ले जनाइन्छ। सूत्रमा हेर्दा,

$$\text{स्तरीय भिन्नताको गुणाङ्क} = \frac{\sigma}{\bar{X}} \quad (१०)$$

$$\text{विचलनको गुणाङ्क} = \frac{\sigma}{\bar{X}} \times १०० \quad (११)$$

माथिको उदाहरणमा मध्यक २६.५४ र स्तरीय भिन्नता १९.९ आएको थियो।

$$\text{स्तरीय भिन्नताको गुणाङ्क} = \frac{\sigma}{\bar{X}} = \frac{१९.९}{२६.५४} = ०.७५ \text{ हुन्छ।}$$

$$\text{विचलनको गुणाङ्क (CV)} = \frac{\sigma}{\bar{X}} \times १०० = ०.७५ \times १०० = ७५\% \text{ हुन्छ।}$$

१२.४.६ भेरियन्स (Variance)

स्तरीय भिन्नतामा वर्गमूल निकाल्नु पर्ने भएकोले यसलाई अझै सरल बनाउन वर्ग गरेर भेरियन्स निकाल्ने गरिन्छ। तथ्याङ्क विश्लेषणमा भेरियन्सको कुनै अर्थपूर्ण व्याख्या नभए पनि विचलनको मापन गर्नुका अतिरिक्त थप गणितीय क्रियाहरूलाई यसको प्रयोग गरिन्छ। कतिपय तथ्याङ्कीय सूत्रहरूमा भेरियन्सको प्रयोग भएको हुन्छ। यसलाई σ^2 ले जनाइन्छ। सूत्रमा,

$$\text{भेरियन्स } (\sigma^2) = (\text{स्तरीय भिन्नता})^2$$

उदाहरण: स्तरीय भिन्नता १९.९ भएको श्रेणीको भेरियन्स कति हुन्छ ?

$$\text{भेरियन्स } (\sigma^2) = (१९.९)^2 = ३९६.०१$$

१२.५ विचलन मापनका विशेषताहरू

विचलनका विभिन्न मापनका गुण, अवगुण तथा प्रयोगका क्षेत्रहरू तलको तालिका १२.४ मा दिइएको छ ।

तालिका १२.४: विचलन मापनका प्रमुख विशेषताहरू

क्र.सं.	औसत	गुण	अवगुण	प्रयोगका क्षेत्र
१	स्क्याटर डायग्राम	<ul style="list-style-type: none"> चित्रात्मक विधि भएकाले आकर्षक र बुझ्न सजिलो हुन्छ । 	<ul style="list-style-type: none"> विचलनको यकिन गरी मान पत्ता लगाउन सकिन्न । 	<ul style="list-style-type: none"> कुनै पनि तथ्याङ्क प्रशोधन गर्नुपूर्व कुनै रेखाचित्रमा तथ्याङ्क हेर्नु राम्रो मानिन्छ ।
२	विस्तार	<ul style="list-style-type: none"> स्पष्ट रूपले परिभाषित छ, हेर्न, बुझ्न र निकाल्न सजिलो छ, 	<ul style="list-style-type: none"> दुईवटा आँकडामा मात्र निर्भर गर्दछ, अति ठुलो र अति सानो मानले निकै प्रभावित गर्दछ, नमुना छनोटको फेर बदलले निकै प्रभावित गर्दछ, खुला अन्त भएका अन्तराल समूहको तथ्याङ्कमा प्रयोग गर्न सकिँदैन, थप गणितीय कार्यका लागि उपयुक्त छैन । 	<ul style="list-style-type: none"> थोरैमात्र परिवर्तन हुने क्षेत्र जस्तै: स्टक मार्केट, मुद्राको विनिमय दर, तापक्रम, मौसम आदिमा प्रयोग गर्न सकिन्छ ।
३	मध्य विचलन	<ul style="list-style-type: none"> स्पष्ट रूपले परिभाषित छ, हेर्न, बुझ्न र निकाल्न सजिलो छ, सबै आँकडाहरूमा आधारित छ, 	<ul style="list-style-type: none"> यो निकाल्दा नकारात्मक मानलाई सकारात्मक बनाउनु नै यसको प्रमुख अवगुण हो, बहुलक आफैँमा स्पष्ट परिभाषित नभएका कारण यसमा आधारित मध्य विचलन सन्तोषजनक हुँदैन, 	<ul style="list-style-type: none"> अर्थशास्त्र र व्यापार शास्त्रमा यसको प्रयोग गर्न सकिन्छ, सम्पत्तिसम्बन्धी अध्ययनमा यसको प्रयोग गरिन्छ ।

	<ul style="list-style-type: none"> • तुलनात्मक रूपमा अति तुलो र अति सानो मानले कम प्रभावित हुन्छ, • यथार्थ र सत्य विचलनको प्रतिनिधित्व गर्दछ, 	<ul style="list-style-type: none"> • खुला अन्तराल समूहमा व्यक्त तथ्याङ्कबाट यसको गणना गर्न सकिँदैन । 	
४	<p>स्तरीय भिन्नता</p> <ul style="list-style-type: none"> • स्पष्ट रूपले परिभाषित छ, • सबै आँकडाहरूमा आधारित छ, • तुलनात्मक रूपमा अति तुलो र अति सानो मानले कम प्रभावित हुन्छ, • यथार्थ र सत्य विचलनका प्रतिनिधित्व गर्दछ, • मध्य विचलनका कमजोरीहरू निराकरण गर्दछ, • नमुना छनोटको फेर बदलले कम प्रभावित हुन्छ, • थप गणितीय कार्यको लागि उपयुक्त हुन्छ । 	<ul style="list-style-type: none"> • यो निकाल्न केही भन्भटिलो हुन्छ, • खुला अन्तराल समूहमा व्यक्त तथ्याङ्कबाट यसको गणना गर्न सकिँदैन । 	<ul style="list-style-type: none"> • स्तरीय भिन्नतासँग असल विचलनको मापनका अधिकांश गुणहरू भएको हुनाले सामान्यतया सबै जसो तथ्याङ्कीय अध्ययनमा यसको प्रयोग गरिन्छ ।
५	<p>भेरियन्स</p> <ul style="list-style-type: none"> • स्तरीय भिन्नताजस्तै 	<ul style="list-style-type: none"> • स्तरीय भिन्नताजस्तै 	<ul style="list-style-type: none"> • स्तरीय भिन्नताजस्तै

१२.६ जेड-स्कोर (z-score)

दिइएका आँकडाहरूको उद्गम बिन्दु तथा स्केल परिवर्तन गरेर आउने मानलाई जेड स्कोर भनिन्छ । यसको कुनै मापन एकाइ हुँदैन । यसलाई स्तरीय स्कोर (Standard Deviation) पनि भनिन्छ । सूत्रमा व्यक्त गर्दा,

$$z - score = \frac{X - \bar{X}}{\sigma}, \quad (१३)$$

जहाँ, \bar{X} भनेको मध्यक र σ भनेको स्तरीय भिन्नता हो । जेड स्कोरको प्रयोग तथ्याङ्क क्षेत्रमा निकै महत्त्वपूर्ण छ । दुई भिन्न तथ्याङ्कको आँकडाको सापेक्ष तुलना अध्ययन गर्दा यसको प्रयोग हुन्छ । नमुना छनोटको आकार पत्ता लगाउन पनि यसको प्रयोग हुन्छ । सबै भन्दा महत्त्वपूर्ण त यसको प्रयोग नर्मल डिस्ट्रिब्युसनको सम्भाव्यता बक्रारेखा अन्तर्गतको क्षेत्रफल निकाल्नमा प्रयोग हुन्छ । यस्तो क्षेत्रफल परिकल्पना परीक्षण अध्ययन गर्दा आवश्यक हुन्छ । यी विषयहरू हामी पाठ १७ मा संक्षेपमा चर्चा गर्नेछौं ।

१२.७ लरेन्ज बक्रारेखा र जिनीको गुणाङ्क (Lorenz Curve and Gini Coefficient)

रेखाचित्रको माध्यमबाट विचलनको अध्ययन गर्ने विधिहरू मध्येको एक विधि लरेन्ज बक्रारेखा पनि हो । खासगरी आर्थिक असमानताको अध्ययनमा यसको अत्याधिक प्रयोग भएको हुन्छ । जस्तै धनी र गरिब बिचको असमानता अध्ययनमा लरेन्ज बक्रारेखा र जिनीको गुणाङ्क प्रयोग गरिन्छ । यसका अतिरिक्त व्यवसायमा जस्तै: तलब ज्यालाको असमानता अध्ययन गर्न, नाफा नोक्सानको वितरणको असमानता अध्ययन गर्न, उत्पादनको असमान वितरण अध्ययन गर्न, बिक्री वितरणको असमानता अध्ययन गर्न, जनसंख्याको असमान वितरण अध्ययन गर्न पनि यिनीहरूको प्रयोग गरिन्छ । जिनीको गुणाङ्क लरेन्ज बक्रारेखाको सहायताले र सूत्रको प्रयोग गरेर निकाल्न सकिन्छ । यहाँ, हामी बक्रारेखाको सहायताले जिनीको गुणाङ्क निकाल्ने विधिका बारेमा अभ्यास गर्दै छलफल गर्दौं ।

१२.७.१ लरेन्ज बक्रारेखा बनाउने तरिका

लरेन्ज बक्रारेखाले आँकडा र बारम्बारताका सुरुका मानको सट्टा सञ्चित मानहरूको प्रयोग गर्दछ । यो नै यसको प्रमुख विशेषता हो । लरेन्ज बक्रारेखा बनाउन निम्नानुसार गर्नुपर्दछ ।

१. आँकडाको मान र बारम्बारता दुवैलाई “भन्दा-सानो सञ्चित मान” मा लैजाने ।
२. आँकडा र बारम्बारताको छुट्टाछुट्टै कुल जम्मालाई १०० मानेर तिनीहरूको सञ्चित मानलाई प्रतिशतमा लैजाने ।
३. अब, X-अक्षमा सञ्चित बारम्बारताको प्रतिशत र Y-अक्षमा सञ्चित आँकडा मानको प्रतिशतलाई वितरण गरेर बक्रारेखा बनाउने (चित्र नं. १२.२) । यही बक्रारेखालाई लरेन्ज बक्रारेखा भनिन्छ ।
४. बिन्दु (०,०) र (१००,१००) जोड्ने एउटा विकर्ण OP खिच्ने । यो OP रेखाले X-अक्षसँग ४५ डिग्रीको कोण बनाएको हुन्छ । यो रेखालाई समान वितरणको रेखा भनिन्छ ।

उदाहरणका लागि नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ मा छनोट भएका र दैनिक ज्यालामा काम गर्ने श्रमिकहरूले दैनिक पाउने ज्याला र श्रमिकको संख्या तालिका १२.५ मा दिइएको छ। यस तालिकामा दैनिक ज्याला (रु) (कोलम) १ र श्रमिक संख्या कोलम ५ दिइएको आँकडा हो। कोलम २ मा दैनिक ज्याला अन्तराल समूहको मध्यमान निकालिएको छ, कोलम ३ मा मध्यमानको सञ्चित मान हो र कोलम ४ मध्यमानको सञ्चित मानको प्रतिशत हो। त्यसै गरी श्रमिक संख्याको सञ्चित बारम्बारता (कोलम ६) र श्रमिक संख्याको सञ्चित बारम्बारताको प्रतिशत हो। थप स्पष्ट हुनका लागि तिनीहरूको एउटा एउटा गणितीय गणना यसप्रकार छ।

कोलम २, मध्यमान: पहिलो अन्तराल समूहको लागि, $\frac{0+40}{2} = २५$

कोलम ३, मध्यमानको सञ्चित मान: दोस्रो अन्तराल समूहको लागि, $२५ + ७५ = १००$,
चौथो अन्तराल समूहको लागि, $१०० + १२५ = २२५$,

कोलम ४ मध्यमानको सञ्चित मानको प्रतिशत: पहिलो अन्तराल समूहको लागि,

$$\frac{२५}{७०००} \times १०० = ०.४\%$$

दोस्रो अन्तराल समूहको लागि, $\frac{१००}{७०००} \times १०० = १.४ \approx १\%$,

यसै गरी कोलम ६ र ७ का मानहरू श्रमिक संख्या वा बारम्बारताका लागि निकालिएका छन्।

अब, X-अक्षमा सञ्चित बारम्बारताको प्रतिशत र Y-अक्षमा सञ्चित मध्यमानको प्रतिशतलाई वितरण गरेर बक्ररेखा बनाउँदा चित्र १२.२ को लरेन्ज बक्ररेखा तयार हुन्छ।

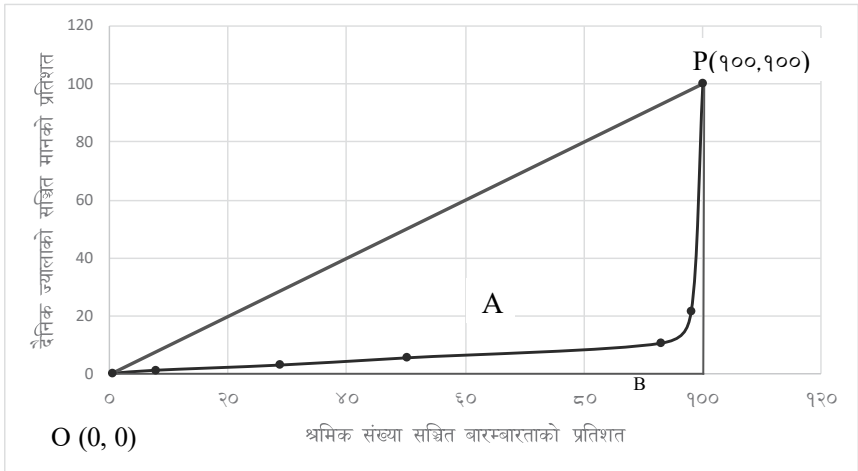
तालिका १२.५: नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ छनोटमा परेका श्रमिकहरूको दैनिक ज्याला आम्दानी वितरण

दैनिक ज्याला (रु)	मध्यमान (M)	मध्यमानको सञ्चित मान	मध्यमानको सञ्चित मानको प्रतिशत	श्रमिक संख्या (f)	श्रमिक संख्याको सञ्चित बारम्बारता	श्रमिक संख्याको सञ्चित बारम्बारताको प्रतिशत
१	२	३	४	५	६	७
० देखि ५०	२५	२५	०.४०%	११	११	०.५०%
५० देखि १००	७५	१००	१.००%	१६९	१८०	८.००%
१०० देखि १५०	१२५	२२५	३.००%	४८९	६६९	२९.००%
१५० देखि २००	१७५	४००	६.००%	४९९	१,१६८	५०.००%
२०० देखि ५००	३५०	७५०	११.००%	९९८	२,१६६	९३.००%
५०० देखि १०००	७५०	१,५००	२१.००%	११९	२,२८५	९८.००%
१००० देखि १००००	५,५००	७,०००	१००.००%	४६	२,३३१	१००.००%
जम्मा	७,०००			२,३३१		

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

चित्र १२.२ मा रेखा OP ले ज्यालाको समान वितरण बुझाउँछ । बाङ्गो रेखाले लरेन्ज बक्ररेखा देखाउँछ । यदि यो लरेन्ज बक्ररेखा OP सँग खिप्टियो भने त्यस अवस्थामा दैनिक ज्यालामा पूर्ण समानता भन्ने बुझिन्छ तर व्यवहारमा यस्तो अवस्था हुँदैन । यदि लरेन्ज बक्ररेखा OC र CP मा खिप्टियो भने त्यस अवस्थालाई ज्यालामा पूर्ण असमानता भन्ने बुझिन्छ । यस्तो अवस्था पनि व्यवहारमा सामान्तया हुँदैन । त्रिभुज OCP भित्र दुई भाग छन् । एउटा बक्ररेखाभन्दा माथि जसलाई A ले जनाइएको छ भने अर्को बक्ररेखाभन्दा मुनि जसलाई B ले जनाइएको छ । जति जति A को क्षेत्रफल बढी भयो त्यति त्यति असमानता बढी देखिन्छ । तसर्थ, यस चित्रमा पनि श्रमिकको दैनिक ज्यालामा निकै ठुलो असमानता प्रष्टै देखिन्छ ।

चित्र १२.३: नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ छनोटमा परेका श्रमिकहरूको दैनिक ज्याला आम्दानीको लरेन्ज बक्ररेखा



स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

१२.७.२ जिनीको गुणाङ्क

जिनीको गुणाङ्क असमानताको स्तर बताउने एउटा सूचक हो । यो एकाइविहीन हुने हुनाले दुई वा दुई भन्दा बढी तथ्याङ्क वितरणको असमानता तुलना गरी अध्ययन गर्नका लागि पनि निकै उपयोगी हुन्छ । यसलाई निकाल्ने प्रमुख दुई विधिहरू छन् । एउटा विधि भनेको गणितीय सूत्र प्रयोगको विधि हो, जुन निकै प्राविधिक भएको कारण यस पुस्तकमा समावेश गरिएको छैन । अर्को विधि भनेको लरेन्ज बक्ररेखा प्रयोग गर्ने विधि हो । यस विधिमा निम्नानुसारको सूत्र प्रयोग गरिन्छ ।

$$\text{जिनीको गुणाङ्क} = \frac{A \text{ को क्षेत्रफल}}{A \text{ को क्षेत्रफल} + B \text{ को क्षेत्रफल}}, \quad (१३)$$

जहाँ, A को क्षेत्रफल भन्नाले चित्र १२.३ मा देखाइएको जस्तो लरेन्ज बक्ररेखाभन्दा माथि र समान वितरण रेखा OP भन्दा तलको क्षेत्रफल बुझनुपर्दछ र B को क्षेत्रफल भन्नाले लरेन्ज बक्ररेखा भन्दा तलको क्षेत्रफल भन्ने बुझनुपर्दछ। यस्ता क्षेत्रफल गणना गर्दा कम्प्युटरका सफ्टवेयरहरूको प्रयोग गर्नुपर्ने भएकोले यहाँ यसका बारेमा चर्चा गरिएको छैन। उदाहरणका लागि नेपालको विभिन्न समयमा जिनीको गुणाङ्क कति थियो भन्ने बारे तलको तालिका हेरौं।

तालिका १२.६: नेपालको लागि उपभोग खर्चको असमानता जनाउने जिनीको गुणाङ्क

विभिन्न जीवनस्तर सर्वेक्षण	आर्थिक वर्ष	जिनीको गुणाङ्क
पहिलो सर्वेक्षण	१९९५/९६	०.३२२
दोस्रो सर्वेक्षण	२००३/०४	०
तेस्रो सर्वेक्षण	२०१०/११	०.३२८

स्रोत: केन्द्रीय तथ्याङ्क विभाग^६

जिनीको गुणाङ्क जति जति १ को नजिक भयो उति उति असमानता छ भन्ने बुझिन्छ भने जति जति ० को नजिक भयो उति उति समानता छ भन्ने बुझिन्छ। नेपालको सन्दर्भमा माथि तालिका १२.६ मा प्रस्तुत गरिएबमोजिम आर्थिक वर्ष २०१०/११ मा उपभोग खर्चको असमानता जनाउने जिनीको गुणाङ्क लगभग ०.३३ छ। जसको अर्थ उपभोग खर्चमा केही असमानता छ भन्ने बुझिन्छ। तुलनात्मक हिसाबले आर्थिक वर्ष १९९५/९६ को तुलनामा नगण्य रूपमा यो असमानता बढी भए तापनि आर्थिक वर्ष २००३/०४ तुलनामा धेरै नै कम छ भन्ने बुझिन्छ।

१२.८ अभ्यासका लागि प्रश्नहरू

क) तलका मध्ये कुनचाहिँ भनाइ गलत छ ?

- १) विचलन मापन स्पष्ट रूपमा परिभाषित भएको हुनुपर्दछ।
- २) विचलन मापन सबै आँकडाहरू आधारित भएर निकालिएको हुनुपर्दछ।
- ३) विचलन मापन थप गणितीय क्रियाहरूमा उपयोगी हुनुपर्दछ।
- ४) विचलन मापन अत्यन्त तुला वा साना आँकडाले प्रभावित पार्नु हुँदैन।
- ५) माथिका सबै भनाइ गलत छन्।

6 Central Bureau of Statistics. (2068). नेपालमा गरिबी (नेपाल जीवनस्तर सर्वेक्षण - तेस्रो २०६६/६७ मा आधारित संक्षिप्त प्रतिवेदन). Kathmandu, Nepal; Central Bureau of Statistics. Retrieved on December 10, 2021 from: <https://cbs.gov.np/pover-ty-in-nepal-2010-11/>

ख) तलका मध्ये कुनचाहिँ विचलन मापन होइन ?

- १) स्तरीय भिन्नता
- २) भेरियन्स
- ३) मध्यिका
- ४) मध्य विचलन
- ५) माथिका कुनै पनि होइन

ग) खाली ठाउँ भर्नुहोस्

- १) विस्तार = - सबैभन्दा सानो आँकडा
- २) तथ्याङ्क वितरणमा केन्द्रीय मानका बारेमा औसतहरूले बताउँछन् भने कति छरिएर रहेका छन् भनेर ले बताउँछ ।
- ३) असमानता अध्ययन गर्नका लागि र को प्रयोग गरिन्छ ।
- ४) दिइएका आँकडाहरूको उद्गम बिन्दु तथा स्केल परिवर्तन गरेर आउने मानलाई भनिन्छ ।
- ५) भेरियन्स भनेको को वर्ग हो ।

घ) जोडा मिलाउनुहोस् ।

मध्य विचलन	$\frac{x - \bar{x}}{\sigma}$
स्तरीय भिन्नता (σ)	$\frac{\sigma}{\bar{x}} \times 100$
स्तरीय भिन्नताको गुणाङ्क	$\sqrt{\frac{\sum (x - \bar{x})^2}{N}}$
विचलनको गुणाङ्क	$\frac{1}{N} \sum_{i=1}^N X_i - A $
z-score	$\frac{\sigma}{\bar{x}}$

घ) मध्य विचलन र स्तरीय भिन्नताबिचको फरक छुट्याउनुहोस् ।

ङ) विचलन मापनको महत्त्वमाथि प्रकाश पार्दै कुनै एक मध्य विचलनको मापन निकाल्ने विधि उदाहरणसहित प्रस्तुत गर्नुहोस् ।

च) दिइएको तथ्याङ्कका आधारमा स्तरीय भिन्नता निकाल्नुहोस् । साथै, यही तथ्याङ्क प्रयोग गरेर मध्यिकाबाट निकालिएको मध्य विचलन सबैभन्दा सानो हुन्छ भनी देखाउनुहोस् ।

तालिका १२.७: नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ छनोटमा परेका श्रमिकहरूको दैनिक ज्याला आम्दानी वितरण

दैनिक ज्याला (रु)	श्रमिक संख्या (f)
० देखि ५०	११
५० देखि १००	१६९
१०० देखि १५०	४८९
१५० देखि २००	४९९
२०० देखि ५००	९९८
५०० देखि १०००	११९
१००० देखि १००००	४६
जम्मा	२,३३१

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

१२.९ सन्दर्भ सामग्री

विचलनको मापन अध्ययन गर्नका लागि निम्न सामग्री उपयोगी छ ।

१. Fundamentals of Statistics by S.C. Gupta, Sixth Revised & Enlarged Edition, Himalaya Publishing House, 2009



सहसम्बन्ध र यसको प्रयोग

१३.१ परिचय

यस अगाडिका पाठहरूमा हामीले सामान्यतया एकल चरको विशेषताका बारेमा सामान्य चर्चा गरेका छौं। केन्द्रीय प्रवृत्तिको मापन, विचलनको मापनले खासगरी एकल चरको विशेषता व्यक्त गरेका हुन्छन्। यस पाठमा हामी दुई चरहरूले एकआपसमा देखाउने सम्बन्धका बारेमा चर्चा गर्नेछौं। कुनै एक चरमा आउने परिवर्तनले अर्को चरलाई असर पार्नुलाई सहसम्बन्ध हुनु भनिन्छ। अर्को शब्दमा भन्दा सहसम्बन्धले एउटा चरको मान बढ्दा वा घट्दा अर्को चरको मान बढ्छ वा घट्छ वा परिवर्तनविहीनता छ भनेर देखाउँछ।

उदाहरणका लागि तालिका १३.१ मा प्रदेश १ अन्तर्गतका १४ वटा जिल्लामा मकै र गहुँ लागेको क्षेत्रफल तथा तिनीहरूको उत्पादन विवरण प्रस्तुत गरिएको छ। यसमा प्रस्तुत गरिएका चरहरू: बाली लागेको क्षेत्रफल तथा उत्पादन एक आपसमा सम्बन्धित छन् वा छैनन् भनेर सहसम्बन्धको अध्ययनबाट हेर्न सकिन्छ। अर्को शब्दमा भन्नुपर्दा, विभिन्न जिल्लाहरू बिच बाली लागेको क्षेत्रफल र उत्पादन दुवै फरक-फरक छन् तर त्यो फरकपनाले कुनै सम्बन्ध स्थापित गर्छ वा गर्दैन भन्ने कुरा सहसम्बन्धको अध्ययनले देखाउँछ। संक्षेपमा भन्नुपर्दा एउटा चरमा परिवर्तन आउँदा अर्को चरमा पनि परिवर्तन आउँछ भने ती दुईवटा चरहरूलाई सहसम्बन्ध भएको चरहरू भनेर भनिन्छ।

सहसम्बन्धको अध्ययनले दुई चरहरूबिचको सम्बन्धको मान र त्यसको दिशा (Degree and Direction) बताउँछ। अर्थात्, यसको अध्ययनले चरहरू सामान्य रूपमा सम्बन्धित छन् वा प्रगाढ रूपमा सम्बन्धित छन् भनेर भन्दछ। तर यसले कारण र प्रभावका बारेमा भने बताउन सक्दैन। जस्तै: तालिका १३.१ मा प्रस्तुत गरिएका चरहरूमध्ये मकै र गहुँको उत्पादनका दुई चरहरू सहसम्बन्ध भएका चरहरू देखिन्छ। यहाँ, सहसम्बन्ध विश्लेषणले मकैको उत्पादन बढेर गहुँको उत्पादन बढेको हो वा गहुँको उत्पादन बढेर मकैको उत्पादन बढेको हो भनेर भन्न सक्दैन। कहिलेकाहीँ कुनै पनि रूपमा सम्बन्ध नराख्ने चरहरूबिच पनि सहसम्बन्ध विश्लेषणको तथ्याङ्कीय सूत्रले सहसम्बन्ध देखाएको हुनसक्दछ। उदाहरणका लागि सहसम्बन्ध तथ्याङ्कीय सूत्रले नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ को तथ्याङ्कमा बच्चाको तौल र परिवारले मल तथा विषादीमा गरेको खर्चबिच सकारात्मक सहसम्बन्ध देखाउँछ। यस्तो सहसम्बन्ध वास्तवमा अर्थहीन हो र यसमा कुनै सत्यता हुँदैन। सत्यता वा वास्तविकता नभएका गणितीय सहसम्बन्धलाई वाहियात सहसम्बन्ध (Spurious or Non-sense Correlation) भनिन्छ।

तालिका १३.१: प्रदेश १ अन्तर्गत मकै र गहुँको क्षेत्रफल तथा उत्पादन

जिल्ला	मकै		गहुँ	
	बाली लागेको क्षेत्रफल (हे.)	उत्पादन (मे.टन)	बाली लागेको क्षेत्रफल (हे.)	उत्पादन (मे.टन)
ताप्लेजुङ	९,५९८	२६,८७४	१,५३५	३,५४६
संखुवासभा	१२,८५८	२८,१४७	७१५	१६८१
सोलुखुम्बु	१४,०५०	३५,३०९	२,६६९	५,६०३
पाँचथर	१८,९८३	४१,१०३	३,९०४	८,४३३
इलाम	३३,८१५	१,२६,४५८	४,६०८	१३,८२२
तेह्रथुम	१२,११७	३२,६१५	२,४००	५,०६४
धनकुटा	१४,००४	४५,६४७	१,३०२	२,९३०
भोजपुर	३६,१२१	९९,२१३	४००	८२०
खोटाङ	४२,९४५	९८,९८६	५,०३०	११,५६९
ओखलढुङ्गा	१२,४००	३५,८८४	२,१४५	४,५६९
उदयपुर	१७,९०७	४२,३५०	५,१३०	१४,१०८
भद्रापा	३९,३०२	१,५८,९८५	७,१२५	२३,०१४
मोरङ	१७,२५५	६८,२८६	१२,६३९	३९,४३४
सुनसरी	८,९४९	२७,९००	१२,९५९	४२,११७
जम्मा	२,९०,३०४	८,६७,७१८	६२,५६९	१,७६,७०७

स्रोत: कृषि तथा पशुपन्छी विकास मन्त्रालय, नेपाली कृषि सम्बन्धी तथ्याङ्गीय सूचना, २०७६/७७^७

१३.२ सहसम्बन्धका प्रकार र यसका विशेषता

चरहरूको परिवर्तन हुने प्रकृतिअनुसार सहसम्बन्धका प्रकार र तिनीहरूको विशेषता निर्धारण हुन्छ। यसलाई मुख्य रूपमा तिन भागमा विभाजन गरेर अध्ययन गर्न सकिन्छ।

१३.२.१ सकारात्मक र नकारात्मक सहसम्बन्ध

सहसम्बन्ध विश्लेषणले हामीलाई चरहरूबिच रहेको सम्बन्धको दिशा बताउँदछ। यस्तो दिशा सकारात्मक वा नकारात्मक हुन सक्दछ। एउटा चरको मान बढ्दा अर्को चरको मान पनि बढ्छ भने त्यस्तो सम्बन्धलाई सकारात्मक सहसम्बन्ध (Positive Correlation)

7 Ministry of Agriculture and Livestock Development. *Statistical Information on Nepalese Agriculture*. Kathmandu, Nepal; Ministry of Agriculture and Livestock Development. Retrieved on December 3, 2021 from: <https://www.moald.gov.np/publication/Agriculture%20Statistics>

भनिन्छ भने एउटा चरको मान बढ्दा अर्को चरको मान घट्छ भने त्यस्तो सम्बन्धलाई नकारात्मक सहसम्बन्ध (Negative Correlation) भनिन्छ । त्यस्तै, एउटा चरको मानमा परिवर्तन आउँदा अर्को चरमा कुनै असर देखिँदैन भने त्यसलाई सहसम्बन्ध नभएको (No/Zero Correlation) भनिन्छ । जस्तै: सामान्यतया हाम्रो आमदानी बढ्यो भने खर्च पनि बढ्ने गर्छ, यस मानेमा आमदानी र खर्चको सकारात्मक सहसम्बन्ध छ भनेर बुझिन्छ । त्यस्तै: कुनै सामानको मूल्य बढ्यो भने त्यसको माग घट्न सक्छ, यस्तो अवस्थामा सामानको मूल्य र मागको नकारात्मक सहसम्बन्ध हुन्छ । तर अत्यावश्यक सामान जस्तै: औषधीको मूल्य बढ्दा वा घट्दा औषधीको माग बढ्ने वा घट्ने भन्ने नहुन सक्छ, यस्तो अवस्थामा औषधीको मूल्य र मागको बिचमा सहसम्बन्ध नभएको मानिन्छ । तालिका १३.१ मा प्रस्तुत भएका चरहरूमा मकै वा गहुँको बाली लागेको क्षेत्रफल र तिनीहरूको उत्पादन सकारात्मक सहसम्बन्धमा छन् । अर्थात्, क्षेत्रफल बढ्दा उत्पादन पनि बढेको छ । जल तथा मौसम विज्ञान विभागको तथ्याङ्कअनुसार विभिन्न स्थानको उचाइ र वर्षाको मात्राबिच नकारात्मक सहसम्बन्ध छ । अर्थात्, स्थानको उचाइ बढ्दै जाँदा वर्षाको मात्रा घट्दै गएको छ ।

१३.२.२ रेखीय तथा गैर-रेखीय सहसम्बन्ध

सहसम्बन्धको अध्ययनले दुई चरहरूबिच रहेको सम्बन्धको रूप वा आकार बताउँछ । यदि एउटा चरमा एक एकाइ परिवर्तन हुँदा अर्को चरमा स्थिर परिवर्तन हुन्छ भने ती चरहरूबिचको सहसम्बन्ध रेखीय (Linear Correlation) हुन्छ । यदि एउटा चरमा एक एकाइ परिवर्तन हुँदा अर्कोमा स्थिर परिवर्तन हुँदैन भने ती चरहरूबिचको सहसम्बन्ध गैर-रेखीय (Non - linear Correlation) हुन्छ । गैररेखीय सहसम्बन्ध अध्ययन गर्न जटिल हुने भएकोले यस पाठमा हामी रेखीय सहसम्बन्धका बारेमा मात्र अध्ययन गर्दछौं । सामान्यतया, वास्तविक चरहरूले परिभाषामा भनेको जस्तो पूर्ण रूपमा रेखीय सहसम्बन्ध स्थापित गर्दैनन्, तर हामीले रेखीय सहसम्बन्ध मानेर अध्ययन गर्दछौं । उदाहरणका लागि नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ मा पाँच वर्ष मुनिका बालबालिकाको उचाइ तथा तौल मापन गरिएको थियो । यी दुई चरहरूबिच पूर्ण सकारात्मक सहसम्बन्ध छ । अर्थात्, उचाइ बढ्दै जाँदा तौल पनि बढ्दै गएको छ । यहाँ, उचाइ र तौलको पूर्ण रूपमा रेखीय सहसम्बन्ध नभए तापनि हामीले रेखीय सहसम्बन्ध मानेर अध्ययन गर्न सकिन्छ । रेखीय सहसम्बन्ध भएका चरहरूको रेखाचित्र एउटा सिधा रेखा हुन्छ । त्यसैगरी, उमेर र उचाइ, उमेर र तौलबिचको सहसम्बन्ध पनि रेखीय सहसम्बन्ध मानी अध्ययन गर्न सकिन्छ ।

१३.२.३ सम्बन्धको स्तर (डिग्री)

सहसम्बन्धको अध्ययनले दुई चरहरू बिचको सम्बन्धको डिग्री (Level of Relationship) बारे बताउन सक्छन्। दुई चरहरू पूर्ण रूपमा, बलियो रूपमा, सामान्य रूपमा सम्बन्धित हुन पनि सक्दछन् वा कुनै पनि सम्बन्ध नभएको पनि हुन सक्दछन्।

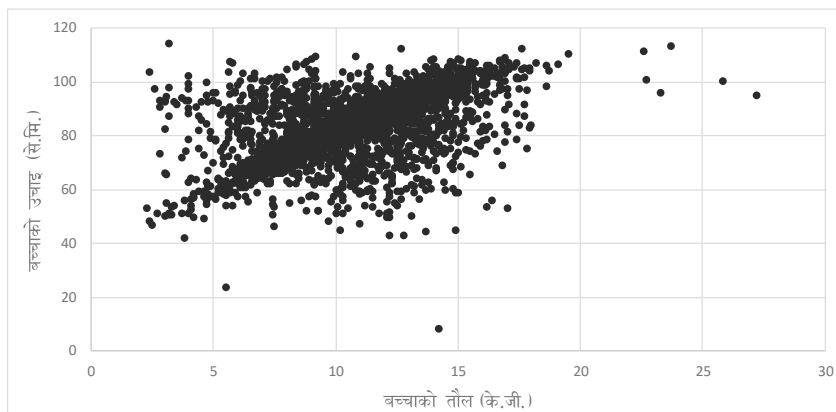
१३.३ सहसम्बन्ध अध्ययन गर्ने विधि

दुई चरहरूबिचको सहसम्बन्ध अध्ययन गर्नका लागि विभिन्न विधिहरूको प्रयोग गर्ने गरिन्छ। यस पाठमा हामी स्क्याटर चित्र र सहसम्बन्ध गुणाङ्कको गणना गर्ने विधिको बारेमा चर्चा गर्नेछौं।

१३.३.१ स्क्याटर चित्रको माध्यमबाट सहसम्बन्ध अध्ययन

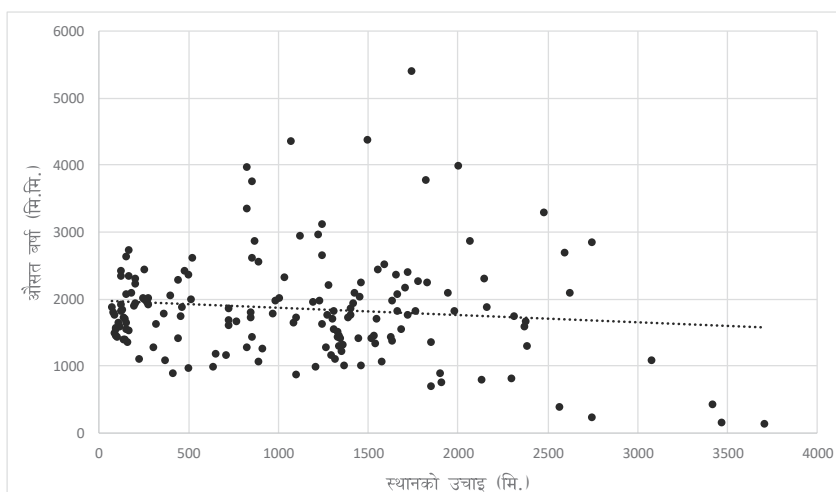
दुई चरहरूको मानलाई ग्राफमा राखेर अध्ययन गरिने विधिलाई स्क्याटर चित्रको प्रयोग गर्ने विधि भनिन्छ। यो विधि सबैभन्दा सरल र सजिलो खालको विधि हो। यस विधिको माध्यमबाट सामान्यतया सकारात्मक वा नकारात्मक सहसम्बन्ध कस्तो छ भनेर पनि छुट्याउन सकिन्छ। सहसम्बन्ध रेखीय हो वा गैरेखीय हो भनेर पनि हेर्न सकिन्छ। त्यसैगरी सहसम्बन्ध कतिको बलियो छ भनेर पनि अनुमान गर्न सकिन्छ। तर यकिनका साथ भन्नका लागि गुणाङ्कको प्रयोग गर्नुपर्दछ। उदाहरणका लागि केही चित्र १३.१ (क) र १३.१ (ख) मा विभिन्न चरहरूबिचका विभिन्न प्रकृतिका सहसम्बन्धहरू देखाइएको छ। चित्र १३.१ (क) मा दुई चरहरूबिच बलियो सकारात्मक सहसम्बन्ध देखाइएको छ। यस चित्रमा बच्चाको तौल बढ्दै जाँदा उचाइ पनि बढेको छ। अर्थात् उचाइ बढ्दा तौल पनि बढेको छ। समग्र स्क्याटर चित्र तलबाट माथि तर्फ ढल्केको देखिन्छ। तसर्थ, यस चित्रमा तौल र उचाइको बलियो सकारात्मक सहसम्बन्ध देखिन्छ। अर्कोतर्फ चित्र १३.१ (ख) मा स्थानको उचाइ र औसत वर्षाबिच कमजोर नकारात्मक सहसम्बन्ध देखाइएको छ। यहाँ स्थानको उचाइ बढ्दै जाँदा सूक्ष्म तरिकाले वर्षाको मात्रा घटिरहेको छ। समग्र चित्र देब्रेबाट दाहिनेतर्फ हल्का रूपमा माथिबाट तलतिर ढल्किएको छ। कतिपय चित्रहरूबाट हेरेकै भन्ना सहसम्बन्धको विशेषता यकिन गरेर भन्न नसकिने पनि हुन्छ। जस्तै चित्र १३.१ (ग) मा परिवारमा अनुपस्थित व्यक्तिको उमेर र उनीहरूले पठाएको विप्रेषणको स्क्याटर चित्र देखाइएको छ। चित्रलाई सामान्य रूपमा हेर्दा कस्तो किसिमको सहसम्बन्ध छ भनेर यकिन गर्न कठिन हुन्छ। तथापि यी दुई चरहरूबिच निकै कमजोर नकारात्मक सहसम्बन्ध छ। यस्तो अवस्थामा यिनीहरूबिच सहसम्बन्ध छैन भन्दा पनि फरक पर्दैन।

चित्र नं. १३.१ (क): पाँच वर्ष मुनिका बच्चाको तौल र उचाइको स्क्याटर चित्र



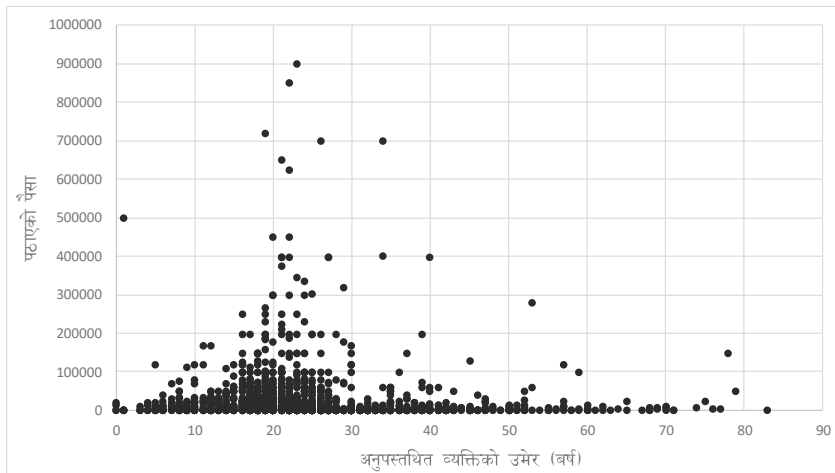
स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

चित्र नं. १३.१ (ख): स्थानको उचाइ र तिस वर्षको औसत वर्षको स्क्याटर चित्र



स्रोत: जल तथा मौसम विज्ञान विभाग

चित्र नं. १३.१ (ग): परिवारमा अनुपस्थित व्यक्तिको उमेर र उनीहरूले पठाएको पैसाको स्क्वाटर चित्र



स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

१३.३.२ सहसम्बन्ध गुणाङ्कको गणना

सहसम्बन्ध अध्ययन गर्ने अर्को प्रमुख विधि भनेको यसको गुणाङ्क (Correlation Coefficient) निकाल्ने विधि हो। यो विधि निकै प्रचलनमा रहेको छ। यस विधिबाट चरहरूबीच सकारात्मक वा नकारात्मक, बलियो वा कमजोर कस्तो सहसम्बन्ध छ भनेर यकिनका साथ बताउन सकिन्छ। त्यस्तै सहसम्बन्धको स्तरका बारेमा पनि परिमाणात्मक रूपले व्याख्या गर्न सकिन्छ। सहसम्बन्धको गुणाङ्कलाई अंग्रेजी सानो अक्षर “r” ले जनाउने गरिन्छ। यस पाठमा हामी कार्ल पियर्सनको सहसम्बन्ध गुणाङ्क (Karl – Pearson’s Correlation Coefficient) निकाल्ने तरिका अध्ययन गर्दछौं।

कुनै दुईवटा चरहरू X र Y छन् भने उक्त चरहरूको सहसम्बन्ध गुणाङ्क निकाल्ने सूत्र यसप्रकार छ।

$$r = \frac{\sum(X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum(X - \bar{X})^2} \sqrt{\sum(Y - \bar{Y})^2}} \quad (9)$$

जहाँ, (\bar{X}) भनेको X को औषत र (\bar{Y}) भनेको Y को औषत हो। यसको निकाल्ने विधिको बारेमा उदाहरणसहित अभ्यास गरौं।

उदाहरण: माथि तालिका १३.१ मा दिइएको मकैको बाली लागेको क्षेत्रफल र मकैको उत्पादनबिचको सहसम्बन्ध निकालौं।

तालिका १३.१: प्रदेश १ अन्तर्गत मकै र गहुँको क्षेत्रफल तथा उत्पादन

जिल्ला	मकै		गहुँ	
	बाली लागेको क्षेत्रफल (हे.)	उत्पादन (मे.टन)	बाली लागेको क्षेत्रफल (हे.)	उत्पादन (मे.टन)
ताप्लेजुड	९,५९८	२६,८७४	१,५३५	३,५४६
संखुवासभा	१२,८५८	२८,१४७	७१५	१,६८१
सोलुखुम्बु	१४,०५०	३५,३०९	२,६६९	५,६०३
पाँचथर	१८,९८३	४१,१०३	३,९०४	८,४३३
इलाम	३३,८१५	१,२६,४५८	४,६०८	१३,८२२
तेह्रथुम	१२,११७	३२,६१५	२,४००	५,०६४
धनकुटा	१४,००४	४५,६४७	१,३०२	२,९३०
भोजपुर	३६,१२१	९९,२१३	४००	८२०
खोटाङ	४२,९४५	९८,९८६	५,०३०	११,५६९
ओखलढुङ्गा	१२,४००	३५,८८४	२,१४५	४,५६९
उदयपुर	१७,९०७	४२,३५०	५,१३०	१४,१०८
भापा	३९,३०२	१,५८,९८५	७,१२५	२३,०१४
मोरङ	१७,२५५	६८,२८६	१२,६३९	३९,४३४
सुनसरी	८,९४९	२७,९००	१२,९५९	४२,११७
जम्मा	२,९०,३०४	८,६७,७१८	६२,५६१	१,७६,७०७

स्रोत: कृषि तथा पशुपन्छी विकास मन्त्रालय, नेपाली कृषि सम्बन्धी तथ्याङ्कीय सूचना, २०७६/७७^४

१३.४ सहसम्बन्ध गुणाङ्कको व्याख्या र प्रयोग

सहसम्बन्धको गुणाङ्क निकाल्ने विधिका बारेमा जानकारी गरिसकेपछि यसको व्याख्या र प्रयोग कसरी गर्ने भन्ने बारेमा जानकारी राख्नुपर्ने हुन्छ। सहसम्बन्धको गुणाङ्कको व्याख्या गर्दा यसका मानहरू कति हुनसक्दछन् भन्नेबारे जानकारी हुनुपर्दछ। सहसम्बन्धको गुणाङ्क अर्थात्, r को मान १ देखि -१ सम्ममात्र हुन्छ। यदि r को मान १ भन्दा ठुलो वा -१ भन्दा सानो आयो भने हिसाब बिग्रियो भन्ने बुझेर पुनः हिसाब गर्नुपर्दछ। यहाँ r को मान व्याख्या गर्नका लागि उपयुक्त हुने केही दृष्टान्तहरू दिइएका छन्।

क) यदि r को मान १ आयो भने ती दुई चरहरूबिच पूर्ण रूपको सकारात्मक सहसम्बन्ध (Positive Correlation) छ भन्ने बुझ्नुपर्दछ।

ख) यदि r को मान -1 आयो भने ती दुई चरहरूबिच पूर्ण रूपको नकारात्मक सहसम्बन्ध (Negative Correlation) छ भन्ने बुझ्नुपर्दछ ।

ग) यदि r को मान शून्य आयो भने ती दुई चरहरूबिच कुनै पनि सहसम्बन्ध (No Correlation) छैन भन्ने बुझ्नुपर्दछ ।

घ) r को मान जति जति शून्य नजिक आयो त्यति त्यति ती दुई चरहरूबिच कमजोर सहसम्बन्ध छ भन्ने बुझ्नुपर्दछ र जति जति 1 को नजिक आयो त्यति त्यति ती दुई चरहरूबिच बलियो सहसम्बन्ध छ भन्ने बुझ्नुपर्दछ ।

ङ) कतिपय अवस्थामा कुनै पनि सम्बन्ध नभएका चरहरूबिच पनि संयोगवश r को मान शून्य नआई अन्य कुनै अंक आउन सक्दछ । यस्तो अवस्थामा सहसम्बन्ध अर्थहीन हुन्छ ।

माथि उल्लिखित बुँदाहरूलाई ध्यानमा राखेर सहसम्बन्ध गुणाङ्कको प्रयोग गर्नुपर्दछ । सामान्यतया 0.7 भन्दा माथि आउनुलाई बलियो सकारात्मक सहसम्बन्ध भनेर बुझिन्छ भने -0.7 भन्दा तल आउनुलाई बलियो नकारात्मक सहसम्बन्ध भनेर बुझिन्छ । माथिको उदाहरणमा r को मान 0.91 आएको हुनाले मकै बाली लागेको क्षेत्रफल र उत्पादनबिच बलियो सकारात्मक सहसम्बन्ध छ भनेर भन्न सकिन्छ । हिसाब गर्न र बुझ्न साजिलो भएकोले जुनसुकै क्षेत्रमा पनि कुनै चरहरूको सम्बन्ध हेर्नुपर्दा सहसम्बन्धको मानाङ्क निकाल्ने गरिन्छ । पारिमाणात्मक आँकडामा सबै क्षेत्रहरूमा यो उक्तकै उपयोगी देखिन्छ । अन्य तथ्याङ्क विश्लेषणको विधि छनौटको लागि समेत यो उपयोगी हुन्छ ।

नाममात्र तहमा मापन गरिएका आँकडामा सहसम्बन्धको मानाङ्क प्रयोग गर्नु उचित हुँदैन । उदाहरणका लागि सर्वेक्षणमा छानिएका व्यक्तिहरूको कोड र उनीहरूको उमेरबिचको सहसम्बन्धको अध्ययनले कुनै अर्थ राख्दैन । क्रमिक तहमा मापन गरिएका आँकडामा भने सहसम्बन्धको प्रयोग गर्न सकिन्छ । यस किसिमको तथ्याङ्कमा सहसम्बन्ध अध्ययन गर्ने विधिलाई क्रम सहसम्बन्ध विधि (Rank Correlation Method) । यस विधिको बारेमा यस पुस्तकमा चर्चा गरिएको छैन ।

१३.५ कोअफिसियन्ट अफ डिटरमिनेशन

कोअफिसियन्ट अफ डिटरमिनेशन बुझ्नका लागि पहिला दुई प्रकारका चरहरूका बारेमा बुझौं । सहसम्बन्धित चरहरूलाई एउटालाई निर्भर चर (Dependent Variable) र अर्कोलाई स्वतन्त्र चर (Independent Variable) को रूपमा लिन सकिन्छ । स्वतन्त्र चरको मानमा आएको परिवर्तनले निर्भर चरमा पनि परिवर्तन आउँछ । जस्तै: बाली

लागेको क्षेत्रफललाई स्वतन्त्र चर मान्ने हो भने मकैको उत्पादनलाई निर्भर चर मान्न सकिन्छ। किनभने बाली लागेको क्षेत्रफल घटबढ गराउँदा मकैको उत्पादनमा पनि घटबढ हुन्छ। हुन त मकैको उत्पादन घटबढ बनाउने अरू चरहरू पनि छन् जस्तै प्रयोग गरेको मलको मात्रा, वर्षाको मात्रा, सिँचाई, गोडमेल, रोगव्याधि इत्यादि। यीमध्ये बाली लागेको क्षेत्रफलमा भएको परिवर्तनले उत्पादनमा कति फरक पऱ्यो भन्नका लागि कोअफिसियन्ट अफ डिटरमिनेशनको प्रयोग गरिन्छ। यसलाई निकाल्न सहसम्बन्धको गुणाङ्कलाई वर्ग गर्नुपर्दछ र यसलाई R^2 वा r^2 (R Squared) भनिन्छ। अर्थात्,

$$\text{Coefficient of Determination} = r^2. \quad (२)$$

उदाहरणः माथिको उदाहरणमा मकै लागेको क्षेत्रफल र मकैको उत्पादनबिच सहसम्बन्धको गुणाङ्क ०.९१ आएको थियो। यसलाई वर्ग गर्दा, ०.८२८१ हुन आउँछ। यसबाट के भन्न सकिन्छ भने मकैको उत्पादनमा भएको ८२.८१% विचलन (Variation) वा परिवर्तन मकै लागेको क्षेत्रफलमा भएको विचलनले गर्दा हो। कोअफिसियन्ट अफ डिटरमिनेशनको अधिक प्रयोग प्रतिगमन विश्लेषणमा (Regression Analysis) गरिन्छ। यसले एउटा चरले अर्को चरलाई कुन हदसम्म असर गरिरहेको छ भनेर सामान्य धारणा दिन्छ।

१३.६ अभ्यासका लागि प्रश्नहरू

क) खाली ठाउँ भर्नुहोस्।

- १) दुईवटा पूर्ण रूपमा असम्बन्धित चरहरूबिच पनि सहसम्बन्धको गुणाङ्क सकारात्मक वा नकारात्मक आयो भने त्यसलाई भनिन्छ।
- २) सहसम्बन्धको गुणाङ्कको वर्गलाई भनिन्छ।
- ३) एउटा चरमा आएको परिवर्तनले अर्को चरमा पनि परिवर्तन आउँछ भने ती चरहरूबिच भएको मानिन्छ।
- ४) वस्तुको मूल्य बढ्दा उपभोगको वा मागको मात्रा घट्दछ। यस्तो अवस्थामा ती दुई चरहरूबिचको सहसम्बन्ध हुन्छ।
- ५) सहसम्बन्धले एउटा चरको मान हुँदा अर्को चरको मान पनि हुन्छ भनेर देखाउँछ।

ख) तलका भनाइ सही या गलत के छन्, कारणसहित बताउनुहोस्।

- १) r को मान ८.६ आउनुलाई अर्थहीन सहसम्बन्ध भनिन्छ।
- २) एउटा चरको मान बढ्दा अर्को चरको मानमा कुनै असर पर्दैन भने ती दुई चरहरूबिच नकारात्मक सहसम्बन्ध छ।

- ३) R^2 ले एउटा चरले अर्को चरलाई कुन हदसम्म असर गरिरहेको छ भनेर सामान्य धारणा दिन्छ ।
- ४) सहसम्बन्ध विश्लेषणको सहयोगले कारण र असरका बारेमा अध्ययन गर्न सकिन्छ ।
- ५) कतिपय अवस्थामा कुनै पनि सम्बन्ध नभएका चरहरूबिच पनि संयोगवश r को मान शून्य नआई अन्य कुनै अंक आउन सक्दछ ।
- ग) तलका चरहरूबिच सामान्यतया सकारात्मक, नकारात्मक वा शून्य सहसम्बन्ध के हुन्छ, कारणसहित उल्लेख गर्नुहोस् ।
- १) वस्तुको मूल्य र उपभोग
 - २) मौसमको तापक्रम र बाक्लो लुगाको बिक्री
 - ३) कृषि क्षेत्रको उत्पादन र कुल गार्हस्थ्य उत्पादन
 - ४) सरकारले विकासमा गरेको खर्च र कुल गार्हस्थ्य उत्पादन
 - ५) परिवारको आम्दानी र खर्च
- घ) छोटकरीमा व्याख्या गर्नुहोस् ।
- १) सहसम्बन्ध
 - २) सहसम्बन्धको गुणाङ्क
 - ३) R^2
 - ४) अर्थहीन सहसम्बन्ध
- ङ) परिवार नियोजनका साधनको प्रयोगकर्ता बढ्दा जन्मदर घट्दछ भने यी दुई चरहरूमा कुनचाहिँ स्वतन्त्र चर र कुनचाहिँ निर्भर चर होलान् ?
- च) आर्थिक गणना, २०७८ बमोजिम संखुवासभा जिल्लामा कृषि, वन तथा मत्स्य क्षेत्रका प्रतिष्ठानको संख्या र त्यहाँ कार्यरत मानिसहरूको संख्या तालिका १३.३ मा दिइएको छ ।

तालिका १३.३ संखुवासभा जिल्लामा कृषि, वन तथा मत्स्य क्षेत्रका प्रतिष्ठान र त्यहाँ कार्यरत मानिसहरूको संख्या

स्थानीय तहको नाम	प्रतिष्ठानको संख्या	संलग्न व्यक्तिहरूको संख्या
भोटखोला गाउँपालिका	२९३	१,४३४
मकालु गाउँपालिका	४३७	१,१४८
सिलीचोड गाउँपालिका	४०२	२,१०४
चिचिला गाउँपालिका	१९९	४७९

सभापोखरी गाउँपालिका	२२०	५४७
खाँदबारी गाउँपालिका	१,९१३	६,५९७
पाँचखपन नगरपालिका	५९१	१,३३१
चैनपुर नगरपालिका	१,०३५	२४३४
मादी नगरपालिका	४२१	१,१८१
धर्मदेवी नगरपालिका	५२७	१,९७२
जम्मा	६,०३८	१९,२२७

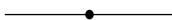
स्रोत: केन्द्रीय तथ्याङ्क विभाग, आर्थिक गणना, २०१८

उक्त तालिकामा प्रस्तुत गरिएका प्रतिष्ठान संख्या र संलग्न व्यक्ति संख्याबिच सहसम्बन्ध छ वा छैन, सहसम्बन्धको गुणाङ्क निकालेर व्याख्यात्मक टिप्पणीसहित उत्तर दिनुहोस् । साथै, R Squared को मान निकालेर व्याख्या गर्नुहोस् ।

१३.७ सन्दर्भ सामग्री

सहसम्बन्धको विस्तृत अध्ययनका लागि निम्न सामग्री उपयोगी छन् ।

१. Statistical Techniques in Business & Economics by Douglas A. Lind, William G. Marchal & Samuel A. Wathen, Seventeenth Edition, McGraw-Hill Education, 2018
२. Fundamentals of Statistics by S.C. Gupta, Sixth Revised & Enlarged Edition, Himalaya Publishing House, 2009



समय श्रृङ्खला विश्लेषण

१४.१ परिचय

समय श्रृङ्खला भनेको तथ्याङ्कको एक प्रकारको संरचना हो । यसलाई थप व्याख्या गर्नुपूर्व तथ्याङ्कको संरचनाका आधारमा यसका प्रकारका बारेमा संक्षिप्त जानकारी लिऔं ।

क) क्रस-सेक्शन तथ्याङ्क: समयको कुनै एक बिन्दुमा लिइएको वा प्रस्तुत गरिएको तथ्याङ्क क्रस-सेक्शन तथ्याङ्क (Cross-Section Data) हो । अर्को शब्दमा भन्नुपर्दा, प्रत्येक आँकडा समयको कुनै एक बिन्दुमा व्यक्ति, फर्म, घरपरिवार, सहर, देश, वा अन्य एकाइबाट लिइएको सूचना हो । खास समयमा सञ्चालन गरिएका गणना, सर्वेक्षण वा अन्य स्रोतबाट प्राप्त तथ्याङ्क यसका उदाहरण हुन् । जस्तै: सन् २०१८ मा सञ्चालन गरिएको नेपालको आर्थिक गणनाका तथ्याङ्कहरू, सन् २०१६ मा सञ्चालन गरिएको नेपाल जनसांख्यिक तथा स्वास्थ्य सर्वेक्षणका तथ्याङ्कहरू आदि । तालिका १४.४ मा क्रस-सेक्शन तथ्याङ्कको उदाहरण दिइएको छ ।

तालिका १४.१: नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ को तथ्याङ्कको सानो अंश

जिल्लाको नाम	स्थानीय तहको कोड	परिवारको कोड	वार्षिक आमदानी
संखुवासभा	१३	१	२,२७,७३४
इलाम	१५	१५	२,१४,५७२
धनकुटा	१५	६	८५,८००
उदयपुर	४३	२	१,४८,११०
भापा	१२	४	१,५७,६७२
मोरङ	१३	१३	२,९९,३४०

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

ख) समय श्रृङ्खला: एक वा एकभन्दा बढी चरहरूको मान विभिन्न समय बमोजिम व्यक्त गरिएका तथ्याङ्क समय श्रृङ्खला (Time Series) हुन् । अर्थात्, समय अनुसार क्रमबद्ध गरिएका तथ्याङ्क नै समय श्रृङ्खला हुन् । उदाहरणका लागि तालिका १४.४ मा समय श्रृङ्खला तथ्याङ्क प्रस्तुत गरिएको छ ।

तालिका १४.२: कुल गार्हस्थ्य उत्पादन तथा वृद्धिदर समय श्रृङ्खला

आर्थिक वर्ष	कुल गार्हस्थ्य उत्पादन (चालु मूल्यमा)	कुल गार्हस्थ्य उत्पादन (स्थिर मूल्यमा)*	कुल गार्हस्थ्य उत्पादनको वृद्धिदर (चालु मूल्यमा)	कुल गार्हस्थ्य उत्पादनको वृद्धिदर (स्थिर मूल्यमा)
२०३१/३२	१६,६०१	३,४९,५२३	-	-
२०३२/३३	१७,३९४	३,६१,६४६	४.७८	३.४७
२०३३/३४	१७,२८०	३,६५,२९९	-०.६६	१.०१
२०३४/३५	१९,७२७	३,७६,७२५	१४.१६	३.१३
२०३५/३६	२६,१२८	३,८४,७५०	३२.४५	२.१३
२०३६/३७	२३,३५१	३,७८,९६३	-१०.६३	-१.५
२०३७/३८	२७,३०७	४,१६,९७८	१६.९४	२७,३०७
२०३८/३९	३०,९८८	४,३५,३७३	१३.४८	४.४१
२०३९/४०	३३,८२१	४,३७,१४७	९.१४	०.४१
२०४०/४१	३९,२९०	४,७५,६०५	१६.१७	८.८
२०४१/४२	४६,५८७	५,०१,२०१	१८.५७	५.३८
२०४२/४३	५५,७३४	५,२४,०८५	१९.६३	४.५७
२०४३/४४	६३,८६४	५,३२,९९३	१४.५९	१.७
२०४४/४५	७६,९०६	५,७४,०१६	२०.४२	७.७
२०४५/४६	८९,२७०	५,९८,८५७	१६.०८	४.३३
२०४६/४७	१,०३,४१६	६,२६,६१५	१५.८५	४.६४
२०४७/४८	१,२०,३७०	६,६६,५०८	१६.३९	६.३७
२०४८/४९	१,४९,४८७	६,९३,८८८	२४.१९	४.११
२०४९/५०	१,७१,४७४	७,२०,५८१	१४.७१	३.८५
२०५०/५१	१,९९,२७२	७,७९,८०७	१६.२१	८.२२
२०५१/५२	२,१९,१७५	८,०६,८५४	९.९९	३.४७
२०५२/५३	२,४८,९१३	८,४९,९२१	१३.५७	५.३४
२०५३/५४	२,८०,५१३	८,९४,६३५	१२.७	५.२६
२०५४/५५	३,००,८४५	९,२०,९५६	७.२५	२.९४
२०५५/५६	३,४२,०३६	९,६२,२४९	१३.६९	४.४८
२०५६/५७	३,७९,४८८	१०,२१,०९५	१०.९५	६.१२

२०५७/५८*	४,४१,५१९	१०,७८,५६७	१६.३५	५.६३
२०५८/५९	४,५९,४४३	१०,७९,८६३	४.०६	०.१२
२०५९/६०	४,९२,२३१	११,२२,४६५	७.१४	३.९५
२०६०/६१	५,३६,७४९	११,७५,०२५	९.०४	४.६८
२०६१/६२	५,८९,४१२	१२,१५,९०५	९.८१	३.४८
२०६२/६३	६,५४,०८४	१२,५६,८१५	१०.९७	३.३६
२०६३/६४	७,२७,८२७	१२,९९,६९३	११.२७	३.४१
२०६४/६५	८,१५,६५८	१३,७९,०३४	१२.०७	६.१
२०६५/६६	९,८८,२७२	१४,४१,५४८	२१.१६	४.५३
२०६६/६७	११,९२,७७४	१५,१०,९७९	२०.६९	४.८२
२०६७/६८**	१५,६२,६८१	१५,५९,२२३	-	-
२०६८/६९	१७,५८,३७९	१६,३२,०४०	१२.५२	४.६७
२०६९/७०	१९,४९,२९५	१६,८९,५७२	१०.८६	३.५३
२०७०/७१	२२,३२,५२५	१७,९१,१४१	१४.५३	६.०१
२०७१/७२	२४,२३,६३८	१८,६२,३५७	८.५६	३.९८
२०७२/७३	२६,०८,१८४	१८,७०,४२४	७.६१	०.४३
२०७३/७४	३०,७७,१४५	२०,३८,३३७	१७.९८	८.९८
२०७४/७५	३४,५५,९४९	२१,९३,७०६	१२.३१	७.६२
२०७५/७६	३८,५८,९३०	२३,३९,७४३	११.६६	६.६६
२०७६/७७ R	३९,१४,७०१	२२,९०,८८०	१.४५	-२.०९
२०७७/७८ P	४२,६६,३२१	२३,८२,७०८	८.९८	४.०१

*आधार वर्ष २०५७/५८, **आधार वर्ष २०६७/६८ R परिमार्जित, P प्रारम्भिक

स्रोत: केन्द्रीय तथ्याङ्क विभाग

ग) पुल्ड क्रस-सेक्सन: विभिन्न समयका क्रस-सेक्सन तथ्याङ्क समूहलाई एउटै तथ्याङ्क समूहमा मिलाएर बनाइएको तथ्याङ्कलाई पुल्ड क्रस-सेक्सन तथ्याङ्क (Pooled Cross-Section Data) भनिन्छ। तालिका १४.३ मा विभिन्न वर्षमा सञ्चालन भएका जीवनस्तर सर्वेक्षणका तथ्याङ्कको सानो अंश प्रस्तुत गरिएको छ। यो तालिकाको तथ्याङ्क पुल्ड क्रस-सेक्सन तथ्याङ्क हो।

तालिका १४.३: परिवारले तिरेको मासिक भाडा

क्र.सं.	वर्ष	मासिक घरभाडा (रु)
१	१९९५/९६	४०
२	१९९५/९६	८६
३	१९९५/९६	५९
४	१९९५/९६	६५
५	१९९५/९६	१४९
.....		
२५१	२००३/०४	४४०
२५२	२००३/०४	४०८
२५३	२००३/०४	३७
२५४	२००३/०४	५५
२५५	२००३/०४	१५४
.....		
५६१	२०१०/११	१,४५९
५६२	२०१०/११	१,२३७
५६३	२०१०/११	१,३८४
५६४	२०१०/११	१,३९०
५६५	२०१०/११	७८०
.....		

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

घ) प्यानल डाटा: यसलाई लङ्गिच्युडिनल डाटा (Longitudinal Data) पनि भनिन्छ। प्यानल डाटा भनेको (Panel Data) विभिन्न समयका क्रस-सेक्सनका उही एकाइहरूको तथ्याङ्क समावेश गरी बनाइएको एउटै तथ्याङ्क समूह हो। तालिका १४.४ मा तिनवटा जीवनस्तर सर्वेक्षणबाट केही अंश तथ्याङ्क लिई प्रस्तुत गरिएको छ। यस तालिकामा उही परिवारबाट तिनवटा सर्वेक्षणहरूमा तथ्याङ्क संकलन गरी मिलाएर राखिएको छ। व्यक्तिगत तथ्याङ्क गोप्य राख्नुपर्ने कानुनी प्रावधानका कारण परिवारमूलीको नाम नलेखी काल्पनिक कोड राखिएको छ। यसरी एउटै एकाइबाट विभिन्न समयमा संकलन गरिएको तथ्याङ्क समूहलाई प्यानल डाटा भनिन्छ।

तालिका १४.४: खाद्य पदार्थमा परिवारले गरेको खर्च

क्र.सं.	परिवारमूलीको नाम कोड	वर्ष	खाद्य पदार्थमा गरेको वार्षिक खर्च
१	१	१९९५/९६	१२,३६७
२	१	२००३/०४	२७,६७७
३	१	२०१०/११	४१,६९८
४	२	१९९५/९६	७,७९८
५	२	२००३/०४	८,२८४
६	२	२०१०/११	१८,८४५
.....			
५६१	१५६	१९९५/९६	३३,७७१
५६२	१५६	२००३/०४	५१,७१५
५६३	१५६	२०१०/११	८०,७६०
.....	१९९५/९६

स्रोत: केन्द्रीय तथ्याङ्क विभाग, विभिन्न समयमा सञ्चालन भएका जीवनस्तर सर्वेक्षणहरू

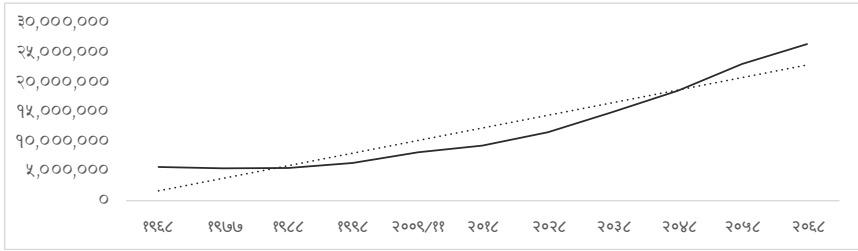
यसरी संरचनाका आधारमा तथ्याङ्कलाई चार किसिमले विभाजन गर्न सकिन्छ। तीमध्ये समय श्रृङ्खला तथ्याङ्कको चर्चा यस पाठमा गरिएको छ।

१४.२ समय श्रृङ्खलाका अवयवहरू

कुनै चरहरू समयअनुसार सँधै बढिरहने प्रकृतिका हुन्छन्, कुनै घटिरहने प्रकृतिका हुन सक्छन् भने कुनै चरहरू कहिले घट्ने कहिले बढ्ने प्रकृतिका हुन्छन्। समयअनुसार परिवर्तन हुँदा फरक-फरक प्रकृति देखाउन सक्ने हुँदा समय श्रृङ्खलामा आउने असङ्गतिलाई ४ भागमा विभाजन गरेर हेर्न सकिन्छ।

१. प्रवृत्ति (Secular Trend): समय परिवर्तन हुँदा लगातार घटिरहने वा बढिरहने चरको प्रवृत्तिलाई प्रवृत्ति भनिन्छ। जस्तै: नेपालको जनसंख्यालाई दश-दश वर्षको अन्तरालमा हेर्दा नेपालको जनसंख्या बढ्दो प्रवृत्तिमा देखिन्छ। चित्र १४.१ मा विगत १०० वर्षको नेपालको जनसंख्यालाई रेखाचित्रमा प्रस्तुत गरिएको छ। चित्रको गाढा बक्रले वास्तविक जनसंख्या प्रतिनिधित्व गरेको छ भने सिधा बिन्दु रेखाले जनसंख्याको प्रवृत्ति प्रतिनिधित्व गरेको छ। यसबाट हामीले नेपालको जनसंख्या बढ्दो प्रवृत्तिमा रहेको स्पष्टसँग देख्न सक्दौं।

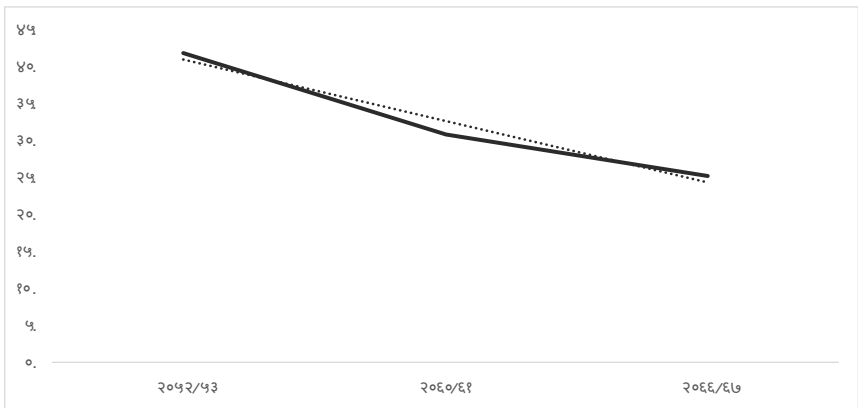
चित्र नं. १४.१: नेपालको जनसंख्या समय श्रृङ्खला



स्रोत: केन्द्रीय तथ्याङ्क विभाग

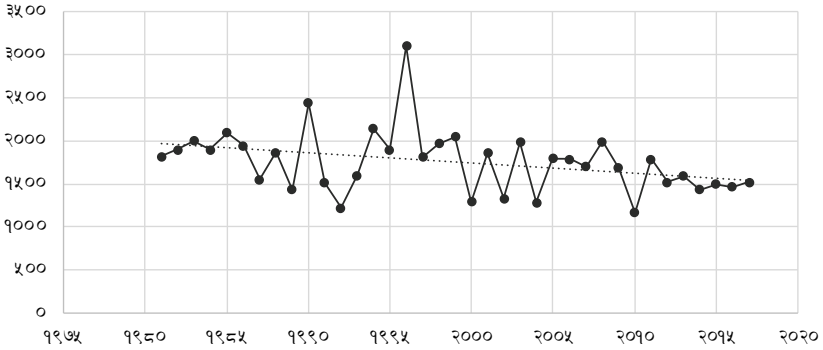
त्यस्तै, गरिबीको रेखामुनि भएको जनसंख्याको प्रतिशत घट्दो प्रवृत्तिमा देखिन्छ (चित्र नं. १४.२)। चित्र नं. १४.३ मा दैलेख जिल्लाको वार्षिक वर्षाको मात्रा प्रस्तुत गर्ने समय श्रृङ्खला देखाइएको छ। यस चित्रमा पनि उक्त जिल्लामा वर्षाको परिमाण घट्दो प्रवृत्तिमा रहेको देखिन्छ।

चित्र नं. १४.२: गरिबीको रेखामुनि भएको जनसंख्याको प्रतिशत



स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण

चित्र नं. १४.३: दैलेख जिल्लाको वार्षिक वर्षा परिमाण समय श्रृङ्खला



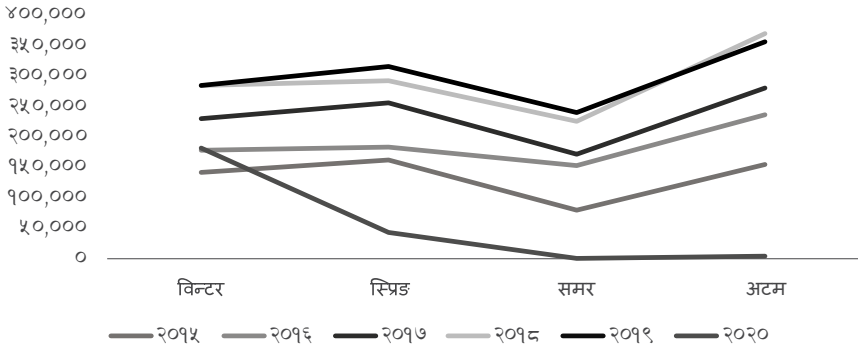
स्रोत: जल तथा मौसम विज्ञान विभाग

२. आवधिक चलायमान वा छोटो अवधिको उतारचढाव (Periodic Movements or Short-term Movement): कतिपय चरहरू समय अन्तरालसँगै छोटो अवधिमा उतारचढाव गरेर लामो अवधिमा कुनै निश्चित प्रवृत्ति देखाएका हुन्छन्। अर्थात्, आवधिक रूपमा चरित्र दोहोरिने गरी कतिपय चरहरूले उतारचढाव गरिरहेका हुन्छन्। यस्तो उतारचढाव छोटो अवधिमा देखापर्दछ। यस किसिमको उतारचढावलाई दुई भागमा वर्गीकरण गरी हेर्न सकिन्छ।

क) मौसमी विचलन (Seasonal Variations)

मौसमअनुसार घटबढ हुने चरको प्रकृतिलाई सिजनल भेरिएसन्स भनिन्छ। जस्तै: कुनै पसलेको पसलमा हरेक महिना छाता कति बिक्री हुन्छ भनेर हेर्दा वर्षा मौसममा बढी व्यापार हुन्छ भने जाडो मौसममा कम व्यापार हुन्छ। यो क्रम हरेक वर्ष दोहोरिन्छ। त्यस्तै, कुनै किसानले मासिक कति अम्दानी गर्छ भनेर हेर्दा उसले अन्न उब्जनी गरेर बेच्ने समयमा उसको आम्दानी बढी हुनसक्छ भने अरू समयमा कम अम्दानी हुनसक्छ। हरेक वर्ष नेपालमा आउने पर्यटकको सङ्ख्या सिजनअनुसार हेर्ने हो भने चित्र १४.४ जस्तो देखिन्छ। यहाँ, हेर्दा प्रत्येक वर्ष समर र विन्टर समयमा पर्यटकको आगमन कम देखिएको छ भने मध्यम मौसम हुने समयमा आगमन बढी देखिन्छ। सन् २०२० मा भने कोभिड-१९ कारण पर्यटन घटेको हो।

चित्र नं. १४.४: ऋतुअनुसार पर्यटन आगमन



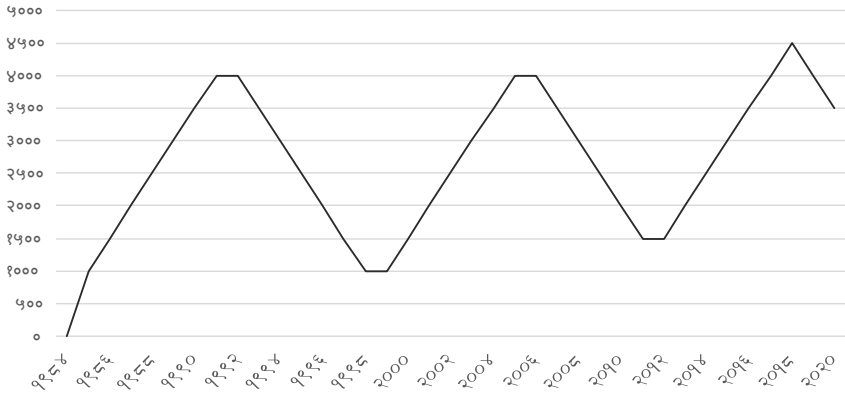
स्रोत: संस्कृति, पर्यटन तथा नागरिक उड्डयन मन्त्रालय^९

ख) चक्रीय विचलन (Cyclical Variations)

एक वर्षभन्दा लामो समयसम्म एउटै प्रकृतिमा घट्ने वा बढ्ने भएर त्यसपछि प्रकृति परिवर्तन हुने र समय समयमा यस्तो प्रकृति दोहोरिरहने चरको प्रकृतिलाई चक्रीय विचलन भनिन्छ । जस्तै: कुनै कम्पनीले एउटा नयाँ सामान बजारमा बेच्न सुरु गर्‍यो भने सुरुमा उसको व्यापार शून्यबाट सुरु हुन्छ, बिस्तारै बढ्दै जान्छ । केही समयको अन्तरालमा उसको अरू प्रतिस्पर्धी कम्पनीहरूले त्यस्तै सामान अझै राम्रो र सस्तो दरमा बजारमा बेच्न सुरु गरे भने उसको व्यापार घट्न थाल्छ । घट्दै जाने क्रममा उसले आफ्नो सामानलाई नयाँ ढाँचामा परिवर्तन गरेर अरूसँग प्रतिस्पर्धा गरेर व्यापार बढाउन सक्छ । त्यसको केही समयपछि फेरि अरू प्रतिस्पर्धीको सामानको कारणले व्यापार घट्दै जानसक्छ । यसरी नै उक्त कम्पनीको व्यापार घटबढ भइरहन सक्छ । यही प्रकृतिको समय श्रृङ्खलालाई नै चक्रीय विचलन भनिन्छ । चक्रीय विचलन भएको चरलाई समयसँगै ग्राफमा हेर्‍यो भने चित्र १४.५ जस्तो देखिन्छ ।

9 Ministry of Culture, Tourism and Civil Aviation. *Tourism Statistics*. Kathmandu, Nepal; Ministry of Culture, Tourism and Civil Aviation. Retrieved on December 4, 2021 from: <https://www.tourism.gov.np/statistic>

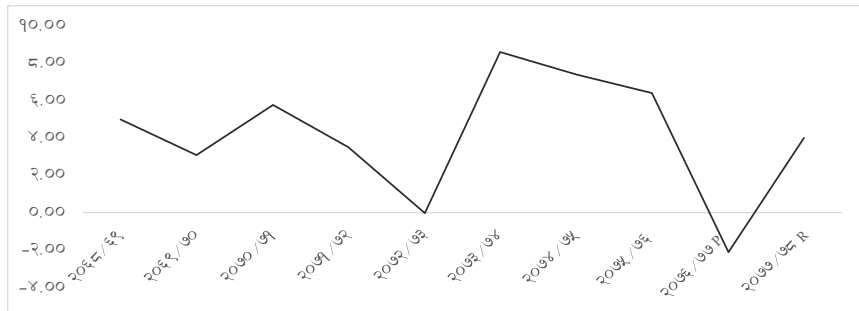
चित्र नं. १४.५: चक्रिय विचलन भएको समय श्रृङ्खलाको नमुना चित्र



स्रोत: संस्कृति, पर्यटन तथा नागरिक उड्डयन मन्त्रालय⁸

३. अनियमित उतारचढाव (Random Fluctuations): माथि उल्लेखित तिनवटा श्रृङ्खलामध्ये कुनै पनि श्रृङ्खलासँग नमिल्ने अर्थात् कुनै पनि समयमा घटबढ भइरह्न सक्ने प्रकृतिको श्रृङ्खलालाई अनियमित उतारचढाव भनिन्छ। चित्र १४.६ मा नेपालको कुल गार्हस्थ्य उत्पादन वृद्धिदर (आधार मूल्यमा) दिइएको छ। यो एउटा अनियमित उतार चढावको उदाहरण हो।

चित्र नं. १४.६: कुल गार्हस्थ्य उत्पादन वृद्धिदर (आधार मूल्यमा)



नोट: P प्रारम्भिक, R परिमार्जित

स्रोत: केन्द्रीय तथ्याङ्क विभाग

१४.३ समय श्रृङ्खलाको विश्लेषण तथा महत्त्व

कुनै पनि चरका विशेषताहरू समयअनुसार अध्ययन गर्ने विधिलाई समय श्रृङ्खला विश्लेषण भनिन्छ। कुनै पनि चरको मान समयअनुसार कसरी परिवर्तन भएको छ भनेर हेर्न यसको प्रयोग गरिन्छ। जस्तै: कुनै गाउँपालिकाले हरेक वर्ष विकास खर्च कति गर्‍यो भनेर

हेर्नको लागि यसको प्रयोग गर्न सकिन्छ । नेपालको विकासमा वैदेशिक लगानीले कस्तो प्रभाव पारेको छ भनेर पनि यसको प्रयोग गर्न सकिन्छ । त्यस्तै, प्रविधिको आगमनले नेपालको कुल गार्हस्थ्य उत्पादनमा कसरी परिवर्तन ल्याएको छ भनेर पनि समय श्रृङ्खलाको विश्लेषणबाट अध्ययन गर्न सकिन्छ । यस्ता कैयौँ विषयहरूमा समय श्रृङ्खलाको महत्त्व रहेको छ । समय श्रृङ्खला विश्लेषणमा मुख्यतया निम्न कुराहरू पर्दछन् ।

क) विभिन्न बाह्य प्रभाव तथा कारणहरूले समय श्रृङ्खलामा विचलन ल्याएको हुन्छ, त्यस्ता बाह्य कारणहरू पत्ता लगाउने काम समय श्रृङ्खला विश्लेषणले गर्दछ ।

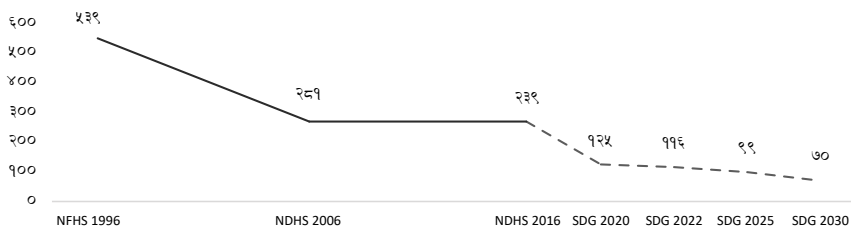
ख) बाह्य प्रभाव तथा कारण दुई वा दुई भन्दा बढी हुन सक्दछन् । समय श्रृङ्खला विश्लेषणका माध्यमबाट तिनीहरूको अलग अलग अध्ययन गरी असर पत्ता लगाउन सकिन्छ ।

समय श्रृङ्खला अध्ययनबाट व्यवसाय तथा आर्थिक क्षेत्रका साथसाथै सामाजिक, वातावरणीय, लगायतका विभिन्न क्षेत्रमा भइरहने परिवर्तनका बारेमा थाहा पाउन सकिन्छ । यसका केही प्रमुख महत्त्वहरू निम्नानुसार छन् ।

क) समय श्रृङ्खला विश्लेषणका माध्यमबाट विगतका विविध घटनाहरूका विशेषता जानकारी हुन्छ । त्यस्ता घटनाको प्रकार र प्रकृतिका बारेमा जानकारी प्राप्त हुन्छ । जस्तै नेपालको जनसंख्याको समय श्रृङ्खलाको अध्ययनबाट विगतमा नेपालको जनसंख्या कसरी बढिरहेको थियो, संरचना कस्तो थियो जस्ता कुराहरू थाहा पाउन सकिन्छ ।

ख) समय श्रृङ्खलाको विश्लेषणबाट कुनै क्षेत्रमा विचलन ल्याउने बाह्य तत्त्वहरू पत्ता लगाएर नीति निर्माण, योजना तर्जुमा तथा अनुगमन र मूल्याङ्कनमा प्रयोग गर्न सकिन्छ । उदाहरणका लागि माथि चित्र १४.३ मा दैलेख जिल्लाको वर्षाको परिमाण क्रमश घटेको देखिन्छ । वर्षाको मात्रा कम हुँदै जानु भनेको जलवायु परिवर्तनको असर हो र यसलाई समाधान गर्न जलवायु परिवर्तनका कारक तत्त्वहरूका बारेमा सचेत हुनुपर्छ । यसका लागि नीतिनिर्माताहरूले वातावरणीय पक्षको नीति निर्माण गर्दा वनजङ्गल संरक्षण गर्ने, वृक्षरोपण गर्ने, आगलागी रोक्ने, प्रदूषण रोक्ने जस्ता कार्यक्रमहरू तय गर्न सक्दछन् ।

चित्र १४.७: मातृ मृत्यु अनुपात १९९६-२०१६ र २०३० सम्मको लक्ष्य



स्रोत: नेपाल परिवार स्वास्थ्य सर्वेक्षण, १९९६ (NFHS १९९६), नेपाल जनसांख्यिक स्वास्थ्य सर्वेक्षण २००६ र २०१६ (NDHS २००६, २०१६), दिगो विकास लक्ष्य (SDG)

घ) समय श्रृङ्खलाको विश्लेषणबाट पूर्वानुमान वा प्रक्षेपण गर्न सकिन्छ। यसको एउटा उदाहरणका रूपमा तालिका १४.५ मा नेपालको जनसंख्या प्रक्षेपणको तथ्याङ्क दिइएको छ।

तालिका १४.५: नेपालको जनसंख्या प्रक्षेपण

साल	जनसंख्या	साल	जनसंख्या	साल	जनसंख्या
२०११	२,६४,९४,५०४	२०१८	२,९२,१८,८६७	२०२५	३,१८,१४,१६१
२०१२	२,६८,७३,०६६	२०१९	२,९६,०९,६२३	२०२६	३,२१,४४,९२१
२०१३	२,७२,५७,३४७	२०२०	२,९९,९६,४७८	२०२७	३,२४,६२,५१८
२०१४	२,७६,४६,०५३	२०२१	३,०३,७८,०५५	२०२८	३,२७,६६,४८३
२०१५	२,८०,३७,९०४	२०२२	३,०७,५२,१४६	२०२९	३,३०,५६,८६१
२०१६	२,८४,३१,४९४	२०२३	३,११,१६,९१३	२०३०	३,३३,३३,६९३
२०१७	२,८८,२५,७०९	२०२४	३,१४,७१,२१३	२०३१	३,३५,९७,०३२

स्रोत: केन्द्रीय तथ्याङ्क विभाग^{१०}

ङ) विभिन्न घटनाहरूको विभिन्न समय र स्थानमा भएका परिवर्तनहरूको अध्ययन गरेर त्यस्ता परिवर्तनहरूको तुलनात्मक अध्ययन गर्न सकिन्छ।

१४.४ समय श्रृङ्खला विश्लेषणको लागि गणितीय मोडेलहरू

समय श्रृङ्खलाका अवयवहरू पत्ता लगाउनका लागि मुख्यतया तिनवटा विधिहरू प्रयोग गरिन्छन्।

10 Central Bureau of Statistics. (2014). *National Population and Housing Census 2011 (Population Projection 2011 – 2031)*. Kathmandu, Nepal; Central Bureau of Statistics. Retrieved on December 5, 2021 from: <https://cbs.gov.np/wp-content/uploads/2018/12/PopulationProjection2011-2031.pdf>

क) योगफलमा आधारित मोडल: यस मोडलमा समय श्रृङ्खलाका सबै अवयवहरूलाई जोडेर समय श्रृङ्खला उल्लेख गरिन्छ । जस्तै, समय श्रृङ्खलालाई Y ले, प्रवृत्तिलाई T ले मौसमी विचलनलाई S ले, चक्रीय विचलनलाई C ले र अनियमित विचलनलाई I ले जनाउँदा यस मोडललाई निम्नअनुसार व्यक्त गर्न सकिन्छ ।

$$Y = T + S + C + I \quad (१)$$

ख) गुणनफलमा आधारित मोडल: यस मोडलमा समय श्रृङ्खलाका सबै अवयवहरूलाई गुणन गरेर समय श्रृङ्खला उल्लेख गरिन्छ । जस्तै, माथि समीकरण (१) मा उल्लेख भएका अवयवहरूलाई निम्न समीकरणमा व्यक्त गरिन्छ ।

$$Y = T \times S \times C \times I \quad (२)$$

ग) मिश्रित मोडल: यस मोडलमा माथिका दुवै मोडल समावेश भएका हुन्छन् । जस्तै:

$$Y = T \times S \times C + I \quad (३)$$

$$Y = T \times C + S \times I \quad (४)$$

$$Y = T + S \times C \times I \quad (५)$$

$$Y = T + S + C \times I \quad (६)$$

१४.५ समय श्रृङ्खला अवयवहरूको मापन

माथि समीकरण (१) देखि (६) मा प्रयोग भएका समय श्रृङ्खलाका अवयवहरूको कसरी मापन गर्ने भन्ने बारेमा यहाँ संक्षिप्त चर्चा गरिएको छ ।

१४.५.१ प्रवृत्ति (Trend = T)

प्रवृत्ति मापनका लागि निम्न विधिहरू प्रचलनमा छन् ।

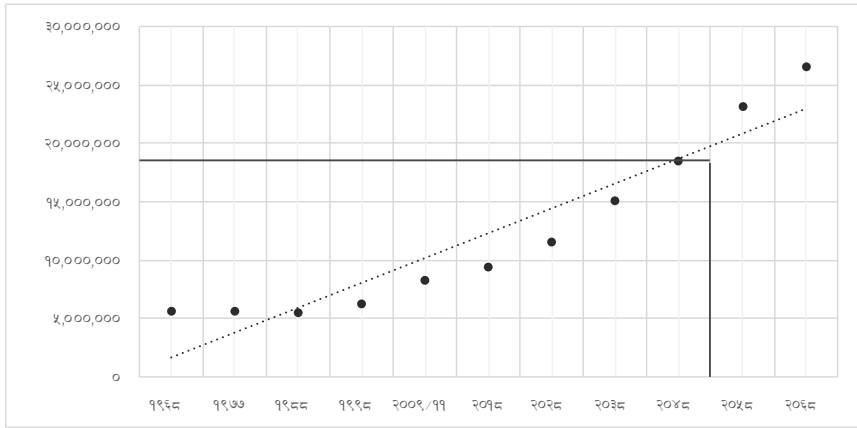
- १) रेखाचित्र विधि
- २) आधा-औसत विधि
- ३) लिस्ट स्क्वायरको सिद्धान्तमा आधारित बक्ररेखा विधि
- ४) चलायमान औसत विधि

यी विधिहरूमध्ये रेखाचित्र विधि र चलायमान औसत विधि (Moving Average Method) को मात्र यहाँ थप व्याख्या गरिएको छ ।

१) रेखाचित्र विधि: यस विधिमा पहिला दिइएका समय श्रृङ्खलालाई ग्राफ पेपरमा बिन्दुहरूले प्रतिनिधित्व गराइन्छ । त्यसपछि समग्र बिन्दुहरू प्रतिनिधित्व हुनेगरी एउटा

सिधा रेखा खिचिन्छ । आजकल कम्प्युटरको सहयोगबाट प्रवृत्ति रेखा सजिलै खिच्न सकिन्छ । जस्तै: तलको चित्र नं. (१४.८) मा विभिन्न वर्षको नेपालको जनसंख्या गाढा थोप्लाले जनाएर हिस्टोरिग्राम (Historigram) बनाइएको छ । तिनै थोप्लाहरूलाई प्रतिनिधित्व हुने गरी एउटा सिधा थोप्ला रेखा खिचिएको छ, जसलाई प्रवृत्ति रेखा भनिन्छ । प्रवृत्ति रेखाहरू विभिन्न प्रकारका हुन्छन् । यहाँ प्रयोग गरिएको प्रवृत्ति रेखा सिधा-रेखीय प्रवृत्ति रेखा हो ।

चित्र नं. १४.८: नेपालको जनसंख्या समय श्रृङ्खला



स्रोत: केन्द्रीय तथ्याङ्क विभाग

माथिको चित्र नं. १४.८ को प्रवृत्ति रेखाका आधारमा हामीले कुनै पनि वर्षको प्रवृत्ति पत्ता लगाउन सक्दछौं । कुनै समयका लागि ग्राफबाट प्रवृत्ति पत्ता लगाउन सो समय बिन्दु (X अक्षमा) बाट ठाडो रेखा प्रवृत्ति रेखासम्म खिच्ने र जहाँ यो ठाडो रेखाले प्रवृत्ति बिन्दुलाई छुन्छ त्यहाँबाट तेर्सो रेखा Y अक्षसम्म खिच्ने । यसरी तेर्सो रेखाले Y अक्षलाई जहाँ छुन्छ, त्यही बिन्दुको मान प्रवृत्ति हुन्छ । उदाहरणका लागि वि. सं. २०५३ सालको लागि प्रवृत्ति २,००,००,००० हुन्छ । यही तरिकाले अन्य वर्षहरूका लागि प्रवृत्ति पत्ता लगाउन सकिन्छ । जस्तै: २०१८ सालको लागि १,२५,००,००० हुन्छ । वि. सं. २०३३ सालको लागि लगभग १,५५,००,००० हुन्छ ।

२) **चलायमान औसत विधि:** यस विधिमा दिइएका समय श्रृङ्खलालाई निश्चित वर्षको समूहअनुसार औसत निकालेर सो औसत बिचको वर्षमा राख्दै जानुपर्दछ । उदाहरणका लागि तालिका १४.६ को समय श्रृङ्खलाको ३ वर्षे चलायमान औसत निकाल्ने हो भने आर्थिक वर्ष २०३१/३२, २०३२/३३ र २०३३/३४ का मानहरूको औसत निकाल्ने । त्यसपछि सो औसतलाई २०३२/३३ मा लगेर राख्ने । फेरि आर्थिक वर्ष २०३२/३३,

तालिका १४.६: नेपालको निर्यात मूल्य समय श्रृङ्खला तथ्याङ्क

आर्थिक वर्ष	निर्यात (रु दश लाख)	आर्थिक वर्ष	निर्यात (रु दश लाख)	आर्थिक वर्ष	निर्यात (रु दश लाख)
२०३१/३२	८८९.६	२०४७/४८	७,३८७.५०	२०६३/६४	५९,३८३.१०
२०३२/३३	१,१८५.८०	२०४८/४९	१३,७०६.५०	२०६४/६५	५९,२६६.५०
२०३३/३४	१,१६४.७०	२०४९/५०	१७,२६६.५०	२०६५/६६	६७,६९७.५०
२०३४/३५	१,०४६.२०	२०५०/५१	१९,२९३.४०	२०६६/६७	६०,८२४.००
२०३५/३६	१,२९६.८०	२०५१/५२	१७,६३९.२०	२०६७/६८	६४,३३८.५०
२०३६/३७	१,१५०.५०	२०५२/५३	१९,८८१.१०	२०६८/६९	७४,२६१.००
२०३७/३८	१,६०८.७०	२०५३/५४	२२,६३६.५०	२०६९/७०	७६,९१७.१०
२०३८/३९	१,४९१.५०	२०५४/५५	२७,५१३.५०	२०७०/७१	९१,९९१.४०
२०३९/४०	१,१३२.००	२०५५/५६	३५,६७६.३०	२०७१/७२	८५,३१९.१०
२०४०/४१	१,७०३.९०	२०५६/५७	४९,८२२.७०	२०७२/७३	७०,११७.१०
२०४१/४२	२,७४०.६०	२०५७/५८	५५,६५४.१०	२०७३/७४	७३,०४९.१०
२०४२/४३	३,०७८.००	२०५८/५९	४६,९४४.८०	२०७४/७५	८१,३५९.८०
२०४३/४४	२,९९१.४०	२०५९/६०	४९,९३०.६०	२०७५/७६	९७,१०९.५०
२०४४/४५	४,११४.५०	२०६०/६१	५३,९१०.७०	२०७६/७७	९७,७०९.१०
२०४५/४६	४,१९५.३०	२०६१/६२	५८,७०५.७०	२०७७/७८	१४१,१२४.१०
२०४६/४७	५,१५६.२०	२०६२/६३	६०,२३४.१०		

स्रोत: भन्सार विभाग

२०३३/३४ र २०३४/३५ का मानहरूको औसत निकाल्ने । त्यसपछि सो औसतलाई २०३३/३४ मा लागेर राख्ने । यही प्रक्रिया अन्तिम ३ वटा वर्षसम्म दोहोर्‍याई रहने ।

अर्थात्,

$$\frac{८८९.६+११८५.८+११६४.७}{३} = १,०८० \text{ लाई } २०३२/३३ \text{ मा,}$$

$$\frac{११८५.८+११६४.७+१०४६.२}{३} = १,१३२.२ \text{ लाई } २०३३/३४ \text{ मा}$$

.....

$$\frac{९७१०९.५+९७७०९.१+१४११२४.१}{३} \text{ लाई } २०७६/७७ \text{ मा राख्ने । यसरी निकालिएको}$$

समय श्रृङ्खला तालिका १४.७ मा दिइएको छ ।

तालिका १४.७: तिन वर्षे चलायमान औसत विधिबाट निकालिएको नियात तथ्याङ्क

आर्थिक वर्ष	नियात (रु दश लाख)	३ वर्षे चलायमान औसत	आर्थिक वर्ष	नियात (रु दश लाख)	३ वर्षे चलायमान औसत
२०३१/३२	८८९.६		२०५५/५६	३५,६७६.३०	३७,६७०.८०
२०३२/३३	१,१८५.८०	१,०८०.००	२०५६/५७	४९,८२२.७०	४७,०५१.००
२०३३/३४	१,१६४.७०	१,१३२.२०	२०५७/५८	५५,६५४.१०	५०,८०७.२०
२०३४/३५	१,०४६.२०	१,१६९.२०	२०५८/५९	४६,९४४.८०	५०,८४३.२०
२०३५/३६	१,२९६.८०	१,१६४.५०	२०५९/६०	४९,९३०.६०	५०,२६२.००
२०३६/३७	१,१५०.५०	१,३५२.००	२०६०/६१	५३,९१०.७०	५४,९८२.३०
२०३७/३८	१,६०८.७०	१,४९६.९०	२०६१/६२	५८,७०५.७०	५७,६१६.८०
२०३८/३९	१,४९१.५०	१,४१०.७०	२०६२/६३	६०,२३४.१०	५९,४४१.००
२०३९/४०	१,१३२.००	१,४४२.५०	२०६३/६४	५९,३८३.१०	५९,६२७.९०
२०४०/४१	१,७०३.९०	१,८५८.८०	२०६४/६५	५९,२६६.५०	६२,११५.७०
२०४१/४२	२,७४०.६०	२,५०७.५०	२०६५/६६	६७,६९७.५०	६२,५९६.००
२०४२/४३	३,०७८.००	२,९३६.७०	२०६६/६७	६०,८२४.००	६४,२८६.७०

२०४३/४४	२,९९१.४०	३,३९४.६०	२०६७/६८	६४,३३८.५०	६६,४७४.५०
२०४४/४५	४,११४.५०	३,७६७.१०	२०६८/६९	७४,२६१.००	७१,८३८.९०
२०४५/४६	४,१९५.३०	४,४८८.७०	२०६९/७०	७६,९१७.१०	८१,०५६.५०
२०४६/४७	५,१५६.२०	५,५७९.७०	२०७०/७१	९१,९९१.४०	८४,७४२.५०
२०४७/४८	७,३८७.५०	८,७५०.१०	२०७१/७२	८५,३१९.१०	८२,४७५.९०
२०४८/४९	१३,७०६.५०	१२,७८६.८०	२०७२/७३	७०,११७.१०	७६,१६१.८०
२०४९/५०	१७,२६६.५०	१६,७५५.५०	२०७३/७४	७३,०४९.१०	७४,८४२.००
२०५०/५१	१९,२९३.४०	१८,०६६.४०	२०७४/७५	८१,३५९.८०	८३,८३९.५०
२०५१/५२	१७,६३९.२०	१८,९३७.९०	२०७५/७६	९७,१०९.५०	९२,०५९.५०
२०५२/५३	१९,८८१.१०	२०,०५२.३०	२०७६/७७	९७,७०९.१०	१,११,९८०.९
२०५३/५४	२२,६३६.५०	२३,३४३.७०	२०७७/७८	१,४९,१२४.१०	
२०५४/५५	२७,५१३.५०	२८,६०८.७०			

स्रोत: भन्सार विभाग

यसरी चलायमान औसत विधिबाट निकालिएको निर्यात तथ्याङ्क सुरूको तथ्याङ्कभन्दा Smooth हुन्छ । तथ्याङ्क विश्लेषणमा यसरी Smooth गरिएको तथ्याङ्क प्रयोग गर्दा अभ्र बढी भरपर्दो निष्कर्ष निकाल्न सकिन्छ ।

नोट: माथिको उदाहरणमा ३ वर्षे चलायमान औसत विधिबाट प्रवृत्तिको मापन गरिएको छ । यदि जोर संख्याका चलायमान औसत विधि प्रयोग गर्ने हो भने प्राप्त भएको औसतलाई मध्य समयमा राख्नुपर्दछ । अर्थात्, आर्थिक वर्ष २०३१/३२ र २०३२/३३ का मानहरूको औसत निकाल्ने । त्यसपछि सो औसतलाई २०३१/३२ र २०३२/३३ को बिच समयमा लगेर राख्ने । यही प्रक्रिया दोहोर्‍याउँदै जाने । स्मरणीय कुरा के छ भने, चलायमान औसत विधि प्रयोग गर्दा सुरूका केही समय र अन्तिमका केही समयका लागि मानहरू उपलब्ध हुँदैनन् । जस्तै: माथि तालिका १४.७ मा आर्थिक वर्ष २०३१/३२ र २०७७/७८ का लागि चलायमान औसत मानहरू उपलब्ध छैनन् । लामो समयको श्रृङ्खला भएको हुनाले दुई तिन वर्षका तथ्याङ्क उपलब्ध नभए तापनि विश्लेषण गरेर निष्कर्ष निकाल्नलाई कुनै बाधा हुँदैन । छोटो अवधिको समय श्रृङ्खलामा भने यो विधि उपयुक्त मानिँदैन । सामान्यतया छोटो अवधिको समय श्रृङ्खलाका लागि रेखाचित्र प्रयोग गर्न सकिन्छ ।

१४.५.२ मौसमी विचलन (Seasonal Variation = S)

मौसमी विचलन मापनका लागि निम्न विधिहरू प्रचलनमा छन् ।

- १) सामान्य औसत विधि
- २) प्रवृत्ति अनुपात विधि
- ३) चलायमान औसत अनुपात विधि
- ४) लिङ्क सापेक्ष विधि

यी विधिहरूमध्ये सामान्य औसत विधिको बारेमा यहाँ उल्लेख गरिएको छ । यसका लागि एउटा उदाहरणसहित चर्चा गरौं । नेपालमा आएका पर्यटनहरूको मासिक तथा वार्षिक विवरण तालिका १४.६ मा प्रस्तुत गरिएको छ । यस तालिकाका तथ्याङ्कहरूलाई निम्न चरणहरू बमोजिम गर्दै जाऔं ।

क) तथ्याङ्कलाई वर्ष र महिनाबमोजिम मिलाएर राख्ने ।

ख) प्रत्येक महिनाको लागि अंकगणितीय मध्यक निकालौं ।

ग) अंकगणितीय मध्यकको पनि मध्यक निकालौं ।

$$\text{अर्थात्, } \frac{६२९४९+८२२५४+\dots+\dots+८५२०३+७९९२०}{१२} = ६७,११९ \text{ हुन्छ ।}$$

घ) अब प्रत्येक महिनाको लागि मौसमी सूचकाङ्क निम्नानुसारको सूत्र प्रयोग गरेर निकालौं ।

$$\text{जनवरीको लागि मौसमी सूचकाङ्क} = \frac{\text{जनवरीको मध्यक}}{\text{मध्यकको पनि मध्यक}} \times १०० \quad (७)$$

यसरी प्रत्येक महिनाको लागि मौसमी सूचकाङ्क निकाल्न सकिन्छ र माथिको उदाहरणको मौसमी सूचकाङ्क तालिका १४.८ को अन्तिम कोलममा दिइएको छ । यो विधि प्रयोग गर्नुपूर्व तथ्याङ्कमा भएको प्रवृत्ति विचलन हटाउनु पर्दछ । यहाँ सामान्य औसत विधिको चरणहरू बुझ्ने प्रयोजनका लागि मात्र प्रवृत्ति विचलन नहटाई चरणहरू मात्र बताइएको हो । पाठकहरूलाई पहिला माथि भने बमोजिम प्रत्येक महिनाको अलग अलग प्रवृत्ति निकालेर उक्त तथ्याङ्कबाट मौसमी सूचकाङ्क निकाल्न आग्रह गरिन्छ ।

तालिका १४.८: नेपालमा आएका पर्यटनहरूको मासिक तथा वार्षिक विवरण

वर्ष	२०१५	२०१६	२०१७	२०१८	२०१९	२०२०	मासिक मध्यक	मौसमी सूचकाङ्क
जनवरी	३८,६१६	४२,२३५	६२,६३२	७३,१८७	८१,२७३	७९,७०२	६२,९४१	९३.८
फेब्रुअरी	५८,५२३	६०,८२१	८४,०६१	८९,५०७	१,०२,४२३	९८,१९०	८२,२५४	१२२.६
मार्च	७९,१८७	७६,४४४	१,०६,२९१	१,२४,६८६	१,२७,३५१	४२,७७६	९२,७८९	१३८.२
अप्रिल	६५,७२९	६०,२१४	८८,५९१	९८,६५०	१,०९,३९९	१४	७०,४३३	१०४.९
मे	१७,५६९	४६,६८३	६२,७७३	६८,८२५	७८,३२९	३१	४५,७०२	६८.१
जुन	१८,३६८	३८,८५२	५५,९५६	६५,१५९	७४,८८३	१०२	४२,२२०	६२.९
जुलाई	२२,९६७	४८,११५	४२,२४०	७३,२८१	७०,९१६	१९६	४२,९५३	६४
आगस्ट	३८,६०६	६६,३४१	७३,७७८	८७,६७९	९४,७४९	२६७	६०,२३७	८९.७
सेप्टेम्बर	३९,०५०	७४,६७०	६८,६३४	९१,८७४	९२,६०४	५८४	६१,२३६	९१.२
अक्टोबर	५६,५८४	८९,३८१	१,१२,४९२	१,३०,७४५	१,३४,०९६	२,०२५	८७,५३७	१३०.४
नोभेम्बर	५८,३०४	७२,९९०	९९,८०४	१,४७,८५९	१,३०,३०२	१,९५६	८५,२०३	१२६.९
डिसेम्बर	४५,४६७	७६,३५६	८२,९६६	१,२१,६२०	१,००,८६६	४,२४२	७१,९२०	१०७.२

स्रोत: पर्यटन तथा नागरिक उड्ययन मन्त्रालय

माथिको तालिका १४.८ मा मौसमी सूचकाङ्क निकालेर राखिएको छ । यी सूचकाङ्कहरू दिइएको समय श्रृङ्खलाबाट मौसमी विचलन प्रभाव हटाउनका लागि प्रयोग गरिन्छ । यदि दिइएको समय श्रृङ्खलालाई योगमा आधारित मोडल मान्ने हो भने माथिको सूत्र (१) बाट मौसमी विचलन घटाएर हटाउन सकिन्छ ।

$$Y = T + S + C + I$$

अर्थात्, $Y - S = T + C + I$ यसरी प्राप्त हुने समय श्रृङ्खलालाई मौसमी विचलन नभएको समय श्रृङ्खला मानिन्छ ।

यदि दिइएको समय श्रृङ्खलालाई गुणामा आधारित मोडल मान्ने हो भने माथिको सूत्र (२) बाट मौसमी विचलन घटाएर हटाउन सकिन्छ ।

$$Y = T \times S \times C \times I$$

अर्थात्, $\frac{Y}{\text{Seasonal Effect}}$ जहाँ, Seasonal Effect भनेको मौसमी प्रभाव हो र यसलाई निम्नानुसार परिभाषित गरिन्छ ।

$$\text{Seasonal Effect} = \frac{\text{मौसमी सूचकाङ्क}}{१००}$$

यसरी प्राप्त हुने समय श्रृङ्खलालाई मौसमी विचलन नभएको समय श्रृङ्खला मानिन्छ । उदाहरणका लागि सन् २०२० जनवरीका लागि,

$$\text{Seasonal Effect} = \frac{\text{मौसमी सूचकाङ्क}}{१००} = \frac{९३.८}{१००} = ०.९३८ \text{ हुन्छ ।}$$

$$\text{मौसमी विचलन नभएको मान} = \frac{७९,७०२}{\text{Seasonal Effect}} = \frac{७९,७०२}{०.९३८} = ८४,९७० \text{ हुन्छ ।}$$

यही प्रक्रियाबाट अरू मानहरूको पनि मौसमी विचलन नभएको मान निकालेर समय श्रृङ्खला तयार गर्न सकिन्छ ।

नोट: मासिक समय श्रृङ्खलाको सबै महिनाको मौसमी सूचकाङ्कको योग १२०० हुनुपर्दछ भने त्रैमासिक समय श्रृङ्खलाको सबै त्रैमासिकको मौसमी सूचकाङ्कको योग ४०० हुन्छ ।

१४.५.३ चक्रीय तथा अनियमित विचलन (Seasonal and Irregular Variation = S)

चक्रीय तथा अनियमित विचलन मापन गर्ने कच्चा विधिको रूपमा शेष विधि (Method of Residual) प्रयोग गरिन्छ । किनभने अनियमित विचलन वास्तवमा अनुमान गर्न पनि सकिन्न र यो भैपरी आउने घटनाहरूले निर्देशित हुन्छ । जस्तै: भूकम्पका कारण कुल गार्हस्थ्य उत्पादन घट्नु (आर्थिक वर्ष २०१५/१६ मा कुल गार्हस्थ्य उत्पादन वृद्धिदर ०.००३% मात्र थियो), कोभिडका कारण पर्यटन घट्नु (सन् २०२० को अप्रिल, मे,

जुन महिनामा क्रमशः १४, ३१, १०२ जना पर्यटकमात्र नेपाल आएका थिए (तालिका १४.८) आदि। यस्ता किसिमका विचलनहरू अनुमान नै गर्न नसकिने किसिमका हुन्छन्। तसर्थ, शेष विधि प्रयोग गरेर चक्रीय तथा अनियमित विचलन संयुक्त रूपमा मापन गरिन्छ। यदि अलग अलग निकाल्नु परेमा अनियमित विचलनलाई नगण्य मानेर चक्रीय विचलन निकाल्ने पनि गरिन्छ। यस विधिको लागि माथिको समीकरण (२) बाट हेरौं।

$$Y = T \times S \times C \times I$$

यसलाई पुन यसरी लेख्न सकिन्छ।

$$CI = \frac{Y}{TS} \quad (८)$$

समीकरण (८) बाट चक्रीय तथा अनियमित विचलन संयुक्त रूपमा निकाल्न सकिन्छ किनभने प्रवृत्ति (T) र मौसमी विचलन (S) हामीले माथि निकालिसकेका छौं भने Y दिइएको समय श्रृङ्खला नै भयो। यदि अनियमित विचलनको असर नगण्य छ भन्ने मान्ने हो भने समीकरण (८) लाई निम्नानुसार लेख्न सकिन्छ।

$$C = \frac{Y}{TS} \quad (९)$$

१४.६ समय श्रृङ्खलामा अनुमान

कुनै पनि चर समयअनुसार कसरी परिवर्तन भएको छ भन्ने आधारमा उक्त चरको भविष्यमा मान कति हुन सक्छ भनेर अनुमान लगाउने विभिन्न तरिकाहरू हुन्छन्। माथि उल्लेखित चार प्रकारका समय श्रृङ्खलामध्ये अनियमित विचलन भएको चरमा अनुमान लगाउन गाह्रो हुन्छ भने बाँकी ३ वटामा भने विभिन्न विधिहरूको प्रयोग गरेर भविष्यमा हुनसक्ने मानको अनुमान लगाउन सकिन्छ। उक्त विधिहरू सबै बुझ्नको लागि गहिरो तथ्याङ्कीय ज्ञानको आवश्यकता पर्दछ। सिधा रेखामा लगातार घटिरहने वा बढिरहने चरको भने सिधा रेखीय प्रक्षेपण (Linear Forecasting) को सहायताले सजिलै अनुमान लगाउन सकिन्छ। यहाँ सिधा रेखीय प्रक्षेपण कसरी प्रयोग गर्ने भनेर व्याख्या गरिएको छ।

१४.६.१ समय श्रृङ्खलाको सिधा रेखीय प्रक्षेपण

सामान्य सिधा रेखीय प्रतिगमन विश्लेषण (Simple Linear Regression Analysis) मा जसरी नै समय श्रृङ्खलाको आँकडामा पनि प्रतिगमनको समीकरण निकाल्न सकिन्छ। उक्त समीकरणलाई प्रयोग गरेर भविष्यमा हुनसक्ने मानहरू अनुमान लगाउन सकिन्छ। सिधा रेखीय समय श्रृङ्खलाको समीकरणलाई निम्नानुसार प्रस्तुत गर्न सकिन्छ।

$$Y = A + BT \quad (१०)$$

यहाँ, Y भनेको अनुमान गर्नुपर्ने चर अथवा निर्भर चर (Variable to be Predicted/ Dependent Variable) हो, T भनेको समय/स्वतन्त्र चर (Time/Independent Variable) हो, A भनेको अध्ययन समय सुरु हुनुभन्दा १ एकाइ समय अगाडिको Y को मान (Intercept) हो, B भनेको १ एकाइ समय बढदा Y मा आउने परिवर्तनको दर (Slope/Rate of Change) हो । यो समीकरणमा A र B को अनुमान पत्ता लगाउन सक्ने हो भने विभिन्न समयको लागि Y को मान अनुमान वा प्रक्षेपण गर्न सकिन्छ । A र B को अनुमान पत्ता लगाउने कार्यलाई अनुमान भनिन्छ । दिइएको समय अन्तराल भित्रको कुनै समयको लागि Y को मान निकाल्ने विधिलाई इन्टरपोलेसन (Interpolation) भनिन्छ भने दिइएको समय अन्तराल बाहिरको कुनै समयको लागि Y को मान निकाल्ने विधिलाई एक्स्ट्रापोलेशन (Extrapolation) भनिन्छ । भविष्यको समयको लागि Y को मान निकाल्ने विधिलाई प्रक्षेपण (Projection) भनिन्छ । तसर्थ, प्रक्षेपण पनि एक्स्ट्रापोलेसनकै एउटा रूप हो । यहाँ समीकरण (१०) लाई अनुमान गर्ने एउटा विधि उल्लेख गरिएको छ ।

अनुमान विधि:

- क) आँकडामा भएका निर्भर र स्वतन्त्र चरहरू (Y / T) लाई तालिकामा राख्ने,
 ख) सबैभन्दा अगाडिको समयलाई T=1, त्यसपछिकोलाई T=2 गर्दै सबै समयलाई अङ्क दिने,
 ग) Y / T को मानहरूलाई गुणन (TY) गर्ने,
 घ) T को वर्ग (T²) निकाल्ने,
 ङ) Y, T, TY / T² का मानहरूको योगफल निकाल्ने,
 च) आएका मानहरू तलको सूत्रमा राखेर A / B को मान निकाल्ने

$$\hat{A} = \frac{(\sum y)(\sum t^2) - (\sum t)(\sum ty)}{n(\sum t^2) - (\sum t)^2} \quad (91)$$

$$\hat{B} = \frac{n(\sum ty) - (\sum t)(\sum y)}{n(\sum t^2) - (\sum t)^2} \quad (92)$$

नोट: सूत्र (११) र (१२) निकाल्नको लागि भ्रमांश (Error) को वर्गको योगफललाई न्यूनीकरण गर्नेगरी गणितीय विधि प्रयोग गरिएको छ । यहाँ भ्रमांश भन्नाले Y को वास्तविक मान र अनुमानित मानबिचको फरक भन्ने बुझ्नुपर्दछ ।

छ) \hat{A} र \hat{B} को मान माथिको समीकरण (१०) मा राख्ने ।

ज) अब, समीकरण (१०) लाई यसरी लेख्न सकिन्छ ।

$$\hat{Y} = \hat{A} + \hat{B}T \quad (93)$$

यहाँ, \hat{Y} भनेको Y को अनुमान हो । समीकरण १३ मा T का विभिन्न मानहरू राखेर चाहेको वर्षको लागि Y को अनुमान निकाल्न सकिन्छ अर्थात्, इन्टरपोलेसन, एक्स्ट्रापोलेसन वा प्रक्षेपण गर्न सकिन्छ । यहाँ ध्यान दिनुपर्ने कुरा के छ भने यो विधि समय श्रृङ्खला विश्लेषणको प्रारम्भिक चरणो विधि हो । यस विधिबाट छोटो समयका लागि प्रक्षेपण गर्न सकिन्छ तर लामो अवधिका लागि प्रक्षेपण गर्नका लागि गहन गणितीय तथा तथ्याङ्गीय ज्ञान र सीप चाहिन्छ । यद्यपि समय श्रृङ्खलामा लामो समयको लागि प्रक्षेपण गर्नु त्यति बुद्धिमता मानिँदैन । समयसँगै आउने विभिन्न प्रकारका आकस्मिक घटनाहरूले समय श्रृङ्खलाहरूलाई उल्लेखनीय असर गरिरहेका हुन्छन् जसका कारण प्रक्षेपण नमिल्न सक्दछ ।

समीकरण (१३) को प्रयोग गरेर प्रक्षेपण गर्नका लागि चाहिने \hat{A} र \hat{B} को मान कम्प्युटरबाट सजिलै निकाल्न सकिने हुनाले यहाँ, यो विधिबमोजिम उदाहरण दिइएको छैन । हामीले अन्य पाठहरूमा प्रतिगमन विश्लेषण र कम्प्युटरको प्रयोगसम्बन्धी यस बारेमा विस्तृत चर्चा गर्नेछौं । यद्यपि कम्प्युटरको STATA बाट कसरी नतिजा प्राप्त हुन्छ भनेर जस्ताको तस्तै तल प्रस्तुत गरिएको छ ।

उदाहरण तालिका १४.९ को तथ्याङ्क प्रयोग गरेर २०७८/७९ सालको लागि कुल गार्हस्थ्य उत्पादन (स्थिर मूल्यमा) प्रक्षेपण गर्नुहोस् ।

समीकरण (१३) को \hat{A} र \hat{B} का मानहरू STATA को प्रयोग गरी निकाल्दा, चित्र नं. १४.९ बमोजिम पाइन्छ । यसबाट \hat{A} र \hat{B} मानहरू राख्दा,

$$\hat{Y} = \hat{A} + \hat{B}T$$

$$\hat{Y} = १३,३८,५५६ + ७५,२४३ \times T \text{ हुन्छ ।}$$

अब, T को विभिन्न मानहरू राखेर जिडिपी प्रक्षेपण गर्न सकिन्छ । जस्तै: २०७८/७९ सालको लागि T को मान १२ राख्दा प्रक्षेपित जिडिपी २२,४१,४७३ (दश लाख) हुन आउँछ । आर्थिक वर्ष २०७८/७९ दिइएको समय श्रृङ्खलामा नभएका कारण यसको मान अनुमान गर्ने विधिलाई एक्स्ट्रापोलेसन पनि भनिन्छ । यस उदाहरणमा समय श्रृङ्खलाका विभिन्न अवयवहरूको उपचार नगरी सिधै प्रक्षेपणका लागि अनुमान गरिएका कारण भरपर्दो प्रक्षेपण भने नहुन सक्दछ ।

तालिका १४.९: नेपालको कुल गार्हस्थ्य उत्पादन

समय (T)	वर्ष	कुल गार्हस्थ्य उत्पादन (जिडिपी (Y)) (रु १०,००,०००)
१	२०६७/६८	१४,३६,०७३
२	२०६८/६९	१५,०७,१७२
३	२०६९/७०	१५,५३,५०२
४	२०७०/७१	१६,४२,७११
५	२०७१/७२	१७,००,४०५
६	२०७२/७३	१७,००,४४८
७	२०७३/७४	१८,४६,५०६
८	२०७४/७५	१९,८२,६५३
९	२०७५/७६	२१,०९,२६३
१०	२०७६/७७	२०,६४,६००
११	२०७७/७८R	१,४६,८२४
१२	२०७८/७९P	? ?

नोट: R परिमार्जित, P प्रारम्भिक

स्रोत: केन्द्रीय तथ्याङ्क विभाग

चित्र नं. १४.९: जिडिपी समय श्रृङ्खलाको अनुमान

Source	SS	df	MS	Number of obs	-	11
Model	6.2277e+11	1	6.2277e+11	F(1, 9)	-	265.30
Residual	2.1127e+10	9	2.3474e+09	Prob > F	-	0.0000
				R-squared	-	0.9672
				Adj R-squared	-	0.9635
Total	6.4389e+11	10	6.4389e+10	Root MSE	-	48450

वर्ण	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
समय	75243.05	4619.555	16.29	0.000	64792.89 85693.21
const	1338556	31331.35	42.72	0.000	1267680 1409433

१४.७ वृद्धिदर

समय श्रृङ्खलासम्बन्धी अध्ययन गर्दा विभिन्न चरहरूको वृद्धिदर पत्ता लगाउनु निकै महत्वपूर्ण मानिन्छ । जस्तै: जनसंख्याको वृद्धिदर, कुल गार्हस्थ्य उत्पादन वृद्धिदर, बिरामीको वृद्धिदर, मूल्य वृद्धिदर आदि । यस खण्डमा हामीले समयसँगै बढेका वा घटेका चरहरूको वृद्धिदर कसरी निकाल्ने भनेर संक्षिप्त चर्चा गर्दछौं । वृद्धिदर निकाल्नका

लागि प्रमुख रूपमा अङ्कगणितीय, ज्यामितीय, एक्सपोनेन्सियल, र लगायतदिमिक गरी चार प्रकारका विधिहरू प्रचलनमा रहेका छन् । तिनीहरूको प्रयोग देहायबमोजिम गरिन्छ ।

क) अंकगणितीय वृद्धिदर: कुनै चरको निश्चित समय अवधिमा स्थिर रूपमा मान बढ्दै जान्छ भने यो वृद्धिदरको प्रयोग गरिन्छ । यो वृद्धिदर निकाल्नका लागि निम्नानुसारको सूत्र प्रयोग गरिन्छ ।

$$r_a = \frac{(P_t - P_0)}{P_0} \times 100 \quad (१४)$$

जहाँ, r_a भनेको अंक गणितीय वृद्धिदर हो, P_t भनेको पछिल्लो समयको चरको मान हो, P_0 भनेको अघिल्लो समयको चरको मान हो, t भनेको समयावधि हो ।

उदाहरण: नेपालको जनसंख्या २०५८ सालमा २,३१,५१,४२३ रहेको थियो र २०६८ सालमा २,६४,९४,५०४ रहेको थियो भने जनसंख्याको वार्षिक अंकगणितीय वृद्धिदर कति हुन्छ ?

यहाँ, r_a निकाल्नु पर्ने मान हो,

P_t भनेको २,६४,९४,५०४ हो,

P_0 भनेको २,३१,५१,४२३ हो, र

t भनेको १० वर्ष हो ।

अब, यी मानहरूलाई सूत्र (१४) मा राख्दा,

$$r_a = \frac{\left(\frac{२,६४,९४,५०४ - २,३१,५१,४२३}{१०} \right)}{२,३१,५१,४२३} \times १००$$

$$= १.४४\%$$

तसर्थ, वार्षिक अंकगणितीय वृद्धिदर १.४४% हुन्छ ।

ख) ज्यामितीय वृद्धिदर: वास्तवमा कतिपय चरहरू मान स्थिर रूपमा नबढी तिनीहरूको वृद्धिदर तोकिएको समयअवधिमा स्थिर रूपमा बढ्ने हुन्छ । अर्को शब्दमा भन्दा, अंकगणितीय वृद्धिमा चरका सम्मुख समय (Adjacent Time) का मानहरूको फरक स्थिर हुन्छ भने ज्यामितीय वृद्धिमा चरका सम्मुख समयका मानहरूको अनुपात स्थिर हुन्छ । जस्तै: बैंकको ब्याज यदि चक्रीय ब्याजदर हो भने ब्याजको पनि ब्याज हुँदै बढ्दै जान्छ । यस्तो अवस्थामा ज्यामितीय वृद्धिदर प्रयोग गर्न उपयुक्त हुन्छ । जनसंख्याको हकमा

पनि स्थिर जनसंख्या घटबढ नभई जनसंख्याको वृद्धिदर नै बढ्ने वा घट्ने हुन्छ । यस्तो अवस्थामा ज्यामितीय वृद्धिदर प्रयोग गरिन्छ । यसका लागि निम्नानुसारको सूत्र प्रयोग गर्नुपर्दछ ।

$$r_g = \left[\sqrt[t]{\frac{P_t}{P_0}} - 1 \right] \times 100, \quad (१५)$$

जहाँ, r_g भनेको ज्यामितीय वृद्धिदर हो, P_t भनेको पछिल्लो समयको चरको मान हो, P_0 भनेको अघिल्लो समयको चरको मान हो, t भनेको समयावधि हो ।

उदाहरण: नेपालको जनसंख्या २०५८ सालमा २,३१,५१,४२३ रहेको थियो र २०६८ सालमा २,६४,९४,५०४ रहेको थियो भने जनसंख्याको वार्षिक ज्यामितीय वृद्धिदर कति हुन्छ ?

यहाँ, r_g निकाल्नु पर्ने मान हो,

P_t भनेको २,६४,९४,५०४ हो,

P_0 भनेको २,३१,५१,४२३ हो, र

t भनेको १० वर्ष हो ।

अब, यी मानहरूलाई सूत्र (१५) मा राख्दा,

$$r_g = \left[\sqrt[10]{\frac{२६४९४५०४}{२३१५१४२३}} - १ \right] \times १००$$

$$= १.३६\%$$

तसर्थ, वार्षिक ज्यामितीय वृद्धिदर १.३६% हुन्छ ।

ग) **एक्सपोनेन्सियल वृद्धिदर:** कतिपय चरहरूको मान कुनै समय खण्डमा मात्र नभएर प्रत्येक पल बढिरहेको हुन्छन् । जनसंख्या वृद्धि यसको उदाहरण हो । प्रत्येक पल जनसंख्या परिवर्तन भइरहेको हुन्छ । यस्तो अवस्थामा एक्सपोनेन्सियल वृद्धिदर प्रयोग गर्नुपर्दछ । ज्यामितीय वृद्धिले हरपलभन्दा केही लामो समय अन्तरको वृद्धिलाई समेटेको हुन्छ भने एक्सपोनेन्सियल वृद्धिले हरपलको वृद्धिलाई समेटेको हुन्छ । यो वृद्धिदर निकाल्न निम्नानुसारको सूत्र प्रयोग गर्नुपर्दछ ।

$$r_e = \frac{\ln\left(\frac{P_t}{P_0}\right)}{t} \times 100,$$

जहाँ, r_e भनेको एक्सपोनेन्सियल वृद्धिदर हो, P_t भनेको पछिल्लो समयको चरको मान हो, P_0 भनेको अघिल्लो समयको चरको मान हो, \ln भनेको प्राकृतिक लग (Natural

Log) हो, t भनेको समयावधि हो ।

उदाहरण: नेपालको जनसंख्या २०५८ सालमा २,३१,५१,४२३ रहेको थियो र २०६८ सालमा २,६४,९४,५०४ रहेको थियो भने जनसंख्याको वार्षिक एक्सपोनेन्सियल वृद्धिदर कति हुन्छ ?

यहाँ, r_e निकाल्नु पर्ने मान हो,

P_t भनेको २,६४,९४,५०४ हो,

P_0 भनेको २,३१,५१,४२३ हो, र

t भनेको १० वर्ष हो ।

अब, यी मानहरूलाई सूत्र (१६) मा राख्दा,

$$r_e = \frac{\ln\left(\frac{2,64,94,504}{2,31,51,423}\right)}{10} \times 100$$

$$= 1.35\%$$

तसर्थ, वार्षिक एक्सपोनेन्सियल वृद्धिदर १.३५% हुन्छ ।

घ) लगायिदमिक वृद्धिदर: कतिपय चरहरू अनन्तसम्म बढिरहने हुँदैन । जस्तै: जनसंख्याकै उदाहरणमा पनि अनन्त कालसम्म जनसंख्या बढिरहने भन्ने हुँदैन । तसर्थ, माथिल्लो सीमा पनि परिभाषित गर्ने गरी निकालिने वृद्धिदरलाई लगायिदमिक वृद्धिदर भनिन्छ । यो निकाल्न गहन गणितीय ज्ञान तथा सीपको आवश्यकता पर्ने तथा अभै धेरै तथ्याङ्क चाहिने हुनाले यसको बारेमा यस पुस्तकमा थप चर्चा गरिएको छैन ।

माथिको वृद्धिदर सम्बन्धी छलफलमा हामीले विभिन्न किसिमका वृद्धिदर कसरी निकाल्ने भनेर चर्चा गर्नु । तिनै सूत्रहरूबाट यदि हामीलाई वृद्धिदर थाहा छ भने तोकिएको समयको लागि जनसंख्या निकाल्न वा प्रक्षेपण गर्न सकिन्छ ।

उदाहरण: नेपालको जनसंख्या २०६८ सालमा २,६४,९४,५०४ रहेको थियो र वृद्धिदर १.३५% छ भने २०७८ सालको जनसंख्या कति होला ?

यहाँ, r_e भनेको १.३५% अर्थात् ०.०१३५ हो,

P_t २०७८ सालको जनसंख्या हो, जुन निकाल्नु पर्नेछ,

P_0 भनेको २,६४,९४,५०४ हो ।

यी मानहरूलाई सूत्रमा राख्नुपूर्व सूत्र (१६) लाई पुनः परिभाषित गरौं।

$$r_e = \frac{\ln\left(\frac{P_t}{P_0}\right)}{t} \times 100$$

$$t * r_e = \ln\left(\frac{P_t}{P_0}\right) \times 100$$

$$\frac{t * r_e}{100} = \ln\left(\frac{P_t}{P_0}\right)$$

यहाँ, r को मानलाई प्रतिशतबाट हटाएपछि सूत्रको १०० हटेर जान्छ। त्यसपछि दुवैतर्फ एक्सपोनेन्सियल लिने र \ln हटाउने। अर्थात्,

$$e^{t r_e} = \frac{P_t}{P_0}$$

$$P_0 e^{t r_e} = P_t$$

अब, यो सूत्रमा मानहरू राखौं।

$$P_t = २६४९४५०४ \times e^{१० \times ०.०१३५} = २६४९४५०४ \times १.१४४५ \\ = ३,०३,२३,९३४$$

तसर्थ, २०७८ सालको प्रक्षेपित जनसंख्या ३,०३,२३,९३४ हुन्छ।

जनसंख्याको वृद्धिदरका कुरा गर्दा जनसंख्या दोब्बर हुन लाग्ने समय पनि निकै महत्त्वपूर्ण रूपमा हेरिन्छ। जनसंख्या दोब्बर हुन लाग्ने समय निकाल्न निम्न बमोजिमको सूत्र प्रयोग गरिन्छ।

$$t = \frac{\ln 2}{r}, \quad (१७)$$

जहाँ t ले जनसंख्या दोब्बर हुन लाग्ने समय प्रतिनिधित्व गरेको छ भने r ले जनसंख्या वृद्धिदर प्रतिनिधित्व गरेको छ। नेपालको जनसंख्या वृद्धिदर १.३५% छ भने जनसंख्या दोब्बर हुन कति वर्ष लाग्ला, हेरौं।

$$t = \frac{\ln 2}{r} = \frac{०.६९३४१७}{०.०१३५} = ५१.४ \text{ वर्ष।}$$

तसर्थ, नेपालको जनसंख्या दोब्बर हुनका लागि ५१.४ वर्ष लाग्दछ।

१४.८ अभ्यासका लागि प्रश्नहरू

क) संक्षिप्त व्याख्या गर्नुहोस् ।

- १) समय श्रृङ्खला
- २) मौसमी विचलन
- ३) एक्टापोलेसन

ख) ठिक वा बेठिक छुट्याउनुहोस् ।

- १) विभिन्न समय अवाधिमा मिलाएर राखेको तथ्याङ्कलाई क्रस-सेक्सन तथ्याङ्क भनिन्छ ।
- २) जनसंख्या वृद्धिदर निकाल्न अंकगणितीय वृद्धिदर प्रयोग गर्नु उपयुक्त हुन्छ ।
- ३) समय श्रृङ्खलाको प्रवृत्ति ग्राफको सहयोगबाट पनि पत्ता लगाउन सकिन्छ ।

ग) खाली ठाउँ भर्नुहोस् ।

- १) समयको कुनै एक बिन्दुमा लिइएको वा प्रस्तुत गरिएको तथ्याङ्क तथ्याङ्क हो ।
- २) कुनै पनि समयमा घटबढ भइरहन सक्ने वा आकस्मिक घटनाका कारण परिवर्तन हुने प्रकृतिको श्रृङ्खलालाई उतारचढाव भनिन्छ ।
- ३) विभिन्न समयका क्रस-सेक्सन तथ्याङ्क समूहलाई एउटै तथ्याङ्क समूहमा मिलाएर बनाइएको तथ्याङ्कलाई तथ्याङ्क भनिन्छ ।

घ) जोडा मिलाउनुहोस् ।

जनसंख्या दोब्बर हुन लाग्ने समय $\frac{(P_t - P_0)}{P_0} \times 100$

एक्पोनेन्सियल वृद्धिदर $[\sqrt[t]{\frac{P_t}{P_0}} - 1] \times 100$

ज्यामितीय वृद्धिदर $\frac{\ln(\frac{P_t}{P_0})}{t} \times 100$

अंकगणितीय वृद्धिदर $\frac{\ln 2}{r}$

ङ) समय श्रृङ्खला विश्लेषणको महत्त्वका बारेमा उदाहरणसहित व्याख्या गर्नुहोस् ।

च) आर्थिक वर्ष २०७६/७७ मा नेपालको कुल गार्हस्थ्य उत्पादन आधार तथा स्थिर मूल्यमा रू. २०,६४,५९९.९ (दश लाखमा) थियो र २०७७/७८ मा रू. २१,४६,८२४.३ (दश लाखमा) थियो भने उक्त अवधिको आर्थिक वृद्धिदर अंकगणितीय, ज्यामितीय, र एक्पोनेन्सियल विधिबाट निकालेर तुलना गर्नुहोस् । साथै, नेपालको कुल गार्हस्थ्य उत्पादन दोब्बर हुन कति समय लाग्ला निकाल्नुहोस् ।

१४.९ सन्दर्भ सामग्री

समय श्रृङ्खलासम्बन्धी थप अध्ययन गर्नका लागि निम्नानुसारका सामग्री उपयोगी छन् ।

१. Statistical Techniques in Business & Economics by Douglas A. Lind, William G. Marchal & Samuel A. Wathen, Seventeenth Edition, McGraw-Hill Education, 2018
२. Fundamentals of Statistics by S.C. Gupta, Sixth Revised & Enlarged Edition, Himalaya Publishing House, 2009



पाठ १५

नर्मल डिस्ट्रिब्युसन

१५.१ परिचय

नर्मल डिस्ट्रिब्युसन अविच्छिन्न चरको सम्भाव्यता वितरण अध्ययन गर्ने एउटा महत्वपूर्ण डिस्ट्रिब्युसन हो । यस किसिमको डिस्ट्रिब्युसन अध्ययनका लागि सम्भाव्यता (Probability) को आधारभूत धारणा आवश्यकता हुन्छ । यस पुस्तकको दायराभिन्न सम्भाव्यता नपरेको हुनाले सम्भाव्यता भनेको के हो त्यसको सामान्य अर्थमात्र हेरौं ।

सम्भाव्यता: सम्भावना भएका सम्पूर्ण जनसंख्याबाट कुनै एक किसिमको घटना घट्ने दरलाई सो घटनाको सम्भाव्यता भनिन्छ । मानौं, A ले कुनै एउटा घटनालाई प्रतिनिधित्व गरेको छ । यो घटनामा पर्ने सम्भावना भएका सम्पूर्ण जनसंख्याको समूहलाई N ले जनाउने हो भने, सो समूहबाट कुनै पनि सदस्यले उक्त घटनामा पर्न सक्ने सम्भावनालाई निम्नानुसार व्यक्त गरिन्छ ।

$$P(A) = \frac{m}{N}, \quad (१)$$

जहाँ, P(A) भनेको घटना A को सम्भाव्यता हो, m भनेको घटना A को अनुकुल हुनेगरी घट्ने घटनाको संख्या हो भने N भनेको जम्मा नतिजाहरूको समूह हो । उदाहरणका लागि कुनै गाउँपालिकामा एक वर्ष अर्धमा गर्भवती महिलाहरूको संख्या २०० छ, र उनीहरूमध्ये ११० जनाले छोरी जन्माए र बाँकी ९० जनाले छोरा जन्माए भने, छोरी जन्मने सम्भाव्यता $११०/२०० = ०.५५$ अर्थात्, ५५% हुन्छ । सम्भाव्यताको मान ० देखि १ सम्म मात्र हुन्छ । सम्भाव्यता शून्य भनेको असम्भव घटना हो । जस्तै कुनै समुदायमा ५ वटा आँखा हुने मान्छेको सम्भाव्यता शून्य हुन्छ । सम्भाव्यता १ हुनु भनेको निश्चित घटना हो । जस्तै: १० जना केटाहरूको समूहबाट एक जना छानियो भने केटा नै पर्ने सम्भावना १ हुन्छ अर्थात् निश्चित रूपमा केटा नै छानिन्छ ।

थप उदाहरण: इलाम जिल्लाको सूर्योदय नगरपालिको आर्थिक वर्ष २०७६/७७ को अनुमानित गर्भवती संख्या १,५४६ जना थियो र अनुमानित जीवित जन्म संख्या १,३११ जना थियो ।¹¹ यी आँकडाका आधारमा उक्त नगरपालिकामा कुल अनुमानित गर्भवतीबाट जीवित बच्चा जन्माउने हुने सम्भाव्यता निम्नानुसार निकाल्न सकिन्छ ।

11 Department of Health Services. (2021). Annual Report 2076/77 (2019/20). Kathmandu, Nepal; Department of Health Services. Retrieved from December 14, 2021 from: <http://dohs.gov.np/annual-report-2076-77-2019-20/>

$$\begin{aligned} \text{सूर्योदय नगरपालिकाका पुरुषले बच्चा जन्माउने सम्भावना} &= \frac{१३,११}{१५,४६} \\ &= ०.८५ \end{aligned}$$

असम्भव घटना: सूर्योदय नगरपालिकाका पुरुषले बच्चा जन्माउने सम्भावना = ०

निश्चित घटना: सूर्योदय नगरपालिकामा गर्भवती महिलाहरूको समूहबाट एकजना छान्दा गर्भवती महिला नै पर्ने सम्भावना = १

सम्भाव्यताको सामान्य जानकारी प्राप्त भएपछि नर्मल डिस्ट्रिब्युसनका बारेमा थप बुझ्न सजिलो हुन्छ। यसमा कुनै अविच्छिन्न चरले दिइएको अन्तरालका मानहरू लिनै सम्भाव्यता कति हुन्छ, त्यस्ता सबै सम्भावित सम्भाव्यताहरूको वितरण जसका केही अद्वितीय विशेषताहरू हुन्छन् भने त्यस्तो सम्भाव्यताहरूको वितरणलाई नर्मल डिस्ट्रिब्युसन भनिन्छ। जस्तै: नेपालमा जन्मने बच्चाहरूको तौल एउटा अविच्छिन्न चरको उदाहरण हो। यो चरले लिनै मान २ के.जी. देखि २.५ के.जी. सम्म हुने सम्भाव्यता कति हुन्छ भनेर हेर्नका लागि उक्त चरको सम्भाव्यताको वितरण (Probability Distribution) आवश्यक पर्दछ। यदि यो चरको सम्भाव्यताको वितरणसँग नर्मल डिस्ट्रिब्युसनका विशेषतासँग मिल्ने गुणहरू छन् भने माथि उल्लेख गरिए बमोजिमको अन्तरालका मान (२ के.जी. देखि २.५ के.जी. सम्म) लिनै सम्भाव्यता निकाल्न सकिन्छ र यस्ता वितरणलाई नर्मल डिस्ट्रिब्युसन भनिन्छ। त्यस्ता अद्वितीय प्रमुख विशेषताहरू निम्नानुसार छन्।

क) यसको प्रोब्याबिलिटी डेन्सिटी फङ्सन (Probability Density Function) निम्नानुसार हुन्छ:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty, \quad (२)$$

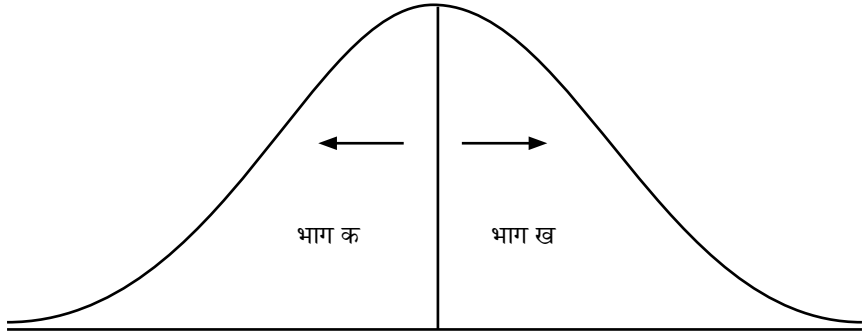
जहाँ, x भनेको कुनै अविच्छिन्न चर हो, σ भनेको त्यो चरको स्तरीय भिन्नता हो, μ भनेको त्यो चरको मध्यक हो। $\pi = 22/7$, $e = 2.71828$ (प्राकृतिक लगको आधार) हो। यहाँ, x ले वास्तविक संख्या (Real Line) का कुनै पनि मान लिन सक्दछ। तर तल्लो सीमा $-\infty$ र माथिल्लो सीमा ∞ का संख्याहरू कहिल्यै अस्तित्वमा हुँदैनन्। तर व्यवहारमा यदि तथ्याङ्कको आकार (N) ठुलो भयो भने त्यसलाई अनन्त मानेर अन्य कार्यहरू गरिन्छ।¹² रोचक कुरा के छ भने, x ले कुनै एक निश्चित संख्या मान लिनै सम्भावना शून्य हुन्छ किनभने वास्तविक संख्या रेखामा अनगिन्ती संख्याहरू हुन्छन्, ती संख्याहरूले

12 कसैले ३० भन्दा ठुलोलाई, कसैले ५० भन्दा ठुलोलाई, लाई ठुलो मानेका हुन्छन्। यो विषय हेरी हुन्छ।

कुनै एउटा संख्यालाई भाग गर्दा लगभग शून्य नै हुन्छ । अर्थात्, सम्भाव्यता निकाल्दा $1/∞ = 0$ हुन्छ । तसर्थ, x ले वास्तविक संख्या रेखाको कुनै अन्तराल लिने सम्भाव्यता 0 र 1 को बिचको कुनै संख्या हुन्छ ।

ख) यसको प्रोब्याबिलिटी डेन्सिटी फङ्सनको बक्ररेखा घण्टी आकार (Bell Shape) को हुन्छ । जुन बराबर उस्तै दुई भागमा विभाजन गर्न सकिने (Symmetrical) हुन्छ चित्र नं. १५.१ (क) मा देखाइएका दुई भाग क र ख दुवै उस्तै र बराबर छन्, जसलाई एउटा सिधा ठाडो रेखाले विभाजन गरेको छ । तसर्थ, यो बक्ररेखा सिमेट्रिकल (Symmetrical) छ ।

चित्र नं. १५.१ (क): नर्मल प्रोब्याबिलिटी डेन्सिटी फङ्सनको बक्ररेखाको आकार



ग) नर्मल डिस्ट्रिब्युसन हुने चरको मध्यक, मध्यिका र बहुलक (रीत) सबै बराबर हुन्छन् । तसर्थ, नर्मल डिस्ट्रिब्युसनका लागि,

मध्यक = मध्यिका = बहुलक अर्थात्, Mean = Median = Mode लेख्न सकिन्छ ।

घ) चर x ले वास्तविक संख्या a र b ($a < b$) बिचको मान लिने सम्भावनालाई निम्नानुसार निकाल्न सकिन्छ ।

$$p(a < x < b) = \int_a^b f(x)dx = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx,$$

जहाँ, \int लाई इन्टिग्रेसन (Integration) भनिन्छ । यो गणितीय कार्य गर्नका लागि गणितमा गहन ज्ञान हुनुपर्ने हुनाले यसका बारेमा यहाँ थप चर्चा गरिएको छैन । यद्यपि, समीकरण (३) ले प्रोब्याबिलिटी डेन्सिटी फङ्सनअन्तर्गत बिन्दु a र b बिचको क्षेत्रफल दिन्छ । यो क्षेत्रफल चर x ले वास्तविक संख्या a र b ($a < b$) बिचको मान लिने सम्भाव्यतालाई जनाउँछ । माथि उल्लेख गरे बमोजिम चर x ले कुनै निश्चित बिन्दु लिने

सम्भाव्यता शून्य हुन्छ किनभने एउटा मात्र निश्चित बिन्दुमा बक्ररेखा मुनिको क्षेत्रफल शून्य हुन्छ ।

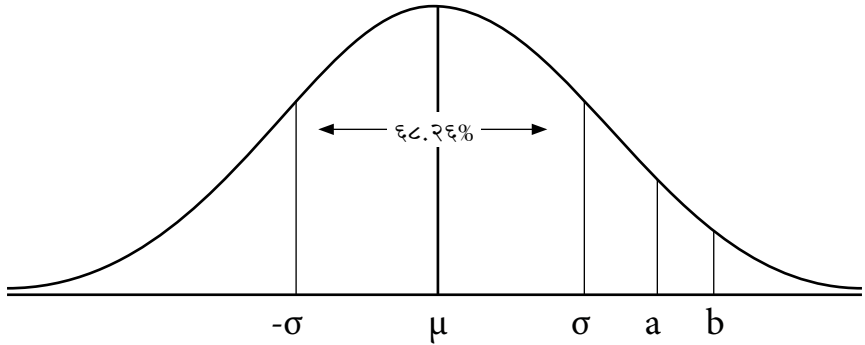
नोट: नर्मल बक्ररेखा र X अक्षबिचको केही महत्त्वपूर्ण क्षेत्रफल विवरण

$\mu + "0"$ ले कुल क्षेत्रफलको लगभग ६८.२६ % समेट्छ ।

$\mu + "2\sigma"$ ले कुल क्षेत्रफलको करिब ९५.४४% समेट्छ ।

$\mu + "3\sigma"$ ले कुल क्षेत्रफलको लगभग ९९.७३ % समेट्छ ।

चित्र नं. १५.१ (ख): नर्मल डिस्ट्रिब्युसन फङ्सनको बक्ररेखा नमुना



सम्भाव्यता वितरणहरूमध्ये तथ्याङ्कीय क्षेत्रमा सबैभन्दा बढी प्रयोग हुने विषय नर्मल डिस्ट्रिब्युसन नै हो । किनभने अनुमानात्मक तथ्याङ्कको अध्ययनमा, परिकल्पना परीक्षणको अध्ययनमा, नमुना आकार अनुमान गर्ने कार्यमा तथा अन्य विविध अध्ययनहरूमा यसको व्यापक प्रयोग भएको हुन्छ । कतिपय अवस्थामा खण्डित तथ्याङ्क वितरणका आँकडा संख्या वा आवृत्ति ठुलो हुँदै गएमा तिनीहरूलाई अविच्छिन्न श्रेणीमा रूपान्तरण गरेर नर्मल डिस्ट्रिब्युसनमा लगिन्छ । उदाहरणका लागि जनगणनामा व्यक्तिको उमेर खण्डित श्रेणीको अवधारणामा संकलन गरिएको हुन्छ । तर कतिपय अवस्थामा पछि यी उमेरका तथ्याङ्कलाई अविच्छिन्न श्रेणीमा लगेर विश्लेषण गरिन्छ । तसर्थ, यसलाई ठुला संख्याको सिद्धान्त (Law of Large Numbers) भन्ने पनि गरिन्छ । यो डिस्ट्रिब्युसको अध्ययन गर्न तथ्याङ्कशास्त्रको व्यापक अवधारणा आवश्यक पर्ने भए तापनि पाठकवर्गलाई सामान्य जानकारी दिने उद्देश्यले यस पुस्तकमा यसका बारेमा परिचयात्मक वर्णन मात्र गरिएको हो ।

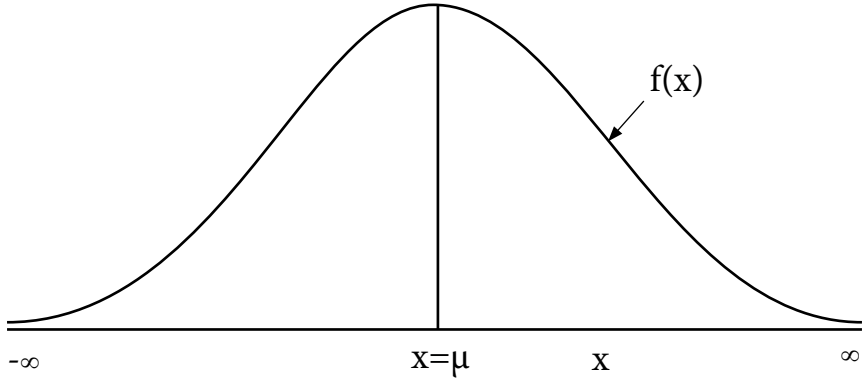
१५.२ नर्मल डिस्ट्रिब्युसन बक्ररेखाका विशेषता र समीकरण

कुनै अविच्छिन्न चर x ले नर्मल डिस्ट्रिब्युसन फलो गर्दछ भने यसका मानहरूलाई x अक्षमा र समीकरण (३) मा दिइएको $f(x)$ का मानलाई Y अक्षमा राखेर बनाइएको

बक्ररेखालाई नर्मल डिस्ट्रिब्युसन बक्ररेखा भनिन्छ । यहाँ, $f(x)$ भनेको माथि समीकरण (२) मा परिभाषित गरेको प्रोब्याबिलिटी डेन्सिटी फङ्सन नै हो । यसको मद्दतले x ले विभिन्न वास्तविक संख्या रेखाको अन्तराल लिने सम्भाव्यता निकाल्न सकिन्छ ।

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty, \quad (३)$$

चित्र नं. १५.२ (क): नर्मल डिस्ट्रिब्युसन फङ्सनको बक्ररेखा



यो बक्ररेखाका विशेषताहरू देहायबमोजिम छन् ।

क) यो घण्टी आकारको हुन्छ ।

ख) यो बराबर दुई उस्तै भागमा विभाजन गर्न सकिने हुन्छ । यस्तो दुई भाग लगाउने सिधा रेखा मध्यक (μ) बाट खिच्न सकिन्छ ।

ग) यो बक्ररेखाको तल्लो सीमा र माथिल्लो सीमा अनन्त (∞) मा हुन्छ, जुन कहिल्यै भेट्न सकिँदैन ।

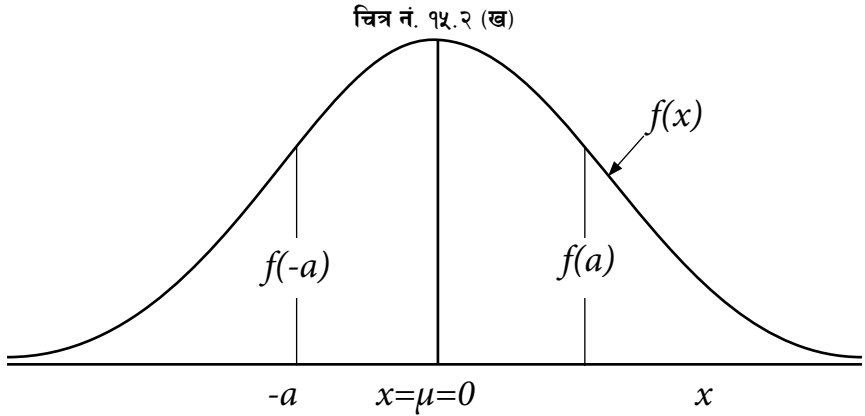
घ) यो बक्ररेखाको तल्लो सीमा र माथिल्लो सीमाले X अक्षलाई छुन छुन खोज्दछ (asymptotic to x -axis), तर कहिल्यै छुँदैन अर्थात्, अनन्तमा गएर मात्र छुन्छ ।

ङ) यो बक्ररेखा मुनि र x अक्ष माथिको पूरा क्षेत्रफल १ हुन्छ । तसर्थ, मध्यक रेखाले यसलाई बराबर उस्तै दुई भागमा क्षेत्रफल ०.५/०.५ हुनेगरी विभाजन गर्दछ ।

च) यसले x अक्षसँग बनाउने क्षेत्रफल चर x को सम्भाव्यता मापन गर्न प्रयोग गरिन्छ ।

छ) मध्यक शून्य हुने र यही बिन्दुमा Symmetrical हुने प्रोब्याबिलिटी डेन्सिटी

फड्सनमा $f(-a) = f(a)$, (वास्तविक संख्या रेखाका सबै संख्या "a" को हकमा लागु हुने) हुन्छ (चित्र नं. १५.२ (ख)) ।



१५.३ स्टाण्डार्ड नर्मल डिस्ट्रिब्युसन

तथ्याङ्कीय कार्यहरू जस्तै: अनुमानात्मक तथ्याङ्कको अध्ययन गर्न, परिकल्पना परीक्षणको अध्ययन गर्न, नमुना आकार अनुमान गर्न प्रयोग गरिने सूत्रहरूमा स्टाण्डार्ड नर्मल चरको प्रयोग हुन्छ । यो चरको प्रोब्याबिलिटी डेन्सिटी फड्सनको बक्ररेखा र यससँग सम्बन्धित सम्भाव्यताको प्रयोग गर्न सजिलो हुन्छ । किनभने यो चरको मध्यक शून्य र स्तरीय भिन्नता १ हुन्छ जसले अध्ययनलाई सहज बनाउँछ । नर्मल डिस्ट्रिब्युसनको मध्यक र स्तरीय भिन्नता यसरी स्थिर नहुने र स्टाण्डार्ड नर्मल डिस्ट्रिब्युसनको स्थिर हुने हुनाले तथ्याङ्कीय कार्यहरूमा स्टाण्डार्ड नर्मल डिस्ट्रिब्युसनको प्रयोग ठुलो महत्त्व छ । नर्मल डिस्ट्रिब्युसन भनेको एउटा विशेष प्रकारको नर्मल डिस्ट्रिब्युसन हो । यसमा नर्मल डिस्ट्रिब्युसनको चर x लाई देहायबमोजिम रूपान्तरण गरी नयाँ चर z मा परिभाषित गरिन्छ र यो z लाई स्टाण्डार्ड नर्मल भेरियन्ट भनिन्छ ।

$$z = \frac{x - \mu}{\sigma}, \quad (४)$$

जहाँ, x भनेको नर्मल डिस्ट्रिब्युसनको चर हो, σ भनेको x स्तरीय भिन्नता हो, μ भनेको x को मध्यक हो । यसरी परिभाषित गरेको z लाई स्टाण्डार्ड नर्मल चर (Standard Normal Variable) भनिन्छ । नयाँ परिभाषित z को मध्यक ० हुन्छ भने स्तरीय भिन्नता १ हुन्छ । यस अवस्थामा यसको प्रोब्याबिलिटी डेन्सिटी फड्सन निम्नानुसारको हुन्छ ।

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \quad -\infty < z < \infty. \quad (५)$$

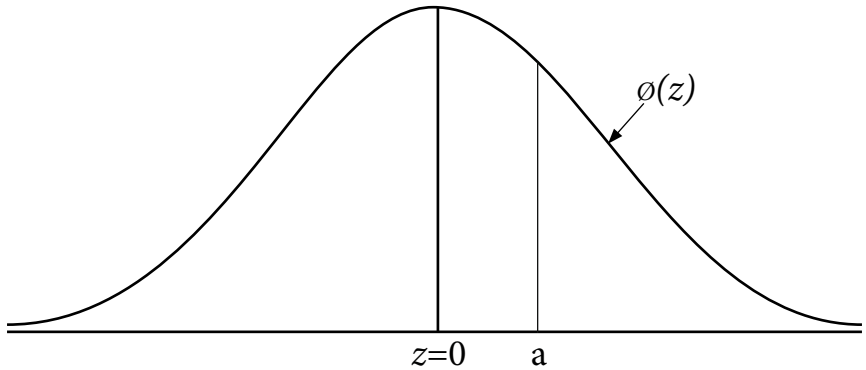
समीकरण (५) को प्रयोगबाट z ले लिने अन्तरालको सम्भाव्यता पत्ता लगाउन सकिन्छ । चर x को जस्तै z को पनि कुनै निश्चित बिन्दुमा सम्भाव्यता शून्य हुन्छ । तर अन्तराल सम्भाव्यता भने तलको सूत्रबाट निकालिन्छ ।

$$p(a < z < b) = \int_a^b \phi(z) dz = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz. \quad (६)$$

समीकरण (६) का सन्दर्भमा पनि माथि भनिए जस्तै गहन गणितीय ज्ञान आवश्यक पर्ने हुनाले थप चर्चा गरिएको छैन ।

समीकरण (५) बाट x अक्षमा z का मान र y अक्षमा $\phi(z)$ का मान राखेर बनाइएको बक्ररेखालाई स्टाण्डर्ड नर्मल बक्ररेखा (Standard Normal Curve) भनिन्छ । तथ्याङ्कविद्हरूले यो मानक सामान्य चरको वास्तविक सम्भावना पत्ता लगाउन सामान्य सम्भावना बक्र विकास गरेका छन् । यसलाई चित्र १५.३ मा देखाइएको छ ।

चित्र १५.३: स्टाण्डर्ड नर्मल बक्ररेखाको नमुना



१५.४ स्टाण्डर्ड नर्मल सम्भाव्य बक्ररेखा अन्तर्गतको क्षेत्रफल

तथ्याङ्कका विविध कार्यहरूमा स्टाण्डर्ड नर्मल सम्भाव्य बक्ररेखा अन्तर्गतको क्षेत्रफल आवश्यक पर्दछ । जस्तै: परिकल्पना परीक्षण गर्दा, अनुमान गर्दा, नमुना छनोटमा नमुनाको आकार पत्ता लगाउँदा आदि । स्टाण्डर्ड नर्मल सम्भाव्य बक्ररेखा अन्तर्गतको क्षेत्रफलबाट z को मान पत्ता लगाउन होस् वा z को मानबाट तोकिएको क्षेत्रफल पत्ता लगाउनुहोस्, यसको प्रयोग महत्त्वपूर्ण हुन्छ । यसका लागि हामीसँग z तालिका उपलब्ध हुन्छ । यस्ता तालिका विभिन्न प्रकारले तयार गरी राखिएका हुन्छन् । तसर्थ, तालिका प्रयोग गर्नुपूर्व त्यसको बारेमा पूर्ण जानकारी राख्नुपर्ने हुन्छ । कुनै तालिकामा $P(-\infty < z < a)$ बमोजिम

z ले a भन्दा सानो हुने सम्भाव्यता दिइएको हुन्छ भने कुनैमा $P(0 < z < a)$ बमोजिम z ले 0 भन्दा ठुलो तर a भन्दा सानो हुने सम्भाव्यता दिइएको हुन्छ। हामीले यस पुस्तकमा प्रयोग गरेको तालिका (अनुसूची २) ले $P(-\infty < z < a)$ बमोजिमको क्षेत्रफल दिएको छ। अर्थात्, अनुसूची २ को तालिकामा भएका मानहरू भनेको स्टान्डर्ड नर्मल भेरियट z ले $-\infty$ देखि कुनै दिइएको मान (a) सम्म मान लिने सम्भाव्यता हो। जस्तै: तालिकामा z ले $-\infty$ देखि 0 सम्म मान लिने सम्भाव्यता 0.5 पाइन्छ। यही तरिकाले अन्य अन्तरालको सम्भाव्यता पनि निकाल्न सकिन्छ। स्टान्डर्ड नर्मल डिस्ट्रिब्युसनको Symmetrical विशेषताको कारण यदि हामीलाई $P(0 < z < a)$ आवश्यक पर्यो भने $P(0 < z < a) = P(-\infty < z < a) - 0.5$ गरेर निकाल्न सकिन्छ। आजकल त यस्ता तालिका अनलाइनमा अन्तरक्रियात्मक विशेषतासहित पनि पाइन्छन्।¹³

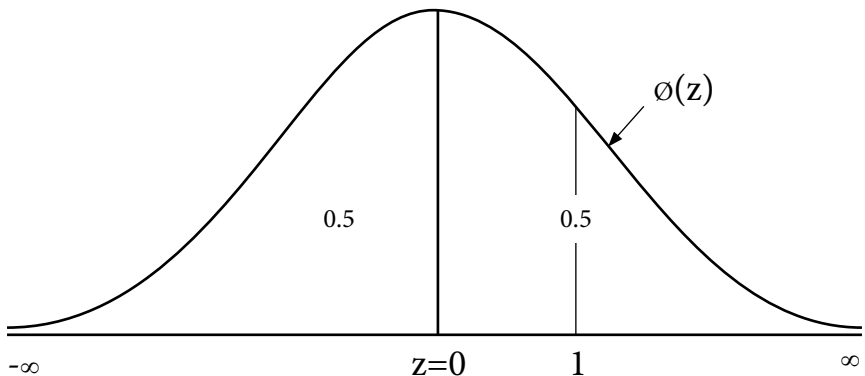
उदाहरण: अनुसूची २ मा दिइएको तालिका प्रयोग गरेर z ले 0 भन्दा ठुलो तर 1 भन्दा सानो मान लिने सम्भाव्यता निकाल्नुहोस्।

यसका लागि पहिला तालिकाको प्रयोग गरेर z ले $-\infty$ भन्दा ठुलो तर 1 भन्दा सानो मान लिने सम्भाव्यता निकालौं।

$P(-\infty < z < 1) = 0.2420$ हुन्छ (अनुसूची २ को तालिका बाट लिइएको)। अब, $P(0 < z < 1) = P(-\infty < z < 1) - 0.5 = 0.2420$ हुन्छ।

तसर्थ, z ले 0 भन्दा ठुलो तर 1 भन्दा सानो मान लिने सम्भाव्यता 0.2420 हुन्छ।

चित्र १५.४: स्टान्डर्ड नर्मल बक्ररेखाको क्षेत्रफल

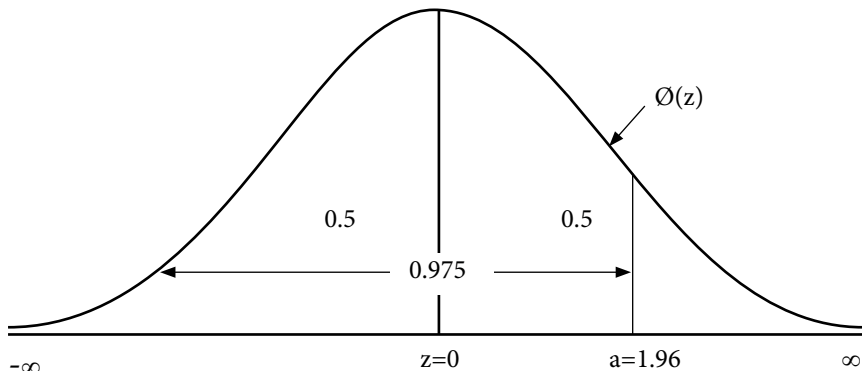


13 Maths is Fun. Retrieved on December 6, 2021 from: <https://www.mathsisfun.com/data/standard-normal-distribution-table.html>

उदाहरण: यदि $P(-\infty < z < a) = 0.975$ छ भने, a को मान कति हुन्छ ?

यसका लागि तालिकामा ०.९७५ संख्या पत्ता लगाऔं । यो संख्या तालिकाको १.९ तेर्सो लहर र ०.०६ ठाडो लहरमा छ । तसर्थ, a को मान १.९६ हुने भयो ।

चित्र १५.५: स्टाण्डर्ड नर्मल बक्ररेखाको क्षेत्रफल



१५.५ नर्मल डिस्ट्रिब्युसनको महत्त्व र प्रयोग

यो तथ्याङ्कमा सबैभन्दा महत्त्वपूर्ण सम्भाव्यता वितरण हो किनभने यो धेरै प्राकृतिक घटनाहरूमा उपयुक्त हुन्छ । उदाहरणका लागि, उचाइ, रक्तचाप, मापन त्रुटि, र आई. क्यू. (IQ) स्कोरहरूले यस किसिमको सामान्य वितरणलाई पछ्याउँछन् । यसलाई गौसियन वितरण (Gaussian Distribution) पनि भनिन्छ । यसलाई मध्यबाट मानक विचलनहरूको निर्दिष्ट संख्याभिन्न पर्ने मानहरूको अनुपात निर्धारण गर्न प्रयोग गर्न सकिन्छ । उदाहरणका लागि, नर्मल डिस्ट्रिब्युसनमा ६८.२६ % अवलोकनहरू (मानहरू) औसतबाट ± 1 मानक विचलनभिन्न आउँछन् । अर्थात्, $(\mu - 1)$ देखि $(\mu + 1)$ सम्मको अन्तरालमा ६८.२६% आँकडाहरू रहेका हुन्छन् । यसका अतिरिक्त आर्थिक तथा सामाजिक घटनाहरूसँग सम्बन्धित तथ्याङ्कमा यसको व्यापक प्रयोग हुन्छ ।

१५.६ अभ्यासका लागि प्रश्नहरू

क) खाली ठाउँ भर्नुहोस् ।

- १) नर्मल डिस्ट्रिब्युसनले चरको सम्भाव्यता वितरणको अध्ययन गर्दछ ।
- २) सम्भाव्यताको मान देखि सम्म मात्र हुन्छ ।
- ३) निश्चित भएको घटनाको सम्भाव्यता हुन्छ ।
- ४) असम्भव घटनाको सम्भाव्यता हुन्छ ।
- ५) स्टाण्डर्ड नर्मल पछ्याउने चरको मध्यक र ? हुन्छ ।

ख) फरक छुट्याउनुहोस् ।

- १) नर्मल डिस्ट्रिब्युसन र स्टाण्डर्ड नर्मल डिस्ट्रिब्युसन
- २) निश्चित घटना र असम्भव घटना
- ३) अविच्छिन्न र खण्डित चर

ग) ठिक बेठिक छुट्याउनुहोस् ।

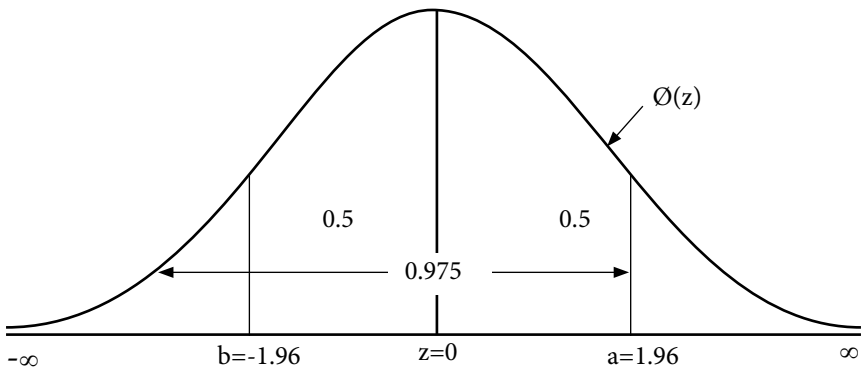
- १) $P(-\infty < z < 2) = 0.975$ हुन्छ ।
- २) नर्मल बक्ररेखा र X अक्षबिच $\mu + "2\sigma"$ ले कुल क्षेत्रफलको करिब ९५.४४% समेट्छ ।
- ३) सम्भाव्यताको मान नकारात्मक पनि हुन सक्दछ ।
- ४) पाँच वर्षभरिमा कुनै एउटा फेब्रुअरी महिनामा ३१ वटा दिन हुने सम्भावना शून्य हुन्छ ।
- ५) स्टाण्डर्ड नर्मल बक्ररेखा र X अक्षबिचको पूरा क्षेत्रफल १ हुन्छ ।

घ) अनुसूची २ मा दिइएको तालिका प्रयोग गरेर यदि $P(-0 < z < a) = 0.975$ छ भने, a को मान कति हुन्छ, सचित्र वर्णन गर्नुहोस् ।

ङ) तथ्याङ्कीय गतिविधिमा नर्मल डिस्ट्रिब्युसन किन महत्त्वपूर्ण मानिन्छ, उल्लेख गर्नुहोस् ।

च) नर्मल प्रोब्याबिलिटी बक्ररेखाका प्रमुख विशेषताहरू उल्लेख गर्नुहोस् ।

छ) तलको चित्र अध्ययन गरेर सोधिएका प्रश्नहरूको जवाफ दिनुहोस् ।



चित्र १५.५ स्टाण्डर्ड नर्मल बक्ररेखासम्बन्धी अभ्यास

- १) यो चित्रमा $P(a < z < \infty)$ कति हुन्छ ?
- २) यो चित्रमा $P(-\infty < z < \infty)$ कति हुन्छ ?

- ३) यो चित्रमा $P(-\infty < z < b)$ कति हुन्छ ?
 ४) यो चित्रमा $P(b < z < \infty)$ कति हुन्छ ?
 ५) यो चित्रमा $P(z=1.96)$ कति हुन्छ ?
 ६) यो चित्रमा $P(-\infty < z < 0)$ कति हुन्छ ?
 ७) $P(-\infty < z < b)$ र $P(a < z < \infty)$ जनाउने क्षेत्रफललाई चित्रमा छायौं पारेर देखाउनुहोस् । साथै यी दुई क्षेत्रफलहरू बराबर छन् वा छैनन् तर्क प्रस्तुत गर्नुहोस् ।

१५.६ सन्दर्भ सामग्री

नर्मल डिस्ट्रिब्युसनका बारेमा थप अध्ययन गर्न देहायको सामग्री उपयुक्त छ ।

१. Fundamentals of Statistics by S.C. Gupta, Sixth Revised & Enlarged Edition, Himalaya Publishing House, 2009.



अनुमान तथा नमुनासम्बन्धी आधारभूत जानकारी

१६.१ अनुमानको सिद्धान्त

अध्ययन गरिने समग्र क्षेत्र वा जनसंख्याका विशेषताहरू थाहा पाउन गणना वा नमुना सर्वेक्षण गरिन्छ । गणना वा नमुना सर्वेक्षणबाट संकलन गरिएका तथ्याङ्क प्रशोधन गरेर सूचनाहरू प्राप्त गरिन्छ र तिनै सूचनाको आधारमा समग्र जनसंख्याको विशेषता बताउन सकिन्छ । यसरी प्राप्त हुने सूचना गणनाबाट प्राप्त भएका हुन् भने सोभै अध्ययन दायरा वा क्षेत्र भित्रको समग्र जनसंख्याको विशेषताका बारेमा बताउन सक्ने हुन्छन् । तर, नमुना सर्वेक्षणबाट प्राप्त भएका सूचनाले समग्र जनसंख्याको प्रतिनिधित्व मात्र गरेका हुन्छन् वा जनसंख्याका विशेषताको बारेमा अनुमान लगाउन सक्दछन् । तसर्थ, नमुना छनोट विधिबाट संकलन गरिएका तथ्याङ्कीय चरहरूको सहयोगले समग्र अध्ययन गरिने क्षेत्र (Population) को विवरण वा विशेषता प्राप्त गर्ने विधिलाई अनुमानको सिद्धान्त भनिन्छ । उदाहरणका लागि नेपाल श्रमशक्ति सर्वेक्षण २०१७/१८ बाट प्राप्त सूचनाको आधारमा नेपालमा बेरोजगार दर ११.४% प्रतिशत छ भन्न सकिन्छ ।¹⁴ यस उदाहरणमा समग्र नेपालको बेरोजगार दर नेपाल श्रमशक्ति नमुना सर्वेक्षणको नतिजाबाट प्राप्त सूचनाबाट आँकलन गरिएको हो ।

अनुमानसम्बन्धी अध्ययन गर्दा केही शब्दावलीहरूको परिभाषा बुझ्नुपर्ने हुन्छ । जस्तै: जनसंख्या (Population), नमुना (Sample), जनसंख्या मान (Population Value), स्टाटिस्टिक (Statistic), आँकलन (Estimation), आँकलक (Estimator) आदि । यी शब्दावलीहरूको संक्षिप्त जानकारी अनुसूचीमा दिइएको छ । अनुमानसम्बन्धी अध्ययन स्वभावतः निकै प्राविधिक विषय हो । तसर्थ, यस पाठमा हामीले तथ्याङ्कीय साक्षरताका लागि अनुमानको परिचयात्मक चर्चा मात्र गर्नेछौं ।

१६.२ अनुमानका गुणहरू

जनसंख्याका विशेषता अनुमान गर्नुपर्ने भएकोले सकेसम्म अनुमान वा आँकलकले दिने मान जनसंख्याको मानसँग नजिक पुग्ने किसिमको हुनुपर्दछ । यदि आँकलकले निम्नानुसारका गुण धारण गर्दछन् भने त्यस्ता आँकलकलाई असल मान्न सकिन्छ ।

14 Central Bureau of Statistics. (2019). *Report on the Nepal Labour Force Survey 2017/18*. Kathmandu, Nepal; Central Bureau of Statistics. Retrieved on December 6, 2021 from: <https://cbs.gov.np/wp-content/uploads/2019/05/Nepal-Labour-Force-Survey-2017-18-Report.pdf>

- क) निष्पक्षता (Unbiasedness)
 ख) सामञ्जस्यता (Consistency)
 ग) दक्षता (Efficiency)
 घ) पर्याप्तता (Sufficiency)

क) निष्पक्षता: यदि आँकलकको अपेक्षित मान जनसंख्याको मानसँग बराबर हुन्छ भने त्यस्तो आँकलकलाई निष्पक्ष आँकलक भनिन्छ । नमुना मध्यक, नमुना मध्यिका निष्पक्ष आँकलक हुन् । नमुना भेरियन्स (Sample Variance) भने निष्पक्ष हुँदैन ।

ख) सामञ्जस्यता: यदि नमुना छनोटको आकार बढाउँदा आँकलकको मान जनसंख्याको मानको नजिक हुने सम्भावना बढ्दछ भने त्यस्तो गुणलाई सामञ्जस्यता भनिन्छ । नमुना मध्यक, नमुना मध्यिका सामञ्जस्यता भएका आँकलक हुन् ।

ग) दक्षता: जनसंख्याको मान अनुमान गर्ने दुई वा दुईभन्दा बढी निष्पक्ष आँकलकहरूमध्ये सबैभन्दा कम विचरण वा प्रसरण (Variance) भएको आँकलकलाई दक्ष आँकलक भनिन्छ । नमुना मध्यक र नमुना मध्यिकामध्ये नमुना मध्यक बढी दक्ष आँकलक हो ।

घ) पर्याप्तता: छानिएको नमुनाले दिनसक्ने सबै सूचनाहरू उपयोग गर्ने वा सबै सूचनाहरू धारण गर्ने आँकलकलाई पर्याप्त आँकलक भनिन्छ । नमुना मध्यक पर्याप्त आँकलक हो । पर्याप्तता भएका आँकलक निष्पक्ष नहुन पनि सक्दछन् ।

१६.३ स्याम्प्लिङ डिस्ट्रिब्युसन अफ स्टाटिस्टिक्स

कुनै एउटा जनसंख्याबाट दुई वा दुईभन्दा बढी नमुनाहरू छनोट गर्न सकिन्छ । त्यसरी छानिएका नमुनाहरूका स्टाटिस्टिकहरूको वितरणलाई नै स्याम्प्लिङ डिस्ट्रिब्युसन अफ स्टाटिस्टिक्स भनिन्छ । उदाहरणका लागि कुनै जनसंख्याबाट केही संख्याका अलग अलग नमुना s_1, s_2, s_3, \dots । छानेर ती प्रत्येक नमुनाबाट नमुना मध्यक $\bar{X}_1, \bar{X}_2, \bar{X}_3, \dots$ निकाल्न सकिन्छ । ती नमुना मध्यकहरूको वितरण $\bar{X}_1, \bar{X}_2, \bar{X}_3, \dots$ लाई स्याम्प्लिङ डिस्ट्रिब्युसन अफ स्टाटिस्टिक्स भनिन्छ । यस किसिमको वितरणको प्रयोग परिकल्पना परीक्षण गर्दा, अनुमान गर्दा, तथा अन्य तथ्याङ्कीय अध्ययनमा प्रयोग गरिन्छ ।

१६.४ सेन्ट्रल लिमिट सिद्धान्त र यसको प्रयोग

अध्ययनको लागि लिइने चरले नर्मल डिस्ट्रिब्युसन पछ्याउने बारेको सैद्धान्तिक ज्ञान हामीले यसभन्दा अगाडिको पाठमा हासिल गरिसकेका छौं । यस्तो नर्मल डिस्ट्रिब्युसन पछ्याउने जनसंख्याबाट लिइने नमुनाले सँधै नर्मल डिस्ट्रिब्युसन नै पछ्याउने गर्दछन् । तर यदि अन्य कुनै डिस्ट्रिब्युसन पछ्याउने जनसंख्याबाट पर्याप्त मात्रामा ठुलो आकारको

नमुना छनोट गरियो भने, छानिएको नमुनाले नर्मल डिस्ट्रिब्युसन पछ्याउँदछ भन्ने सिद्धान्तलाई सेन्ट्रल लिमिट सिद्धान्त (Central Limit Theorem) भनिन्छ । यही सिद्धान्तको प्रयोग गरेर जनसंख्याको डिस्ट्रिब्युसन अज्ञात वा अन्य कुनै भए तापनि ठुलो नमुना छ भने नमुनालाई नर्मल डिस्ट्रिब्युसन पछ्याएको मानेर अनुमान, परिकल्पना परीक्षण, प्रतिगमन विश्लेषणजस्ता तथ्याङ्कीय अध्ययन कार्य गर्न सकिन्छ ।

उदाहरणका लागि नेपाल जीवनस्तर सर्वेक्षणमा छानिएका परिवारले गरेको वार्षिक खर्च, वार्षिक आमदानी आदिको मध्यकलाई पूरै जनसंख्याको परिवारले गरेको वार्षिक खर्च, वार्षिक आमदानी आदिको मध्यकको आँकलनको रूपमा प्रयोग गर्दछौं । त्यसैगरी जनसंख्या मानका बारेमा परिकल्पना परीक्षण गर्दा पनि छानिएको नमुनालाई नर्मल डिस्ट्रिब्युसन पछ्याएको मानिन्छ । यस्तो अनुमान गर्ने कार्य वा परिकल्पना परीक्षण गर्ने कार्यका लागि वास्तविक जनसंख्याले पछ्याउने डिस्ट्रिब्युसन थाहा नभए तापनि नमुनाले पछ्याउने डिस्ट्रिब्युसनलाई नर्मल डिस्ट्रिब्युसन मानेका हुन्छौं । यस सर्वेक्षणमा छानिने नमुना पर्याप्त मात्रामा ठुलो आकारको हुन्छ । त्यसैगरी अन्य सर्वेक्षण जस्तै, नेपाल श्रमशक्ति सर्वेक्षण, नेपाल बहुसूचक सर्वेक्षण, नेपाल जनसांख्यिक तथा स्वास्थ्य सर्वेक्षणका नमुनाहरूले नर्मल डिस्ट्रिब्युसन पछ्याएको मान्न सकिन्छ ।

समग्रमा भन्नुपर्दा, सेन्ट्रल लिमिट सिद्धान्तले देहायका प्रमुख दुई कुरा बताउँदछ:

- १) पर्याप्त मात्रामा ठुलो आकारका नमुनाहरूबाट प्राप्त हुने नमुना मध्यकको स्याम्प्लड डिस्ट्रिब्युसनले नर्मल डिस्ट्रिब्युसन पछ्याउँदछ र
- २) यस्तो स्याम्प्लड डिस्ट्रिब्युसनको मध्यक र स्तरीय भिन्नता (Mean and Standard Deviation of Sampling Distribution) क्रमशः $(\bar{X}) = \mu$ र $s = \frac{\sigma}{\sqrt{n}}$ हुन्छ । यहाँ, μ भनेको जनसंख्या मध्यक, σ भनेको जनसंख्या स्तरीय भिन्नता र n भनेको नमुना आकार हो ।

१६.४.१ सेन्ट्रल लिमिट सिद्धान्तको महत्त्व तथा प्रयोग

नमुना मध्यकको प्रयोगबाट जनसंख्या मध्यकको बारेमा निष्कर्ष निकाल्न तथा नमुना स्तरीय भिन्नताको प्रयोगबाट जनसंख्या स्तरीय भिन्नताको बारेमा निष्कर्ष निकाल्नका लागि सेन्ट्रल लिमिट सिद्धान्तको प्रमुख महत्त्व रहेको छ । तथ्याङ्क विश्लेषण तथा व्याख्याका कार्यहरू जस्तै: अनुमान गर्न, परिकल्पना परीक्षण गर्न, प्रतिगमन विश्लेषण गर्न यस सिद्धान्तको प्रयोग हुन्छ ।

१६.४.२ सेन्ट्रल लिमिट सिद्धान्तले प्रयोग गर्नका लागि ध्यान दिनुपर्ने कुराहरू

- १) समान सम्भाव्यता छनोटमा आधारित पर्याप्त मात्रामा ठुलो आकारको नमुना हुनुपर्दछ ।
- २) नमुनामा छानिने एकाइहरू स्वतन्त्र हुनुपर्दछ । अर्थात्, एउटा एकाइ छानिएको घटनाले अर्को एकाइ छानिने सम्भाव्यतामा कुनै असर गर्नु हुँदैन ।

१६.५ अनुमानका प्रकारहरू

अनुमान दुई प्रकारका हुन्छन् । ती बिन्दुगत अनुमान (Point Estimation) र अन्तराल अनुमान (Interval Estimation) हुन् ।

क) बिन्दुगत अनुमान: यदि जनसंख्याको मानलाई नमुनाको एकल मानले अनुमान गरिन्छ भने त्यस्तो अनुमान गर्ने प्रविधिलाई बिन्दुगत अनुमान भनिन्छ । उदाहरणका लागि नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ अनुसार नेपालको गरिबीको रेखामुनिको जनसंख्या २५.१६% अनुमान गरिएको थियो ।

ख) अन्तराल अनुमान: यदि जनसंख्याको मानलाई केही अन्तरालमा गणना गरिन्छ भने त्यस्तो अनुमान गर्ने प्रविधिलाई अन्तराल अनुमान भनिन्छ । उदाहरणका लागि यदि नेपालको गरिबीको रेखामुनिको जनसंख्यालाई २५.१६% नभनी २३-२६ प्रतिशतसम्म छ, भनेको भए अन्तराल अनुमान हुन्थ्यो ।

१६.६ विश्वस्तता अन्तराल

अन्तराल अनुमान गर्दा सामान्यतया सम्भाव्यताको भाषामा बोल्ने गरिन्छ । अर्को शब्दमा भन्नुपर्दा, छानिएको नमुनाबाट प्राप्त सूचनाका आधारमा अन्तराल आँकलक निर्माण गर्दा जनसंख्याको वास्तविक मान त्यसभित्र पर्ने निश्चित सम्भावना भएको एउटा अन्तराल तयार गरिन्छ । यस्तो अन्तरालमा दुई निश्चित संख्याको सीमा निर्धारण गरिएको हुन्छ । उक्त अन्तरालभित्र जनसंख्याको मान पर्ने सम्भाव्यतालाई विश्वस्तताको स्तर (Confidence Level) भनिन्छ र यसलाई $(1-\alpha)$ ले जनाइन्छ । जहाँ, α ले भ्रमांशको सम्भाव्यता प्रतिनिधित्व गरेको हुन्छ र यसलाई Level of Significance भनिन्छ । उदाहरणका लागि नेपालको गरिबीको रेखामुनि रहेका जनसंख्या २३ देखि २६ प्रतिशत हुनेमा ९५ प्रतिशत ढुक्क छु भनियो भने ९५ प्रतिशत विश्वस्तताको स्तर भयो भने बाँकी ५ प्रतिशत भनेको सो भनाइमा सत्यता नहुनसक्ने सम्भावना भयो । यस आधारमा विश्वस्तताको अन्तराल निर्धारण गर्न सकिन्छ । जनसंख्याको मध्यमानलाई नमुनाको मध्यमानले अन्तराल अनुमान गर्दा तय गरिने विश्वस्तताको अन्तराल निम्न सूत्रबाट व्यक्त गर्न सकिन्छ ।

$$P \left\{ \left(\bar{x} - z_{\alpha/2} \frac{s}{\sqrt{n}} \right) < \mu < \left(\bar{x} + z_{\alpha/2} \frac{s}{\sqrt{n}} \right) \right\} = 1 - \alpha \quad (9)$$

जहाँ, μ र \bar{x} भनेका क्रमशः जनसंख्याको मध्यक र नमुनाको मध्यक हो, s भनेको नमुनाको स्तरीय भिन्नता हो, n भनेको नमुनाको आकार हो र $z_{\alpha/2}$ भनेको विश्वसनीयता गुणक (Reliability Factor) हो, $1 - \alpha$ भनेका विश्वस्तताको स्तर हो, P ले सम्भाव्यता जनाएको छ। यहाँ, \bar{x} , s र n को मान नमुनाबाट प्राप्त हुन्छ भने $z_{\alpha/2}$ को मान जेड तालिका (अनुसूची २) बाट प्राप्त हुन्छ। केही बारम्बार प्रयोगमा आउने विश्वसनीयता गुणक यसप्रकार छन्:

क) ९०% विश्वस्तता अन्तराल: $z_{0.05} = 1.65$

ख) ९५% विश्वस्तता अन्तराल: $z_{0.025} = 1.96$

ग) ९९% विश्वस्तता अन्तराल: $z_{0.005} = 2.58$

विश्वस्तता अन्तराल व्याख्या गर्ने तरिकाको एउटा उदाहरण यसप्रकार छ। जनसंख्या मानको ९० प्रतिशत विश्वस्तता अन्तराल भनेको यदि धेरैपटक उस्तै विधिसे नमुनाहरू छानियो र प्रत्येक नमुना मानका आधारमा छुट्टाछुट्टै जनसंख्या मानका लागि अन्तराल निर्माण गरियो भने ९० प्रतिशत यस्ता विश्वस्तताका अन्तरालभित्र जनसंख्याको वास्तविक मान पर्दछ। अर्थात्, ९० प्रतिशत विश्वस्तताको अन्तरालले जनसंख्या मान समेटेका हुन्छन् भन्ने कुरामा ९० प्रतिशत दुक्क हुन सकिन्छ।

१६.७ नमुना छनोटसम्बन्धी अवधारणा

यस पाठमा हामीले नमुना मानहरूका बारेमा चर्चा गर्थौं। यसबाट अनुमानसम्बन्धी कार्यमा नमुना छनोट एक अपरिहार्य काम हो भन्नेमा दुई मत रहेन। यसका अतिरिक्त अन्य तथ्याङ्कीय क्रियाकलाप जस्तै: प्रतिगमन विश्लेषण, परिकल्पना परीक्षण आदिमा पनि नमुना छनोट कार्य अनिवार्य हुन्छ। तसर्थ, यस खण्डमा हामीले नमुना छनोटसम्बन्धी संक्षिप्त परिचय लिनेछौं। सबैभन्दा पहिला नमुना छनोट अध्ययनका महत्त्व बारेमा केही प्रमुख बुँदाहरू हेरौं।

१६.७.१ नमुना छनोट अध्ययनका फाइदा तथा बेफाइदा

फाइदा

क) खर्च तथा समयको किफायत हुन्छ।

ख) जनशक्ति कम लाग्दछ तर दक्षता आवश्यक हुन्छ।

ग) विश्वसनीय नतिजा प्राप्त हुन्छ।

घ) कतिपय अध्ययन नमुना छनोट गर्नुको विकल्प नै हुँदैन। जस्तै: हाम्रो शरीरको रगतमा ग्लुकोजको मात्रा मापन गर्नुपर्ने भन्ने रगतको नमुना लिएर अध्ययन

गर्नुको विकल्प नै हुँदैन ।

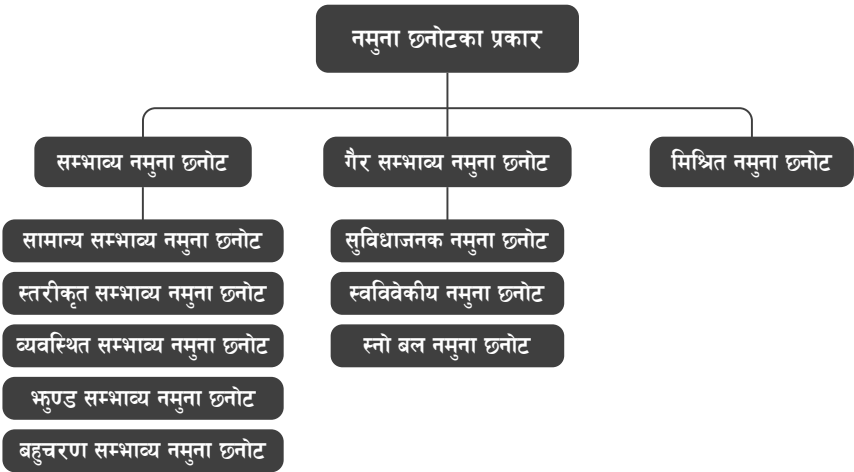
- ड) अत्यन्त ठुलो अर्थात् अनन्त जनसंख्यामा नमुना छनोटको विकल्प हुँदैन । जस्तै: फेवातालको माछाको औसत तौल पत्ता लगाउनु परेमा नमुना छनोटकै आधार लिनुपर्दछ ।
- च) विनाशकारी अध्ययनमा नमुना छनोट अनिवार्य हुन्छ । जस्तै: बन्दुकको गोली बनाउने कारखानाले गोली पड्किन्छ वा पड्किन्न भनेर सबै गोलीको जाँच गर्न मिल्दैन ।

बेफाइदा

- क) तथ्याङ्कको गहन ज्ञान भएका जनशक्ति आवश्यक पर्दछ ।
- ख) नमुनाका मानहरूले जनसंख्याका सही मान नदिन पनि सक्दछन् ।
- ग) नमुना छनोटसम्बन्धी भ्रमांश समावेश हुन सक्दछ ।

१६.७.२ नमुना छनोटका प्रकार

चित्र नं. १६.१: नमुना छनोटका प्रकारहरू



क) सामान्य सम्भाव्य नमुना छनोट (Simple Random Sampling): जनसंख्याका सम्पूर्ण एकाइहरूको नमुनामा पर्ने समान सम्भावना रहने गरी छनोट गरिने विधिलाई सामान्य सम्भाव्य नमुना छनोट भनिन्छ । यसका लागि जनसंख्याको सम्पूर्ण एकाइहरूको सूची आवश्यक पर्दछ । प्रत्येक एकाइहरूलाई अलग अलग क्रम संख्या प्रदान गरी चिठ्ठा वा दैवी छनोट तालिकाको प्रयोगबाट यसप्रकारको नमुना छनोट गरिन्छ । आजकल कम्प्युटरका सफ्टवेयरहरूले यस किसिमका नमुना छनोट गरिदिन्छन् । अध्येता वा

अनुसन्धानकर्तासँग जनसंख्याको अध्यावधिक सूची छ वा प्राप्त गर्न सकिन्छ भने यो नमुना छनोट विधि अवलम्बन गर्न सकिन्छ ।

ख) स्तरीकृत सम्भाव्य नमुना छनोट (Stratified Random Sampling): यस किसिमको नमुना छनोट विधिमा समग्र जनसंख्यालाई समान विशेषता सहितका विभिन्न संख्यामा समूहमा विभाजन गरिन्छ । यस्ता समूहलाई स्ट्राटा (Strata) भनिन्छ । त्यसपछि प्रत्येक समूहबाट चाहिएको संख्यामा सामान्य सम्भाव्य नमुना छनोट विधि प्रयोग गरेर प्रतिनिधिमूलक एकाइहरू छनोट गरिन्छ । यस विधिमा पनि जनसंख्याका सबै एकाइहरूको सूची आवश्यक पर्दछ । उदाहरणका लागि नेपालका जिल्लाहरूलाई हिमाल, पहाड र तराई गरेर तिनवटा समूहमा विभाजन गरौं । ती तिनवटा समूहलाई पनि ग्रामीण र सहरी गरेर दुई दुई भागमा विभाजन गरी जम्मा छ वटा समूहमा अर्थात् स्ट्राटामा बनाऔं । यी ६ वटा समूहलाई पनि पुनः विभिन्न विशेषताका आधारमा अझै धेरै स्ट्राटामा विभाजन गर्न पनि सकिन्छ । यस किसिमले स्ट्राटा बनाइसकेपछि प्रत्येक स्ट्राटाबाट तोकिएको संख्याका नमुनाहरू सम्भाव्य नमुना विधिबाट छनोट गरीयो भने यो नमुना छनोटलाई स्तरीकृत सम्भाव्य नमुना छनोट भनिन्छ । नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ मा यही विधिको प्रयोग गरिएको थियो ।

ग) व्यवस्थित सम्भाव्य नमुना छनोट (Systematic Random Sampling): यस किसिमको नमुना छनोट विधिमा जनसंख्याका एकाइहरूको विशेष प्रकारको सूची बनाइन्छ । सूची बनाउँदा सामान्यतया कुनै विशेषताका आधारमा क्रमबद्ध गरिन्छ । त्यसपछि जनसंख्याको आकारलाई नमुनाको आकारले भाग गरेर नमुना अन्तराल (Sample Interval) पत्ता लगाइन्छ । त्यसपछि समग्र सूचीको पहिलो नमुना अन्तरालभित्र पर्ने गरी पहिलो एकाइ सम्भाव्य छनोट विधिबाट छानेर बाँकी सबै एकाइहरू नमुना अन्तरालपछि क्रमशः छनोट गरिन्छ । यसमा पहिलो छानिएको एकाइपछि नमुना अन्तराल छाडेर अर्को एकाइ छनोट गरिन्छ । फेरि दोस्रो छानिएको एकाइपछि नमुना अन्तराल छाडेर अर्को एकाइ छनोट गरिन्छ । एवं रीतले आवश्यक संख्याका एकाइहरू छानेर नमुना तयार गरिन्छ । मानौं, कुनै जनसंख्याको आकार १०० छ र २० वटा एकाइको नमुना छनोट गर्नुछ भने, $100/20 = 5$ नमुना अन्तराल हुन्छ । त्यसपछि जनसंख्याको सूचीबाट पहिलो पाँचवटा एकाइबाट सम्भाव्य नमुना छनोटबमोजिम एउटा एकाइ छनोट गर्नुपर्दछ र त्यसपछि त्यसमा क्रमशः ५ जोड्दै बाँकी १९ वटा एकाइहरू छनोट गर्नुपर्दछ । यदि पहिलो छानिएको एकाइ तेस्रो नम्बरको पच्यो भने, दोस्रो एकाइ आठौं नम्बरको, तेस्रो एकाइ १३ औं नम्बरको हुँदै २० वटा एकाइ छान्नुपर्दछ । नेपाल जीवनस्तर सर्वेक्षण, नेपाल श्रमशक्ति सर्वेक्षणमा यस विधिको पनि प्रयोग गरिएको छ ।

घ) भ्रुण्ड सम्भाव्य नमुना छनोट (Cluster Random Sampling): यस विधिमा समग्र जनसंख्यालाई सकेसम्म असमान विशेषता सहितका समूहमा विभाजन गर्नुपर्दछ । यस्ता समूहलाई भ्रुण्ड (Cluster) भनिन्छ । त्यसपछि सम्भाव्य नमुना छनोटबाट केही संख्याका भ्रुण्डहरू छानेर भ्रुण्डका सबै एकाइहरूको विवरण संकलन गर्नुपर्दछ । जनसंख्याका एकाइहरूको सूची प्राप्त नभएको र असमान विशेषता भएका एकाइको जनसंख्याका लागि यो विधि उपयुक्त मानिन्छ । उदाहरणका लागि नेपालका जङ्गलहरूमा पाइने प्रमुख रूखहरूका बारेमा अध्ययन गर्न नमुना छनोट गर्नका लागि विभिन्न स्थानका जङ्गलहरूलाई अलग अलग समूह वा क्लस्टरमा विभाजन गरेर ती समूहमध्य केही समूहलाई सम्भाव्य नमुना छनोट विधिबाट छानेर छानिएका जङ्गलका सबै रूखहरूको अध्ययन गर्ने विधि भ्रुण्ड सम्भाव्य नमुना छनोट विधि अन्तर्गत पर्दछ ।

ङ) बहुचरण सम्भाव्य नमुना छनोट (Multistage Random Sampling): यस विधिमा नमुना एकाइमा पुनका लागि विभिन्न चरणमा नमुना छनोट गर्ने गरिन्छ । तसर्थ, समग्र जनसंख्याको विभिन्न तहमा नमुना छनोट गर्दै जाने भएकोले यसलाई बहुचरण सम्भाव्य विधि भनिएको हो । उदाहरणका लागि कुनै जिल्लाको परिवारमा विवरण संकलन गर्नु छ भने पहिलो चरणमा प्रतिनिधिमूलक स्थानीय तहको छनोट सम्भाव्य नमुना छनोट विधि प्रयोग गरेर गर्नुपर्दछ । यो पहिलो चरणो छनोट भयो । त्यसपछि छानिएका स्थानीय तहबाट केही संख्याका प्रतिनिधिमूलक वडा सम्भाव्य नमुना छनोट विधि प्रयोग गरेर छानिन्छ । यो दोस्रो चरण भयो । छानिएका वडाबाट फेरि सम्भाव्यक नमुना छनोट विधि प्रयोग गरेर प्रतिनिधिमूलक परिवार छानिन्छ । यो तेस्रो चरण भयो । यसरी बहुचरणमा सम्भाव्य नमुना छनोट गरी प्रतिनिधिमूलक नमुना छान्न सकिन्छ ।

च) आकारसँग सम्भाव्य समानुपातिक हुनेगरी नमुना छनोट (Probability Proportional to Size Sampling): खासगरी भ्रुण्ड सम्भाव्य नमुना छनोट गर्दा यदि भ्रुण्डहरूको आकार समान छैनन् भने यो नमुना छनोट विधि अबलम्बन गरिन्छ । यस्तो अवस्थामा भ्रुण्डको आकार अनुसार नमुनामा छनोट हुने वा नहुने सम्भाव्यता निर्धारण हुन्छ । यसका लागि प्रत्येक भ्रुण्डका एकाइहरूको संख्या सूचीकृत गर्नुपर्ने हुन्छ । त्यसपछि एकीकृत गरिएको समग्रबाट निश्चित संख्याका एकाइहरू सम्भाव्यताको आधारमा छनोट गरिन्छ । यसरी छनोट भएका एकाइहरूले कुन भ्रुण्ड छान्ने र छानिएका भ्रुण्डबाट कति एकाइ नमूनमा समावेश गर्ने भन्ने विषय निर्देशित गर्दछन् । यस्तो गर्दा ठुलो आकारका भ्रुण्ड छानिने सम्भावना धेरै हुन्छ भने सानो आकारका भ्रुण्ड छानिने सम्भावना थोरै हुन्छ । यसका अतिरिक्त समूहीकृत गरिएका जनसंख्याबाट नमुना छनोट गर्दा समूहको आकारबमोजिम उक्त समूहबाट नमुनामा समावेश हुने एकाइहरूको सम्भाव्यता

समानुपातिक हुनेगरी छनोट गरिने नमुनाहरूलाई आकारसँग सम्भाव्य समानुपातिक हुनेगरी नमुना छनोट भनिन्छ । उदाहरणका लागि कुनै गाउँपालिकाले आफ्नो कार्य क्षेत्रभित्र पिउने पानीको अवस्था अध्ययन गर्ने पर्यो भने सबै वडाबाट बराबर संख्याको नमुना एकाइ लिनुभन्दा बढी जनसंख्या भएको वडाबाट धेरै र कम जनसंख्या भएको वडाबाट थोरै एकाइ छनोट गर्नु उत्तम हुन्छ । यसका लागि यो विधिको प्रयोग गर्न सकिन्छ । स्मरण रहोस्, यस विधिमा समग्र जनसंख्या भित्रको समूहहरूको आकारसँग सम्भाव्य समानुपातिक हुनेगरी सामान्य सम्भाव्य अथवा व्यवस्थित सम्भाव्य विधिको प्रयोग गरी नमुना छनोट गरिएको हुन्छ ।

छ) सुविधाजनक नमुना छनोट (Convenient Sampling): यस विधिमा अध्ययनकर्ताको सुविधाका लागि अध्येता स्वयंले छानिनु पर्ने एकाइहरूको निक्यौल गर्दछन् । जहाँ जुन एकाइलाई नमुनामा छनोट गर्न उपयुक्त हुन्छ त्यसैलाई नमुनाको रूपमा छनोट गर्ने विधिलाई सुविधाजनक नमुना छनोट भनिन्छ । यो विधिमा छानिने एकाइहरूको कुनै सम्भाव्यता निहित रहेको हुँदैन । साधारणतया सीमित उद्देश्यका अध्ययनहरूमा मात्र यो विधिको प्रयोग गरिन्छ । यो विधिबाट तथ्याङ्कीय महत्त्वका अनुमानहरू गर्न सकिँदैन ।

ज) स्वविवेकीय नमुना छनोट (Judgement Sampling): यस किसिमको नमुना छनोट विधिमा सम्भाव्यताको प्रयोग नगरी अध्येताले स्वविवेकको प्रयोग गरेर जनसंख्याका एकाइहरूमध्येबाट प्रतिनिधित्व गर्न सक्ने एकाइहरू छनोट गर्दछन् ।

झ) स्नो बल नमुना छनोट (Snow Ball Sampling): समग्र जनसंख्याको यथार्थ पहिचान गर्न गाह्रो हुने प्रकृतिका अध्ययनहरूमा यस विधिको प्रयोग गरिन्छ । यसमा कुनै माध्यमबाट एउटा एकाइको छनोट गरिन्छ र सो एकाइबाट गोप्य तरिकाले अर्को एकाइको पहिचान पत्ता लगाइन्छ । उदाहरणका लागि लागु औषध प्रयोगकर्ताको सर्वेक्षण गर्नुपर्ने भने यो विधिको प्रयोग गर्न सकिन्छ ।

ञ) मिश्रित नमुना छनोट (Mixed Sampling): यस विधिमा सम्भाव्य र गैर सम्भाव्य दुवै विधिको मिश्रण गरेर नमुना छनोट गरिन्छ । उदाहरणका लागि नेपालका स्थानीय तहको बारेमा अध्ययन गर्नका लागि केही जिल्लाहरू सम्भाव्य नमुना छनोट विधिबाट छानेर छानिएका जिल्लाका स्थानीय तह स्वविवेकीय नमुना छनोट विधिबाट छानियो भने मिश्रित विधि हुन्छ ।

माथि उल्लेख गरिएका बाहेक नमुना छनोट गर्ने अन्य विधिहरू पनि छन् । यस पुस्तकमा नमुना छनोटको सामान्य परिचय मात्र उल्लेख गरिएको हुनाले सबै विधिहरूका बारेमा विस्तृत वर्णन गरिएको छैन ।

१६.७.३ नमुना सर्वेक्षण सञ्चालनका प्रमुख चरण

नमुना सर्वेक्षणबाट अध्ययन गर्नुपूर्व यसका प्रमुख चरण बारेमा पूर्ण रूपमा जानकारी राख्नुपर्दछ। यसका लागि प्रमुख रूपमा निम्न बमोजिमका चरण अपनाउनु उपयुक्त हुन्छ।

क) सर्वेक्षणको उद्देश्य: सर्वेक्षणको पहिलो चरण भनेको यसको उद्देश्य तय गर्नु हो। अध्येता अथवा अनुसन्धानकर्ताले सर्वेक्षण गर्नुको मुख्य उद्देश्यका बारेमा स्पष्ट हुनुपर्दछ। यसबाट सर्वेक्षणका सम्पूर्ण विधिहरू निर्धारित हुन्छन्। नमुनाको आकार निर्धारण गर्दैरिख सर्वेक्षणका समग्र योजना बनाउन यसको उद्देश्य स्पष्ट हुनुपर्दछ। उदाहरणका लागि सर्वेक्षणबाट कुनै चरको मध्यक अनुमान गर्न र कुनै विशेषताको समानुपात अनुमान गर्न फरक-फरक आकारका नमुनाहरू आवश्यक पर्दछन्।

ख) अध्ययन गरिने जनसंख्याको निर्धारण: जुन समग्र समूहबाट प्रतिनिधिमुलक नमुना लिएर अध्ययन गरिन्छ, त्यो जनसंख्या राम्रोसँग निर्धारण गरी स्पष्टसँग परिभाषित भएको हुनुपर्दछ। जुनसुकै सर्वेक्षणहरू पनि जनसंख्या मानको बारेमा सूचना प्राप्त गर्नका लागि सञ्चालन गरिएका हुन्छन्। तसर्थ, नमुना सर्वेक्षण गर्नुपूर्व कुन जनसंख्याको बारेमा अध्ययन गरिँदैछ, सो को बारेमा स्पष्ट हुनुपर्दछ। उदाहरणका लागि कुनै गाउँपालिकामा दलित समुदायको साक्षरता अवस्था अध्ययन गर्नका लागि लिइने जनसंख्या भनेको सो गाउँपालिकाको सबै दलित परिवारको समूह हो। गाउँपालिकाका सबै दलित परिवारलाई जनसंख्या मानेर उक्त जनसंख्याबाट प्रतिनिधिमुलक उपसमूह छनोट गरी समग्र दलित समुदायको साक्षरता सम्बन्धी विभिन्न सूचकहरूको अनुमान गर्न सकिन्छ।

ग) संकलन गरिने तथ्याङ्क: अध्ययन गरिने जनसंख्या र सोबाट छनोट भएको उपसमूहको यकिन भइसकेपछि छनोटमा परेका तथ्याङ्कीय एकाइहरूबाट के कस्ता विवरण संकलन गर्ने भन्ने विषयमा निर्णय लिनुपर्दछ। यसका लागि सर्वेक्षणको उद्देश्य मार्गदर्शन हुन्छ। के कस्ता सूचकहरू अनुमान गर्ने हो, सो बमोजिमका तथ्याङ्क संकलन हुने कुरामा विश्वस्त हुनुपर्दछ। सर्वेक्षणको उद्देश्यभन्दा बाहिर गएर तथ्याङ्क संकलन गर्नुहुँदैन।

घ) आवश्यक शुद्धताको तह: नमुना सर्वेक्षणबाट उपलब्ध हुने नमुना मानहरूले अध्ययन गरिएको जनसंख्याको विशेषता जनाउने सूचक अनुमान गर्दछन्। यसरी अनुमान गर्दा कुनै न कुनै भ्रमांशहरू समावेश हुने हुनाले शतप्रतिशत शुद्ध आकलक प्राप्त गर्न सकिँदैन। तसर्थ, सर्वेक्षण गर्नुपूर्व नै कतिम्मको तहमा शुद्ध आँकलक अपेक्षा गरिएको हो, सो बारेमा स्पष्ट हुनुपर्दछ।

ङ) तथ्याङ्क संकलक र प्रश्नावली: सर्वेक्षणको सबैभन्दा महत्त्वपूर्ण कार्य भनेको तथ्याङ्क

संकलन हो । सर्वेक्षणको उद्देश्यअनुरूप सही तथ्याङ्क संकलन गर्नका लागि गुणस्तरीय प्रश्नावली र तालिम प्राप्त तथ्याङ्क संकलक हुनु अनिवार्य हुन्छ ।

च) नमुना छनोट विधि: सर्वेक्षणको उद्देश्य र जनसंख्याको प्रकृतिबमोजिम उपयुक्त नमुना छनोट विधि अबलम्बन गर्नुपर्दछ । जनसंख्याका एकाइहरूको विशेषता लगभग समान प्रकृतिको भएमा सामान्य सम्भाव्य विधि प्रयोग गर्न सकिन्छ भने त्यस्ता एकाइहरूको विशेषतामा विषमता रहेको छ भने स्तरीकृत, भुण्ड, व्यवस्थित वा बहुचरण नमुना विधि अनुसरण गर्न सकिन्छ । साधन र स्रोतको अवस्था हेरेर गैर-सम्भाव्य नमुना छनोट विधिको पनि अवलम्बन गर्न सकिन्छ । तर गैर-सम्भाव्य नमुना छनोट विधिको प्रयोग गर्दा प्राप्त हुने नमुना मानले जनसंख्या मानलाई सही प्रतिनिधित्व नगर्न पनि सक्दछ । यद्यपि अनुसन्धानकर्ताको बुद्धिमता क्षमता प्रयोग गरेर जनसंख्याका समग्र एकाइहरूको प्रतिनिधित्व हुनेगरी नमुना एकाइ छनोट गर्न सकियो भने जनसंख्या मानको नजिक नजिक पुग्न पनि सकिन्छ ।

छ) नमुना एकाइ: नमुना छनोट विधि बमोजिम नमुना एकाइहरू छानिन्छन् र छानिएका एकाइहरूबाट आवश्यक विवरण संकलन गरिन्छ । नमुना छनोट गर्दा जनसंख्याका सबै एकाइहरूको छानिने सम्भावना समान भयो भने समान सम्भाव्यताको नमुना छनोट भनिन्छ र यदि एकाइहरूको छानिने सम्भावना फरक-फरक भयो भने आकारसँग सम्भाव्यता समानुपातिक नमुना छनोट भनिन्छ ।

ज) पूर्व परीक्षण: सर्वेक्षण सञ्चालनका लागि तय गरिएको समग्र योजना व्यवस्थापन तथा तथ्याङ्क संकलनका लागि तयार गरिएका प्रश्नावली तथा विधिहरूको उपयुक्तता परीक्षणका लागि पूर्व परीक्षण गर्ने गरिन्छ । लागत तथा गुणस्तर सुधार गर्न पूर्व परीक्षण निकै प्रभावकारी मानिन्छ ।

झ) फिल्ड कार्यको संगठन: तथ्याङ्क संकलन र सो कार्यको अनुगमन तथा मूल्याङ्कन गर्नका लागि गरिएको समग्र व्यवस्थापनलाई फिल्ड कार्यको संगठन भनिन्छ । यसबाट गुणस्तरीय तथ्याङ्क संकलन गर्न ठुलो मद्दत मिल्दछ ।

ञ) तथ्याङ्क प्रशोधन तथा विश्लेषण: तथ्याङ्क संकलनको फिल्ड कार्य समाप्त भइसकेपछि संकलित तथ्याङ्कको सम्पादन तथा प्रशोधनको कार्य गर्नुपर्दछ । यस अन्तर्गत संकलित तथ्याङ्क अस्वाभाविक किसिमको भएमा सम्बन्धित तथ्याङ्कीय एकाइमा पुनः विवरण यकिन गरिन्छ । आवश्यकता अनुसार कोडिङ, डिकोडिङ, वर्गीकरण, तालिकीकरण, प्रस्तुतिकरण र विश्लेषणका कार्यहरू यस चरणमा गरिन्छ ।

ट) भविष्यमा गरिने सर्वेक्षणका लागि पृष्ठपोषण: कुनै पनि सर्वेक्षण सञ्चालन गर्दा

त्यससँग सम्बन्धित सबल पक्ष, कमजोर पक्ष, चुनौती र अवसरहरू अनुभूत गरिएको हुन्छ। यी सबै विषयहरू भविष्यमा सञ्चालन गरिने अन्य सर्वेक्षणहरूका लागि निकै उपयोगी हुन्छन्। तसर्थ, सर्वेक्षणको नतिजासँगै सिकाइ उपलब्धलाई पनि संस्थागत गरेर राख्नुपर्दछ।

१६.७.४ नमुनाको आकार (Sample Size)

नमुना छनोटका लागि कतिवटा एकाइ छनोट गर्ने भन्ने प्रश्न निकै महत्त्वपूर्ण हुन्छ। उपयुक्त नमुना आकार भएन भने त्यस्ता नमुनाहरूबाट प्राप्त गरिने सूचना जनसंख्याका विशेषता व्यक्त गर्न असफल हुन्छन्। तसर्थ, नमुना आकार निर्धारण गर्ने विधिको बारेमा साक्षिप्त चर्चा गरौं। जनसंख्याका एकाइहरूको विशेषता र आकार, अध्ययनको उद्देश्य, निकाल्नु पर्ने सूचकको विशेषता, साधन स्रोतको उपलब्धता, विश्वस्तताको तह, अनुमानको शुद्धताको स्तर लगायतका विभिन्न विषयमा नमुना आकार निर्भर गर्दछ। उदाहरणका लागि जनसंख्याको मध्यक अनुमान गर्नका लागि आवश्यक पर्ने नमुनाको आकार निम्नानुसारको सूत्रबाट निकालिन्छ।

$$n = \left(\frac{\sigma z_{\alpha/2}}{E} \right)^2 \quad (२)$$

जहाँ, n भनेको आवश्यक नमुना आकार हो, $z_{\alpha/2}$ भनेको विश्वसनीयता गुणक (Reliability Factor) हो, E भनेको स्वीकार गर्न सक्ने भ्रमांशको तह हो र α भनेको लेभल अफ सिजिनीफिक्यान्स (Level of Significance) हो। α को मान निकाल्न १ बाट विश्वस्तताको स्तर घटाउनु पर्दछ। σ भनेको जनसंख्याको स्तरीय भिन्नता हो।

उदाहरण: यदि हामीलाई ९५% विश्वस्तताको तह कायम हुने गरेर स्वीकार गर्न सक्ने भ्रमांश $+३$ देखि -३ सम्म राख्ने हो भने जनसंख्याको मध्यक अनुमान गर्ने प्रयोजनार्थ जनसंख्याको स्तरीय भिन्नता ७ छ भने हामीलाई निम्न अनुसारको नमुना आकार आवश्यक पर्दछ।

९५% विश्वस्तताको तह भन्नाले $1-\alpha = 0.95$

तसर्थ, $\alpha = 0.05$ र $\frac{\alpha}{2} = 0.025$

$z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$ हुन्छ।

अब, सूत्र (२) बमोजिम नमुना आकार निकाल्दा,

$$n = \left(\frac{\sigma z_{\alpha/2}}{E} \right)^2 = \left(\frac{7 \times 1.96}{3} \right)^2 = 20.915 = 21$$

त्यसैगरी जनसंख्याको कुनै विशेषताको समानुपात (Proportion) अनुमान गर्नका लागि,

जनसंख्याको आकार निकै ठुलो भएमा:
$$n = \frac{z^2 * P (1-P)}{e^2}, \quad (३)$$

जनसंख्याको आकार निकै ठुलो भएमा:
$$n' = \frac{n}{1 + \frac{n'}{N}}, \quad (४)$$

जहाँ, n भनेको आवश्यक नमुना आकार हो, z भनेको विश्वसनियता गुणक (Reliability Factor) हो, e भनेको स्वीकार गर्न सक्ने भ्रमांशको तह हो र P भनेको जनसंख्याको कुनै विशेषताको समानुपात हो । उदाहरणका लागि मानौं, औसतमा वार्षिकरूपमा लगभग ५० प्रतिशत नेपालीहरूले पर्यटकीय भ्रमण गर्दछन् । यस अवस्थामा नेपालीहरूको भ्रमणसम्बन्धी विशेषता अध्ययन गर्नका लागि ५ प्रतिशतसम्म भ्रमांश स्वीकार गर्ने गरी निम्नानुसारको नमुना आकार आवश्यक पर्दछ ।

$$n = \frac{z^2 * P (1-P)}{e^2},$$

$$n = \left(\frac{1.96}{0.05} \right)^2 \times 0.5 (1 - 0.5)$$

$$n = 384.16 \approx 385$$

यी बाहेक अन्य सूचकहरू अनुमान गर्नका लागि नमुना आकार निर्धारण गर्ने अरू सूत्रहरू पनि छन् तर यस पुस्तकमा ती सबैको उल्लेख गरिएको छैन । आजकल नमुना आकार पत्ता लगाउन अनलाइनमा विभिन्न सफ्टवेयर पाइन्छ, जहाँ विश्वस्तताको तह, स्वीकार गर्न सकिने भ्रमांशको मान र जनसंख्याको आकार राखेर नमुनाको आकार पत्ता लगाउन सकिन्छ ।

१६.८ अभ्यासका लागि प्रश्नहरू

क) संक्षिप्त परिभाषा दिनुहोस् ।

- १) जनसंख्या (Population)
- २) नमुना (Sample)
- ३) जनसंख्या मान (Population Value)
- ४) स्टार्टिस्टिक (Statistic)
- ५) आँकलन (Estimation)
- ६) आँकलक (Estimator)

ख) फरक छुट्याउनुहोस् ।

- १) सामान्य सम्भाव्यता नमुना छनोट र स्तरीकृत सम्भाव्य नमुना छनोट
- २) बिन्दुगत अनुमान र अन्तराल अनुमान
- ३) निष्पक्षता (Unbiasedness) र सामञ्जस्यता (Consistency)

ग) ठिक बेठिक छुट्याउनुहोस् ।

१) पर्याप्तता भएका सबै आँकलकहरू निष्पक्ष हुन्छन् ।

२) नमुना छनोटको अध्ययनले कम खर्चमा विश्वासिलो नतिजा पाइन्छ ।

३) स्नो बल नमुना छनोट एउटा सम्भाव्य नमुना छनोटको उदाहरण हो ।

घ) असल आँकलकमा हुनुपर्ने गुणहरू उल्लेख गर्नुहोस् ।

ङ) विश्वस्तताको तह ९५% कायम गरेर नमुना मध्यकबाट बढीमा ३ सम्म फरक पर्ने गरी जनसंख्याको मध्यक अनुमान गर्नुपर्ने भएमा नमुनाको आकार न्यूनतम कति हुनुपर्दछ ?

१६.९ सन्दर्भ सामग्री

अनुमान तथा नमुनासम्बन्धी थप जानकारीका लागि निम्न बमोजिमका सामग्री उपयोगी छन् ।

१. Statistical Techniques in Business & Economics by Douglas A. Lind, William G. Marchal & Samuel A. Wathen, Seventeenth Edition, McGraw-Hill Education, 2018

२. Fundamentals of Statistics by S.C. Gupta, Sixth Revised & Enlarged Edition, Himalaya Publishing House, 2009

१६.१० पदावलीहरू र तिनीहरूको अर्थ

१. **जनसंख्या (Population)**: अध्ययन गर्न खोजिएको सम्पूर्ण वस्तु वा वस्तुहरूको समष्टिलाई जनसंख्या भनिन्छ । जस्तै: नेपालमा सञ्चालित उद्योगहरूको समूह, कीर्तिपुर नगरपालिकाभित्र बसोबास गर्ने महिलाहरूको समूह, नेपालमा अक्सर बसोबास गर्ने मानिसहरूको समूह आदि ।

२. **नमुना (Sample)**: कुनै पनि जनसंख्याको समूहबाट, सो जनसंख्यालाई प्रतिनिधित्व गर्ने उपसमूहलाई नमुना भनिन्छ ।

३. **जनसंख्या मान (Population Value)**: जनसंख्याको विशेषता जनाउने वा जनसंख्याका सबै एकाइहरूको लक्षण बताउने सूचक वा मानलाई जनसंख्या मान भनिन्छ । जनसंख्या मानले जनसंख्याको विशेषताको वर्णन गर्दछन् । यसलाई प्यारामिटर (Parameter) पनि भन्ने गरिन्छ । उदाहरणका लागि जनसंख्या मध्यक, जनसंख्या एकाइहरूको कुलसंख्या, जनसंख्या स्तरीय भिन्नता, जनसंख्याका कुनै गुणको प्रतिशत आदि । यिनीहरूलाई ग्रीक भाषाका संकेतहरूले जनाउने गरिन्छ । जस्तै: जनसंख्याको मध्यकलाई μ ले जनाइन्छ भने स्तरीय भिन्नतालाई σ ले जनाइन्छ ।

४. **स्टाटिस्टिक (Statistic):** जनसंख्या मानले जनसंख्याको विशेषता जनाएजस्तै स्टाटिस्टिकले छानिएको नमुनाको विशेषता वा नमुनाका सबै एकाइहरूको लक्षण बताउँदछन् । अर्थात्, स्टाटिस्टिकले नमुनाको विशेषताको वर्णन गर्दछन् । यसलाई सङ्ख्यक पनि भन्ने चलन छ । उदाहरणका लागि नमुना मध्यक, नमुना स्तरीय भिन्नता आदि । यिनीहरूलाई अंग्रेजी भाषाका साना अक्षरका संकेतले जनाउने गरिन्छ । जस्तै: नमुना मध्यकलाई \bar{x} , स्तरीय भिन्नतालाई s ले जनाइन्छ ।

५. **आँकलन (Estimation):** नमुना छनोटको आधारमा गरिएको अध्ययन, अनुसन्धान वा अवलोकनबाट प्राप्त सूचनाको आधारमा समग्र जनसंख्याको बारेमा जानकारी लिने कार्यलाई अनुमान वा आँकलन भनिन्छ ।

६. **आँकलक (Estimator):** समग्र जनसंख्याको विशेषता वा कुनै लक्षण वा गुण विशेषको मान अनुमान गर्न नमुना छनोटबाट प्राप्त सूचनाको सूचना उपयोग गर्ने गणितीय सूत्रलाई आँकलक भनिन्छ । उदाहरणका लागि जनसंख्याको मध्यक आँकलन गर्न प्रयोग गरिने नमुना मध्यकको देहायको सूत्रलाई आँकलक भनिन्छ ।

$$\bar{x} = \frac{\sum x}{n}, \quad (9)$$

जहाँ, \bar{x} भनेको नमुना मध्यक हो, $\sum x$ भनेको नमुनामा छानिएका निश्चित लक्षणका मानहरूको योग हो, n भनेको नमुनाको आकार हो ।

७. **आँकलन भ्रमांश (Estimation Error):** अनुमान र जनसंख्या मानको बिचको फरक नै आँकलन भ्रमांश हो । मध्यकको भ्रमांश निकाल्दा, $\mu - E(x)$ हेर्न सकिन्छ । जहाँ, $E(x)$ भनेको जनसंख्या मध्यकको अनुमान हो ।



परिकल्पना परीक्षण

१७.१ परिचय

पाठ १६ मा हामीले विश्वसनीय अन्तरालको निर्माण गरी जनसंख्याका मान रहने दायरा अनुमान गर्ने विधिको बारेमा चर्चा गर्थौं। अब यस पाठमा हामीले जनसंख्याका मानका बारेमा लिइएको वा राखिएको धारणा वा कल्पना कति हदसम्म सही छ वा छैन भन्ने बारेमा तथ्याङ्कीय विधिको प्रयोग गरेर परीक्षण गर्ने सम्बन्धमा चर्चा गर्नेछौं। सामान्य अर्थमा परिकल्पना (Hypothesis) भनेको जनसंख्या मानका बारेमा गरिएको एक अप्रमाणित अनुमान हो। अध्येता वा शोधकर्ताले यस्ता अप्रमाणित अनुमानहरूलाई परिकल्पना परीक्षणको माध्यमबाट तथ्याङ्कीय रूपमा सार्थक प्रमाण पुऱ्याएर प्रमाणित गराउन चाहन्छन्। यसलाई केही निर्दिष्ट समूहको घटनाको व्याख्याको रूपमा प्रस्तुत गरिएको प्रस्ताव वा प्रस्तावहरूको समूहको रूपमा पनि परिभाषित गर्न सकिन्छ। परिकल्पनालाई प्रमाणीकरणको पर्खाइमा रहेको जनसंख्या मानहरूको बारेमा भनिएको कथन भनेर भन्न पनि सक्नेछन्। त्यस्ता कथनहरूको प्रमाणीकरणका लागि परिकल्पना परीक्षण गरिन्छ। अर्को शब्दामा भन्नुपर्दा, अनुसन्धान परिकल्पना एक भविष्यवाणी कथन हो, जसलाई वैज्ञानिक विधिहरूद्वारा परीक्षण गरी प्रमाणीकरण गरिन्छ। तसर्थ, परिकल्पना परीक्षण (Hypothesis Testing) भनेको सम्भाव्यता सिद्धान्तको प्रयोग गरेर नमुनाबाट प्राप्त सूचनाका आधारमा परिकल्पना उचित कथन हो वा होइन भनेर निर्धारण गर्ने विधि हो।

१७.२ तथ्याङ्कीय परिकल्पना र यसका प्रकार

तथ्याङ्कीय परिकल्पना भनेको के हो, भन्ने बारेमा हामीले जानकारी पाइसकेका छौं। यस बारेमा थप स्पष्ट हुनका लागि प्राध्यापक एस.सी. गुप्ताको भनाइ जस्ताको तस्तै यहाँ दिइएको छ:

“तथ्याङ्कीय परिकल्पना भनेको जनसंख्याको बारेमा वा समान रूपमा दिइएको जनसंख्याको विशेषता हुने सम्भाव्यता वितरणको बारेमा भनिएको वा दाबी गरिएको केही अनुमान वा कथन हो, जुन सही हुन पनि सक्छ वा नहुन पनि सक्दछ तर यसलाई हामी अनियमित नमुना (Random Sample) बाट प्रमाणको आधारमा परीक्षण गर्न चाहन्छौं।”

- प्राध्यापक एस.सी. गुप्ता

यदि यस्ता परिकल्पनाले जनसंख्यालाई पूर्ण रूपले निर्दिष्ट (Specify) गर्दछ भने त्यसलाई सामान्य परिकल्पना (Simple Hypothesis) भनिन्छ। यदि जसंख्यालाई पूर्ण रूपले निर्दिष्ट गर्दैन भने त्यसलाई जटिल परिकल्पना (Composite Hypothesis) भनिन्छ।

उदाहरण: नर्मल जनसंख्या (Normal Population) का दुईवटा प्यारामिटर (Parameters: मध्यक (μ) र भेरियन्स (σ^2)) हुन्छन् । यदि तथ्याङ्कीय परिकल्पनालाई निम्नानुसार उल्लेख गरियो भने, त्यो सामान्य परिकल्पना हुन्छ ।

सामान्य परिकल्पना, $H: \mu = \mu_0$ and $\sigma^2 = \sigma_0^2$ (यसले नर्मल डिस्ट्रिब्युसनलाई पूर्ण रूपमा निर्दिष्ट गरेको छ ।)

यदि तथ्याङ्कीय परिकल्पनालाई निम्नानुसार उल्लेख गरियो भने, त्यो जटिल परिकल्पना हुन्छ ।

जटिल परिकल्पना, $H: \mu = \mu_0$ (यसले नर्मल डिस्ट्रिब्युसनलाई पूर्ण रूपमा निर्दिष्ट गरेको छैन किनभने यो परिकल्पनाले भेरियन्सको बारेमा केही बोलेको छैन ।)

त्यसैगरी देहायका परिकल्पनाहरू जटिल प्रकारका परिकल्पना हुन् ।

$H: \sigma^2 = \sigma_0^2$ (यो परिकल्पनाले मध्यको बारेमा केही बोलेको छैन ।)

$H: \mu < \mu_0$ and $\sigma^2 = \sigma_0^2$ (यो परिकल्पनाले मध्यलाई पूर्ण रूपमा निर्दिष्ट गरेको छैन ।)

$H: \mu = \mu_0$ and $\sigma^2 > \sigma_0^2$ (यो परिकल्पनाले भेरियन्सलाई पूर्ण रूपमा निर्दिष्ट गरेको छैन ।)

सामान्य वा जटिल जस्तो प्रकारको भए तापनि परिकल्पना तय गर्दा सामान्यतया निम्न कुराहरूमा ध्यान दिनुपर्दछ ।

- यो स्पष्ट र सटिक हुनुपर्दछ,
- यो परीक्षण गर्न मिल्ने खालको हुनुपर्दछ,
- यसले चरहरू बिचको सम्बन्धलाई प्रतिनिधित्व गरेको हुनुपर्दछ,
- यसको दायरा सीमित हुनुपर्छ र विशिष्ट हुनुपर्छ,
- यो सबैभन्दा साधारण शर्तहरूमा भनिएको कथन हुनुपर्छ
- यसले वस्तुगत व्याख्याको आवश्यकता बोध गराउन सक्नुपर्दछ ।

तथ्याङ्कीय परिकल्पनालाई दुई वर्गमा विभाजन गरिएको हुन्छ: शून्य परिकल्पना (Null Hypothesis) र वैकल्पिक परिकल्पना (Alternative Hypothesis) ।

क) शून्य परिकल्पना: शोधकर्ताको दाबीको बारेमा तटस्थ रहेको परिकल्पनालाई शून्य परिकल्पना भनिन्छ । शून्य परिकल्पना जनसंख्याका मानका बारेमा गरिएको दाबी वा कथन हो । जस्तै: कसैले नेपालको गरिबी १५ प्रतिशत छ भन्ने दाबी गरे भने यो एउटा परिकल्पना हुन्छ । यसलाई तटस्थ भएर तथ्याङ्कीय परीक्षणमा लैजानका लागि शून्य परिकल्पनाको रूपमा, नेपालको गरिबी १५ प्रतिशत छ भन्दछौँ र तथ्याङ्कीय परीक्षण अवधिभर यसलाई सत्य नै

मान्दछौं । सामान्यतया शून्य परिकल्पनालाई हामीले अस्वीकार गर्न खोजिरहेका हुन्छौं । यसलाई H_0 (एच नट) ले जनाइन्छ । उदाहरणका लागि नेपालीहरूको औसत उमेर ३० वर्ष हो भन्ने कसैको दाबीलाई सत्य हो वा होइन भनेर प्रमाणिकरण गर्न शोधकर्ता तटस्थ भएर शून्य परिकल्पनाको रूपमा यस कथनलाई निम्नानुसार लेख्न सकिन्छ ।

$H_0 : \mu = 30$ (नेपालीहरूको औसत उमेर ३० वर्ष हो ।)

स्मरणीय कुरा के छ भने शून्य परिकल्पना लेख्दा जनसंख्या मानका बारेमा “=” (= वा \geq वा \leq) चिन्हको अनिवार्य प्रयोग गरेर लेख्नुपर्दछ । अर्थात्, “=” प्रयोग नभएको कल्पना शून्य परिकल्पना हुन सक्दैन ।

उदाहरण:

$H_0 : \mu = 30$ (नेपालीहरूको औसत उमेर ३० वर्ष हो ।) अथवा

$H_0 : \mu \geq 30$ (नेपालीहरूको औसत उमेर ३० वर्ष वा सोभन्दा बढी छ ।) अथवा

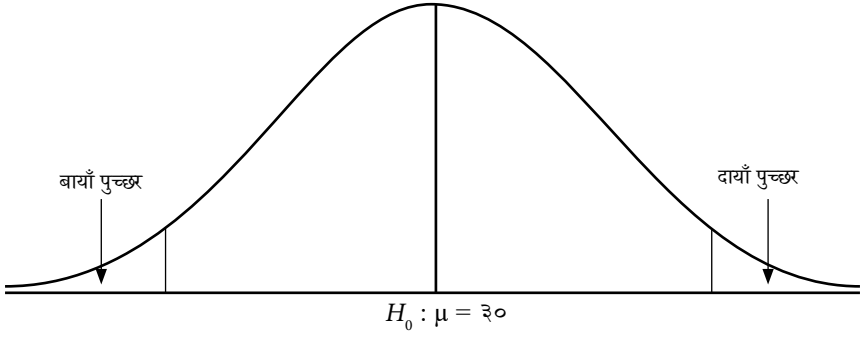
$H_0 : \mu \leq 30$ (नेपालीहरूको औसत उमेर ३० वर्ष वा सोभन्दा कम छ ।) यी कथनहरू शून्य परिकल्पना हुन् ।

तर $H_1 : \mu > 30$ (नेपालीहरूको औसत उमेर ३० वर्षभन्दा बढी छ ।) यो कथन शून्य परिकल्पना हुन सक्दैन किनभने यहाँ “=” को प्रयोग भएको छैन ।

ख) वैकल्पिक परिकल्पना: शोधकर्ताको दाबीको बारेमा तटस्थ नहुने परिकल्पनालाई वैकल्पिक परिकल्पना भनिन्छ । यसलाई H_1 (एच वान) ले जनाइन्छ । वैकल्पिक परिकल्पनालाई शोधकर्ताले प्रमाण पुऱ्याएर सत्य सावित गर्न खोजिरहेका हुन्छन् । तथ्याङ्कले शून्य परिकल्पना गलत छ भन्ने प्रमाण पुऱ्याएको खण्डमा वैकल्पिक परिकल्पना स्वीकार गरिन्छ । माथिको उदाहरणको लागि वैकल्पिक परिकल्पना निम्नानुसार लेख्न सकिन्छ । यहाँ वैकल्पिक परिकल्पनाका तिनवटा प्रकारहरू उल्लेख गरिएको छ । यीमध्ये कथनको दाबीको आवश्यकताअनुसार कुनै एउटामात्र वैकल्पिक परिकल्पनाका तय गर्नुपर्दछ । उदाहरणका लागि नेपालीहरूको औसत उमेर ३० होइन भन्ने हो भने पहिलो, ३० वर्षभन्दा बढी छ भन्ने हो भने दोस्रो र ३० वर्षभन्दा कम छ भन्ने हो भने तेस्रो वैकल्पिक परिकल्पना तय गर्नुपर्दछ ।

१) $H_1 : \mu \neq 30$ (दुईतर्फी परीक्षण - Two-tailed Test)

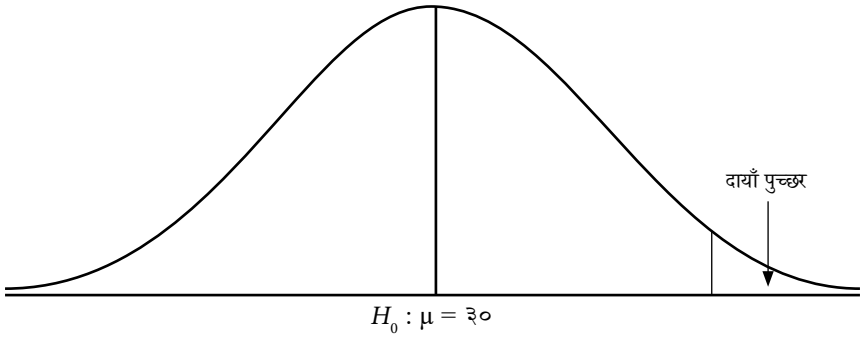
चित्र नं. १७.१: दुईतर्फी परिकल्पना परीक्षण



वा

२) $H_1 : \mu > 30$ (दायाँतर्फी परीक्षण - Right-tailed Test)

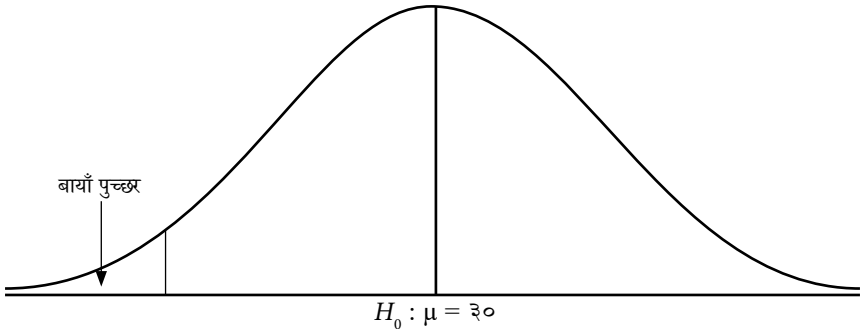
चित्र नं. १७.२: दायाँतर्फी परिकल्पना परीक्षण



वा

३) $H_1 : \mu < 30$ (बायाँतर्फी परीक्षण - Left-tailed Test)

चित्र नं. १७.३: बायाँतर्फी परिकल्पना परीक्षण



१७.३ परिकल्पना परीक्षणमा हुने त्रुटि र तिनीहरूको प्रकार

परिकल्पना परीक्षण गर्ने क्रममा नमुनाबाट प्राप्त सूचना वा प्रमाणका आधारमा तथ्याङ्कीय विधिबाट निर्णय लिइन्छ । कहिलेकाहीँ पर्याप्त सूचनाको अभावले गलत निर्णय लिने सम्भावना पनि रहेको हुन्छ । जस्तो कि जनसंख्याका बारेमा गरिएको उचित वा सही कथनलाई पनि प्रमाण नपुगेर अस्वीकार गर्ने निर्णय गर्नु वा गलत कथनलाई अस्वीकार गर्ने पर्याप्त प्रमाणको अभावमा अस्वीकार गर्न नसक्नु । उदाहरणका लागि नेपालीको जन्मदाको औसत आयु सन् २०११ मा ६६.६ वर्ष थियो भनेर कसैले दाबी गरेको कथनलाई प्रमाणको अभावमा अस्वीकार गरियो भने यो गलत निर्णय हुन्छ किनभने उक्त समयमा नेपालीको जन्मदाको औसत आयु ६६.६ वर्ष नै थियो । त्यसैगरी हाल नेपालीको जन्मदाको औसत आयु ६६.६ वर्ष छ भनेर कसैले दाबी गरेको कथनलाई प्रमाणको अभावमा अस्वीकार गर्न सकिएन भने पनि यो गलत निर्णय हुन्छ किनभने नेपालीको जन्मदाको औसत आयु ६६.६ वर्ष भनेको सन् २०११ को हो हालको होइन । यी दुवै परिकल्पना परीक्षणका सम्भावित त्रुटिहरू (Errors in Hypothesis Testing) हुन् । यी दुई किसिमका त्रुटिका सम्भाव्यताहरूलाई क्रमशः α र β ले जनाइन्छ । अर्थात्,

क) पहिलो प्रकारको त्रुटि (Type I Error)

$$P(\text{सही कथनलाई अस्वीकार गर्नु}) = \alpha, \quad (१)$$

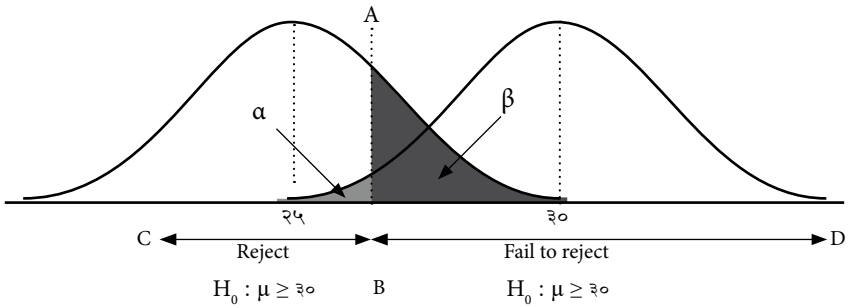
ख) दोस्रो प्रकारको त्रुटि (Type II Error)

$$P(\text{गलत कथनलाई अस्वीकार गर्न नसक्नु}) = \beta, \quad (२)$$

जहाँ, P ले सम्भाव्यता भन्ने बुझाउँदछ । अर्थशास्त्रीय दृष्टिबाट पहिलो प्रकारको त्रुटिलाई उत्पादकको जोखिम (Producer's Risk) र दोस्रो प्रकारको त्रुटिलाई उपभोक्ताको जोखिम (Consumer's Risk) पनि भनिन्छ । उदाहरणका लागि कुनै औषधी कम्पनीले उत्पादन गरेको गुणस्तरीय औषधी अस्वीकार गरियो भने त्यो उत्पादकको लागि हानिकारक हुन्छ । अर्कोतर्फ गुणस्तरहीन औषधी पनि अस्वीकार गर्न सकिएन भने त्यो उपभोक्ताको लागि हानिकारक हुन्छ । तसर्थ, परिकल्पना परीक्षण निकै संवेदनशील विषय हो । अर्थात्, सकेसम्म माथिका दुवै किसिमका त्रुटिहरूलाई न्यूनीकरण गरेर मात्र निर्णयमा पुग्न उत्तम हुन्छ । तर एक किसिमको त्रुटि कम गर्दछु भन्दा आपसे आफ अर्को किसिमको त्रुटि बढ्न जाने हुनाले प्रायजसो पहिलो प्रकारको त्रुटि (α) लाई स्थिर गरेर (साधारण तया १%, ५%, १०%) दोस्रो प्रकारको त्रुटि (β) कम गर्ने गरिन्छ । उदाहरणका लागि चित्र नं. १७.४ मा α र β ले प्रतिनिधित्व गर्ने क्षेत्रफलहरू देखाइएको छ, जहाँ देब्रेतिरको बक्ररेखाले जनसंख्याको मध्यकको वास्तविक मान २५ हुँदा नमुना मध्यकको

वास्तविक सम्भाव्यता वितरण जनाएको छ भने दाहिनेतिरको बक्ररेखाले जनसंख्याको मध्यकको मान ३० मान्दा (शून्य परिकल्पनालाई सत्य मान्दा) नमुना मध्यकको काल्पनिक सम्भाव्यता वितरण जनाएको छ । यस चित्रमा ठाडो रेखा AB ले α र β को सम्बन्ध स्थापित गरिदिएको छ । यदि यो ठाडो रेखालाई दायाँतर्फ सार्ने हो भने α बढ्दछ र β घट्दछ । यदि यदि यो ठाडो रेखालाई बायाँतर्फ सार्ने हो भने α घट्दछ र β घट्दछ । यहाँ, $(1 - \alpha)$ लाई विश्वस्तताको तह (Confidence Level) र $(1 - \beta)$ लाई परीक्षणको शक्ति (Power of Test) भनिन्छ । विश्वस्तताको तहले सही शून्य परिकल्पनालाई अस्वीकार नगर्ने सम्भाव्यता जनाउँदछ भने परीक्षणको शक्तिले गलत शून्य परिकल्पनालाई अस्वीकार गर्ने सम्भाव्यता जनाउँदछ ।

चित्र १७.४: परिकल्पना परीक्षणका त्रुटिहरू



नोट: चित्र नं. १७.४ मा दुईवटा नर्मल सम्भाव्यता बक्ररेखा देखाइएको छ । यस उदाहरणमा मानौं हामीले कुनै एउटा वास्तविक जनसंख्या लिएका छौं जसको मध्यक २५ रहेको छ । यस अवस्थामा नमुना मध्यकको वास्तविक सम्भाव्यता वितरण बायाँतर्फको बक्रले जनाउँदछ । कुनै एउटा कथन जसले सो जनसंख्याको मध्यक ३० वा सोभन्दा बढी भनेर दाबी गरेको अवस्थामा यस परिकल्पना ($H_0: \mu \geq 30$) लाई परीक्षण गर्न नमुना लिएर नमुना मध्यकको काल्पनिक सम्भाव्यता वितरण जनाउन दायाँतर्फको बक्र तयार गरिएको छ । तसर्थ, परिकल्पना परीक्षण गर्नका लागि ठाडो रेखा AB भन्दा बायाँको भाग अस्वीकार्य क्षेत्र हुन्छ र दायाँतर्फको बक्रले बायाँतर्फ बनाएको क्षेत्रफल (फिक्का खैरो) α हुन्छ । त्यसैगरी, परिकल्पना परीक्षण गर्नका लागि ठाडो रेखा AB भन्दा दायाँको भाग अस्वीकार गर्न नसकिने क्षेत्र हुन्छ । तर बायाँतर्फको वास्तविक सम्भाव्यता बक्रलाई हेर्दा ठाडो रेखा AB भन्दा दायाँको भाग (गाढा खैरो रङ लगाएको क्षेत्र) अस्वीकार्य क्षेत्र हुनुपर्ने हो तर शून्य परिकल्पनालाई सत्य मान्दा यो क्षेत्र अस्वीकार गर्न नसकिने क्षेत्र भइदिएको हुनाले यसको क्षेत्रफल β हुन्छ । यस अवस्थामा गलत कथनलाई पनि अस्वीकार गर्न सकिँदैन ।

१७.४ लेभल अफ सिजिनीफक्यान्स

पहिलो प्रकारको त्रुटिको सम्भाव्यतालाई लेभल अफ सिजिनीफक्यान्स (Level of Significance) भनिन्छ । यसलाई ग्रीक संकेत α ले जनाइन्छ । माथि उल्लेख गरिए बमोजिम तुलनात्मक रूपमा दोस्रो प्रकारको त्रुटिभन्दा पहिलो प्रकारको त्रुटि कम हानिकारक हुने हुन्छ । पहिलो प्रकारको त्रुटिले शून्य परिकल्पना सही हुँदाहुँदै पनि यसलाई अस्वीकार गर्ने अवस्था बताउँदछ भने दोस्रो प्रकारको त्रुटिले गलत परिकल्पनालाई पनि अस्वीकार गर्न नसक्ने अवस्था बताउँदछ । कुनै पनि अवस्थामा गलत विषयलाई अस्वीकार गर्न नसक्नु भनेको ठुलो क्षति व्यहोर्नु हो । जस्तै: कुनै औषधी खराब हुँदाहुँदै पनि यसलाई अस्वीकार गर्न सकिएन भने अथवा कुनै सामान खराब हुँदाहुँदै पनि अस्वीकार गर्न सकिएन भने त्यसले ठुलो हानि गर्दछ । तर सही कुरालाई अस्वीकार गर्दा तुलनात्मक रूपमा कम हानि गर्दछ ।¹⁵ तसर्थ, पहिलो प्रकारको त्रुटिलाई कुनै निर्दिष्ट तहमा स्थिर कायम गरी दोस्रो प्रकारको त्रुटि कम गर्नेतर्फ परिकल्पना परीक्षण केन्द्रित हुन्छ । सामान्यतया पहिलो प्रकारको त्रुटिलाई १%, ५%, वा १०% कायम गर्ने गरिन्छ । पहिलो प्रकारको त्रुटिलाई कुन तहमा स्थिर गर्ने भन्ने विषय निकै जटिल विषय हो । यो मुख्यतया: अध्ययनको क्षेत्रअनुसार निर्भर गर्दछ । उदाहरणका लागि अर्थशास्त्रका भनाइहरूमा ५% वा १०% त्रुटिलाई सहज रूपमा स्वीकार गर्न सकिन्छ भने चिकित्सा विज्ञानका भनाइहरूमा ज्यादै कम त्रुटिको अपेक्षा गरिएको हुन्छ । कसैले यो औषधी १०% सम्म कम प्रभावकारी हुने सम्भावना छ भनेमा हामीलाई उक्त औषधी प्रयोग गर्न मन लाग्दैन वा प्रयोग गर्न सक्दैनौं । तर अर्को वर्ष नेपालको आर्थिक वृद्धिदर ७% हुनेछ र यो भनाइ १०% त्रुटिपूर्ण हुनसक्दछ भनेमा पनि खासै फरक पर्दैन ।

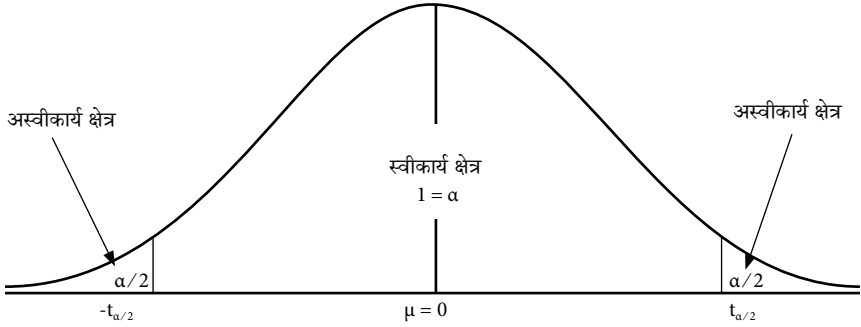
१७.५ क्रिटिकल मान तथा क्रिटिकल क्षेत्र

परिकल्पना परीक्षणमा निर्णय लिनका लागि कुनै सन्दर्भ आधार हुनुपर्दछ । यसका लागि हामीले स्टाण्डर्ड नर्मल डिस्ट्रिब्युसन फङ्सनको बक्ररेखाको सहायताले बुझ्न सक्दछौं । तल प्रस्तुत गरिएको चित्र नं. १७.५ मा उदाहरणका लागि दुईतर्फी परीक्षणको नमुना देखाइएको छ । पहिलो प्रकारको त्रुटिलाई बराबर दुई भागमा बाँडेर नर्मल बक्ररेखाको दुवै तर्फ वितरण गरिएको छ । जसलाई $\alpha/2$ ले जनाइएको छ । यी दुई भागलाई अस्वीकार्य क्षेत्र (Region of Rejection) भनिन्छ । यदि परीक्षण गर्ने स्टाटिस्टिक यी क्षेत्रमा पर्यो भने हामीले शून्य परिकल्पनालाई अस्वीकार गर्दछौं । उदाहरणका लागि तलको

15 पहिलो प्रकारको त्रुटिलाई उत्पादकको जोखिम पनि भनिन्छ किनभने सही उत्पादन हुँदाहुँदै पनि यो त्रुटिका कारण अस्वीकार हुनपुग्दछ भने दोस्रो प्रकारको त्रुटिलाई उपभोक्ताको जोखिम पनि भनिन्छ किनभने गलत सामान हुँदाहुँदै पनि यो त्रुटिका कारण अस्वीकार गर्न सकिएको हुँदैन ।

चित्र १७.५ मा यदि परीक्षण गर्ने स्टाटिस्टिकको मान हिसाब गर्दा $t_{\alpha/2}$ भन्दा ठुलो वा $-t_{\alpha/2}$ भन्दा सानो आयो भने शून्य परिकल्पनालाई अस्वीकार गरिन्छ । यी दुई क्षेत्रभन्दा बाँकी रहेको क्षेत्रलाई स्वीकार्य क्षेत्र (Region of Acceptance) भनिन्छ । जसलाई $(1 - \alpha)$ ले जनाइएको छ । यदि परीक्षण गर्ने स्टाटिस्टिक यो क्षेत्रमा पर्यो भने हामीले शून्य परिकल्पनालाई अस्वीकार गर्न सक्दैनौं । अर्थात्, यदि परीक्षण गर्ने स्टाटिस्टिकको मान हिसाब गर्दा $t_{\alpha/2}$ भन्दा सानो र $-t_{\alpha/2}$ भन्दा ठुलो आयो भने शून्य परिकल्पनालाई अस्वीकार गर्न सकिँदैन । यस्ता अस्वीकार क्षेत्र छुट्याउने मान (जस्तै: चित्र १७.५ मा $-t_{\alpha/2}$ र $t_{\alpha/2}$) लाई क्रिटिकल मान भनिन्छ । यही क्रिटिकल मानका आधार मा अस्वीकार गर्ने वा नगर्ने भन्ने निर्णय लिने गरिन्छ । क्रिटिकल मानहरू विभिन्न तालिका (जेड टेबल, टी टेबल आदि) मा उपलब्ध हुन्छन् ।

चित्र १७.५: क्रिटिकल मान तथा क्रिटिकल क्षेत्र



१७.६ टेष्ट स्टाटिस्टिक्स र यसको पी-मान

परिकल्पना परीक्षण गर्नका लागि नमुनाका मानको प्रयोग गरिएको हुन्छ । त्यस्ता नमुनाका मानलाई सान्दर्भिक सम्भाव्यता वितरणमा (Relevant Probability Distribution) रूपान्तरण गरी व्यक्त गर्ने स्टाटिस्टिक्स नै टेष्ट स्टाटिस्टिक्स हो । यस्ता सम्भाव्यता वितरणहरूमा स्टाण्डर्ड नर्मल डिस्ट्रिब्युसन, टी-डिस्ट्रिब्युसन, काई-डिस्ट्रिब्युसन, एफ-डिस्ट्रिब्युसन आदि प्रचलित छन् । यस खण्डमा हामी नर्मल अर्थात् जेड-डिस्ट्रिब्युसन र टी-डिस्ट्रिब्युसनको प्रयोगको मात्र उदाहरण प्रस्तुत गर्दछौं । वास्तवमा स्टाण्डर्ड नर्मल डिस्ट्रिब्युसन र टी-डिस्ट्रिब्युसन ठुलो भिन्नता छैन । यदि जनसंख्याको स्तरीय भिन्नता (σ) थाहा छ भने स्टाण्डर्ड नर्मल डिस्ट्रिब्युसन प्रयोग गरिन्छ, थाहा छैन भने टी-डिस्ट्रिब्युसनको प्रयोग गरिन्छ । टी-डिस्ट्रिब्युसन प्रयोग गर्ने बेलामा स्वतन्त्रताको तह

(Degrees of Freedom = df) निर्धारण गर्नुपर्ने हुन्छ ।¹⁶ टी-डिस्ट्रिब्युसनको हकमा स्वतन्त्रताको तह निकाल्न नमुना आकारबाट १ घटाउनु पर्दछ । उदाहरणका लागि नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ मा जम्मा २८,६७० जनाको उमेर सोधिएको थियो, यस अवस्थामा स्वतन्त्रताको तह $२८,६७० - १ = २८६६९$ हुन्छ ।

उदाहरण: माथि उल्लेख गरिएको परिकल्पनामा प्रयोग हुने टेष्ट स्टाटिस्टिक्स हेरौं । यहाँ, शून्य परिकल्पना र वैकल्पिक परिकल्पना निम्न बमोजिम छन् ।

$H_0: \mu = ३०$ (नेपालीहरूको औसत उमेर ३० वर्षभन्दा फरक छैन ।)

$H_1: \mu \neq ३०$ (नेपालीहरूको औसत उमेर ३० वर्षभन्दा फरक छ: दुईतर्फा परीक्षण

(Two-tailed Test)

क) यदि जनसंख्याको स्तरीय भिन्नता (σ) थाहा छ भने,

$$\text{टेष्ट स्टाटिस्टिक्स } (z_{cal}) = \frac{x - \mu}{\sigma / \sqrt{n}}, \quad (३)$$

ख) यदि जनसंख्याको स्तरीय भिन्नता (σ) थाहा छैन भने,

$$\text{टेष्ट स्टाटिस्टिक्स } (t_{cal}) = \frac{\bar{x} - \mu}{s / \sqrt{n}}, \quad (४)$$

जहाँ, \bar{X} भनेको नमुनाको मध्यक हो, जसको मान २६.९२७०७ (नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ को नमुनाको मध्यक) रहेको छ । μ भनेको जनसंख्याको मध्यक हो, जसको मान शून्य परिकल्पनाको आधारमा ३० राखिन्छ । σ भनेको जनसंख्याको स्तरीय भिन्नता हो, जुन हामीलाई थाहा छैन । तसर्थ, यो परिकल्पना परीक्षणमा टी-टेष्टको प्रयोग गर्नुपर्दछ । s भनेको नमुनाको उमेरको स्तरीय भिन्नता हो, जुन २०.२०२२८ (नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ को नमुनाको उमेरको स्तरीय भिन्नता) रहेको छ । n भनेको नमुनामा समावेश भएका सदस्यहरूको जम्मा संख्या हो, जुन २८,६७० (नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ को उमेर लिइएको मानिसको संख्या) रहेको छ । यी मानहरू सूत्र ४ मा राख्दा,

16 स्वतन्त्रताको तह (Degrees of Freedom): कुनै पनि डाटामा स्वतन्त्र रूपले छनोट गर्न सकिने मानहरूको संख्यालाई स्वतन्त्रताको तह भनिन्छ । उदाहरणका लागि लमजुङ जिल्लाका ८ वटा स्थानीय तहमध्ये ३ वटा छान्नु परेको छ । त्यसरी छान्दा ती तिनवटाको जम्मा जनसंख्या ६०००० भन्दा बढी हुनुपर्दछ भनियो भने दुईवटा स्थानीय तह स्वतन्त्र रूपले छान्न सकिन्छ । तर तेस्रो स्थानीय तह छान्नका लागि स्वतन्त्र हुन सकिँदैन किनभने तेस्रो छानेर तिनवटाको जनसंख्या ६०००० भन्दा बढी बनाउनुपर्ने छ । यस्तो अवस्थामा स्वतन्त्रताको तह २ हुन्छ ।

$$\text{टेष्ट स्टाटिस्टिक्स } (t_{cal}) = \frac{२६.९२७०७ - ३०}{२०.२०२२८} = \frac{-३.०७२९३}{०.११९३१३} = -२५.७५ \text{ हुन्छ।}$$

परिकल्पना परीक्षणमा बारम्बार प्रयोग हुने केही अन्य टेष्ट स्टाटिस्टिक्सहरू देहायबमोजिम छन् ।

ग) दुईवटा जनसंख्याका भेरियन्सहरू तुलना गर्न,

$$F_{cal} = \frac{s_1^2}{s_2^2}, \quad (५)$$

जहाँ, F भनेको एफ स्टाटिस्टिक्स हो, s_1 / s_2 क्रमशः दुईवटा जनसंख्याहरूबाट लिइएका नमुनाहरूको स्तरीय भिन्नता हुन् ।

घ) दुईवटा जनसंख्याको मध्यकबिच कुनै फरक छैन भनेर परीक्षण गर्न,

१) यदि ती दुई जनसंख्याको स्तरीय भिन्नता थाहा छ भने,

$$Z_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}, \quad (६)$$

जहाँ, \bar{x}_1 र \bar{x}_2 भनेको दुईवटा जनसंख्याबाट लिइएका नमुनाहरूको मध्यकहरू हुन्, σ_1^2 र σ_2^2 भनेको तिनीहरूको आ-आफ्नो भेरियन्स हो । n_1 र n_2 भनेको सम्बन्धित नमुनाका आकारहरू हुन् ।

२) यदि ती दुई जनसंख्याको स्तरीय भिन्नता थाहा छ भने,

$$t_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}, \quad (७)$$

जहाँ, s_p^2 भनेको पुल्ड भेरियन्स हो र यसलाई निम्नानुसार निकालिन्छ ।

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}, \quad (८)$$

जहाँ, s_1 र s_2 भनेका सम्बन्धित नमुनाहरूका स्तरीय भिन्नता हुन् ।

ङ) जनसंख्याको कुनै गुणको समानुपात (Proportion) का बारेमा परीक्षण गर्दा,

$$Z_{cal} = \frac{p - P}{\sqrt{\frac{P(1-P)}{n}}}, \quad (९)$$

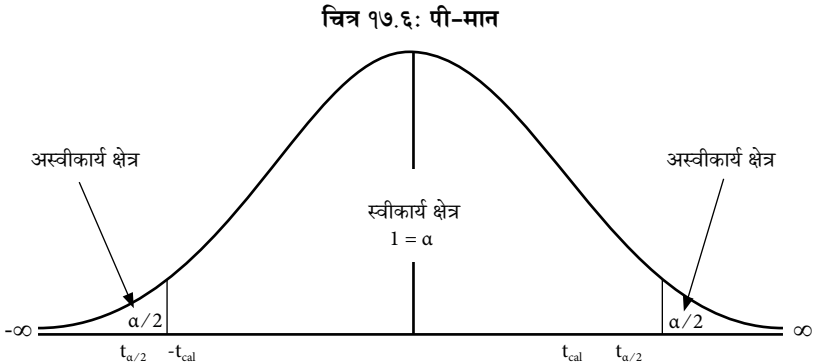
जहाँ, P भनेको नमुनामा भएको कुनै गुणको समानुपात हो भने P भनेको उक्त गुणको जनसंख्यामा भएको समानुपात हो । सूत्र (९) अनुसार टेस्ट स्टाटिस्टिक्स निकाल्नको लागि जनसंख्याको समानुपात P को मान शून्य परिकल्पनामा अनुमान गरिएको मान राखेर हिसाब गर्नुपर्दछ । n भनेको नमुनाको आकार हो ।

माथि उल्लेख गरिएका बाहेक धेरै टेस्ट स्टाटिस्टिक्सहरू हुन्छन् । यस पुस्तकको दायराभित्र सबै नपर्ने हुनाले यहाँ उल्लेख गरिएको छैन ।

परिकल्पना परीक्षणमा निर्णय लिने दुई विधिहरू छन् । एउटा टेष्ट स्टाटिस्टिक्सको हिसाब गरेर आएको मानलाई तालिकाको क्रिटिकल मानसँग दाँजेर निर्णय लिने र अर्को भनेको टेष्ट स्टाटिस्टिक्सको पी-मान (P-Value) निकालेर α सँग तुलना गरेर निर्णय लिने । निर्णय प्रक्रियामा पी-मानको प्रयोग गर्ने प्रचलन निकै लोकप्रिय छ । तसर्थ, यसका बारेमा संक्षिप्त जानकारी लिऔं । हिसाब गरेर निकालिएको टेष्ट स्टाटिस्टिक्सको मान देखि ∞ सम्म सम्बन्धित डिस्ट्रिब्युसनको बक्ररेखाले बनाएको क्षेत्रफल नै पी-मान हो । अर्थात्, एकतर्फी पुच्छर परीक्षण (One-tailed Test) का लागि,

$$\text{पी-मान (P-Value)} = p(t > t_{\text{cal}}), \quad (१०)$$

जहाँ, t_{cal} भनेको हिसाब गरेर निकालिएको टेष्ट स्टाटिस्टिक्सको मान हो ।



चित्र १७.६ मा t_{cal} भन्दा दायँतर्फ ∞ सम्म बक्ररेखाले X अक्षसँग बनाएको क्षेत्रफल पी-मान हो । निर्णय प्रक्रियामा पी-मान α भन्दा सानो भएमा मात्र हामीले शून्य परिकल्पना अस्वीकार गर्न सक्दछौं । किनभने पी-मान α भन्दा सानो भएमा मात्र हाम्रो टेष्ट स्टाटिस्टिक्स अस्वीकार्य क्षेत्रमा पर्दछ ।

१७.७ परिकल्पना परीक्षणका प्रमुख चरणहरू

परीक्षण परिकल्पनाका लागि चाहिने आधारभूत जानकारी प्राप्त गरिसकेपछि यसका

प्रमुख चरणहरूका बारेमा संक्षिप्त बुँदागत रूपमा स्मरण गरौं।

चरण १: शून्य परिकल्पना तथा वैकल्पिक परिकल्पना तय गर्ने।

चरण २: पहिलो प्रकारको त्रुटि हुने सम्भाव्यता (α) निर्धारण गर्ने।

चरण ३: टेष्ट स्टाटिस्टिक्स ($z, t, F, Chi\text{-square}$ मध्ये कुनै एक) पहिचान गर्ने।

चरण ४: क्रिटिकल मान वा पी-मानबाट निर्णय नियम तयार गर्ने।

चरण ५: नमुना छनोट गर्ने र टेष्ट स्टाटिस्टिक्सको हिसाब गरेर क्रिटिकल मान वा पी-मानबाट निर्णयमा पुग्ने।

चरण ६: नतिजाको व्याख्या गर्ने।

१७.८ एकल मध्यक परीक्षण गर्ने उदाहरण

नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ ले छनोटमा परेका परिवारका २८,६७० सदस्यहरूको औसत उमेर २६.९२७०७ वर्ष र त्यसको स्तरीय भिन्नता २०.२०२२८ वर्ष देखाएको छ। यी विवरणका आधारमा ९५ प्रतिशत ढुक्क हुनेगरी नेपालीहरूको औसत उमेर ३० वर्ष हुन सक्दैन भनी परिकल्पना परीक्षण गर्नुहोस्।

यसका लागि माथिका ६ चरणहरू अनुसरण गरौं।

चरण १: शून्य परिकल्पना तथा वैकल्पिक परिकल्पना तय गर्ने।

$$H_0 : \mu = 30$$

$$H_1 : \mu \neq 30 \text{ (दुईतर्फी परीक्षण - Two-tailed Test)}$$

चरण २: पहिलो प्रकारको त्रुटि हुने सम्भाव्यता (α) निर्धारण गर्ने।

यहाँ, ९५% ढुक्क हुनका लागि $\alpha = 5\% = 0.05$ र $\frac{\alpha}{2} = 0.025$ हुन्छ।

चरण ३: टेष्ट स्टाटिस्टिक्स ($z, t, F, Chi\text{-square}$ मध्ये कुनै एक) पहिचान गर्ने।

औसत उमेरका बारेमा परीक्षण गर्नुपर्ने र जनसंख्याको स्तरीय भिन्नता थाहा नभएको हुनाले हामीले टी-टेष्ट प्रयोग गर्दछौं, जसको टेष्ट स्टाटिस्टिक्स यसप्रकार हुन्छ।

$$\text{टेष्ट स्टाटिस्टिक्स } (t_{cal}) = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

चरण ४: क्रिटिकल मान वा पी-मानबाट निर्णय नियम तयार गर्ने।

क) क्रिटिकल मानबाट गरिने निर्णय नियम: नमुना आकार २८,६७० भएको हुनाले स्वतन्त्रताको तह २८,६६९ हुन्छ र तालिकामा सबैभन्दा ठुलो १००० भएको हुनाले हामीले यही लहरको मान हेर्नुपर्दछ। अर्को कुरा, अनुसूची २ मा दिइएको तालिका एकतर्फी टी-तालिका हो र हाम्रो परीक्षण दुईतर्फी हो। यस्तो अवस्थामा ०.०५ लेभल अफ सिग्नीफिकेन्सका लागि एकतर्फी टी-तालिकाको ०.०२५ लेभल अफ

सिग्नीफिकेन्समा क्रिटिकल मान हेर्नुपर्दछ। तसर्थ, हाम्रो यस उदाहरणका लागि क्रिटिकल मान १.९६२ हुन्छ।

निर्णय नियम

यदि $t_{cal} > १.९६२$ वा $t_{cal} < -१.९६२$ भयो भने हामीलाई शून्य परिकल्पना अस्वीकार गर्ने प्रमाण पुग्दछ। यस्तो अवस्थामा वैकल्पिक परिकल्पना सत्य सावित हुन्छ।

यदि $१.९६२ \geq t_{cal} \geq -१.९६२$ भयो भने हामीलाई शून्य परिकल्पना अस्वीकार गर्ने प्रमाण पुग्दैन। यस्तो अवस्थामा शून्य परिकल्पना अस्वीकार गर्न सक्दैनौं।

ख) पी-मानबाट गरिने निर्णय नियम: यसका लागि पहिला पी-मान (P-Value) निकाल्नुपर्दछ। त्यसपछि निकालिएको P-Value लाई α सँग तुलना गर्ने।

यदि P-Value $< \alpha$ भयो भने हामीलाई शून्य परिकल्पना अस्वीकार गर्ने प्रमाण पुग्दछ। यस्तो अवस्थामा वैकल्पिक परिकल्पना सत्य सावित हुन्छ।

यदि P-Value $> \alpha$ भयो भने हामीलाई शून्य परिकल्पना अस्वीकार गर्ने प्रमाण पुग्दैन। यस्तो अवस्थामा वैकल्पिक परिकल्पना अस्वीकार गर्न सक्दैनौं।

नोट: P-Value निकाल्नको लागि अलि बढी तथ्याङ्कीय ज्ञान र सीप आवश्यक पर्ने तथा आजकल यसको मान निकाल्न अनलाइनमा विभिन्न सफ्टवेयर समेत पाइने हुनाले¹⁷ यहाँ, यसलाई निकाल्ने प्रयास गरिएको छैन।

चरण ५: नमुना छनोट गर्ने र टेष्ट स्टाटिस्टिक्सको हिसाब गरेर क्रिटिकल मान वा पी-मानबाट निर्णयमा पुग्ने।

नेपाल जीवनस्तर सर्वेक्षणबाट नमुनाका मानहरू लिएको हुनाले यसैलाई नमुना मानिएको छ। अब, टेष्ट स्टाटिस्टिक्सको हिसाब गर्दा,

$$\text{टेष्ट स्टाटिस्टिक्स } (t_{cal}) = \frac{२६.९२७०७ - ३०}{\frac{२०.२०२२८}{\sqrt{२८,६७०}}} = \frac{-३.०७२९३}{०.११९३१३} = -२५.७५ \text{ हुन्छ।}$$

चरण ६: नतिजाको व्याख्या गर्ने।

निर्णय नियम बमोजिम $t_{cal} < -१.९६२$ अर्थात्, $-२५.७५ < -१.९६२$ भएको हुनाले शून्य परिकल्पना अस्वीकार गर्ने प्रशस्त प्रमाण पुगेको छ। तसर्थ, नेपालीहरूको औसत उमेर ३० वर्ष होइन भन्नेमा ९५% ढुक्का हुन सकिन्छ।

17 Social Science Statistics. Retrieved on December 7, 2021 from: <https://www.socscistatistics.com/pvalues/tdistribution.aspx>

अनलाइन सफ्टवेयरबाट P -Value हेर्दा $p < 0.00001$ पाइयो र यो मान 0.05 भन्दा सानो भएको हुनाले शून्य परिकल्पना अस्वीकार गर्न सकिने भयो ।

१७.९ दुई मध्यकबिचको फरक परीक्षण गर्ने उदाहरण

नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ ले तनहुँ जिल्लाको घाँसीकुवा र आँबुखैरेनी गाउँपालिकाहरूको नमुना औसत प्रतिव्यक्ति आय, प्रतिव्यक्ति आयको नमुना स्तरीय भिन्नता, र नमुना आकार निम्नानुसार देखाएको छ ।

घाँसीकुवा:

नमुना मध्यक: $\bar{X}_1 = \text{रु. } ७४९०९.३६$

नमुना स्तरीय भिन्नता: $s_1 = \text{रु. } ७११८९.५१$

नमुना आकार: $n_1 = \text{रु. } १२$

आँबुखैरेनी:

नमुना मध्यक: $\bar{X}_2 = \text{रु. } ६५९९३.८७$

नमुना स्तरीय भिन्नता: $s_2 = \text{रु. } ८५६१६.३३$

नमुना आकार: $n_2 = १२$

घाँसीकुवा र आँबुखैरेनी गाउँपालिकाहरूको औसत प्रतिव्यक्ति आयमा कुनै फरक छैन भन्ने कथन वा दाबीलाई परिकल्पना परीक्षणबाट अस्वीकार गर्न सकिन्छ वा सकिँदैन ?

चरण १: शून्य परिकल्पना तथा वैकल्पिक परिकल्पना तय गर्ने ।

$H_0 : \mu_1 = \mu_2$ (घाँसीकुवा र आँबुखैरेनी गाउँपालिकाहरूको औसत प्रतिव्यक्ति आयमा कुनै फरक छैन)

$H_0 : \mu_2 \neq \mu_1$ (घाँसीकुवा र आँबुखैरेनी गाउँपालिकाहरूको औसत प्रतिव्यक्ति आयमा कुनै फरक छैन: दुईतर्फी परीक्षण- Two-tailed Test)

चरण २: पहिलो प्रकारको त्रुटि हुने सम्भाव्यता (α) निर्धारण गर्ने ।

यहाँ, ९५% ढुक्क हुनका लागि $\alpha = 5\% = 0.05$ र $\frac{\alpha}{2} = 0.025$ हुन्छ ।

चरण ३: टेष्ट स्टाटिस्टिक्स ($z, t, F, \text{Chi-square}$ मध्ये कुनै एक) पहिचान गर्ने ।

औसत प्रतिव्यक्ति आयको फरकका बारेमा परीक्षण गर्नुपर्ने र जनसंख्याको स्तरीय भिन्नता थाहा नभएको हुनाले हामीले टी-टेष्ट प्रयोग गर्दछौं जसको टेष्ट स्टाटिस्टिक्स यसप्रकार हुन्छ ।

$$\text{टेष्ट स्टाटिस्टिक्स } (t_{cal}) = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

जहाँ, s_p^2 भनेको पुल्ड भेरियन्स हो र यसलाई निम्नानुसार निकालिन्छ ।

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

चरण ४: क्रिटिकल मानबाट निर्णय नियम तयार गर्ने ।

यदि $t_{cal} > १.९६२$ वा $t_{cal} < -१.९६२$ भयो भने हामीलाई शून्य परिकल्पना अस्वीकार गर्ने प्रमाण पुग्दछ । यस्तो अवस्थामा वैकल्पिक परिकल्पना सत्य सावित हुन्छ ।

यदि $१.९६२ \geq t_{cal} \geq -१.९६२$ भयो भने हामीलाई शून्य परिकल्पना अस्वीकार गर्ने प्रमाण पुग्दैन । यस्तो अवस्थामा शून्य परिकल्पना अस्वीकार गर्न सक्दैनौं ।

चरण ५: नमुना छनोट गर्ने र टेष्ट स्टाटिस्टिक्सको हिसाब गरेर क्रिटिकल मानबाट निर्णयमा पुग्ने ।

यो चरणमा लागि आवश्यक पर्ने नमुनाका मानहरू माथि दिइएको हुनाले ती मानहरूलाई सूत्रमा राखेर हिसाब गरौं । पहिला पुल्ड भेरियन्स निम्नानुसार निकालिन्छ ।

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$s_p^2 = \frac{(१२ - १)७११८९.५१^२ + (१२ - १)८५६१६.३३^२}{१२ + १२ - २}$$

$$s_p^2 = \frac{(११)५०६७९५७७२४.३७ + (११)७३३०१५५९६२.६७}{२२}$$

$$s_p^2 = \frac{१३६३७९२५०५५७.४१}{२२}$$

$$s_p^2 = ११३६४९३७५४६.४५$$

अब, टेष्ट स्टाटिस्टिक निकाल्ने,

$$t_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

$$t_{cal} = \frac{७४९०९.३६ - ६५९९३.८७}{\sqrt{११३६४९३७५४६.४५ \left(\frac{1}{१२} + \frac{1}{१२} \right)}}$$

$$t_{cal} = \frac{८९१५.४९}{४३५२१.९१}$$

$$t_{cal} = ०.२०४९$$

चरण ६: नतिजाको व्याख्या गर्ने ।

यदि $१.९६२ \geq (t_{cal} = ०.२०४९) \geq -१.९६२$ भएको हुनाले हामीलाई शून्य परिकल्पना अस्वीकार गर्ने प्रमाण पुग्दैन । यस्तो अवस्थामा शून्य परिकल्पना अस्वीकार गर्न सकिँदैन । तसर्थ, घाँसीकुवा र आँबुखैरेनी गाउँपालिकाहरूको औसत प्रतिव्यक्ति आयमा कुनै फरक छैन भन्ने कथनलाई अस्वीकार गर्न सकिएन ।

१७.१० अभ्यासका लागि प्रश्नहरू

क) ठिक बेठिक छुट्याउनुहोस् ।

- १) शून्य परिकल्पना जनसंख्याका मानका बारेमा गरिएको दाबी वा कथन हो ।
- २) सही कथनलाई अस्वीकार गर्नुलाई परीक्षणको शक्ति भनिन्छ ।
- ३) जनसंख्याको स्तरीय भिन्नता थाहा छ र यसको मध्यकको बारेमा परिकल्पना परीक्षण गर्नुछ भने जेड परीक्षण गरिन्छ ।
- ४) क्रिटिकल मान र क्रिटिकल क्षेत्र भनेको एउटै कुरा हो ।

ख) फरक छुट्याउनुहोस् ।

- १) शून्य परिकल्पना र वैकल्पिक परिकल्पना
- २) पहिलो प्रकारको त्रुटि र दोस्रो प्रकारको त्रुटि
- ३) लेभल अफ सिजिनीफक्यान्स र कन्फिडेन्स लेभल
- ४) टी-टेस्ट र जेड-टेस्ट

ग) खाली ठाउँ भर्नुहोस् ।

- १) दुईवटा जनसंख्याको भेरियन्स तुलना गर्नको लागि प्रयोग गरिने टेस्ट स्टाटिस्टिक्स हो ।
- २) $(1 - \alpha)$ लाई भनिन्छ ।
- ३) P (गलत कथनलाई अस्वीकार गर्न नसक्नु) =
- ४) लेख्दा जनसंख्या मानका बारेमा “=” (= वा \geq वा \leq) चिन्हको अनिवार्य प्रयोग गरेर लेख्नुपर्दछ ।

घ) परिकल्पना परीक्षणका छवटा चरणहरू उल्लेख गर्नुहोस् ।

ङ) नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ अनुसार छानिएका ५,९८८ परिवारको औसत वार्षिक आमदानी रू. ४,५६,३९२.२ छ र यसको स्तरीय भिन्नता ४८,९०,६२४ रहेको छ । यदि कसैले नेपाली परिवारको सरदर वार्षिक आमदानी ४,५०,००० छ भनेमा ९०% विश्वस्त हुने गरी अस्वीकार गर्नुहुन्छ कि अस्वीकार गर्न सक्नुहुन्छ ?

१७.११ सन्दर्भ सामग्री

परिकल्पना परीक्षणका लागि थप अध्ययन गर्न देहायको सामग्री उपयुक्त छ ।

१. Statistical Techniques in Business & Economics by
Douglas A.Lind, William G.Marchal & Samuel

A. Wathen, Seventeenth Edition, McGraw–Hill Education, 2018



प्रतिगमन विश्लेषण

१८.१ परिचय

केन्द्रीय प्रवृत्ति र फैलावट एकल वितरणमा आधारित तथ्याङ्कीय विश्लेषण हुन् । सम्बन्धको मापन अर्थात् सहसम्बन्धको विश्लेषण द्विविभाजन वितरणमा मात्र सम्भव छ । तथापि सहसम्बन्धले दुई चरबिचको सम्बन्धको दिशा र सम्बन्धको तहमात्र बताएको हुन्छ । सहसम्बन्धले कारण र असरका बारेमा बताउन सक्दैन । यसका अतिरिक्त सहसम्बन्धले एउटा चरमा आएको एकाइ परिवर्तनले अर्को चरमा कति परिवर्तन आउँछ भनेर व्यक्त गर्न सक्दैन । तथ्याङ्क विश्लेषणमा सहसम्बन्धले बताउन नसक्ने कारण र असर तथा एक चरले अर्को चरलाई गर्ने असरको परिमाणात्मक अध्ययनका लागि प्रतिगमन विश्लेषणको अध्ययन गरिन्छ ।

यदि एक चर X को प्रत्येक एकाइ मापनको लागि हामीसँग दोस्रो चर Y को अनुरूप मान छ भने, मानहरूको परिणामस्वरूपका जोडीको वितरणलाई द्विविभाजन वितरण (Bivariate Distribution) भनिन्छ । यसका अतिरिक्त, हामीसँग तेस्रो चर वा चौथो चर वा सोभन्दा बढी अनुरूप मान पनि हुन सक्दछन् र यस्ता एकाइ समूहहरूको वितरणलाई जोडीहरूको नतिजा मानहरूलाई बहुविभाजन वितरण (Multivariate Distribution) भनिन्छ । हामी प्रायः डाटामा दुई वा बढी चरहरूको एकअर्कासँगको सम्बन्ध जान्न चाहन्छौं । उदाहरणका लागि अध्ययन घण्टा, शिक्षकको गुणस्तर, अध्ययन सामग्री आदिसँग परीक्षामा प्राप्त अंकहरूको सम्बन्ध; उमेर, र उचाइसँग तौलको सम्बन्ध; सरकारी खर्च, वैदेशिक लगानीसँग कुल गार्हस्थ्य उत्पादनको सम्बन्ध आदि ।

प्रतिगमन (Regression) शब्दको शाब्दिक अर्थ औसततर्फ फर्कनु हो । जब चरका दुई वा बढी सेटहरू नजिकबाट सम्बन्धित हुन्छन्, यी चरहरू बिचको सम्बन्धको रूपको निर्धारणलाई प्रतिगमन विश्लेषण भनिन्छ । अर्थात्, एउटा चरका मानहरू अर्को चरका मानहरूसँग औसतमा कति सम्बन्धित छन् र कसरी असर गरिरहेका छन् भनेर प्रतिगमन विश्लेषणबाट हेर्ने गरिन्छ । सामान्यतया प्रतिगमनको अध्ययन अर्थमिति (Econometrics) मा अधिक मात्रामा गरिन्छ ।

प्रतिगमन विश्लेषणको प्रयोग निकै व्यापक छ । प्रतिगमन विश्लेषण स्वतन्त्र चरको ज्ञात मानहरूबाट एउटा आश्रित वा निर्भर चरको अज्ञात मानहरूको अनुमान वा भविष्यवाणी गर्ने प्रक्रियाहरूसँग सम्बन्धित छ । यसै कारणले तथ्याङ्क विश्लेषणमा यसको निकै महत्त्व रहेको छ । यसका केहि प्रमुख प्रयोगका क्षेत्र निम्नानुसार रहेका छन् ।

- क) दुई वा दुईभन्दा बढी चरहरू बिचको सम्बन्धको तह अनुमान गर्न,
 ख) सिद्धान्त वा परिकल्पना परीक्षण गर्न,
 ग) चरहरूको पूर्वानुमान वा प्रक्षेपण गर्न,
 घ) नीतिहरूको मूल्याङ्कन गर्न ।

१८.२ सहसम्बन्ध र प्रतिगमनबिचको भिन्नता

सहसम्बन्ध र प्रतिगमन दुवैले चरहरूबिचको सम्बन्धका बारेमा अध्ययन गर्दछन् । तथापि यी दुई तथ्याङ्कीय प्रविधिबिच केही भिन्नता रहेको छ । यिनीहरूको भिन्नताका बारेमा तलको तालिकामा प्रस्तुत गरिएको छ ।

तालिका १८.१: सहसम्बन्ध र प्रतिगमनबिचको भिन्नता

तुलनाका आधार	सहसम्बन्ध	प्रतिगमन
अर्थ, परिभाषा	दुई चरबिचको सम्बन्धका बारेमा बताउने तथ्याङ्कीय मापनलाई सहसम्बन्ध भनिन्छ ।	दुई वा बढी चरहरू बिचको औसत गणितीय सम्बन्धमा आधारित सांख्यिकीय प्रविधिलाई प्रतिगमन भनिन्छ ।
प्रयोग	चरहरूबिचको रेखीय सम्बन्ध प्रतिनिधित्व गर्न ।	एउटा चरको मानको आधारमा अर्को चरको अनुमान गर्न ।
कस्तो अवस्थामा प्रयोग गर्ने ?	दुई चरबिचको सम्बन्धको दिशा तत्काल चाहिएमा ।	एउटा चरले अर्को चरलाई कसरी र कुन हदसम्म असर वा प्रभाव गरिरहेको छ भन्ने चाहिएमा ।
चरको प्रकृति	कुनै फरक पर्दैन ।	एउटाले प्रभाव पारेको हुन्छ भने अर्को प्रभावित हुन्छ ।
के इञ्जित गर्दछ ?	दुई चरहरूले कुन दिशा र तहसम्मको सम्बन्ध राखेका छन् भन्ने बताउँदछ ।	एउटा स्वतन्त्र चरमा गरिएको एकाइ परिवर्तनको कति असर अर्को निर्भर चरमा परेको छ भन्ने देखाउँदछ ।
उद्देश्य	दुई चरहरूबिचको सम्बन्ध देखाउने संख्यात्मक मान पत्ता लगाउनु ।	स्वतन्त्र चरको मानको आधारमा निर्भर चरको मान अनुमान, प्रक्षेपण गर्नु ।
ग्राफमा कस्तो देखिन्छ ?	एउटा बिन्दुले प्रतिनिधित्व गर्दछ ।	रेखाद्वारा रेखीय प्रतिगमन हेर्न सकिन्छ ।

१८.३ स्वतन्त्र तथा निर्भर चरहरू

प्रतिगमन विश्लेषण सुरु गर्नुअघि हामीले चरहरूको प्रकृति पत्ता लगाउनुपर्छ, यो प्रकृति अनुसार यी चरहरूलाई निम्न २ वर्गहरूमा विभाजन गरिएको छ ।

क) निर्भर चर (Dependent Variable): अन्य चरहरूको परिवर्तनशीलताबाट प्रभावित हुने चरहरूलाई निर्भर चर भनिन्छ । जस्तै: बाली उत्पादन, हिमालमा हिउँको स्तर, सहरहरूमा वायु प्रदूषणको स्तर आदि । बाली उत्पादन मलको मात्रा, वर्षाको मात्रा, किटनाशक औषधीको मात्रा आदिमा निर्भर रहेको हुन्छ । त्यस्तै हिउँको स्तर तापक्रममा निर्भर रहन्छ । सहरमा वायु प्रदूषणको स्तर उद्योगको संख्या, मानिसको संख्या आदिमा निर्भर रहन्छ ।

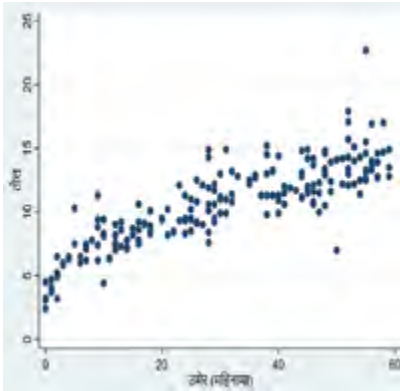
ख) स्वतन्त्र चर (Independent Variable): चरहरू जो आफैमा उत्पन्न हुन्छन् र निर्भर चरहरूलाई प्रभाव पार्छन्, तिनीहरूलाई स्वतन्त्र चर भनिन्छ । जस्तै: वर्षा, मल, सवारी साधनहरूको संख्या आदि हुन् । निर्भर चरहरू पनि सँधै निर्भर रहन्छन् भन्ने हुँदैन । यिनीहरू पनि अध्ययनको क्षेत्र, उद्देश्य वा दायराबमोजिम निर्भर चरका रूपमा परिभाषित गर्न पनि सकिन्छ । जस्तै: सवारी साधनहरूको संख्या जनसंख्याको आकारसँग निर्भर रहन सक्दछ । मलको मात्रा गाईवस्तुको संख्यासँग निर्भर रहन सक्दछ आदि ।

१८.४ सिधा रेखीय प्रतिगमन विश्लेषण

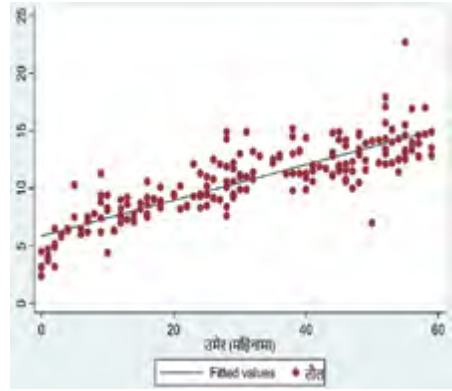
दुई चरहरू बिचको औसत सम्बन्धको वर्णन गर्ने रेखालाई प्रतिगमनको सिधा रेखा (Simple Regression Line) भनिन्छ । प्रतिगमन रेखाले कुनै चरको अर्थसँग मिल्ने अर्को चरको उत्कृष्ट औसत मानहरू देखाउँछ । उदाहरणका लागि नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ को तथ्याङ्कको आधारमा हिमाली क्षेत्रका पाँच वर्षमुनिका १९८ बच्चाहरूको तौल र उमेर (महिनामा) को सम्बन्ध चित्र नं. १८.१ (क) को स्क्याटर डायग्राममा देखाइएको छ । यही सम्बन्ध स्थापित गर्दै उमेर अनुसारको उत्कृष्ट औसत तौल चित्र नं. १८.१ (ख) मा “Fitted Values” नामको सिधा रेखाले प्रतिनिधित्व गरेको छ । यही सिधा रेखालाई प्रतिगमनको सिधा रेखा भनिन्छ ।

चित्र नं. १८.१: पाँच वर्ष मुनिका बच्चा (हिमाली क्षेत्र) को तौल र उमेरको सम्बन्ध

क) तौल र उमेरको स्क्याटर डायग्राम



ख) तौल र उमेरको प्रतिगमन रेखा



स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११

प्रतिगमन रेखाहरू वर्णन गर्ने समीकरणहरूलाई प्रतिगमन समीकरण भनिन्छ । जस्तै: माथिको उदाहरणमा तौल र उमेरको प्रतिगमन रेखा वर्णन गर्नका लागि निम्न बमोजिमको समीकरण प्रयोग गर्न सकिन्छ ।

$$y = \beta_0 + \beta_1 x + u, \quad (१)$$

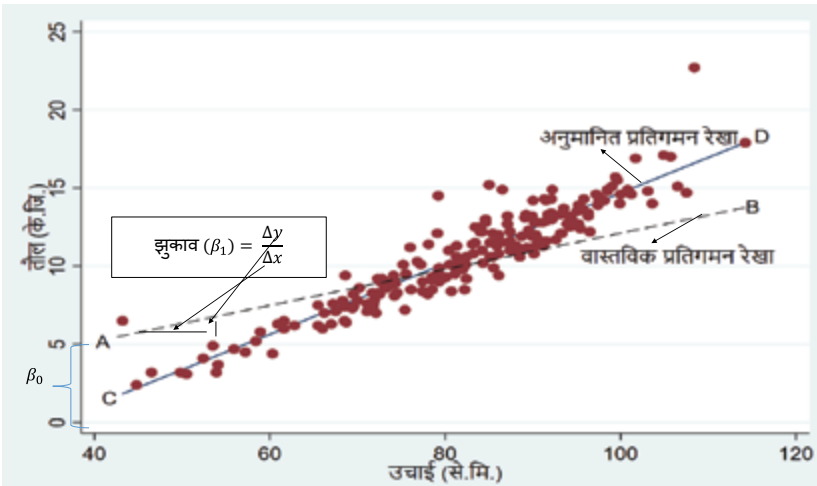
जहाँ, y भनेको निर्भर चर हो, x भनेको स्वतन्त्र चर हो, u भनेको निर्भर चरलाई प्रभाव गर्ने अन्य तत्त्वहरूको प्रतिनिधि हो जसलाई Error Term अथवा Disturbance अथवा Unobservable भनिन्छ। β_0 भनेको Intercept हो र β_1 भनेको भुकाव (Slope) हो । अर्को शब्दमा β_0 भनेको स्वतन्त्र चर तथा अन्य कारक चरहरूलाई शून्य बनाउँदा निर्भर चरको मान हो भने β_1 भनेको स्वतन्त्र चर x ले निर्भर चर y लाई पारेको प्रभाव मापक गुणाङ्क हो । यसले एक एकाइ x बढदा y मा आउने परिवर्तनको मात्रा वा दर बताउँदछ । प्रतिगमन विश्लेषणको प्रमुख उद्देश्य भनेकै यिनै β_0 र β_1 को अनुमान पत्ता लगाउनु हो । समीकरण (१) लाई माथिको उदाहरणको प्रतिनिधिको रूपमा लिइन्छ किनभने वास्तविक चरहरूको सम्बन्ध प्रतिनिधित्व गर्ने प्रतिगमन समीकरण यसप्रकार हुन्छ ।

$$\text{तौल} = \beta_0 + \beta_1 \text{उमेर} + u, \quad (२)$$

समीकरण (१) वा (२) मा β_0 र β_1 का अनुमान पत्ता लगाएपछि हामीले स्वतन्त्र चरले निर्भर चरमा पारेको प्रभाव आँकलन गर्नुका साथै स्वतन्त्र चरका विभिन्न मानको सहायताले निर्भर चरको पूर्वानुमान वा प्रक्षेपण गर्ने सकिन्छ । तलको चित्र नं. १८.२ मा रेखा AB (Dotted line) ले वास्तविक जनसंख्यामा चरहरू (तौल र उचाइ) बिचको प्रतिगमन

रेखा जनाएको छ । यस चित्रमा β_0 (Intercept) र β_1 (Slope = भुकाव) पनि देखाइएको छ । तर यो रेखाको वास्तविक अवस्थिति थाहा हुँदैन किनभने हामीलाई जनसंख्याका सबै मानहरू थाहा हुँदैन । फलस्वरूप यी β_0 र β_1 का वास्तविक मान पनि थाहा हुँदैन । तसर्थ, यिनीहरूको मान अनुमान गर्न उक्त जनसंख्याबाट नमुना छनोट गरी नमुना मानका आधारमा अनुमानित प्रतिगमन रेखा तयार गरीन्छ र β_0 र β_1 का अनुमानहरू पनि पत्ता लगाइन्छ । चित्र १८.२ मा रेखा CD ले नमुनाका मानहरूको आधारमा तयार गरिएको प्रतिगमन रेखा जनाएको छ । यही रेखा CD को Intercept र भुकावलाई जनसंख्याको Intercept र भुकावको आँकलक भनिन्छ । प्रतिगमन विश्लेषणको प्रमुख उद्देश्य भनेको रेखा CD पत्ता लगाएर यसको Intercept र भुकाव पत्ता लगाउनु हो ।

चित्र नं. १८.२: पाँच वर्षमुनिका बच्चा (हिमाली क्षेत्र) को तौल र उमेरको प्रतिगमन



चरहरूबिचको सम्बन्ध अध्ययन गर्नका लागि विभिन्न प्रकारका प्रतिगमन नमुना (Regression Model) हरूको प्रयोग गरिन्छ । खासगरी चरहरूले स्थापित गरेको सम्बन्धको आधारमा केही प्रतिगमनका नमुनाहरू यहाँ दिइएको छ ।

क) सरल प्रतिगमन नमुना (Simple Regression Model): एउटा निर्भर चर र अर्को स्वतन्त्र चर बिचको सम्बन्ध स्थापित गर्ने नमुनालाई सरल वा साधारण प्रतिगमन नमुना भनिन्छ । उदाहरणको लागि माथिको समीकरण (१) सरल प्रतिगमन नमुना हो ।

ख) बहुप्रतिगमन नमुना (Multiple Regression Model): एक निर्भर चर र अन्य धेरै स्वतन्त्र चरहरू बिचको सम्बन्धलाई बहुप्रतिगमन नमुना भनिन्छ, यो प्रतिगमन पूर्वानुमान/मोडलिंग प्रविधिहरूमा सबैभन्दा बढी प्रयोग गरिन्छ । जस्तै:

$$\text{तौल} = \beta_0 + \beta_1 \text{ उमेर} + \beta_2 \text{ आहर} + \beta_3 \text{ रोग} + u, \quad (३)$$

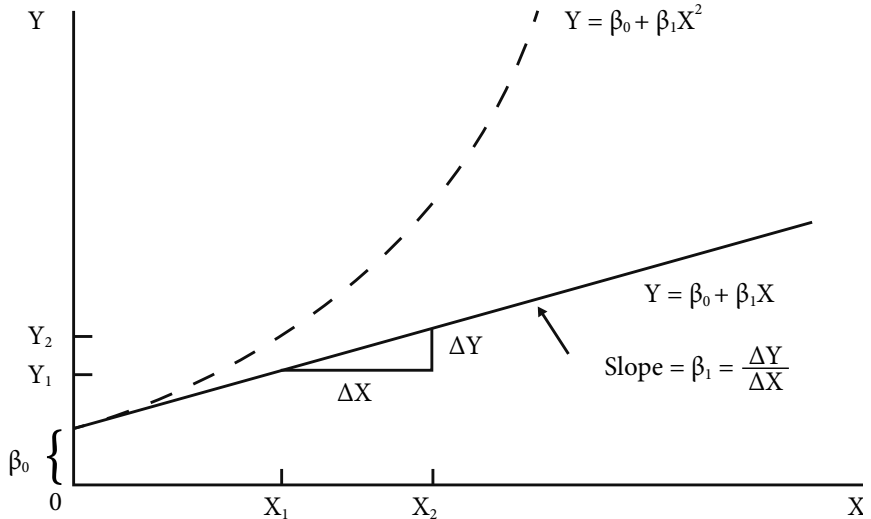
ग) सिधा रेखीय प्रतिगमन नमुना (**Linear Regression Model**): यदि चरहरूबिचको सम्बन्धलाई सिधा रेखाद्वारा व्याख्या गर्न सकिन्छ भने त्यस्तो प्रतिगमनलाई सिधा रेखीय प्रतिगमन नमुना भनिन्छ। उदाहरणको लागि माथिको समीकरण (१) सिधा रेखीय प्रतिगमन नमुना हो।

घ) गैर-रेखीय प्रतिगमन (**Non - Linear Regression Model**): सिधा रेखाद्वारा व्याख्या गर्न नसकिने खण्डमा विभिन्न बक्र रेखाबाट व्यक्त गरिने प्रतिगमनलाई गैररेखीय प्रतिगमन भनिन्छ। जस्तै:

$$\text{तौल} = \beta_0 + \beta_1 \text{ उमेर} + \beta_2 \text{ आहर} + \beta_3 \text{ रोग} + \beta_4 \text{ उमेर}^2 + u, \quad (४)$$

यी विभिन्न प्रकारका प्रतिगमनमध्ये यस पुस्तकमा प्रतिगमनको आधारभूत अवधारणा मात्र प्रस्तुत गर्ने उद्देश्यले सरल सिधा रेखीय प्रतिगमन नमुना (**Simple Linear Regression Model**) को मात्र चर्चा गरिएको छ।

चित्र १८.३ सिधा रेखीय तथा गैर-रेखीय प्रतिगमन नमुना



१८.५ गुणाङ्कहरूको अनुमान तथा व्याख्या

सामान्य सिधा रेखीय प्रतिगमन विश्लेषण (**Simple Linear Regression Analysis**) मा स्वतन्त्र चरले निर्भर चरलाई कस्तो प्रभाव पारिरहेको छ भनेर सेटिर्स परिवस प्रभाव (**Ceteris Paribus Effect**) अध्ययन गर्न सकिन्छ। यसका साथै, उक्त समीकरणलाई

प्रयोग गरेर भविष्यमा हुनसक्ने निर्भर चरका मानहरू समेत अनुमान लगाउन सकिन्छ । उदाहरणका लागि माथिको समीकरण (१) लाई लिऔं ।

$$y = \beta_0 + \beta_1 x + u,$$

यस समीकरणका β_0 र β_1 लाई देहाएका सूत्रहरू प्रयोग गरेर अनुमान गर्न सकिन्छ ।

$$\widehat{\beta}_0 = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2}, \quad (५)$$

$$\widehat{\beta}_1 = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}, \quad (६)$$

जहाँ, β_0 र β_1 मा लगाइएको ह्याट (^) संकेतले तिनीहरूको अनुमानित मान भन्ने बुझाउँदछ । y र x ले निर्भर तथा स्वतन्त्र चरका नमुनाबाट लिइएको मानहरूको प्रतिनिधित्व गर्दछन् भने n ले नमुनाको आकार जनाउँदछ । यसरी नमुनाका मानहरूबाट दुई चरहरूबीचको सम्बन्ध स्थापित गर्ने गुणाङ्कहरूको अनुमान निकालेर गरिने अध्ययनलाई प्रतिगमन विश्लेषण भनिन्छ । प्रतिगमन विश्लेषणमा गुणाङ्कहरूको अनुमान निकालेर तिनीहरू साँच्चिकै महत्त्वपूर्ण हुन् वा होइनन् भनेर परीक्षण समेत गर्नुपर्ने हुन्छ । प्रतिगमन विश्लेषणका लागि कम्प्युटर सफ्टवेयरको व्यापक प्रयोग गरिने तथा यसरी कम्प्युटरको प्रयोग गर्दा अनुमानका गुणस्तर उच्च हुने र सरल र सजिलो हुने हुनाले गणितीय हिसाब प्रयोग गरेर अर्थात्, सूत्र (५) र (६) को हिसाब निकालेर गरिने प्राचीन विधिको यस पुस्तकमा चर्चा गरिएको छैन । साथै, कम्प्युटरको प्रयोगबाट कसरी प्रतिगमन विश्लेषण गरिन्छ भन्ने बारेमा पाठ २२ र २३ मा चर्चा गरिएको हुनाले यस पाठमा कम्प्युटरबाट प्राप्त नतिजाहरू सिधै प्रस्तुत गरिएको छ ।

उदाहरण: नेपालको हिमाली भेगमा बच्चाहरूको तौललाई उमेरले कसरी प्रभाव पारिरहेको छ व्याख्या गर्नुहोस् ।

यसका लागि सबैभन्दा पहिला प्रतिगमन नमुना निर्धारण गरौं । उमेरको प्रभाव तौलमा हेर्न खोजिएको हुनाले यस उदाहरणका लागि समीकरण (२) को साधारण सिधा रेखीय प्रतिगमन नमुना प्रयोग गरौं ।

$$\text{तौल} = \beta_0 + \beta_1 \text{ उमेर} + u,$$

जहाँ, तौल के.जी. मा र उमेर महिनामा मापन गरिएको छ । र यस समीकरणमा तौल निर्भर चर हो भने उमेर स्वतन्त्र चर हो । अर्थात्, तौल उमेरको परिवर्तनसँगै परिवर्तन हुन्छ भने उमेर स्वतः रूपमा बढिरहेको छ ।

दोस्रो चरण भनेको तथ्याङ्क संकलनको चरण हो । समीकरण (२) को अनुमान गर्न आवश्यक पर्ने तथ्याङ्क नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ बाट लिऔं ।

तेस्रो चरणमा कम्प्युटरको सफ्टवेयर प्रयोग गरेर प्रतिगमन विश्लेषण गरौं । यहाँ, यसका लागि STATA बाट प्राप्त प्रतिगमन विश्लेषणको नतिजालाई जस्ताको तस्तै निम्न तालिका १८.२ मा प्रस्तुत गरिएको छ ।

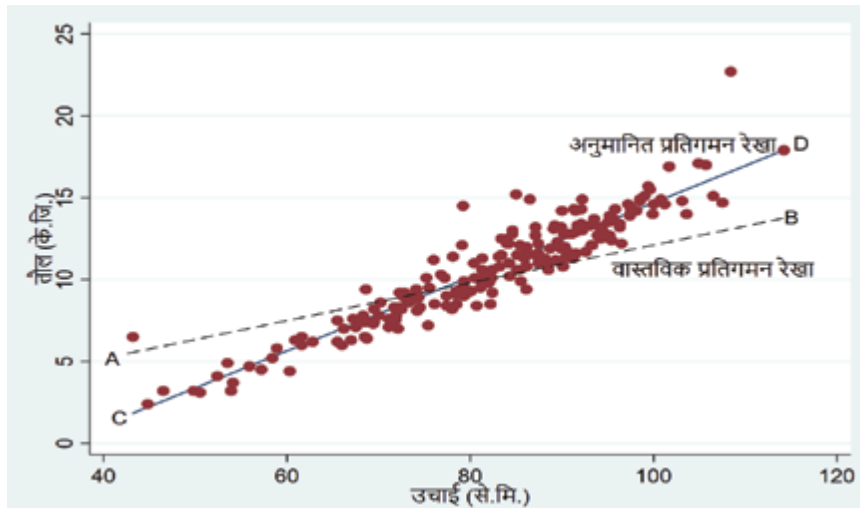
तालिका १८.२ बच्चाको तौल र उमेर बिचको प्रतिगमन विश्लेषण नतिजा

Source	SS	d.f	MS	Number of obs	-	198
Model	1473.2567	1	1473.2567	Prob > chi2	-	495.95
Residual	582.238037	196	2.97060253	Wald chi2	-	0.0000
Total	2055.4948	197	10.4339837	Adj R squared	-	0.7153
				F(1, 196)	-	1.7235

वैर	Coeff.	Std. Err.	t	> t	[95% Conf. Interval]	P-Value
उमेर	1546107	.0069426	22.27	0.000	1.409189 1.683025	
_cons	5.88901	.2472084	23.74	0.000	5.38148 6.35654	

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ को तथ्याङ्क प्रयोग गरी हिमाली क्षेत्रका लागि निकालिएको प्रतिगमन विश्लेषणको नतिजा

चित्र नं. १८.४ बच्चाको तौल र उमेरबिचको प्रतिगमन विश्लेषण नतिजा



चौथो चरणमा प्राप्त नतिजाको व्याख्या गरेर निर्णय गर्नु हो। प्रतिगमन विश्लेषण नतिजाको व्याख्या गर्नुपूर्व प्राप्त नतिजाहरूको तथ्याङ्कीय प्रमाणीकरण आवश्यक हुन्छ। अर्थात् अनुमानहरू तथ्याङ्कीय हिसाबले महत्त्वपूर्ण (Statistically Significant) सावित हुन जरूरी हुन्छ। त्यसका लागि पहिला गुणाङ्कहरू चिनाँ र तिनीहरू अनुमानहरू तथ्याङ्कीय हिसाबले महत्त्वपूर्ण छन् वा छैनन् भन्ने कुराको यकिन गरौं। माथिको तालिका १८.१ मा पुछारका दुई लहरहरूमा β_1 र β_0 का अनुमानहरू दिइएका छन्। उमेरको लहरमा “Coef.” कोलमको गुणाङ्क ०.१५४६१०७ भनेको β_1 हो र “_cons” को लहरमा “Coef.” कोलमको गुणाङ्क ५.८६९०१ भनेको β_0 हो। प्रतिगमन विश्लेषणमा हाम्रो रूचि विशेष गरेर β_1 अर्थात् उमेरको गुणाङ्कमा हुन्छ। अब, यो उमेरको गुणाङ्क साँच्चिकै महत्त्वपूर्ण छ वा ०.१५४६१०७ भन्नु र शून्य भन्नुमा कुनै अन्तर छैन भनेर परीक्षण गर्नुपर्ने हुन्छ। अर्थात्, बच्चाको तौलमा उमेरको प्रभाव महत्त्वपूर्ण हो वा होइन भनेर तथ्याङ्कीय परीक्षण गर्नुपर्दछ। यसका लागि निम्नानुसारको शून्य परिकल्पना र वैकल्पिक परिकल्पना हुन्छन्।

शून्य परिकल्पना: तौलमा उमेरको प्रभाव महत्त्वपूर्ण छैन: $H_0: \beta_1 = 0$

वैकल्पिक परिकल्पना: तौलमा उमेरको प्रभाव महत्त्वपूर्ण छ: $H_1: \beta_1 \neq 0$

यो परिकल्पना परीक्षणका लागि आवश्यक पर्ने पी-मान (P – Value) पनि STATA को नतिजा तालिका १८.१ बाट प्राप्त हुन्छ। यस उदाहरणमा पी-मान शून्यको निकै नजिक (०.०००) रहेको हुनाले जुनसुकै लेभल अफ सिजिनीफक्यान्स (α) मा शून्य परिकल्पना अस्वीकार गर्न पर्याप्त प्रमाण पुग्दछ। अर्थात् अनुमानहरू तथ्याङ्कीय हिसाबले महत्त्वपूर्ण (Statistically Significant) छन्। तसर्थ, हामी तौलमा उमेरको प्रभाव महत्त्वपूर्ण छ भन्ने निर्णयमा पुग्न सक्दछौं। यति भइसकेपछि हामीले अब यो उमेरको गुणाङ्क मान ०.१५४६१०७ लाई यसरी व्याख्या गर्न सक्दछौं।

“नेपालको हिमाली क्षेत्रमा बच्चाको उमेर एक महिनाले बढ्दा उनको तौल ०.१५४६१०७ के.जी. ले बढ्दछ र यो कथन ९९% विश्वस्तताको तहमा तथ्याङ्कीय हिसाबले प्रमाणित छ।”

अब, यस नतिजाबाट हामीले समीकरण (२) लाई देहायबमोजिम लेख्न सक्दछौं।

$$\text{तौल} = ५.८६९०१ + ०.१५४६१०७ \text{ उमेर,}$$

प्रतिगमन विश्लेषण गरिसकेपछि u को मान पनि ५.८६९०१ मा समावेश भएर आएको छ। समीकरण ३ को मद्दतले अब हामीले कुनै पनि उमेरका लागि तौल प्रक्षेपण वा अनुमान

गर्न सकदछौं। तर यस्तो प्रक्षेपण गर्दा अधिल्ला पाठमा भने जस्तै धेरै लामो समयवाधि पछिको लागि प्रक्षेपण वा पूर्वानुमान गर्नु राम्रो मानिँदैन।

यस्ता नतिजाहरू खासगरी स्वास्थ्यकर्मी तथा यो क्षेत्रका नीतिनिर्माताहरूका लागि निकै उपयोगी हुन्छन्। उदाहरणका लागि पाँच वर्ष मुनिका बालबालिकाको वृद्धि कुन दरमा हुनुपर्नेमा कुन दरमा भइरहेको छ भन्ने यो सूचनाले पोषण कार्यक्रमको समीक्षा गर्नुका साथै भावी रणनीति तय गर्नमा ठुलो भूमिका खेलेको हुन्छ।

१८.६ स्पेसिफिकेशन र डाटासम्बन्धी सवाल

माथि उल्लेख गरे बमोजिम सबै चरहरूले सिधा रेखीय सम्बन्ध मात्र बनाउँदैनन्। कहिलेकाहीँ गैर-रेखीय सम्बन्ध राख्ने चरहरूका लागि सोही बमोजिम उपयुक्त प्रतिगमन नमुना (Regression Model) को छनोट गर्नुपर्ने हुन्छ। यस कार्यलाई नमुना स्पेसिफिकेशन (Model Specification) भनिन्छ। प्रतिगमन विश्लेषणका लागि विभिन्न मोडलहरू प्रचलनमा छन्। यस पुस्तकमा प्रतिगमन विश्लेषणको परिचयात्मक जानकारी दिने उद्देश्य भएको हुनाले यस्ता मोडलहरूका बारेमा चर्चा गरिएको छैन। तथापि प्रतिगमन विश्लेषणमा जानुभन्दा पहिला कम्तीमा स्क्वाटर डायग्रामको सहायताले सामान्य रूपमा चरहरूको सम्बन्धको प्रकृति हेर्नु उपयुक्त हुन्छ।

प्रतिगमन विश्लेषणमा एउटा सबैभन्दा महत्त्वपूर्ण पक्ष भनेको प्रयोग गरिने तथ्याङ्कको गुणस्तर हो। नमुनाबाट लिइएका विवरणका आधारमा समग्र जनसंख्याका मानहरू बिचको सम्बन्धका बारेमा धारणा बनाउनुपर्ने भएकोले प्रतिगमन विश्लेषणमा प्रयोग हुने तथ्याङ्क गुणस्तरीय हुन जरूरी छ। तथ्याङ्कको गुणमा हास आउने केही कारणहरूमा मापन त्रुटि, आँकडा छुट हुने, नमुना छनोट गर्दा सम्भाव्यताको अनुसरण नगर्ने जस्ता कुराहरू पर्दछन्। समग्र तथ्याङ्कका गुणहरूका बारेमा हामीले पाठ १९ मा चर्चा गर्नेछौं।

१८.७ हिटेरोस्केडास्टिसिटी

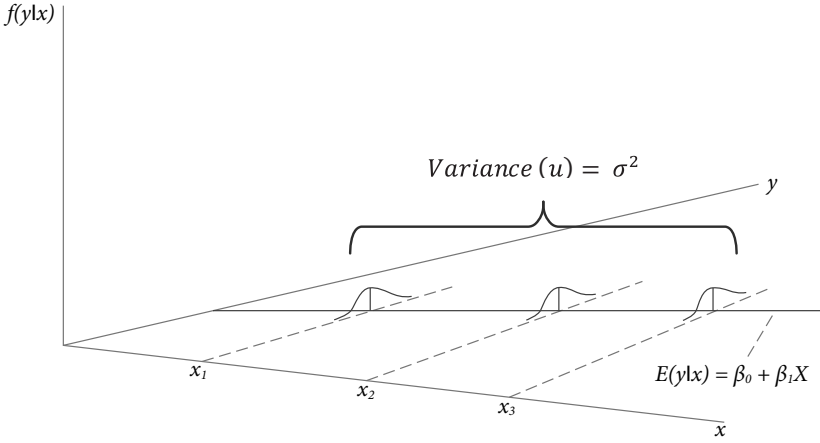
प्रतिगमन विश्लेषणमा बारबार प्रयोग हुने शब्दावली भनेको होमोस्केडास्टिसिटी (Homoskedasticity) र हिटेरोस्केडास्टिसिटी (Heteroskedasticity) पनि हुन्। यी शब्दहरूका बारेमा बुझ्नका लागि माथिको समीकरण (२) लाई हेरौं।

$$\text{तौल} = \beta_0 + \beta_1 \text{उमेर} + u, \quad (२)$$

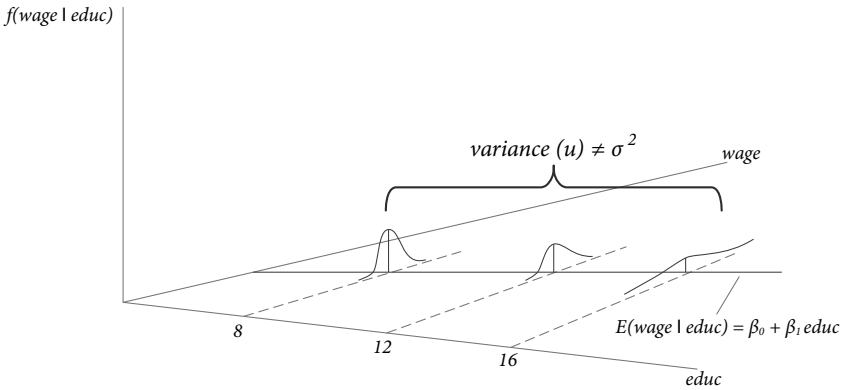
यस समीकरणमा u लाई Error Term or Disturbance Term or Disturbance भन्ने गरिन्छ। u मा खासगरी निर्भर चर तौललाई असर गर्ने उमेर बाहेकका अन्य चरहरू समावेश भएका हुन्छन्। यदि u को भेरियन्स स्थिर छैन भने यस्तो

अवस्थालाई हिटेरोस्केडास्टिसिटी भनिन्छ । यदि u को भेरियन्स स्थिर (Constant) छ भने यस्तो अवस्थालाई होमोस्केडास्टिसिटी भनिन्छ । प्रतिगमन विश्लेषणका लागि होमोस्केडास्टिसिटी आवश्यक शर्त हो किनभने यदि होमोस्केडास्टिसिटी भएन भने अनुमानहरू निष्पक्ष त हुन सक्लान् तर दक्ष (Efficient) हुँदैनन् ।

चित्र नं. १८.५: होमोस्केडास्टिसिटी



चित्र नं. १८.६: हिटेरोस्केडास्टिसिटी



नोट: माथिका चित्र नं. १८.५ र १८.६ Wooldridge, Introductory Econometrics: A Modern Approach, 6e. किताबबाट साभार गरिएको हो ।

१८.८ आर स्क्वायर्ड र गुडनेस अफ फिट

प्रतिगमन विश्लेषणमा आर स्क्वायर्ड (R^2 अथवा r^2) को निकै महत्त्व हुन्छ । यसलाई निर्धारणको गुणाङ्क (Coefficient या Determination) भनिन्छ । दुईवटा चरहरूबिचको सहसम्बन्धको वर्ग नै आर स्क्वायर्ड हो । यसले चरहरूबिच सम्बन्ध देखाउने समीकरणमा निर्भर चरलाई स्वतन्त्र चरले कति प्रतिशत निर्धारण गरेका हुन्छन् भन्ने कुरा बताएको हुन्छ । यस विषयलाई हामी प्रतिगमन विश्लेषण पछि गुडनेस अफ फिट भनेर पनि बुझ्दछौं । उदाहरणका लागि माथिको उदाहरणमा R^2 को मान ०.७१६७ आएको छ (तालिका १८.२) । यसको अर्थ, बच्चाको तौललाई बच्चाको उमेरले ७१.६७% व्याख्या गरेको छ । अर्थात्, बच्चाको तौलमा आउने परिवर्तन ७१.६७% उमेरको कारणले हो, बाँकी २८.३३% अन्य कारणहरूले गर्दा हो भन्ने बुझिन्छ । यस अर्थमा हामीले निर्धारण गरेको प्रतिगमन नमुना स्पेसिफिकेशन पनि ठिकै छ (Goodness of Fit) भन्ने पनि बुझिन्छ ।

तालिका १८.२: बच्चाको तौल र उमेरबिचको प्रतिगमन विश्लेषण नतिजा (माथिको साभार)

Source	SS	df	MS	Number of obs	-	198
Model	1473.2567	1	1473.2567	F(1, 196)	-	495.95
Residual	582.238037	196	2.97060253	Prob > F	-	0.0000
Total	2055.4948	197	10.4339837	R-squared	-	0.7167
				Adj R-squared	-	0.7153
				Root MSE	-	1.7235

Variable	Coeff.	Std. Err.	t	t	[95% Conf. Interval]
उमेर	1546107	.0069426	22.27	0.000	1409189 1683025
_cons	5.86901	.2472084	23.74	0.000	5.38248 6.35654

१८.९ अभ्यासका लागि प्रश्नहरू

क) फरक छुट्याउनुहोस् ।

- १) स्वतन्त्र चर र निर्भर चर
- २) सरल प्रतिगमन नमुना र बहुप्रतिगमन नमुना
- ३) सिधा रेखीय प्रतिगमन र गैर-रेखीय प्रतिगमन
- ४) होमोस्केडास्टिसिटी र हिटेरोस्केडास्टिसिटी

ख) खाली ठाउँ भर्नुहोस् ।

- १) दुईवटा चरहरूबिचको सहसम्बन्धको वर्गलाई भनिन्छ ।
- २) चरहरूको सम्बन्धलाई हेरेर प्रतिगमन नमुना तय गर्ने कार्यलाई भनिन्छ ।
- ३) चरहरूबिचको सम्बन्धलाई सिधा रेखाद्वारा व्याख्या गर्न सकिन्छ भने त्यस्तो

प्रतिगमनलाई भनिन्छ ।

४) नमुनाका मानहरूबाट दुई चरहरूबिचको सम्बन्ध स्थापित गर्ने गुणाङ्कहरूको अनुमान निकालेर गरिने अध्ययनलाई भनिन्छ ।

ग) प्रतिगमन विश्लेषण भनेको के हो र यो किन महत्त्वपूर्ण छ ?

घ) तलको तालिका १८.३ मा नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ को तथ्याङ्क प्रयोग गरी प्रतिव्यक्ति वार्षिक आमदानी र परिवारको सदस्य संख्याबिचको प्रतिगमन विश्लेषणको नतिजा प्रस्तुत गरिएको छ । यस तालिकाको आधारमा सोधिएका प्रश्नहरूको जवाफ दिनुहोस् ।

तालिका १८.३: प्रतिव्यक्ति वार्षिक आमदानी र परिवारको सदस्य संख्याबिचको प्रतिगमन विश्लेषणको नतिजा

Source	SS	df	MS	Number of obs	F(1, 5906)	Prob > F
Model	3.0940e+12	1	3.0940e+12			0.0663
Residual	5.4905e+15	5,986	9.1723e+11			0.0006
Total	5.4936e+15	5,987	9.1759e+11			0.0004
				Root MSE		9.6e+05

Entering pc	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
hhszpc	-9852.3	5364.29	-1.84	0.066	-20368.24 653.6425
cons	153856.5	28352.12	5.43	0.000	96276.11 209436.9

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ को तथ्याङ्क प्रयोग गरी निकालिएको प्रतिगमन विश्लेषणको नतिजा

यहाँ, totinc_{pc} भनेको प्रतिव्यक्ति वार्षिक आमदानी हो र hhszpc भनेको परिवार सदस्य संख्या हो ।

प्रश्नहरू:

- १) $\widehat{\beta}_0$, $\widehat{\beta}_1$ र R^2 का अनुमानहरू कति कति छन् ?
- २) प्रस्तुत प्रतिगमन विश्लेषणको अनुमान गरिएको सरल सिधा रेखीय प्रतिगमन समीकरण लेख्नुहोस् ।
- ३) निम्नलिखित परिकल्पना पी-मानको प्रयोग गरी परीक्षण गर्नुहोस् ।
 शून्य परिकल्पना: प्रतिव्यक्ति वार्षिक आमदानीमा परिवार सदस्य संख्याको प्रभाव महत्त्वपूर्ण छैन: $H_1: \widehat{\beta}_1 = 0$
 वैकल्पिक परिकल्पना: प्रतिव्यक्ति वार्षिक आमदानीमा परिवार सदस्य संख्याको प्रभाव महत्त्वपूर्ण छ: $H_1: \widehat{\beta}_1 \neq 0$

४) प्रतिगमन विश्लेषणको प्राप्त नतिजालाई व्याख्या गर्नुहोस् ।

१८.१० सन्दर्भ सामग्री

प्रतिगमन विश्लेषण थप अध्ययन गर्नका लागि देहायको सामग्री उपयुक्त छ ।

१. Introductory Econometrics: A Modern Approach by
Jeffrey M. Wooldridge, Fifth Edition, South-Western,
Cengage Learning, 2013



भाग 8

**तथ्याङ्कीय क्रियाकलापमा
थप विषयको अध्ययन**

तथ्याङ्कका गुण

१९.१ परिचय

हामीले तथ्याङ्कको परिचयदेखि यसका महत्त्व र प्रयोगका बारेमा अधिल्ला पाठहरूमा आधारभूत जानकारी हासिल गरिसकेका छौं । तथ्याङ्क समाजको ऐना हो र यसले समाजका आर्थिक, सामाजिक, वातावरणीयलगायत अन्य घटना वा विशेषताका बारेमा सूचनाको माध्यमबाट जानकारी दिन्छ । कहिलेकाहीँ पत्रपत्रिकामा र कहिलेकाहीँ चाहिँ सामान्यजनले समेत प्रकाशमा आएका तथ्याङ्क वास्तविक भएन वा मिथ्याङ्क भयो भनेर टिप्पणी गर्ने गरेको हामीले पढ्दै र सुन्दै आएका छौं । यस पाठमा हामी के कस्ता गुण वा विशेषताहरू विद्यमान रहेको अवस्थामा तथ्याङ्क सही मानेमा तथ्याङ्क हुन्छ भनेर चर्चा गर्नेछौं । तथ्याङ्कका गुणहरू विश्वव्यापी रूपमै समान प्रकृतिका हुन्छन् किनभने ती निश्चित आधारहरूमा परस्पर तुलनीय हुनेपछि । तथ्याङ्कले अंकमा यथार्थको प्रकटीकरण गर्ने हुँदा यसले जुन विषय वा क्षेत्रको अध्ययन र अनुसन्धान गर्ने उद्देश्य राखेको हो सोको अधिकतम शुद्ध र विश्वसनीय अंकचित्र प्रस्तुत गर्नेपछि ।

यसका अतिरिक्त राष्ट्रिय तथा अन्तर्राष्ट्रिय विकास लक्ष्यहरू अनुगमन गर्नका लागि पनि तथ्याङ्ककै प्रयोग हुन्छ । यस्तो बृहत् महत्त्व बोकेका औजार गुणस्तरीय हुनु अनिवार्य हुन्छ । गुणस्तरहीन सूचनाले कहिल्यै पनि गन्तव्यमा पुऱ्याउँदैन । तथ्याङ्क संकलनदेखि यसको वितरण वा प्रयोग गर्ने बेलासम्म गुणस्तर कायम गरी सही तथ्याङ्क उत्पादन गर्नु र सही तथ्याङ्कको प्रयोग गर्नु नै तथ्याङ्क व्यवस्थापनको एउटा महत्त्वपूर्ण पाटो हो ।

तसर्थ, तथ्याङ्कमा हुनुपर्ने न्यूनतम गुणहरूका बारेमा यस पाठमा हामीले संक्षिप्त चर्चा गर्नेछौं ।

१९.२ तथ्याङ्क गुणका आयाम

तथ्याङ्कका गुणहरूका बारेमा विभिन्न विद्वानहरूले विभिन्न तरिकाले व्याख्या एवं विश्लेषण गरेको पाइन्छ । गुणस्तरीय तथ्याङ्कमा हुनुपर्ने गुणहरूका बारेमा समग्रमा सबै जनाको एकमत नै छ । यस पुस्तकमा तथ्याङ्कमा हुनुपर्ने न्यूनतम गुणहरूलाई ७ वटा आयामहरूमा विभाजन गरी वर्णन गरिएको छ ।

१) **वैधता वा शुद्धता (Validity or Accuracy):** तथ्याङ्कले जे मापन गर्न खोजेको हो, त्यही मापन गऱ्यो भने त्यसलाई शुद्ध तथ्याङ्क मानिन्छ । यस किसिमका तथ्याङ्कले भ्रमांश (Error) न्यूनीकरण गर्दछन् । तथ्याङ्क संकलनमा खटिने गणकहरूको पूर्वाग्रहबाट हुनसक्ने भ्रमांश, अभिलेखन कार्यमा संलग्न व्यक्तिबाट हुनसक्ने पूर्वाग्रह, तथ्याङ्क

स्थानान्तरण (कागजबाट कम्प्युटरमा सार्ने) गर्दा हुनसक्ने त्रुटि, नमुना छनोटसँग सम्बन्धित भ्रमांश (जस्तै: गलत नमुना छनोट गर्ने, नमुना समूहका सिमानामा रहेका एकाइहरू छुट्टने वा बाहिरको समेटिने आदि) जस्ता भ्रमांशहरूको यथासम्भव न्यून गरिएका तथ्याङ्कलाई शुद्ध तथ्याङ्क भनिन्छ । संक्षेपमा भन्नुपर्दा वस्तु वा घटनासँग सम्बन्धित विशेषताहरूको सही मापन गरी प्रस्तुत गरिएका तथ्याङ्क नै शुद्ध तथ्याङ्क हुन् । उदाहरणका लागि जनगणनामा परिवारका सदस्यहरूमध्ये परिवारमा अक्सर बसोबास गर्ने व्यक्तिको गणना गर्ने भनिएको हुन्छ । यदि परिवारका अनुपस्थित सदस्यको समेत गणना गरियो वा परिवारमा अक्सर बसोबास गर्ने व्यक्ति पनि गणना गरिएन भने त्यस्ता तथ्याङ्कलाई शुद्ध मान्न सकिँदैन ।

२) विश्वसनीयता वा सुसंगतता तथा अन्विति (Reliability or Consistency and Coherence):

तथ्याङ्क उत्पादन कार्यमा व्यक्ति परिवर्तन हुँदा मापन गर्न खोजिएको मानमा परिवर्तन आउँदैन भने त्यस्ता तथ्याङ्कलाई विश्वसनीय वा सुसंगत तथ्याङ्क भनिन्छ । तथ्याङ्क उत्पादनमा प्रयोग हुने औजार वा मेसिन परिवर्तन हुँदा तथ्याङ्कका मानमा परिवर्तन हुँदैन । यस्ता तथ्याङ्कहरू खास प्रकारका अवधारणा, परिभाषा, विधि तथा अभिसन्धिले निर्देशित हुन्छन् र जो जसले उत्पादन गरे पनि उही मान प्रदान गर्दछन् । उदाहरणका लागि परिवारमा अक्सर बसोबास गर्ने व्यक्तिहरूको गणना गर्न कागजी प्रश्नावली होस् वा ट्याबलेटको प्रयोग होस् कुनै परिवारको सदस्य संख्या एउटै आउँछ । त्यस्तै अक्सर बसोबासलाई स्पष्ट परिभाषित गरिएको हुनाले कुनै परिवारमा गणक वा सुपरिवेक्षक वा अधिकृत जो गएर तथ्याङ्क संकलन गर्दा पनि एउटै मान आउनुपर्दछ । तथ्याङ्कमा अन्विति (Coherence) मिलेको हुनुपर्दछ । एउटै परिभाषाभित्र पर्ने गरी दुई भिन्न किसिमका तथ्याङ्क हुनुहुँदैन भने दुई किसिमका भिन्न तथ्याङ्कलाई एउटै परिभाषाले समेट्नु पनि हुँदैन । उदाहरणका लागि परिवार भनेको साधारणतया एउटै घरका बा, आमा, छोरा, छोरी, बुहारी, नाति, नातिनालगायत रगत तथा वैवाहिक सम्बन्धका नाता पर्ने सदस्यहरूको समूह बुझिन्छ । तर तथ्याङ्कीय प्रयोजनमा परिवार भनेको एउटै भान्सामा खाने र आमदानी खर्चको साभेदारी गर्ने नाता पर्ने वा नपर्ने सदस्यहरू भन्ने बुझिन्छ । यसरी परिवार शब्दलाई फरक-फरक तरिकाले बुझिँदा द्विविधा उत्पन्न हुने हुनाले तथ्याङ्कमा प्रयोग भएको परिवारको परिभाषा स्पष्ट रूपले उल्लेख गरी सोही बमोजिम तथ्याङ्क संकलन वा व्यवस्थापन गर्नुपर्दछ ।

३) पूर्णता (Completeness):

तोकिएको क्षेत्र तथा दायरा समेटिएको तथ्याङ्कलाई पूर्ण तथ्याङ्क भनिन्छ । जस्तै: जनगणनामा नेपालभित्र बस्ने सबै व्यक्तिको गणना गर्ने भनिएको छ । तसर्थ, नेपालमा अक्सर बसोबास गर्ने नेपाली नागरिक वा विदेशी नागरिकको

सबैको गणना गरियो भने मात्र पूर्णता हुन्छ। पूर्णतायुक्त तथ्याङ्कमा कुनै स्थान वा आवश्यक विशेषताहरू नछोडी तथ्याङ्क संकलन गरिएको हुन्छ। उदाहरणका लागि जनगणनाबाट पुरुष र महिलाको साक्षरता निकाल्ने अपेक्षा गरिएकोमा दुवैको साक्षरता आउने गरी तथ्याङ्क संकलन गरिएको हुन्छ र यस्तो तथ्याङ्कलाई पूर्णतायुक्त तथ्याङ्क भनिन्छ।

४) **समयबद्धता (Timeliness)**: तथ्याङ्कमा समयको निकै ठुलो महत्त्व हुन्छ। जुन समयको लागि तथ्याङ्क अपेक्षा गरिएको हुन्छ, त्यही समयमा तथ्याङ्क उपलब्ध हुनुपर्दछ। उदाहरणका लागि सरकारलाई कोभिड १९ खोप लगाउने योजना गर्नका लागि कति जनसंख्यालाई खोप दिनुपर्नेछ भन्ने तथ्याङ्क आवश्यक परेको बेला पुरानो जनगणनाको तथ्याङ्कले काम गर्दैन। यस्तो अवस्थामा जनसंख्याको प्रक्षेपण गरेर होस् वा पुनः तथ्याङ्क संकलन गरेर होस् चालु अवधिको अद्यावधिक जनसंख्या नै प्रयोग गर्नुपर्ने हुन्छ। अर्को उदाहरणका लागि सरकारले वि.सं. २०७८ कार्तिक र मंसिरमा राष्ट्रिय जनगणना सञ्चालन गरेको थियो। यदि यो तथ्याङ्क अपेक्षा गरे अनुरूप एक वर्ष भित्रमा नतिजा निकालियो भने मात्र समयबद्धता कायम भएको मानिन्छ। तथ्याङ्क सार्वजनिक गर्न जति विलम्ब भयो त्यति त्यसको महत्त्व र उपयोगिता कम हुँदै जान्छ। अथवा, ढिलो वितरण गरिने तथ्याङ्कको गुणस्तर पूर्ण नभएको मानिन्छ।

५) **इमान्दारी (Integrity)**: राजनीतिक वा व्यक्तिगत कारणले जानाजानी पूर्वाग्रह ढङ्गले हेरफेर हुनबाट सुरक्षित हुनेगरी विकास गरिएको तथ्याङ्क उत्पादन गर्ने प्रणालीबाट उत्पादित तथ्याङ्कलाई इमान्दारी भएको तथ्याङ्क भनिन्छ। यस किसिमको तथ्याङ्क प्रणालीमा कसैको इच्छाले तथ्याङ्कलाई तोडमोड गर्ने अवस्था प्राप्त हुँदैन।

६) **गोपनीयता (Confidentiality)**: व्यक्तिगत विवरणहरू तथ्याङ्कीय प्रयोजन बाहेक सार्वजनिक नहुने गरी व्यवस्थित गरिएको तथ्याङ्कलाई गोपनीयता कायम भएको तथ्याङ्क भनिन्छ। उदाहरणको लागि नेपाल श्रमशक्ति सर्वेक्षण, नेपाल जीवनस्तर सर्वेक्षण, नेपाल बहुसूचक सर्वेक्षण, नेपाल जनसांख्यिक स्वास्थ्य सर्वेक्षण, जनगणना, आर्थिक गणना लगायतका नमुना सर्वेक्षण तथा गणनाहरूबाट उत्पादन हुने तथ्याङ्कमा कुनै पनि व्यक्तिको व्यक्तिगत विवरण खुल्ने गरी तथ्याङ्क वितरण वा प्रकाशन गरिएको हुँदैन।

७) **सान्दर्भिकता (Relevance)**: तथ्याङ्क प्रयोगकर्ताको आवश्यकता परिपूर्ति गर्न सक्षम तथ्याङ्क नै सान्दर्भिकता भएको तथ्याङ्क हो। यसमा तथ्याङ्कको दायरा तथा विषय पूर्ण रूपमा समेटिएको हुन्छ। अध्येतालाई आवश्यक पर्ने सबै क्षेत्रको विवरण उपलब्ध गराउने तथ्याङ्कलाई सान्दर्भिकता भएको तथ्याङ्क भनिन्छ। उदाहरणको लागि नेपालको गरिबी मापन गर्नका लागि नेपाल जीवनस्तर सर्वेक्षणबाट प्राप्त तथ्याङ्क सान्दर्भिक छ। तर

यही तथ्याङ्क देशको वस्तु तथा सेवाको आयात र निर्यातसम्बन्धी अध्ययन गर्नका लागि सान्दर्भिक हुँदैन ।

१९.३ तथ्याङ्कको उपलब्धता र पहुँच

तथ्याङ्क सूचनाको स्रोत भएको हुनाले सबैको चासो र आवश्यकताको विषय हो । यसले समाजका वास्तविक घटनाहरूलाई पारदर्शी तरिकाले प्रस्तुत गरेको हुन्छ । तसर्थ, तथ्याङ्कलाई सार्वजनिक वस्तु (Public Goods) को रूपमा लिइनुपर्दछ र सोही बमोजिम उपलब्धताको सुनिश्चित गर्नुपर्दछ । यसका अतिरिक्त तथ्याङ्क सबैको पहुँचमा हुनुपर्दछ । उपलब्ध तथ्याङ्क पनि सबैको पहुँचमा भएन भने त्यसको सार्थकता हुँदैन । कतिपय मानिसहरूले तथ्याङ्कको उपलब्धता र पहुँचलाई तथ्याङ्कको गुणस्तरकै एउटा भिन्न आयामको रूपमा पनि लिने गर्दछन् । समकालीन समयमा खुला तथ्याङ्क (Open Data) को अवधारणालाई विश्वव्यापी रूपमा अंगिकार गरिँदैछ । यस सन्दर्भमा खुला तथ्याङ्कले सार्वजनिक तथ्याङ्कको सहज र सरल पहुँचको अपेक्षा गरेको हुन्छ । तथ्याङ्कको पहुँचको सन्दर्भमा पनि मेसिनले पढ्ने (Computer Readable) र तथ्याङ्कका प्रयोगकर्ताले चाहेको बखत आफ्नो आवश्यकताबमोजिम विश्लेषण गर्न सक्ने गरी उपलब्ध गराइएको र सबैको पहुँचमा राखिएको तथ्याङ्क भन्ने बुझिन्छ । तसर्थ, विद्युतीय स्वरूपमा सबैले उपयोग गर्न सक्ने गरी राखिएको तथ्याङ्कलाई उपलब्धता र पहुँचयोग्य तथ्याङ्क मानिन्छ । उदाहरणका लागि नेपाल राष्ट्र बैंकको वेबसाइटमा उपलब्ध चालु समीष्टगत आर्थिक तथा वित्तीय अवस्था तालिका (Current Macroeconomic and Financial Situation – Tables) लाई उपलब्धता र पहुँच सुनिश्चित गरिएको तथ्याङ्क मान्न सकिन्छ ।¹ त्यसैगरी केन्द्रीय तथ्याङ्क विभागको राष्ट्रिय तथ्याङ्क प्रोफाइल (National Data Profile) मा उपलब्ध तथ्याङ्कहरू पनि उपलब्धता र पहुँच सुनिश्चित गरिएका तथ्याङ्क हुन् ।²

तथ्याङ्क उपलब्ध र पहुँचयोग्य भएर मात्र पुग्दैन । यसलाई प्रयोगकर्ताले व्याख्या गर्ने क्षमता राख्न सक्ने किसिमको हुनुपर्दछ । यसका लागि तथ्याङ्कको पनि सूचना वा विवरण दिने तथ्याङ्क उपलब्ध गराउनु पर्दछ । तथ्याङ्कको पनि तथ्याङ्कलाई मेटाडाटा (Metadata) भनिन्छ । तसर्थ, सार्वजनिक गरिने तथ्याङ्कसँगै मेटाडाटा पनि उपलब्ध गराइएको हुनुपर्दछ जसले गर्दा सम्बन्धित प्रयोगकर्ताले उपलब्ध तथ्याङ्कको सही व्याख्या र प्रयोग गर्न

1 Nepal Rastra Bank. Retrieved on December 8, 2021 from: <https://www.nrb.org.np/>

2 Central Bureau of Statistics. National Data Profile. Retrieved on December 8, 2021 from: <http://nationaldata.gov.np/>

सकून् । मेटाडाटाअन्तर्गत तथ्याङ्कको स्रोत, तथ्याङ्क संकलनको विधि, तथ्याङ्क संकलन मिति, तथ्याङ्क मापनका एकाइ, सन्दर्भ समय, तथ्याङ्कले सूचित गर्न खोजेको विषयको जानकारी लगायतका विषयहरूको जानकारी पर्दछन् ।

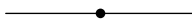
१९.४ अभ्यासका लागि प्रश्नहरू

- क) गुणस्तरीय तथ्याङ्कका विशेषताहरू उल्लेख गर्नुहोस् ।
 ख) तथ्याङ्कको उपलब्धता र पहुँच सुनिश्चित गर्न के के कुरामा ध्यान दिनुपर्दछ ?
 ग) मेटाडाटा भन्नाले के बुझिन्छ ?

१९.५ सन्दर्भ सामग्री

तथ्याङ्कका गुणहरूका सन्दर्भमा थप अध्ययन गर्नका लागि निम्न सामग्रीहरू उपयोगी छन् ।

१. गैह्रे, रामहरि (२०१७) । नेपालको राष्ट्रिय तथ्याङ्क प्रणालीमा तथ्याङ्कका गुणहरूको मापन । केन्द्रीय तथ्याङ्क विभाग, *नेपालको राष्ट्रिय तथ्याङ्क प्रणाली, एक संग्रह (अंग्रेजी संस्करण)* (पृष्ठ २७५-२८६) । काठमाण्डौं, नेपाल: केन्द्रीय तथ्याङ्क विभाग ।
२. The Global Fund, USAID, Measure Evaluation (2008). Data Quality Audit Tool, Guidelines for Implementation. Retrieved on December 8, 2021 from: <https://www.measureevaluation.org/resources/publications/ms-08-29.html>
३. Organisation for Economic Development Co-operation and Development (OECD). (2011). Quality Framework and Guidelines for OECD Statistical Activities. Retrieved on December 8, 2021 from: <https://www.oecd.org/sdd/21687665.pdf>



प्रायोगिक परियोजना सञ्चालनका प्रमुख चरणहरू
तथा तथ्याङ्कमा गोपनीयता र नैतिकता

२०.१ परिचय

प्रायजसो प्रायोगिक परियोजना (Empirical Project) सञ्चालन गर्दा तथ्याङ्कीय क्रियाकलाप सञ्चालन गर्नुपर्ने हुन्छ । अथवा तथ्याङ्क व्यवस्थापन कार्य प्रायोगिक परियोजनामा अपरिहार्य हुन्छ । हामीले यसभन्दा अगाडिका पाठहरूमा तथ्याङ्क व्यवस्थापनका आधारभूत अवधारणाका बारेमा जानकारी लिइसकेका छौं । अब यसको वास्तविक प्रयोगको खाका प्रस्तुत गर्ने क्रममा यसको धेरै प्रयोग हुने क्षेत्र प्रायोगिक परियोजना सञ्चालनका प्रमुख चरण बारेमा चर्चा गर्दा पाठक वर्गलाई तथ्याङ्कीय साक्षरताको सारभूत अर्थ स्पष्ट हुन जान्छ । प्रायोगिक परियोजना सञ्चालन गर्नका लागि विभिन्न चरणहरू पार गर्नुपर्ने हुन्छ । यसका अतिरिक्त शैक्षिक शोध कार्य तथा अन्य अनुसन्धानात्मक कार्यहरूमा पनि निश्चित चरण पूरा गर्नुपर्ने हुन्छ । सामान्यतया सबै किसिमका शोध तथा अनुसन्धानात्मक कार्यका लागि उपयोगी हुने आधारभूत चरणहरूका बारेमा सामान्य जानकारी गराउनु यस पाठको उद्देश्य रहेको छ । यस कार्यका लागि तय गर्नुपर्ने विभिन्न चरणहरूमध्ये यस पाठमा हामीले प्रायोगिक परियोजनाका प्रमुख ६ चरण बारेमा संक्षिप्त छलफल गर्दैछौं । यसका अतिरिक्त तथ्याङ्कमा गोपनीयता र नैतिकताको सवाल पनि निकै महत्त्वपूर्ण विषय भएको हुनाले यस सन्दर्भमा पनि संक्षिप्त चर्चा गर्नेछौं ।

२०.२ प्रायोगिक परियोजना सञ्चालनका प्रमुख चरणहरू

परियोजनाका विभिन्न उद्देश्य हुन्छन् । यहाँ छलफल गर्न खोजिएको परियोजना प्रायोगिक परियोजना शैक्षिक क्षेत्रका उद्देश्यहरूले अध्ययन अनुसन्धानमार्फत् नीति पृष्ठपोषण गर्न परियोजना सञ्चालन गरेका हुन्छन् भने पत्रकार एवं सञ्चार माध्यमहरूले सूचना उत्पादन गर्नका लागि परियोजना सञ्चालन गरेका हुन्छन् । नीतिनिर्माता तथा विकास साभेदारहरूले विकास कार्यक्रमका लागि नीति निर्माणदेखि यसको कार्यान्वयन तथा उपलब्धि अनुगमन गर्नका लागि प्रायोगिक परियोजना सञ्चालन गर्दछन् । यस्ता किसिमका परियोजना सञ्चालन गर्दा वैज्ञानिक विधिको अनुसरण गर्नुका साथै निश्चित केही चरणहरू पार गरेको हुनुपर्दछ । यस्ता चरणहरूमध्ये प्रमुख रूपमा निम्न चरणहरू महत्त्वपूर्ण छन् ।

- १) परियोजना अध्ययनका लागि विषय चयन गर्ने,
- २) सान्दर्भिक साहित्यको समीक्षा गर्ने,
- ३) अध्ययनको विधि र पद्धति निर्धारण गर्ने,
- ४) तथ्याङ्क संकलन गर्ने,

- ५) तथ्याङ्क विश्लेषण गरी, निष्कर्ष निकाल्ने,
६) प्रायोगिक लेख तयार गर्ने ।

१) **विषय चयन (Posing a Question):** प्रायोगिक परियोजनाका विषय चयन गर्दा सञ्चालनकर्ताको उद्देश्य र उनीहरूको क्षेत्र प्रमुख कारक मानिन्छन् । जस्तै: अध्येताहरूले प्रायोगिक परियोजनाको विषय छनोट गर्दा उनीहरूको रूचिको विषय प्राथमिकतामा पर्दछन् भने सञ्चारकर्मीका लागि समकालीन समयमा अधिक माग भएको वा अधिकरूपमा आवश्यक परेका सूचना उत्पादन हुने गरी विषय छनोट गर्नुपर्ने हुन्छ । नीतिनिर्माताको हकमा भने प्रायोगिक परियोजना विषय चयन गर्दा विभिन्न कुराहरूलाई ध्यान दिनुपर्ने हुन्छ । उदाहरणका लागि आवधिक योजनाले निर्दिष्ट गरेका विषय, जनतालाई तत्काल आवश्यक परेका विषय आदि । विषय चयन गर्दा ध्यान दिनुपर्ने प्रमुख कुरा भनेको त्यो विषयमा नयाँपन हुनुपर्दछ र विशिष्टीकृत हुनुपर्दछ । अरूले सञ्चालन गरेर नतिजा निकालिसकेका विषय छनोट गर्नु हुँदैन । यसले साधन र स्रोतको सदुपयोग गराउँदैन । त्यस्तै, सामान्यीकरण गरेका विषय पनि छनोट गर्नुहुँदैन । जस्तै: नेपाल विकासोन्मुख देश (Developing Country) मा प्रवेश गर्नका लागि योगदान पुऱ्याउने क्षेत्रहरू के के हुन् भन्थो भने साह्रै सामान्यीकरण गरेको विषय हुन्छ तर यसैलाई नेपालको आर्थिक विकासमा वैदेशिक लगानीको योगदान कस्तो छ भन्थो भने विशिष्टीकृत विषय हुन्छ । यसका अतिरिक्त परियोजनाका लागि छानिने विषय सान्दर्भिक पनि हुनुपर्दछ । कम महत्त्वका विषयमा भन्दा प्राथमिकता प्राप्त र अत्यावश्यक क्षेत्रका विषयमा प्रायोगिक परियोजना सञ्चालन गर्नुपर्दछ ।

२) **साहित्य समीक्षा (Literature Review):** प्रायोगिक परियोजना सञ्चालन गर्दा सम्बन्धित विषयवस्तुमा अन्य देश विदेशका विज्ञ वा अध्येताहरूको ठम्याइ के रहेको छ भन्ने सन्दर्भमा उपलब्ध सबै साहित्यको गहन समीक्षा गर्नुपर्दछ । यस्ता समीक्षाले परियोजना सञ्चालनकर्तालाई अध्ययनको दायरा वा परिधि निर्धारण गर्न, अध्ययनलाई नयाँपन दिन तथा अध्ययनको विधि तय गर्न सहयोग पुऱ्याउँदछ ।

३) **अध्ययनको विधि र पद्धति (Research Method and Methodology):** परियोजनामा विभिन्न विधि (Method) को प्रयोग गर्न सकिन्छ । यसमा अनुसन्धानका विभिन्न प्रविधि जस्तै: प्रायोगिक विधि (Empirical Method), प्रयोगात्मक विधि (Experimental Method), सर्वेक्षण विधि (Survey Method), परीक्षण विधि (Test Method), अवलोकन विधि (Observation Method) आदि समावेश भएका हुन्छन् । यस्ता विधिहरूमध्ये सम्बन्धित परियोजनाका लागि उत्तम विधि कुन हुनसक्दछ भनेर विधिहरूको विश्लेषण गर्ने कार्यलाई पद्धति (Methodology) भनिन्छ ।

विधि र पद्धति पारदर्शी भएमा मात्र परियोजनाको नतिजा विश्वसनीय र सर्वस्वीकार्य हुन्छ । तसर्थ, कुनै पनि परियोजनामा यस्ता विधि र पद्धतिका बारेमा स्पष्टसँग उल्लेख गर्नुपर्दछ । उदाहरणका लागि नेपालको आर्थिक विकासमा वैदेशिक लगानीको योगदान कस्तो छ भन्ने परियोजना विषयका लागि प्रायोगिक विधि (Empirical Method) उत्तम हुन सक्दछ किनभने यस विधिमा समयशृङ्खला तथ्याङ्कहरूको प्रयोग गरेर उपयुक्त प्रतिगमन नमुना (Regression Model) विश्लेषणमार्फत् नेपालको आर्थिक विकास (निर्भर चर) मा वैदेशिक लगानी (स्वतन्त्र चर) को योगदानको आँकलन गर्न सकिन्छ ।

४) **तथ्याङ्क संकलन (Data Collection):** प्रायोगिक परियोजनाका लागि तथ्याङ्क विश्लेषण अपरिहार्य हुन्छ जसका लागि उपयुक्त तथ्याङ्कको आवश्यकता हुन्छ । तथ्याङ्क प्राप्तिका लागि यसका विभिन्न स्रोतहरू हुन्छन् । त्यस्ता स्रोतहरूबाट उपयुक्त र सान्दर्भिक तथ्याङ्क मात्र संकलन गर्नुपर्दछ । कतिपय प्रायोगिक परियोजनामा माध्यमिक तथ्याङ्क प्रयोग गर्न सकिन्छ । माध्यमिक तथ्याङ्क प्रयोग गर्नका लागि सम्बन्धित स्रोतको आधिकारिकता, विश्वसनीयता र उपलब्ध तथ्याङ्कको सान्दर्भिकताका बारेमा सचेत हुनुपर्दछ । यसका अतिरिक्त हामीले पाठ १९ मा छलफल गरेबमोजिम तथ्याङ्कका गुणहरू भएको हुनुपर्दछ । प्राथमिक तथ्याङ्कको हकमा पनि परियोजना सञ्चालनकर्ताले संकलन गरिने तथ्याङ्कको गुणस्तर सुनिश्चित गर्नुपर्दछ ।

५) **तथ्याङ्क विश्लेषण र निष्कर्ष (Data Analysis and Decision Making):** संकलित तथ्याङ्कहरूको उचित तरिकाले विश्लेषण गरेर सूचनाहरू निर्माण गर्नुपर्दछ । तथ्याङ्क विश्लेषणपूर्व तथ्याङ्कको प्रशोधन गर्न भुल्नु हुँदैन । प्रशोधन अन्तर्गत तथ्याङ्कको सम्पादन, त्रुटि निरूपण, छुटेका तथ्याङ्क (Missing Data) तथा आउटलायर (Outliers) आदिको उपचार जस्ता कार्यहरू पर्दछन् । यसका अतिरिक्त तथ्याङ्क विश्लेषण गर्नुपूर्व तथ्याङ्क सारांश जस्तै मध्यक, स्तरीय भिन्नता, न्यूनतम र अधिकतम मान आदिको अध्ययन गर्नुपर्दछ । तथ्याङ्क विश्लेषणका विभिन्न विधिहरूका बारेमा हामीले यस अधिका पाठहरूमा छलफल गरिसकेका छौं । उदाहरणका लागि अनुमान गर्ने, परिकल्पना परीक्षण गर्ने, प्रतिगमन विश्लेषण गर्ने जस्ता कार्यहरू तथ्याङ्क विश्लेषण कार्य अन्तर्गत पर्दछन् । त्यस्ता विधिहरूको प्रयोग गरेर सान्दर्भिक सूचनाको प्रयोगमार्फत् अर्थपूर्ण र प्रमाणमा आधारित निष्कर्ष निकाल्नुपर्दछ । उदाहरणका लागि यदि प्रतिगमन विश्लेषणको नतिजाले नेपालको आर्थिक विकास (निर्भर चर) मा वैदेशिक लगानी (स्वतन्त्र चर) को योगदान महत्वपूर्ण छ भन्ने प्रमाण पुऱ्यायो भने हामीले वैदेशिक लगानी नेपालको विकासको एउटा महत्वपूर्ण अवयव (Component) हो भनेर भन्न सक्दछौं ।

६) **अनुसन्धानात्मक लेख (Research Paper):** प्रायोगिक परियोजनाको अन्तिम चरण भनेको प्रायोगिक लेख तयार गर्नु हो । प्रायोगिक परियोजना सञ्चालन गरिसकेपछि प्रमाणमा आधारित रहेर स्तरीय लेख तयार गर्नुपर्दछ । सामान्यतया लेखमा निम्नानुसारका विषयवस्तु समावेश गर्नुपर्दछ ।

क) **परिचय:** यस खण्डमा परियोजनाको उद्देश्य, महत्त्व, साहित्य समीक्षाको सारसंक्षेप र संक्षिप्त नतिजा उल्लेख गर्नुपर्दछ ।

ख) **सैद्धान्तिक रूपरेखा (Theoretical Framework):** प्रायोगिक परियोजना कुनै सिद्धान्त परीक्षण गर्न वा सैद्धान्तिक पृष्ठभूमिमा रहेर सञ्चालन गरिएको भएमा लेखको यस खण्डमा खुलाउनुपर्दछ । उदाहरणका लागि यदि मूल्य र माग सम्बन्धमा अध्ययन गरिएको हो भने मागको नियम (Law of Demand) का बारेमा यस खण्डमा उल्लेख गर्नुपर्दछ ।³

ग) **तथ्याङ्क विश्लेषणको नमुना र अनुमानको विधि:** तथ्याङ्क विश्लेषणका लागि प्रयोग गरिएका नमुनाहरू जस्तै अनुमान नमुना (Estimation Model), प्रतिगमन नमुना (Regression Model) आदिका बारेमा यस खण्डमा उल्लेख गर्नुपर्दछ । साथै, र तिनीहरूका सूत्रहरू पनि यस खण्डमा बुझिने गरी उल्लेख गर्नुपर्दछ ।

घ) **तथ्याङ्क:** यस खण्डमा प्रायोगिक परियोजनामा प्रयोग गरिएका तथ्याङ्कको बारेमा संक्षिप्त परिचय दिनुपर्दछ । यस अन्तर्गत तथ्याङ्कका स्रोत, तथ्याङ्क प्राप्तिका विधि, मापनका एकाइ, शुद्धताको स्तर, आकार, सीमाजस्ता कुराहरू पर्दछन् । साथै अध्ययनमा प्रयोग हुने चरका सामान्य विशेषता पनि उल्लेख गर्नुपर्दछ ।

ङ) **नतिजा:** तथ्याङ्क विश्लेषणबाट प्राप्त नतिजाहरूलाई यस खण्डमा प्रस्तुत गर्नुपर्दछ । तथ्याङ्क प्रस्तुत गर्ने विधिका बारेमा हामीले पाठ ७ र ८ मा विस्तृत रूपमा छलफल गरिसकेका छौं । तसर्थ, तथ्याङ्कको प्रकृति हेरेर पाठकवर्गलाई आकर्षण गर्ने गरी तथ्याङ्क प्रस्तुत गर्दै प्राप्त नतिजाका बारेमा यस खण्डमा उल्लेख गर्नुपर्दछ । यसरी प्राप्त भएका तथ्याङ्कीय नतिजाहरू तथ्याङ्कीय हिसाबले सार्थक (Statistically Significant) हुन् वा होइनन् सोसम्बन्धमा पनि यस खण्डमा स्पष्टताका साथ उल्लेख गर्नुपर्दछ । यदि यस परियोजनामा विभिन्न विधिहरूको प्रयोग गरिएको छ भने तिनीहरूको नतिजालाई तुलना गर्नुपर्दछ । साथै, अन्य परिवेशका उस्तै अध्ययन प्रकृतिका नतिजाहरूसँग पनि तुलनात्मक सारसंक्षेप उल्लेख गर्न सकिन्छ ।

3 मागको नियम: कुनै वस्तु वा सेवाको मूल्य बढ्यो भने माग घट्दछ । त्यस्तै खास वस्तु वा सेवाको मूल्य घट्यो भने माग बढ्छ ।

च) निष्कर्षः अनुसन्धानात्मक लेखको अन्तिम भाग भनेको निष्कर्ष हो । यस भागमा नतिजाहरूको सारांश समावेश गर्नुपर्दछ । त्यसबाट निकालिएको निष्कर्ष उल्लेख गर्दै भविष्यको लागि अनुसन्धान गर्ने क्षेत्रका बारेमा पनि उल्लेख गर्न सकिन्छ । यसका अतिरिक्त यस लेखका पाठकसँग छलफल गर्ने गरी केही विषय समावेश गर्न पनि सकिन्छ । स्मरण रहोस्, अनुसन्धानात्मक लेख तयार गर्दा सम्बन्धित लेखकले पाठकसँग अन्तरक्रिया गरिरहेको अनुभूत गरेर लेख्नुपर्दछ । यसबाट लेख पढ्ने व्यक्तिले पनि भौतिक सामीप्यताविना नै लेखकसँग कुराकानी गरिरहेको अनुभूत गर्न सक्दछन् ।

२०.३ तथ्याङ्कमा गोपनीयता र नैतिकता

तथ्याङ्कीय गतिविधिमा गोपनीयता अत्यन्त महत्त्वपूर्ण विषय हो । व्यक्तिगत विवरणहरू मार्फत् तथ्याङ्कको निर्माण हुन्छ । कहिल्यै पनि व्यक्तिगत विवरण सार्वजनिक गरिनु हुँदैन । सबै विवरणहरू तथ्याङ्कीय प्रयोजनका लागि मात्र प्रयोग गरिनुपर्दछ र तथ्याङ्क संकलनमा संलग्न व्यक्ति वा संस्थाले यो कुराको सुनिश्चितता सम्बन्धित उत्तरदाता वा तथ्याङ्क प्रदायकलाई गराउनु पर्दछ । तथ्याङ्कीय गतिविधिको अर्को महत्त्वपूर्ण पक्ष भनेको नैतिकता हो । खासगरी तथ्याङ्क संकलक, प्रदायक तथा प्रयोगकर्ताले नैतिक धरातलभित्र रहेर मात्र तथ्याङ्कको गतिविधि सञ्चालन गर्नुपर्दछ । कतिपय तथ्याङ्क प्रयोगकर्ताले माध्यमिक स्रोत (Secondary Source) का तथ्याङ्क प्रयोग गरेका हुन्छन् । यसरी अन्य व्यक्ति वा संस्थाले उत्पादन गरेका तथ्याङ्क प्रयोग गर्दा सम्बन्धित प्रयोगकर्ताले उपयुक्त तरिकाले तथ्याङ्कको स्रोत उल्लेख गर्नुपर्दछ । तथ्याङ्क प्राप्त गर्ने आधिकारिक विधिको अनुसरण गरेर मात्र तथ्याङ्क प्राप्त गर्नुपर्दछ । गोप्य राख्नुपर्ने तथ्याङ्कको सुरक्षा सुनिश्चित गर्नुपर्दछ । तथ्याङ्कमा कुनै पनि शर्त वा प्रयोजनमा लागि फेरबदल गर्नु हुँदैन । साथै, तथ्याङ्कसँग सम्बन्धित व्यक्ति वा संस्थाले कसैको डर, धम्की, प्रलोभन वा अन्य कुनै प्रभाव वा दबाबमा परेर तथ्याङ्कमा अनधिकृत रूपले सम्पादन गर्ने परिवर्तन गर्ने जस्ता कार्यहरू गर्नु हुँदैन । वास्तवमा यस्ता अनधिकृत रूपमा गरिने तथ्याङ्कको फेरबदलका कार्यलाई तथ्याङ्कीय अपराध (Statistical Crime) मानिन्छ ।

२०.४ गोपनीयता सम्बन्धी कानुनी प्रावधान

तथ्याङ्कमा गोपनीयता निकै महत्त्वपूर्ण विषय भएकोले यससम्बन्धी कानुनी प्रावधान रहेको हुन्छ । नेपालको सन्दर्भमा कुरा गर्दा संविधानको धारा २८ मा तथ्याङ्कको गोपनीयतासम्बन्धी विषय उल्लेख गरिएको छ । त्यसैगरी तथ्याङ्क ऐन, २०१५ को दफा ८ मा व्यक्तिगत विवरण निजको लिखित स्वीकृतिविना प्रकाशन गर्न नपाइने व्यवस्था गरेको छ । यस ऐनमा यदि कसैले व्यक्तिगत विवरण निजको स्वीकृतिविना प्रकाशन गरेमा रू २०० सम्म जरिवाना वा दुई महिनासम्म जेल सजाय वा दुवै सजाय हुने

व्यवस्था गरेको छ ।

२०.५ अभ्यासका लागि प्रश्नहरू

- क) प्रायोगिक परियोजना भन्नाले के बुझ्नुहुन्छ ? उदाहरणसहित व्याख्या गर्नुहोस् ।
- ख) प्रायोगिक परियोजना किन सञ्चालन गर्नुपर्दछ ? पत्रकारहरूलाई यसको के महत्त्व हुन्छ ?
- ग) प्रायोगिक परियोजना सञ्चालन गर्दा तथ्याङ्क संकलन कसरी गरिन्छ ? उल्लेख गर्नुहोस् ।
- घ) प्रायोगिक परियोजना सञ्चालनका प्रमुख चरणहरू बुँदागत रूपमा उल्लेख गर्नुहोस् ।
- ङ) तथ्याङ्कमा गोपनीयताको किन महत्त्व हुन्छ ?

२०.६ सन्दर्भ सामग्री

प्रायोगिक परियोजना सञ्चालन गर्ने विधिका बारेमा विस्तृत अध्ययन गर्नका लागि देहायको सामग्री उपयुक्त छ ।

१. Introductory Econometrics: A Modern Approach by Jeffrey M. Wooldridge, Fifth Edition, South-Western, Cengage Learning, 2013



सर्वेक्षण प्रतिवेदन अध्ययन

२१.१ परिचय

आधिकारिक निकायहरूबाट सञ्चालन गरिएका विभिन्न सर्वेक्षणहरूका प्रतिवेदन तथ्याङ्कका भरपर्दा स्रोत हुन् । कतिपय प्रतिवेदनहरूमा तथ्याङ्क विश्लेषण गरेर सूचकहरू समेत उपलब्ध गराइएका हुन्छन् । यस्ता तथ्याङ्क वा सूचना स्थानीय तह, प्रदेश तथा संघका नीतिनिर्माता, योजनाकारलाई अत्यावश्यक हुन्छन् । साथै, पत्रकार, विद्यार्थी, नागरिक समाज, विकास साभेदार आदिलाई पनि सर्वेक्षण प्रतिवेदनहरू तथ्याङ्कको स्रोतका रूपमा वा सन्दर्भ सामग्रीका रूपमा महत्त्वपूर्ण मानिन्छन् । सर्वेक्षण प्रतिवेदनहरूलाई बुझ्न र उपयोग गर्नका लागि त्यस्ता प्रतिवेदनका विषयवस्तुको गहन अध्ययन गर्नुपर्दछ । यस्ता प्रतिवेदनमा उल्लेख गरिएका तथ्याङ्क वा सूचकहरूको अर्थ बुझ्ने गरी अध्ययन गर्ने कार्यलाई प्रतिवेदन अध्ययन भनिन्छ । यस पाठमा सर्वेक्षण प्रतिवेदन अध्ययन गरेर तिनका सारसंक्षेप बताउन सक्ने हुनका लागि परियोजना कार्यमार्फत् पाठकवर्गमा अध्ययन तथा लेखन शैलीको विकास गर्ने लक्ष्य लिइएको छ ।

२१.२ सर्वेक्षण प्रतिवेदन भनेको के हो ?

कुनै पनि गणना वा नमुना सर्वेक्षण सम्पन्न भइसकेपछि सम्बन्धित निकायले सर्वेक्षणको परिचयसहित यसबाट प्राप्त तथ्याङ्क वा सूचना प्रकाशन गरेका हुन्छन् । यस्ता प्रकाशनहरू हार्डकपी (मुद्रित) वा इलेक्ट्रोनिक कपी (विद्युतीय) वा दुवै हुन सक्दछन् । यसरी सर्वेक्षणका नतिजाहरू प्रस्तुत गरिएका प्रकाशनहरूलाई सर्वेक्षण प्रतिवेदन भनिन्छ । उदाहरणका लागि केन्द्रीय तथ्याङ्क विभागले सञ्चालन गरेको नेपाल श्रमशक्ति सर्वेक्षण, २०१७/१८ को प्रतिवेदन, आर्थिक गणना, २०१८ को प्रतिवेदन, नेपाल बहुसूचक सर्वेक्षण, २०१९ को प्रतिवेदन, स्वास्थ्य मन्त्रालयले सञ्चालन गरेको नेपाल जनसांख्यिक स्वास्थ्य सर्वेक्षण, २०१६ आदि सर्वेक्षण प्रतिवेदनहरू हुन् ।

२१.३ सर्वेक्षण प्रतिवेदन अध्ययन गर्दा ध्यान दिनुपर्ने कुराहरू

सर्वेक्षणका प्रतिवेदनहरू अध्ययन गर्दा विशेष गरी निम्न विषयहरूमा ध्यान दिनुपर्दछ ।

- **सर्वेक्षणको उद्देश्य:** प्रतिवेदनमा सर्वेक्षणका उद्देश्यहरू उल्लेख गरिएको हुन्छ । कुनै खास उद्देश्यका लागि सञ्चालन गरिएका सर्वेक्षणहरूले विशेष गरी त्यही उद्देश्य प्राप्तिका लागि तथ्याङ्क संकलन तथा व्यवस्थापन गरेका हुन्छन् । प्रतिवेदनको प्रयोगकर्ताले प्रतिवेदनमा उल्लेख भएका उद्देश्यसँग आफूले गर्न खोजेको अध्ययन सम्बन्धित छ वा छैन भन्ने कुरा ख्याल गर्नुपर्दछ । तर कतिपय प्रतिवेदनहरूमा प्रत्यक्ष

रूपमा उद्देश्य नलेखिएको पनि हुन सक्दछ । तर सामान्यतया प्रतिवेदनको परिचय खण्डमा सर्वेक्षणको उद्देश्य खुलाइएको हुन्छ । यदि उद्देश्य स्पष्ट खुलाइएको छैन भने पनि अन्य कुनै स्रोतको प्रयोग गरेरै भए पनि सर्वेक्षणको उद्देश्य थाहा पाउनुपर्दछ । उदाहरणका लागि नेपाल श्रमशक्ति सर्वेक्षण, २०१७/१८ को प्रतिवेदनको परिचय खण्डमा तथा दायरा र कभरेज खण्डमा सर्वेक्षणको उद्देश्य खुलाइएको छ । जसअनुसार सो सर्वेक्षणको उद्देश्य भनेको नेपालको श्रम तथ्याङ्क अद्यावधि गर्नु तथा नेपालको श्रम बजारको मुख्य मुख्य नतिजा र प्रवृत्तिहरूको ढाँचाको विश्लेषण गर्नु हो ।

● **सर्वेक्षणको दायरा र कभरेज:** गणनाहरूको दायरा निकै फराकिलो हुन्छ र यिनीहरूले लक्षित जनसंख्या जति सबै समेटेका हुन्छन् । अर्कोतर्फ नमुना सर्वेक्षणहरूको दायरा विशिष्टीकृत हुन्छ र यिनीहरूको कभरेज पनि निश्चित गरिएको हुन्छ । उदाहरणका लागि नेपाल श्रमशक्ति सर्वेक्षण, २०१७/१८ को प्रतिवेदनमा उल्लेख भएबमोजिम यस सर्वेक्षणको दायरा भनेको रोजगार, स्वरोजगार तथा स्वयंसेवासम्बन्धी कामका बारेमा विवरण संकलन गर्नु रहेको थियो । जसका लागि यस सर्वेक्षणले श्रम तथ्याङ्कसम्बन्धी १९ औँ अन्तर्राष्ट्रिय सम्मेलन (19th ICLS) ले सुझाएका अवधारणा तथा परिभाषा अवलम्बन गरेर रोजगार, बेरोजगार तथा अर्ध बेरोजगारसम्बन्धी तथ्याङ्क समेटेको थियो ।

● **लक्षित जनसंख्या र नमुना चयन:** सर्वेक्षणले कुनै निश्चित जनसंख्याको विवरण संकलन गर्ने लक्ष्य लिएको हुन्छ । नमुना सर्वेक्षणको हकमा कुन विधि प्रयोग गरेर नमुना संकलन गरिएको हो सो कुरा पनि प्रतिवेदनमा उल्लेख भएको हुन्छ । तसर्थ, प्रतिवेदन अध्ययन गर्ने व्यक्तिले सम्बन्धित सर्वेक्षणको लक्षित जनसंख्या र नमुना चयनसम्बन्धी विधिका बारेमा पनि जानकारी हासिल गर्नुपर्दछ । उदाहरणका लागि नेपाल श्रमशक्ति सर्वेक्षण, २०१७/१८ को प्रतिवेदनमा उल्लेख भएबमोजिम यस सर्वेक्षणको लक्षित जनसंख्या भनेको काम गर्ने उमेर समूह अर्थात् १५ वर्ष वा सोभन्दा माथिका जनसंख्या हो । यस सर्वेक्षणको नमुना चयन गर्न स्तरीकृत सम्भाव्यता नमुना (Stratified Probability Sampling) प्रयोग गरिएको थियो । यो नमुना चयन बहुचरण नमुना छनोट (Multistage Sampling) मा आधारित थियो । यस सर्वेक्षणको तथ्याङ्क विश्लेषण डोमेन (Domain) १४ वटा रहेका थिए र देशभरिबाट १८,००० परिवारहरू छानिएका थिए ।

● **प्रश्नावली तयार गर्न प्रयोग गरिएको दृष्टिकोण:** सर्वेक्षणमा विवरण संकलन गर्नका लागि प्रश्नावलीको प्रयोग गरिएको हुन्छ । कागजी वा आजकल त ट्याबलेटमा समेत प्रश्नावली तयार गरिन्छ । खास प्रकृतिका सर्वेक्षणहरूका लागि अन्तर्राष्ट्रिय सुझावहरूको अवलम्बन गरेर प्रश्नावलीहरू तयार गरिन्छ । प्रश्नावलीहरू प्रायः प्रतिवेदनको अनुसूचीमा

उपलब्ध गराइएका हुन्छन् । प्रतिवेदन अध्ययन गर्ने व्यक्तिले प्रश्नावलीको पनि अध्ययन गरेर के कस्ता प्रश्नहरू समावेश भएका छन् भन्ने बारेमा जानकारी हासिल गर्नुपर्दछ । प्रश्नावलीकै आधारमा विभिन्न चरहरूको पहिचान गर्न सकिन्छ । यसबाट आफूलाई चाहिएको चरको मात्र तथ्याङ्क हेर्न पनि सकिन्छ । उदाहरणका लागि नेपाल श्रमशक्ति सर्वेक्षण, २०१७/१८ को प्रतिवेदनमा उल्लेख भएबमोजिम यस सर्वेक्षणमा अन्तर्राष्ट्रिय श्रम संगठन (ILO) का सुभावहरू समेटेर १४ वटा खण्डमा १६८ वटा प्रश्नहरू समावेश गरिएका थिए ।

● **सर्वेक्षण कार्यान्वयन प्रक्रिया:** सर्वेक्षण प्रतिवेदन अध्ययन गर्ने क्रममा पाठकले उक्त सर्वेक्षणको कार्यान्वयनको पाटोलाई पनि गहिरिएर नियाल्नुपर्दछ । प्रायजसो प्रतिवेदनहरूमा स्थलगत कार्यको संगठन (Organization of Field Work) भनेर यसका बारेमा उल्लेख गरेका हुन्छन् । यसबाट तथ्याङ्क गुणस्तरको तह थाहा पाउन सकिन्छ । अनुगमन तथा सुपरिवेक्षण र गुणस्तर नियन्त्रणका लागि गरिएका पहलका बारेमा थाहा हुन्छ । यस कुराले प्रतिवेदनमा प्रस्तुत गरिएका सूचनामाथि विश्वास गर्ने आधार तयार हुन्छ । उदाहरणका लागि नेपाल श्रमशक्ति सर्वेक्षण, २०१७/१८ को प्रतिवेदनमा उल्लेख भएबमोजिम यस सर्वेक्षणमा २४ वटा तथ्याङ्क संकलन गर्ने समूह परिचालन गरिएको थियो । प्रत्येक समूहमा एकजना समूह नेता, एकजना सुपरिवेक्षक र अन्य दुईजना गणकहरू थिए । स्थलगत कार्यमा खटिने कर्मचारीलाई मुख्य प्रशिक्षकहरूबाट १८ दिनको प्रशिक्षण गरिएको थियो । सुपरिवेक्षणकर्ताका लागि अलग्गै तालिम प्रदान गरिएको थियो । तालिम अघिमै फिल्ड अभ्यास गरेर प्रशिक्षार्थीहरूको बुझाइको तह र आत्मविश्वास अभिवृद्धि गरिएको थियो । यस सर्वेक्षणमा समूह नेताले दैनिक रूपमा फिल्ड कार्यको गुणस्तर नियन्त्रण गर्दथे भने केन्द्रीय तथ्याङ्क विभाग, राष्ट्रिय योजना आयोग तथा तथ्याङ्क कार्यालयबाट अनुगमन तथा सुपरिवेक्षण भएको थियो । यस्ता सुपरिवेक्षण स्थलगत तथ्याङ्क संकलन कार्यको सुरुका दिनहरूमा निकै गहन रूपमा गरिएको थियो । यसबाट तथ्याङ्क संकलनमा केही कमीकमजोरी भएमा तत्काल सुधार गर्न र केही समस्या भएमा समाधान गर्न सहज भएको थियो ।

● **सन्दर्भ समय:** सर्वेक्षण प्रतिवेदन अध्ययन गर्दा ध्यान दिनुपर्ने अर्को महत्त्वपूर्ण पक्ष भनेको सर्वेक्षणको सन्दर्भ समय हो । कुन समयवाधिलाई आधार मानेर तथ्याङ्क संकलन गरिएको हो सो बारेमा पाठकलाई पूर्ण जानकारी हुनुपर्दछ । सन्दर्भ समयले तथ्याङ्कको सुसंगतता (Consistency) लगायत अन्य गुणहरू कायम गर्न ठुलो मद्दत गर्दछ । उदाहरणका लागि नेपाल श्रमशक्ति सर्वेक्षण, २०१७/१८ को प्रतिवेदनमा उल्लेख भएबमोजिम यस सर्वेक्षणको सन्दर्भ समय विभिन्न चरका लागि विभिन्न तरिकाले लिइएको थियो ।

खासगरी चलायमान हप्ता (Moving Week) को आधारमा रोजगारीसम्बन्धी विवरणका लागि सन्दर्भ समय लिइएको थियो । जसअनुरूप काम गरेको घण्टा र रोजगारीबाट भएको आम्दानीका लागि अन्तर्वार्ता गरिएको दिनदेखिको अघिल्लो हप्तालाई सन्दर्भ समय मानिएको थियो । अर्कोतर्फ पेशा, शिक्षा, वैवाहिक अवस्था जस्ता चरका लागि अन्तर्वार्ता गरिएको दिनलाई सन्दर्भ समय मानिएको थियो । आफ्नै उत्पादनका वस्तु तथा सेवा र स्वयंसेवाका लागि पछिल्लो ३० दिनलाई सन्दर्भ समय मानिएको थियो । बेरोजगारीसम्बन्धी विवरण संकलन गर्नका लागि भने फरक-फरक सन्दर्भ समयको अवधारणा प्रयोग गरिएको थियो । अन्तर्वार्ता भएको दिनदेखिको अघिल्लो हप्ता रोजगार वा व्यवसायमा नभएको तर विगत ३० दिनदेखि काम खोजेको र १५ दिनभित्र काम गर्न उपलब्ध भएका व्यक्तिका लागि बेरोजगारसम्बन्धी विवरण संकलन गरिएको थियो ।

● **अनुत्तरित प्रश्न र यसको व्यवस्थापन (Non-Response and Its Treatment):**

कतिपय सर्वेक्षणहरूमा उत्तरदाताले पूरा वा आंशिक प्रश्नहरूको उत्तर नदिने सक्दछन् । सामान्यतया गणनाहरूमा पूर्ण रूपले उत्तर नदिने भन्ने हुँदैन । तर नमुना सर्वेक्षणहरूमा भने पूर्ण रूपमा उत्तर नदिने उत्तरदातालाई प्रतिस्थापन गर्ने गरी नमुना छनोट विधिमा उल्लेख गरिएको हुन्छ । आंशिक रूपमा अनुत्तरित प्रश्नको हकमा भने तथ्याङ्क व्यवस्थापन गर्दा छुट भएका मान (Missing Values) हरूलाई तथ्याङ्कीय विधिको प्रयोग गरेर उपचार गरिएको हुन्छ । प्रतिवेदनमा तथ्याङ्क प्रस्तुत गर्नुपूर्व यस्ता छुट भएका मानहरूलाई कसरी उपचार गरिएको छ भनेर थाहा पाउनु राम्रो मानिन्छ । तर कतिपय प्रतिवेदनहरूमा यस्ता कुराहरू उल्लेख नगरिएको पनि हुन्छ । उदाहरणका लागि नेपाल श्रमशक्ति सर्वेक्षण, २०१७/१८ को प्रतिवेदनमा यस्ता छुट भएका मानका बारेमा केही उल्लेख गरिएको छैन । तर नेपाल बहुसूचक सर्वेक्षण, २०१९ को प्रतिवेदनमा भने छुट मानको व्यवस्थापन गर्दै अनुमान गर्ने पद्धतिका बारेमा उल्लेख गरिएको छ ।

● **परिणाम र तिनीहरूको प्रयोग:** सर्वेक्षण प्रतिवेदन अध्ययनको सबैभन्दा महत्त्वपूर्ण पक्ष भनेको सर्वेक्षणको नतिजा र प्रयोगका बारेमा जानकारी हासिल गर्नु हो । परिणाम वा नतिजाहरू प्राय तालिका, चित्र वा रेखाचित्रका माध्यमबाट प्रतिवेदनमा उल्लेख गरिएको हुन्छ । उदाहरणका लागि नेपाल श्रमशक्ति सर्वेक्षण, २०१७/१८ को प्रतिवेदनमा उल्लेख भएको एउटा तालिका हेरौं । तालिका २१.१ मा काम गर्ने उमेर समूहका मानिसहरूको श्रमसम्बन्धी विभिन्न सूचनाहरू प्रस्तुत गरिएको छ । आर्थिक वर्ष २०१७/१८ मा काम गर्ने उमेर समूह (१५ वर्ष वा सोभन्दा माथि) का जम्मा २०७,४६,००० जनामध्ये पुरुष ९२,०८,००० र महिला १,१५,३७,००० जना देखिन्छन् । तीमध्ये रोजगारीमा

रहेका ४४,४६,००० पुरुष र २६,४०,००० महिला रहेका छन्। यसैगरी बेरोजगार संख्या, जम्मा श्रमशक्ति, श्रमशक्तिमा नभएका जनसंख्याका बारेमा यस तालिकामा तथ्याङ्क प्रस्तुत गरिएको छ। यी तथ्याङ्कका आधारमा विभिन्न सूचकहरू निर्माण गर्न सकिन्छ। जस्तै: पुरुष बेरोजगार दर १०% र महिला बेरोजगार दर १३% रहेको देखिन्छ।^४ यस्ताबाट महिला र पुरुषबिच रोजगारीको अवस्थाको तुलना पनि गर्न सकिन्छ। नीतिनिर्माताहरूले यही सूचनाको आधारमा महिलाकेन्द्रित रोजगारमूलक कार्यक्रमहरू तर्जुमा गर्न सक्दछन्। यसरी सर्वेक्षणको नतिजा अध्ययन गरेर प्रयोग गर्न सकिन्छ।

तालिका २१.१: काम गर्ने उमेर समूहका जनसंख्याको सूचक र लिङ्गानुसार वितरण

विवरण	पुरुष (०००)	महिला (०००)
रोजगार	४,४४६	२,६४०
बेरोजगार	५११	३९७
श्रमशक्ति	४,९५८	३,०३६
श्रमशक्ति भन्दा बाहिर रहेका	४,२५०	८,५००
काम गर्ने उमेर समूहका जनसंख्या	९,२०८	११,५३७

स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपाल श्रमशक्ति सर्वेक्षण, २०१७/१८^५

२१.४ परियोजना कार्य

यस पुस्तकको मुख्य उद्देश्य भनेको विद्यार्थी, पत्रकार, नीतिनिर्माता, योजनाकार, विकास साभेदारलगायत अन्य सरोकारवालाहरूको तथ्याङ्कीय साक्षरताका लागि आधारभूत ज्ञान प्रदान गर्नु हो। यसअनुरूप हामीले अधिल्ला पाठहरूमा तथ्याङ्कका विविध आयामहरूका बारेमा चर्चा गरिसकेर यस पाठमा सर्वेक्षणका प्रतिवेदनहरू अध्ययन गर्दा के कस्ता कुराहरूमा विशेष ध्यान पुऱ्याउनु पर्दछ भनेर संक्षेपमा छलफल गरिसकेका छौं। यसपछि पाठकहरूले हासिल गरेको ज्ञानको आधारमा केही प्रतिवेदनहरूको अध्ययन गरेर सारांशमा प्रतिवेदनका मूलभूत कुराहरू समेटिएको एक अध्ययन प्रतिवेदन तयार गर्न सक्षम गराउने उद्देश्यले यस पाठमा परियोजना कार्यको परिकल्पना गरिएको हो। त्यस अनुरूप निम्नानुसारका प्रतिवेदनहरू अध्ययन गरेर संक्षिप्त प्रतिवेदन तयार गर्नका लागि सबै

४ बेरोजगार दर = $\frac{\text{बेरोजगार संख्या}}{\text{जम्मा श्रमशक्ति}} \times १००$

५ Central Bureau of Statistics. (2019). *Report on the Nepal Labour Force Survey 2017/18*. Kathmandu, Nepal; Central Bureau of Statistics. Retrieved on December 10, 2021 from: https://cbs.gov.np/wp-content/uploads/2019/05/Nepal-Labour-Force-Survey-2017_18-Report.pdf

पाठकवर्गमा आग्रह गरिएको छ । परियोजना कार्यका लागि छानिएका प्रतिवेदनहरू:

- १) नेपाल जनसंख्या तथा स्वास्थ्य सर्वेक्षण, २०१६^६
- २) नेपाल जीवनस्तर सर्वेक्षण, २०१० (पहिलो भोलुम)^७
- ३) नेपाल बहुसूचक सर्वेक्षण, २०१९^८
- ४) आर्थिक गणना, २०१८ (राष्ट्रिय प्रतिवेदन पहिलो भोलुम)^९

२१.५ सन्दर्भ सामग्री

विभिन्न सर्वेक्षणका प्रतिवेदनहरू अध्ययन गर्नुहोस् ।



6 Ministry of Health and Population. (2016). *Nepal Demographic and Health Survey 2016*. Kathmandu, Nepal; Ministry of Health and Population. Retrieved on December 10, 2021 from: <https://mohp.gov.np/en/publications/2019-09-28-15-00-22/nepal-demographic-health-survey>

7 Central Bureau of Statistics. (2011). *Nepal Living Standard Survey 2010/11 (Statistical Report – Volume One)*. Kathmandu, Nepal; Central Bureau of Statistics. Retrieved on December 10, 2021 from: https://cbs.gov.np/wp-content/uploads/2018/12/Statistical_Report_Vol1.pdf

8 Central Bureau of Statistics. (2020). *Nepal Multiple Indicator Cluster Survey 2019 (Survey Findings Report)*. Kathmandu, Nepal; Central Bureau of Statistics and UNICEF Nepal. Retrieved on December 10, 2021 from: https://cbs.gov.np/wp-content/uploads/2020/12/Nepal_MICS6-2019_Final-Online-Version-15-Dec-2020-v2.pdf

9 Central Bureau of Statistics. (2020). *National Economic Census: National Report on Salaries and Wages*. Kathmandu, Nepal; Central Bureau of Statistics. Retrieved on December 10, 2021 from: <https://cbs.gov.np/wp-content/uploads/2020/09/National-Economic-Census-2018-National-Report-3-1.pdf>

भाग ५

**तथ्याङ्कीय क्रियाकलापमा
कम्प्युटरको प्रयोग**

एक्सेलसम्बन्धी आधारभूत ज्ञान

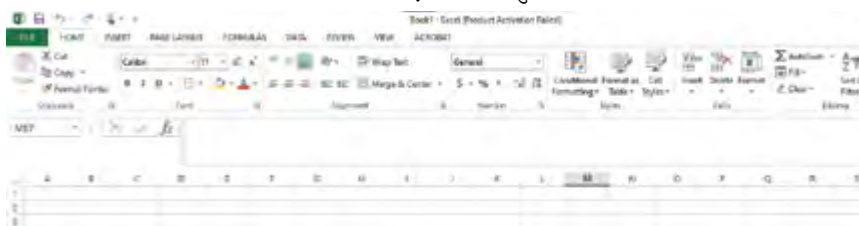
२२.१ परिचय

एक्सेल कम्प्युटरको माइक्रोसफ्ट अफिस अन्तर्गतको एउटा सफ्टवेयर हो । यो सफ्टवेयर तथ्याङ्क व्यवस्थापनका लागि प्रयोग गरिन्छ । यस पाठमा हामीले एक्सेलको प्रयोग गरेर तथ्याङ्क व्यवस्थापनका केही विषयहरूका बारेमा छलफल एवं अभ्यास गर्नेछौं । एक्सेलको सामान्य परिचयका लागि यसको मेनुबारका बारेमा संक्षिप्त जानकारी लिऔं ।

२२.१.१ मेनुबार (Menu Bar)

एक्सेलको मेनुबारमा विभिन्न काम गर्ने ८ वटा शीर्षकहरू (मेनु) रहेका हुन्छन् । प्रत्येक मेनुभित्र विभिन्न सुविधाहरूको सूची हुन्छ । जसरी रेस्पुन्टको मेनुभित्र फरक-फरक परिकारको नाम र मूल्य दिइएको हुन्छ त्यसरी नै एक्सेलको प्रत्येक मेनुभित्र के के प्रकार्य गर्न सकिने सुविधा उपलब्ध हुन्छन् तिनको सूची रहेको हुन्छ । तलको तस्वीरमा मेनुबार देख्न सकिन्छ ।

चित्र नं. १: एक्सेलको मेनुबार



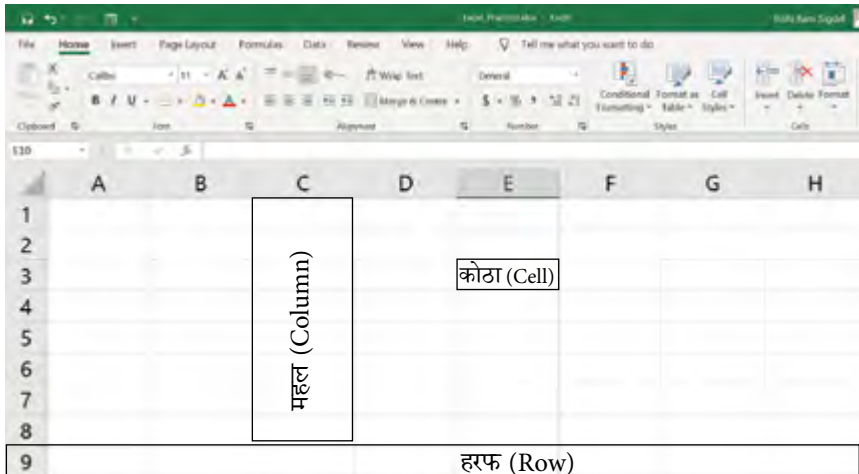
यस चित्रको सिरानमा हेर्नु भन्ने हाइलाइट भएको रंगमा फाइल, त्यसपछि HOME, INSERT, PAGE LAYOUT, FORMULAS, DATA, REVIEW / VIEW देख्न सकिन्छ । यीमध्ये जुनसुकैलाई क्लिक गरेपछि त्यहाँभित्र उपलब्ध सुविधाहरू देख्न सकिन्छ । उदाहरणको लागि स्क्रिनमा होम मेनु थिचिएको अवस्था छ । त्यहाँभित्र अवस्थित सारा सुविधाहरू (जस्तो कि कट, कपी, फरम्याट पेन्टर, बोल्ड, इटालिक, मर्ज एण्ड सेन्टर, इन्सर्ट, डिलिट, फरम्याट लगायतका दर्जनौं सुविधाहरू) स्क्रिनमा देखा परेका छन् ।

२२.१.२ हरफ, महल र कोठा (Row, Column, Cell)

यदि हामीले माथिको स्क्रिनको सबैभन्दा तल्लो भागमा हेर्नु भन्ने बायाँबाट दायाँतिर A, B, C, D, E लगायतका नामहरू र माथिदेखि तलतिर हेर्नु भन्ने १, २, ३, हुँदै गएको देख्न सक्छौं । यी तेर्सोतिरका A, B, C, D, E महल अर्थात् (Column) हुन् भने ठाडो तिरका १, २, ३ हरू हरफ वा (Row) हुन् । स्मरणीय छ एक्सेलको एक सिटमा १०, ४८,

५७६ Row र १६,३८४ Column हुन्छन्। योभन्दा ठुलो खासगरी जनगणनाका तथ्याङ्क एकैपटक एक्सेलले लिन सक्दैन। एक्सेल सिटमा देखिने साना साना कोठाहरूलाई सेल भनिन्छ।

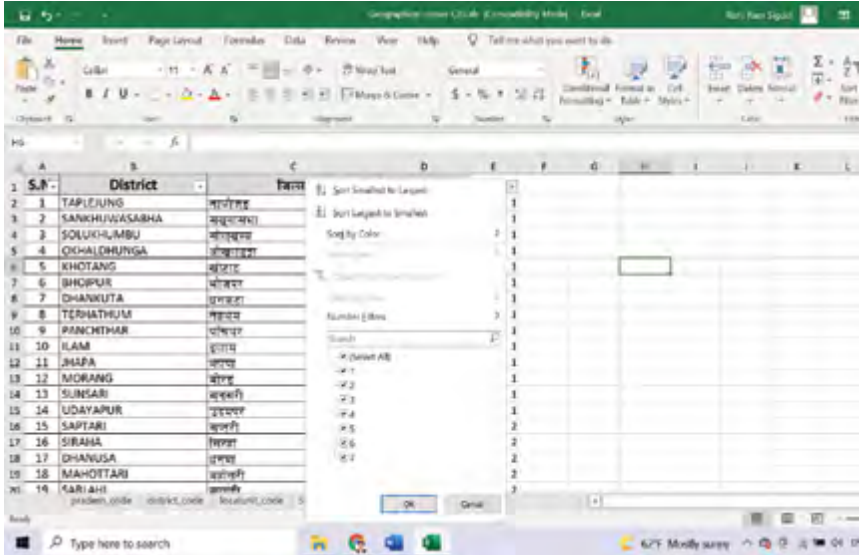
चित्र नं. २: एक्सेलको हरफ, महल र कोठा



२२.१.३ फिल्टर, सर्टिंग र फिल्टर निष्क्रिय गर्ने

एक्सेलको फिल्टर प्रकार्यले भएको डाटासेटबाट आफूलाई चाहिएको खास डाटामात्र खास आधारमा छानेर उही सिटको कुनै स्थानमा वा अर्को सिटमा सिर्जना गर्न मद्दत गर्दछ। कुनै कुनै डाटासेट यति ठुला हुन्छन् कि ती सबैको हामीलाई जरूरत पर्दैन। डाटासेटको एक अंशमात्र छानेर लिनका लागि फिल्टर फंक्सन उपयोग हुन्छ। यसबाट मेनुबारको डाटाभिन्न सोली आकारको आइकन सहित यो सुविधा प्रस्तुत गरिएको छ। तलको चित्र नं. २ मा फिल्टरको प्रयोगबाट नयाँ डायलगबक्स देखिएको छ। जसमा प्रदेश नं. छानेर चाहिएका जिल्लाहरू मात्र राख्न सकिन्छ। फिल्टरलाई निष्क्रिय गर्नका लागि पुनः एकपटक सोली आकारको आइकनलाई क्लिक गर्नुपर्दछ। (अभ्यास: फिल्टरको प्रयोग गरी मधेश प्रदेशका जिल्लाहरूको मात्र लिष्ट तयार गरौं। यसका लागि २ मा ठिक चिन्ह कायम गरेर अरू बटन क्लिक गर्दै हटाउँदै जानुपर्दछ।)

चित्र नं. ३: फिल्टरको प्रयोग



एक्सेलमा तथ्याङ्कलाई सानोदेखि ठुलो वा ठुलोदेखि सानो क्रममा मिलाएर राख्न सकिन्छ। खासगरी दिइएका अवलोकनहरूको योग्यता वा क्रम हेर्नका लागि यो विशेषता निकै महत्त्वपूर्ण हुन्छ।

तालिका २२.१: नेपाल सरकारका कर्मचारीहरूको तलब स्केल, २०७८

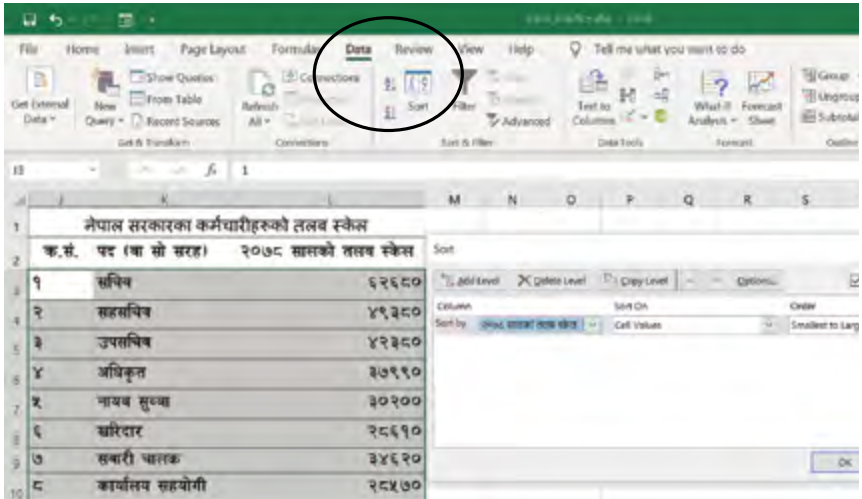
क्र.सं.	पद (वा सो सरह)	२०७८ सालको तलब स्केल
१	सचिव	६२,६८०
२	सहसचिव	४९,३८०
३	उपसचिव	४२,३८०
४	अधिकृत	३७,९९०
५	नायब सुब्बा	३०,२००
६	खरिदार	२८,६१०
७	सवारी चालक	३४,६२०
८	कार्यालय सहयोगी	२८,५७०

स्रोत: अर्थ मन्त्रालय

उदाहरणका लागि तालिका २२.१ मा नेपाल सरकारका कर्मचारीहरूको तलब स्केल दिइएको छ। यो तथ्याङ्कको क्रम सबैभन्दा धेरैबाट थोरै रहेको छ। यसलाई सबैभन्दा थोरैबाट धेरै बनाउन एक्सेलमा Data उपमेनु भित्रको Sort आइकन थिच्नुपर्दछ। यसका

लागि सम्बन्धित तालिकाको कुनै एउटा सेल सेलेक्ट भएको वा पुरै तालिका सेलेक्ट गरेको हुनुपर्दछ। त्यसपछि Column अन्तर्गतको Sort by मा जुन महलका तथ्याङ्कलाई सर्टिग गर्ने हो, त्यही छानेर Order मा आफूलाई चाहिएको Order सेलेक्ट गर्नुपर्दछ (चित्र नं. ४)। त्यसपछि Ok थिचेपछि तथ्याङ्कको क्रम भनेको जस्तै भएर आउँदछ। यसअन्तर्गत अक्षरहरूका सूचीलाई पनि सर्टिगको प्रयोग गरेर वर्णानुक्रमअनुसार क्रम मिलाउन सकिन्छ।

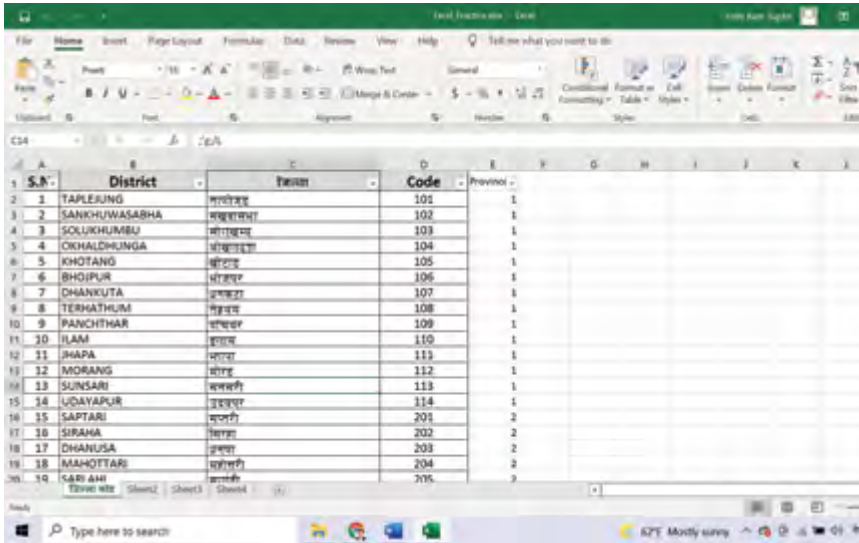
चित्र नं. ४: एक्सेलमा तथ्याङ्क सर्टिङ गर्ने तरिका



२२.१.४ सिट (थप्ने, नाम परिवर्तन गर्ने, नक्कल गर्ने र हटाउने)

वर्कसिटमा नयाँ प्रयोजनका लागि सिट थप्न सकिन्छ। वर्कसिटको सबैभन्दा तलको भागमा उपयोगमा रहेको सिटको नाम Sheet 1, Sheet2, Sheet3 भएर आएको हुन्छ। हामीलाई त्यस्तै अर्को सिटको आवश्यकता परेको खण्डमा त्यस सिटभन्दा ठिक दायाँतिरको + चिन्हमा क्लिक गरेर नयाँ सिट थप्न सकिन्छ। त्यसो गर्दा Sheet4 देखा पर्दछ। कर्सरको सहायताले सिटहरूको क्रमिकतालाई अगाडि पछाडि पनि पार्न सकिन्छ। यदि सिट १ लाई कुनै नयाँ नाम दिनु पर्‍यो भने सम्बन्धित सिटलाई राइट क्लिक गरेर रिनेम (नाम परिवर्तन) गर्न सकिन्छ। राइट क्लिककै सहायताले उक्त सिटको हुबहु नक्कल गर्न वा त्यसलाई हटाउन पनि सकिन्छ। तलको तस्वीरमा Sheet 1 लाई “जिल्ला कोड” नाम दिएको देख्न सकिन्छ।

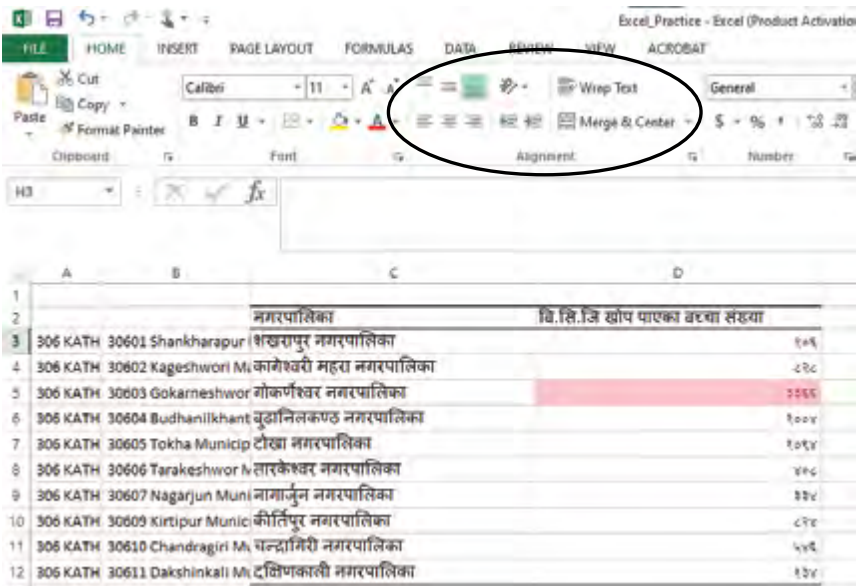
चित्र नं. ५: एक्सेल सिटको नाम परिवर्तन गर्न



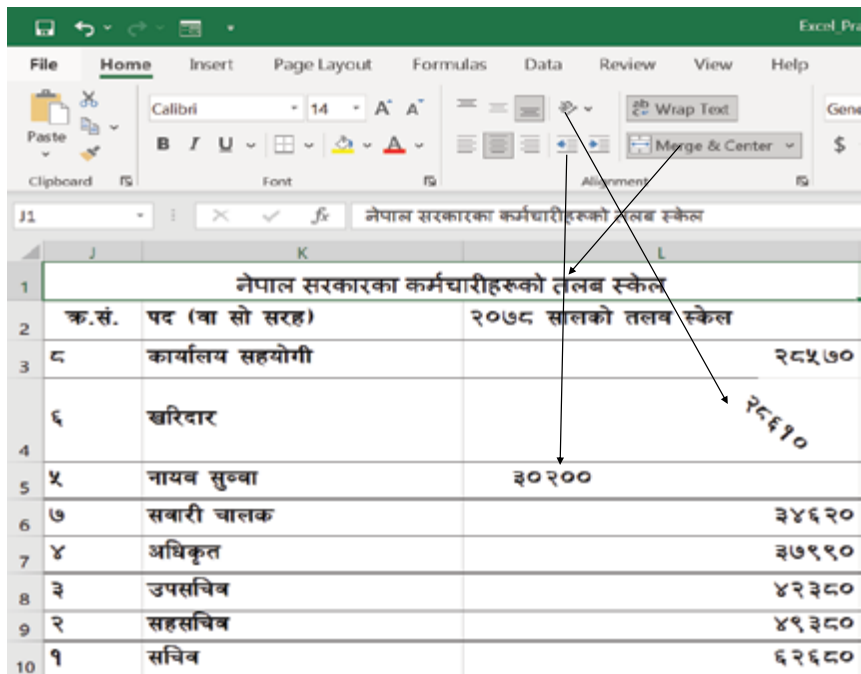
२२.१.५ अक्षरहरूको अवस्थिति मिलान (Alignment), कोठाहरू मर्ज (एकीकरण) र अक्षरलाई हरफीकरण (Wrap) गर्ने

एक्सेलमा प्रविष्ट भएका अक्षर वा अङ्कको अवस्थिति मिलान गर्नका लागि Home उपमेनु भित्रको Alignment समूहका आइकनहरू प्रयोग गर्नुपर्दछ । यसका लागि जुन महल वा हरफका तथ्याङ्क वा अक्षरहरूको अवस्थिति मिलान गर्नुपर्ने हो तिनलाई सेलेक्ट गरेर Top, Middle, Bottom, Left, Centre, Right मध्ये आवश्यकताअनुसार क्लिक गर्नुपर्दछ (चित्र नं. ६) । एक्सेलको कुनै खास कोठामा लेख्दै जाँदा बायाँबाट दायाँतिर जान्छ तर अलि पर पुगेपछि लेखिएका अक्षर हराउँछन् । कोठाको तेर्सो आकार जति सानो भयो पढ्नेलाई उत्तिकै गाह्रो हुन्छ । खास सेलमा यस्तो समस्या देखा परेपछि Wrap Text मेनुभित्रको एलाइनमेन्ट प्रकार्यको प्रयोग गरेर Wrap Text सुविधा लिन सकिन्छ । Alignment भित्रको Wrap Text ले उट्टै कोठामा धेरै अक्षर वा अंकहरू भएमा दोस्रो लहरमा मिलाएर राख्नका लागि मद्दत गर्दछ । Merge and Center ले दुई वा दुईभन्दा बढी कोठाहरूलाई उट्टै बनाएर त्यसभित्रका शब्द वा अंकलाई केन्द्रमा राखिदिन्छ । Orientation ले अक्षर वा अंकलाई तेर्सो, ठाडो वा छड्के बनाएर राख्न मद्दत गर्दछ । Indent ले अक्षर वा अंकलाई दायाँ बायाँ सार्नमा मद्दत गर्दछ ।

चित्र नं. ६: एक्सेलमा प्रविष्ट भएका अक्षर वा अंकको अवस्थिति मिलाउन



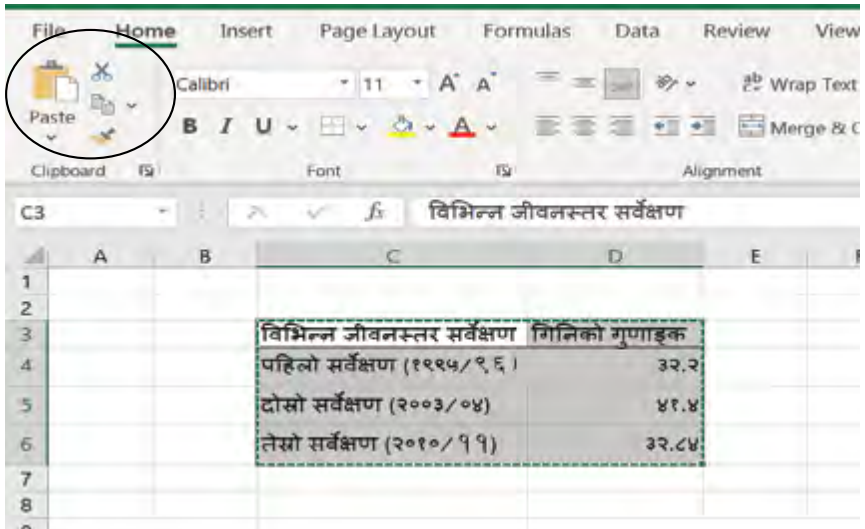
चित्र नं. ७: अक्षरहरूको अवस्थिति मिलाउन



२२.१.६ कपी गरेका कुरा कुतै लगेर टाँस (Paste) गर्ने र फर्मेट गर्ने

यो कार्य माइक्रोसफ्ट वर्डकै जस्तो हो । कुनै एक कोठाको वा धेरै कोठाहरूको सामग्रीलाई अरू कुनै स्थानमा लगेर Paste गर्न सकिन्छ । यो विकल्प पूरै सिट कपी पेस्ट गर्न वा कुनै इमेज, ग्राफ वा चार्ट कपी पेस्ट गर्न पनि उपयोगी हुन्छ । यसका लागि पहिला पेष्ट गर्नुपर्ने सेलहरूलाई सेलेक्ट गर्नुपर्दछ । त्यसपछि कपी आइकन थिचेर जहाँ पेष्ट गर्नु पर्नेछ, त्यहाँको सेल सेलेक्ट गरेर पेष्ट आइकन थिच्नुपर्दछ । यदि कपी गर्ने बेलामा कुँचीजस्तो आइकन थिचेर कट गरियो भने सुरूको ठाउँबाट डाटा वा लेखहरू मेटिएर पेष्ट गरेको ठाउँमा मात्र रहन्छ । यसरी पेष्ट गर्ने बेला यदि पेष्ट स्पेशल गर्ने हो भने विभिन्न विकल्पहरू प्रयोग गर्न पाइन्छ । जस्तै: तेस्रो लहरलाई ठाडो बनाउने, फर्मेटमात्र पेष्ट गर्ने आदि ।

चित्र नं. ८: कपी पेष्ट गर्ने तरिका



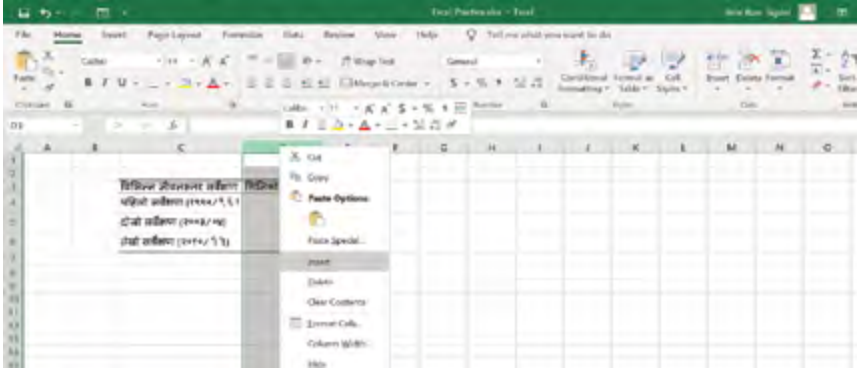
फरम्याट पेन्टरले कुनै सेलमा रहेको अक्षर, एलाइनमेन्ट, रंग, फन्ट, आदिलाई मेमोरीमा राखेर अन्यत्र सोही बमोजिम रूपान्तरण गर्न सहयोग गर्दछ । यसका लागि कुनै सेलमा कर्सर राखेर Home मेनुभित्र रहेको Format Painter मा लगेर क्लिक गर्ने र जुन सेललाई वा सेलहरूलाई अघिल्लो सेलको जस्तै फरम्याटमा ढाल्ने हो त्यस सेलमा लगेर क्लिक गर्नुपर्दछ ।

२२.१.७ हरफ र महल थप्ने

काम गरिरहेको सिटमा आवश्यकताअनुसार बिचमा नयाँ हरफ वा महल थप्नुपर्ने अवस्था आउन सक्छ । यसका लागि जहाँनेर हरफ वा महल थप्नुपर्ने हो, त्यहाँनेरको हरफ वा

महललाई राइट क्लिक (Right Click) गरेर Insert थिच्नु पर्दछ । तल चित्रमा देखाए जस्तै महल D मा राइट क्लिक गरेर Insert थिचेपछि महल C र D को बिचमा नयाँ महल थप हुन्छ । यही प्रक्रिया हरफहरू थप्न पनि प्रयोग गरिन्छ ।

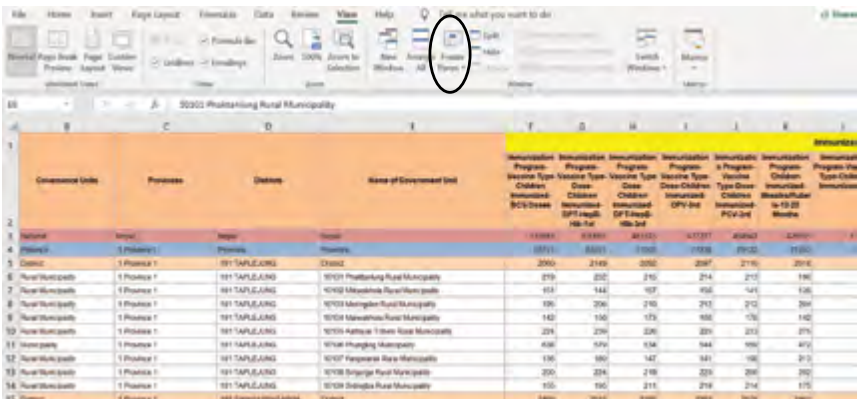
चित्र नं. ९: महल थप्ने तरिका



२२.१.८ महल वा हरफको शीर्षक राखेर अरू भाग लुकाउने (Freeze Panes)

कहिलेकाहीँ काम गरिरहेको सिटका सेलहरूमा यति धेरै विवरण वा डाटा हुन्छन् कि तिनलाई तल तल वा दायाँतिर हेर्दै जाँदा माथि कुन शीर्षक अन्तर्गतको के डाटा हेर्दैछु भनेर थाहा नहुने अवस्था हुन्छ । दायाँतिर जति गए पनि बायाँतिरको शीर्षक देखिई राख्न वा तलतिर जति गए पनि माथिको शीर्षक देखिई राख्न Freeze panes कमाण्ड काम आउँछ । यो सुविधा View मेनुभित्र गइसकेपछि तलको स्क्रिनमा देखिएअनुसार सेलेक्ट गरेर उपयोग गर्न सकिन्छ ।

चित्र नं. १०: फ्रिज प्यान प्रयोग गर्ने तरिका



२२.१.९ एक्सेलमा तथ्याङ्क व्यवस्थापन

आवश्यकताअनुसार एक्सेलबाट तथ्याङ्क प्रविष्टिदेखि विश्लेषणसम्मका कामहरू यसबाट गर्न सकिन्छ । तथ्याङ्क विश्लेषणका लागि कम्प्युटरमा पहिलो पटकका लागि Data Analysis Tool लाई सक्रिय बनाउनुपर्दछ । एकपटक यो टुल सक्रिय बनाएपछि सो कम्प्युटरमा सँधैभरि यसको प्रयोग गर्न पाइन्छ । यो टुल सक्रिय बनाउनका लागि निम्नानुसार गर्नुपर्दछ ।

एक्सेलको फाइलमा
जानुहोस्

More..... मा जानुहोस्

Options मा जानुहोस्

Add-ins मा जानुहोस्

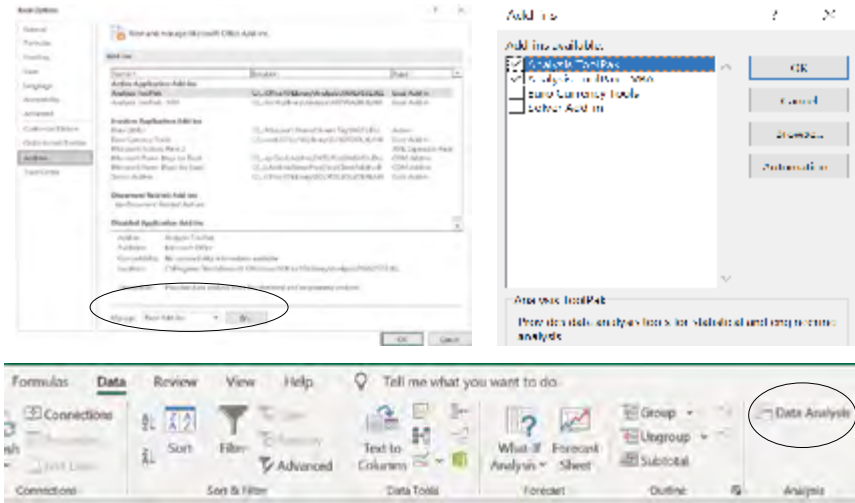
Window को पुछारमा

भएको Excel Add-ins को Go... मा जानुहोस् ।



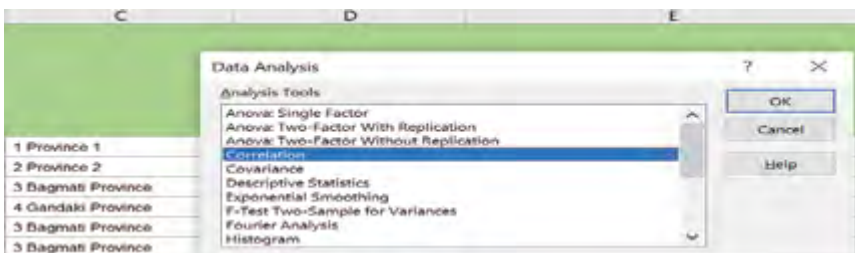
Analysis ToolPak र Analysis ToolPak-VBA लाई सेलेक्ट गरेर OK गरेपछि Data Analysis ट्याब सक्रिय हुन्छ । यसरी एकपटक सेटिङ गरिसकेपछि यो टुल सँधैका लागि प्रयोग गर्न सकिन्छ । यस टुलभिन्न तथ्याङ्क विश्लेषणका विभिन्न सुविधाहरू जस्तै प्रतिगमन विश्लेषण, हिस्टोग्राम बनाउने, सहसम्बन्ध विश्लेषण, परिकल्पना परीक्षणजस्ता कार्यहरू गर्न सकिन्छ ।

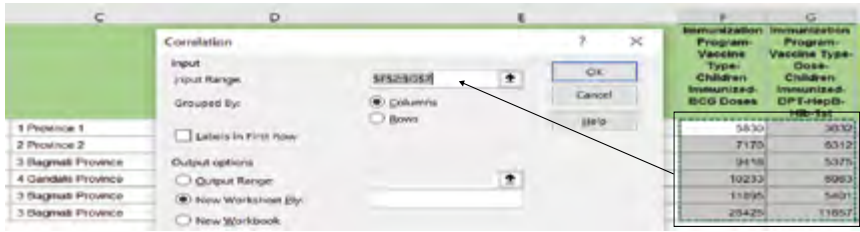
चित्र नं. ११: Data Analysis Tool सक्रिय बनाउने तरिका



उदाहरणका लागि चित्र नं. १२ मा Data Analysis Tool को प्रयोग गरेर सहसम्बन्ध (Correlation) कसरी निकालिन्छ भनेर देखाइएको छ । यसका लागि यस टुलको Correlation सेलेक्ट गरेर OK गरेपछि चरका मानहरू छान्ने विकल्प देखाएको हुन्छ । तलको चित्र नं. १२ मा महल F र G का चरहरूलाई सेलेक्ट गरेर Input Range मा राखिएको छ । त्यसपछि OK थिचेपछि सहसम्बन्धको सानो तालिकासहित मानहरू दिएको हुन्छ । तलको चित्रमा Column1 Column2 भनेर दिइएको छ । सुरुको तथ्याङ्कमा पहिलो महल अर्थात् F लाई Column1 र दोस्रो महल अर्थात् G लाई Column2 भनिएको हो । यी दुई महलमा भएका चरहरूबिचको सहसम्बन्ध ०.९३५२४५ पाइएको छ ।

चित्र नं. १२: एक्सेलबाट सहसम्बन्ध निकाल्ने तरिका





	Column 1	Column 2
Column 1	1	
Column 2	0.935245	1

एक्सेलमा तथ्याङ्क विश्लेषणका लागि अन्य विभिन्न टुलहरूको प्रयोग यही तरिकाले अभ्यास गर्न सकिन्छ । यसको विस्तृत सिकाइ यस पाठ्यसामग्रीको क्षेत्रभन्दा बाहिर भएकोले इच्छुक पाठकहरूले अनलाइनमा उपलब्ध थुप्रै अध्ययन सामग्री (भिडियो, अडियो र अन्य फाइल) प्राप्त गर्न सक्नुहुनेछ ।

२२.१.१० तथ्याङ्क व्यवस्थापन र निर्णय निर्माणमा एक्सेल

दैनिक कार्यसम्पादनको सिलसिलामा उत्पन्न हुने प्रशासनिक अभिलेखहरूको व्यवस्थापनमा एक्सेल अत्यन्त उपयोगी प्रोग्राम मानिन्छ । तथ्याङ्क प्रशोधनका तुला तुला प्रोग्रामिडको ज्ञान नभए पनि एक्सेलको आधारभूत तहको ज्ञान हुनेले तालिका, चार्ट, औसत, प्रतिशतलगायतका सुविधाहरू प्रयोग गरेर आफूसँग भएका आँकडाहरू विश्लेषण गर्ने गरेको पाइन्छ । साना र मध्यम आकारका आँकडाहरू प्रशोधन र प्रस्तुत गर्न एक्सेल निर्णयकर्ताहरूको लागि पर्याप्त सुविधा हो । तुला तुला सर्वेक्षण र गणनाहरूको तथ्याङ्क विश्लेषण र बिग डाटा (Big Data) को प्रशोधनमा भने एक्सेल खासै उपयोगी मानिँदैन । एक्सेलमा असंख्य सुविधाहरू उपलब्ध छन् । ती कतिपय सुविधाहरू हामीलाई जरूरत पर्दैनन् । यहाँ एक्सेललाई विस्तृत हिसाबले सिकाउने उद्देश्य होइन । प्रशासनिक अभिलेखहरूबाट सिर्जित र अन्य स्रोत (प्राथमिक वा माध्यमिक) बाट प्राप्त तथ्याङ्कलाई प्रयोगकर्ताहरूले आफ्नो चाहनाअनुसारका सूचकहरू उत्पादन गर्न वा भइरहेकै तथ्याङ्कलाई सजिलोसँग बुझ्नका लागि तालिका, ग्राफ, चार्ट आदिका स्वरूपमा ढालेर हेर्ने सामान्य तथ्याङ्कीय क्षमता विकास गर्ने उद्देश्य मात्र यहाँ राखिएको छ ।

२२.२ तालिका तथा यसका ढाँचा

एक्सेलको प्रयोग गरेर प्रस्तुत गर्ने अवस्थाका तथ्याङ्कहरूको तालिका फर्मेन्टिड कसरी गर्ने भनेर यस खण्डमा संक्षिप्त जानकारी दिन खोजिएको छ । एक्सेलबाहेक अन्य सफ्टवेयर प्रयोग गरेर विश्लेषण गरिएका तथ्याङ्क पनि एक्सेलको प्रयोग गरेर तालिका

बनाउन सकिन्छ। यसका लागि यहाँ एउटा उदाहरणसहित वर्णन गरिएको छ।
उदाहरण: तलको चित्र नं. १३ को तालिकालाई स्तरीय तालिका बनाउनुहोस्।

चित्र नं. १३: एक्सेलमा तालिका बनाउने

क.सं.	विभिन्न जीवनस्तर सर्वेक्षण	आर्थिक वर्ष	सिनिको मुनाइ
१	पढिलो सर्वेक्षण	१९९४/९६	३२.२
२	दोस्रो सर्वेक्षण	२००३/०४	४१.४
३	तेस्रो सर्वेक्षण	२०१०/११	३२.८४

स्रोत: केन्द्रीय तथ्याङ्क विभाग

क) सुरूको महलमा क्र.सं. राख्नका लागि एउटा महल Insert गर्ने।

ख) नयाँ महलको शीर्षक क्र.सं. लेखी पहिलोमा १ र दोस्रोमा २ टाइप गर्ने। तेस्रोमा ३ टाइप गर्न पनि सकिन्छ। तर अर्को विशेषता हेर्नका लागि हामीले बुँदा (ग) अनुसार अभ्यास गरेका छौं।

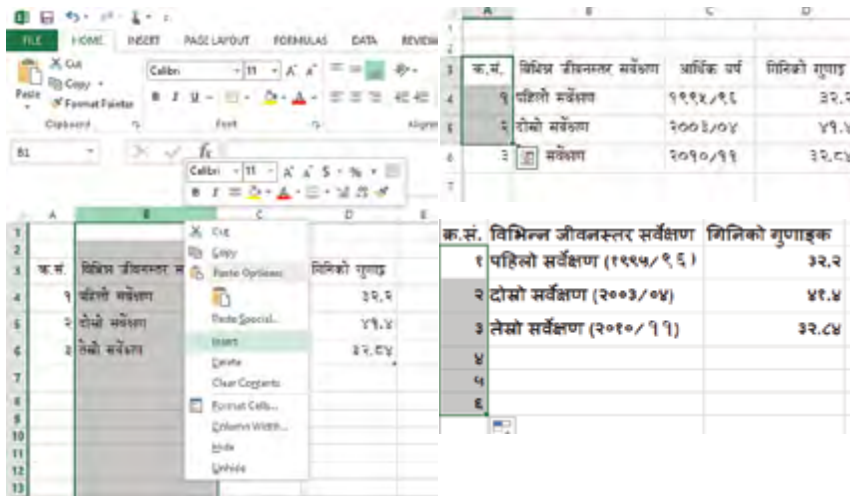
ग) १ र २ लाई सेलेक्ट गरेर ड्रयाग गरेपछि क्र.सं. क्रमशः स्वरूपमा राख्न सकिन्छ। यस उदाहरणमा ३, ४, ५, ६ यही विधिबाट तानिएको छ।

घ) तर क्र.सं. ४, ५, ६ हामीलाई चाहिएको छैन। यी अंकहरू हटाउन विभिन्न तरिका छन्। तिनीहरूलाई सेलेक्ट गरेर सोभै हटाउन पनि सकिन्छ भने, तिनीहरूको लहर वा हरफलाई सेलेक्ट गरेर Right Click गरेर Delete गर्न पनि सकिन्छ।

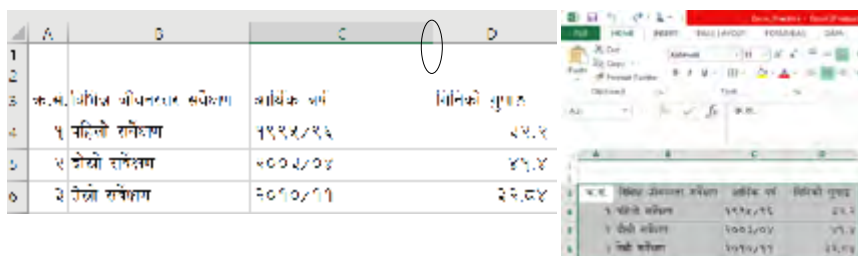
ङ) अब, सर्वेक्षण सञ्चालन वर्षलाई छुट्टै महलमा राख्दा तालिका थप आकर्षक हुनसक्छ। तसर्थ, एउटा महल Insert गरेर ती वर्षहरूलाई छुट्टै राखौं। तलको चित्रमा “आर्थिक वर्ष” को महल निकै चौडा भएको हुनाले कोलम C र D को बिचमा कर्सर राखेर डबल क्लिक गरौं। अन्य महलहरूको पनि चौडाइ मिलाउन यही प्रक्रिया गरौं।

च) तालिकालाई बोर्डर राख्नका लागि तालिका सेलेक्ट गरौं र Home भित्रको बोर्डर आइकनमा जाऔं। त्यसपछि आवश्यकताअनुसारका बोर्डरहरू प्रयोग गरेर तालिका बनाऔं। यहाँ, सामान्यतया प्रयोग गरिने बोर्डरको नमुना तालिका २२.२ मा देखाइएको छ।

चित्र नं. १४: एक्सेलमा तालिका बनाउने



चित्र नं.: १५ एक्सेलमा तालिका बनाउने



तालिका २२.२: जिनीको गुणाङ्क

क्र.सं.	विभिन्न जीवनस्तर सर्वेक्षण	आर्थिक वर्ष	जिनीको गुणाङ्क
१	पहिलो सर्वेक्षण	१९९५/९६	३२.२
२	दोस्रो सर्वेक्षण	२००३/०४	४१.४
३	तेस्रो सर्वेक्षण	२०१०/११	३२.८४

स्रोत: केन्द्रीय तथ्याङ्क विभाग

छ) एक्सेलमा भएका आँकडाहरूलाई तालिकाको फर्मेटमा लैजाने अर्को विधि भनेको Insert मेनुभित्र Table भन्नेआइकनको प्रयोग गर्नु हो । तालिका निर्माणको लागि एक्सेल सजिलो प्रोग्राम मानिन्छ । आफूसँग भएका प्रशासनिक अभिलेख वा कुनै पनि स्रोतबाट प्राप्त अप्रशोधित वा प्रशोधित तथ्याङ्कलाई आफ्नो रुचि र आवश्यकता अनुसारको तालिकामा राखेर प्रस्तुत गर्न सकिन्छ । एक्सेलको Insert उपमेनुभित्र

गएपछि त्यसको ठिक तल Table को आइकन छ । त्यसलाई क्लिक गरेपछि कहाँको कुन डाटा समूहबाट तालिका सिर्जना गर्ने हो भनेर पप-अप बक्स आउँछ । त्यसमा गएर डाटाको Range छान्ने । एउटा तालिका स्वतः बनेर आउँछ र त्यसमा फिल्टर अप्सन पनि उपलब्ध हुन्छ । फिल्टरबाट आफूलाई चाहिएका तथ्याङ्कहरू मात्र छान्न सकिन्छ । त्यसका लागि शीर्षक लहरमा भएको ड्रपडाउन बटनलाई थिच्नुपर्दछ । उदाहरणका लागि तलको चित्रमा यदि हामीलाई प्रदेश १ को मात्र तथ्याङ्क चाहियो भने फिल्टर अप्सनमा गएर Province 1 मात्र सेलेक्ट गर्नुपर्दछ । तालिका नबनाइएको एक्सेलको तथ्याङ्कमा पनि Data भित्रको फिल्टर अप्सनबाट यो सुविधा प्राप्त हुन्छ ।

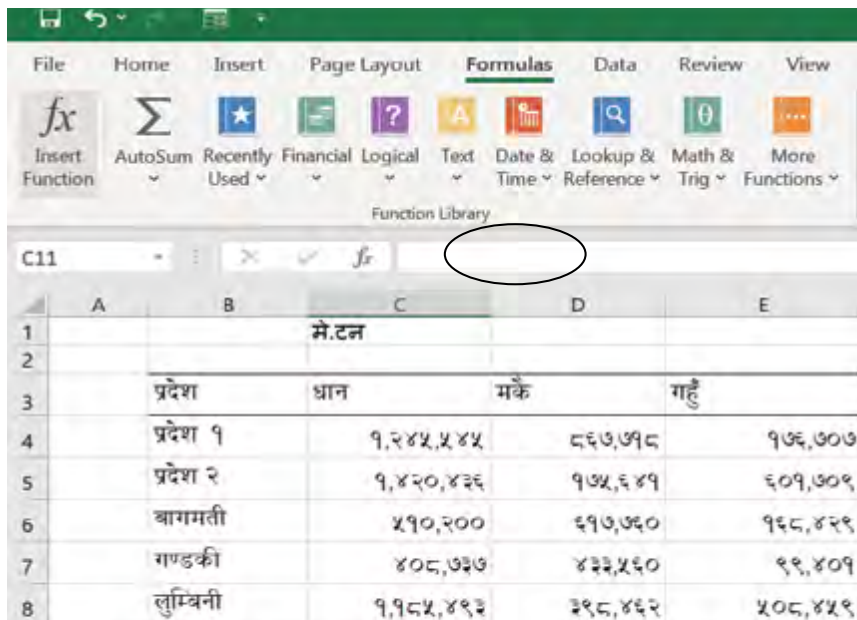
चित्र नं. १६: एक्सेलमा फिल्टरको प्रयोग

Column2	Column3	Column4	Column5
Metropolitan City	1 Province 1	112 KATHMANDU	11214 Kathmandu Metropolitan City
Metropolitan City	2 Province 2	908 PALPA	90804 Palpa Metropolitan City
Metropolitan City	3 Bagmati Province	568 KATHMANDU	56805 Kathmandu Metropolitan City
Metropolitan City	5 Bagmati Province	568 LAJPUR	56807 Lalpur Metropolitan City
Metropolitan City	3 Bagmati Province	513 CHITWAN	51304 Bharatpur Metropolitan City
Metropolitan City	4 Gandaki Province	465 KASKI	46504 Pokhara Metropolitan City
Municipality	1 Province 1	101 TADI PULCH	10105 Tadi Pulch Municipality
Municipality	1 Province 1	102 SARHULWASAL PA	10201 Nandaban Municipality
Municipality	1 Province 1	102 SARHULWASAL PA	10207 Panchakapan Municipality
Municipality	1 Province 1	102 SARHULWASAL PA	10208 Unarapur Municipality
Municipality	1 Province 1	102 SARHULWASAL PA	10209 Masi Municipality
Municipality	1 Province 1	102 SARHULWASAL PA	10210 Unamadai Municipality
Municipality	1 Province 1	103 SOLU KHUMBU	10307 Solu Dudhimesh Municipality
Municipality	1 Province 1	101 DOKHRI TEHRA	10101 Daktari Teahra Municipality
Municipality	1 Province 1	105 BHOJIPUR	10501 Bhojpur Municipality
Municipality	1 Province 1	105 BHOJIPUR	10505 Bhojpur Mahadevpur Municipality
Municipality	1 Province 1	168 BHADKUR	16801 Bhadkaur Municipality
Municipality	1 Province 1	168 BHADKUR	16804 Bhadkaur Municipality
Municipality	1 Province 1	167 DHANRUWA	16701 Madhuni Municipality

२२.३ एक्सेलमा सूत्रको प्रयोग

एक्सेलका सेलहरूमा विभिन्न सूत्रको प्रयोग गरेर गणितीय क्रिया जस्तै जोड, घटाऊ, गुणा, भाग गर्न सकिन्छ । साथै तथ्याङ्कीय सूत्रहरू जस्तै अंकगणितीय मध्यक, ज्यामितीय मध्यक, सहसम्बन्ध आदि प्रयोग गरेर हिसाब गर्न सकिन्छ । मेनुबारको पाँचौं मेनुमा Formulas भन्ने उपमेनु छ । त्यहाँभित्र धेरै प्रकारका सूत्रहरू समावेश गरिएका छन् । एक्सेलका सामान्य प्रयोगकर्ताहरूलाई ती सबै सूत्रहरूको आवश्यकता पर्दैन । यदि हामीले फरमुला उपमेनुलाई क्लिक गर्छौं भने यस्तो स्क्रिन देख्न सकिनेछ जसमा त्यहाँभित्र व्यवस्था गरिएको प्रमुख सुविधाहरू एकै भलकमा थाहा पाउन सकिन्छ । कुनै सेलमा सूत्रको प्रत्यक्ष प्रयोग गर्न पनि सकिन्छ । सूत्र प्रयोग गर्नका लागि जुन सेलमा हिसाब गर्नुछ, त्यहाँ गएर “=” चिन्ह राखेर मात्र सूत्र लेख्नुपर्दछ ।

चित्र नं. १७: एक्सेलमा फरमुला उपमेनु



उदाहरण: तलको चित्र नं. १२ नेपालको आर्थिक गणनाबाट प्राप्त प्रतिष्ठान र संलग्न जनशक्तिको सप्तरी जिल्लाको वडागत विवरणको एउटा अंश हो । यसबाट सप्तरी जिल्लामा जम्मा प्रतिष्ठान संख्या कति रहेछ, निकालौं ।

चित्र नं. १८: वडागत प्रतिष्ठान र संलग्न जनशक्तिको विवरण

District Municipality and Ward in Municipality	Total		A		C		B+D+E+F	
			Agriculture, forestry fishing 3)		Manufacturing		Mining, Electricity, Gas, Water supply & Construction	
	Number of Establish- ments	Number of Persons Engaged	Number of Establish- ments	Number of Persons Engaged	Number of Establish- ments	Number of Persons Engaged	Number of Establish- ments	Number of Persons Engaged
Province 2	117,670	354,994	1,231	4,568	12,654	79,778	342	2,538
201 Saptari District	16,295	42,252	180	817	1,923	8,978	53	207
20101 Saptakoshi Municipality	936	1,973	7	33	79	187	1	1
01	358	880	4	18	47	115	1	1
02	29	70	2	6	2	15	-	-

स्रोत: केन्द्रीय तथ्याङ्क विभाग, आर्थिक गणना, २०१८

जम्मा प्रतिष्ठान संख्या निकाल्ने सेल चयन गरौं । (D122 मा हामी सूत्र प्रयोग गर्दछौं ।) तसर्थ, यस सेलमा = चिन्ह राख्यौं ।

चित्र नं. १९: एक्सेलमा सूत्रको प्रयोग

	A	B	C	D	E	F	G	H
118	320 Saptari	20129 Hanumannagar Kankalini Municipality	13	10	18	-	-	
119	321 Saptari	20130 Hanumannagar Kankalini Municipality	14	111	311	-	-	
120	322 Saptari	20117 Tilathi Kofladi Rural Municipality		561	1,412	1	2	
121	323 Saptari	20118 Chhinnamasta Rural Municipality		526	1,481	5	79	

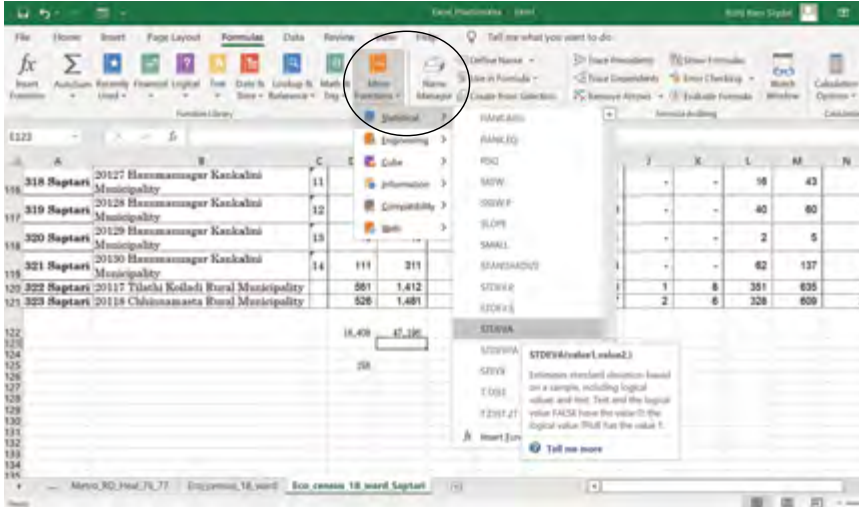
अब, SUM टाइप गरौं र कोष्ठभित्र जोडनुपर्ने तथ्याङ्कको Range टाइप गरौं वा सेलेक्ट गरौं र इन्टर थिचौं । यसरी टाइप गरिएको सूत्र Formula Bar मा देखा पर्दछ । यसरी सूत्रको प्रयोग गरेर हामीले सप्तरी जिल्लामा १८,४०८ वटा प्रतिष्ठान रहेको कुरा पत्ता लगायौं । यही तथ्याङ्कबाट सूत्रको प्रयोग गरेर एक वडामा औसतमा कतिवटा प्रतिष्ठान छन् भन्ने कुरा पनि पत्ता लगाउन सकिन्छ । सूत्र: = AVERAGE(D4:D121)

	A	B	C	D	E	F	G	H
118	318 Saptari	20127 Hanumannagar Kankalini Municipality	11	23	50			
119	319 Saptari	20128 Hanumannagar Kankalini Municipality	12	54	96			
120	320 Saptari	20129 Hanumannagar Kankalini Municipality	13	10	18			
121	321 Saptari	20130 Hanumannagar Kankalini Municipality	14	111	311			
122	322 Saptari	20117 Tilathi Kofladi Rural Municipality		561	1,412			
123	323 Saptari	20118 Chhinnamasta Rural Municipality		526	1,481			
124				18,408				

अर्को तरिका:

Formulas ट्याबमा जाने, AutoSum मा जाने, आवश्यकताअनुसार सूचक छानेर इन्टर थिच्ने । Formulas ट्याब भित्रको More Functions भित्रको Statistical भित्र हामीलाई प्रयोग हुने धेरै सूत्रहरू उपलब्ध छन् । एक्सेलको कुनै तथ्याङ्कमा यी सूत्रहरूको प्रयोग गर्ने अभ्यास गर्न पाठकवर्गमा आग्रह गरिएको छ ।

चित्र नं. २१: एक्सेलमा सूत्रको प्रयोग

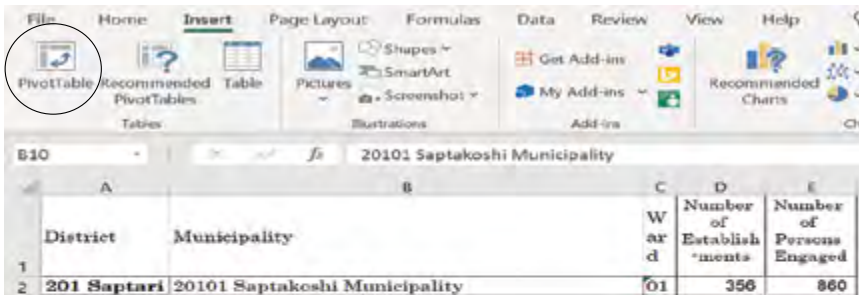


२२.४ एक्सेलमा पिभोट टेबल बनाउने विधि

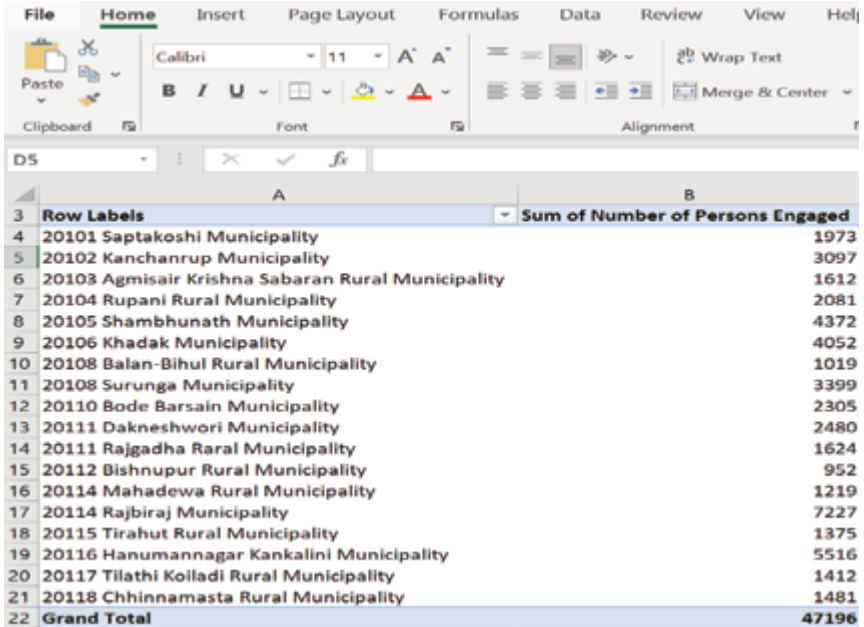
एक्सेलका तथ्याङ्कबाट तालिका बनाउने विधिहरूमध्येको एउटा विधि पिभोट टेबल हो । यसको प्रयोगबाट तथ्याङ्कको सारांश निकालेर तालिकामा प्रस्तुत गर्न सकिन्छ । उदाहरणको लागि तलको चित्र नं. २२ मा सप्तरी जिल्लाका विभिन्न स्थानीय तहका वडाअनुसार प्रतिष्ठान संख्या र सलमन जनशक्ति संख्याको केही अंश देखाइएको छ । यसको एक्सेल फाइलबाट पिभोट टेबलको प्रयोग गरेर वडा अनुसारको तथ्याङ्कलाई स्थानीय तहअनुसार तालिका निम्नबमोजिम गर्न सकिन्छ ।

- १) एक्सेलको तथ्याङ्क भएको कुनै सेल सेलेक्ट गरेर Insert द्याबभिन्न जाने,
- २) पिभोट टेबलमा क्लिक गरेर टेबल Range यकिन गर्ने,
- ३) त्यसपछि आउने नयाँ Dialogue box मा चाहिएका चरहरू सेलेक्ट गर्ने,
- ४) अब, हामीलाई चाहिएको स्थानीय तहअनुसारको तालिका तयार भयो ।

चित्र नं. २२: पिभोट टेबलको प्रयोग गर्ने विधि



चित्र नं. २३: पिभोट टेबलको प्रयोग गर्ने विधि



Row Labels	Sum of Number of Persons Engaged
20101 Saptakoshi Municipality	1973
20102 Kanchanrup Municipality	3097
20103 Agmisair Krishna Sabaran Rural Municipality	1612
20104 Rupani Rural Municipality	2081
20105 Shambhunath Municipality	4372
20106 Khadak Municipality	4052
20108 Balan-Bihul Rural Municipality	1019
20108 Surunga Municipality	3399
20110 Bode Barsain Municipality	2305
20111 Dakneshwori Municipality	2480
20111 Rajgadha Rural Municipality	1624
20112 Bishnupur Rural Municipality	952
20114 Mahadewa Rural Municipality	1219
20114 Rajbiraj Municipality	7227
20115 Tirahut Rural Municipality	1375
20116 Hanumannagar Kankalini Municipality	5516
20117 Tilathi Kolladi Rural Municipality	1412
20118 Chhinnamasta Rural Municipality	1481
Grand Total	47196

२२.५ एक्सेलमा विभिन्न ग्राफ तथा चित्र बनाउने विधि

चार्टहरू तालिकाभन्दा बढी रूचिकर हुन्छन् । तालिकामा सहजै देख्न नसकिने तथ्याङ्कका विशेषताहरू चार्टमा सजिलै देख्न र बुझ्न सकिन्छ । तालिकाभित्र रहेका अंकहरूले कतिपय पाठक वा प्रयोगकर्तालाई थकान महसुस गराउन सक्छन् । चार्टले दिने सन्देश तस्विरले दिने सन्देशजस्तो सहजै बोध गर्न सकिने खालको हुन्छ । हुन त कुनै कुनै चार्टहरू बुझ्नका लागि गहिरो प्राविधिक ज्ञानको आवश्यकता पनि पर्न सक्दछ तर आम सरोकारवालाहरूको लागि बनाइने चार्टहरू साधारण हुन्छन् ।

एक्सेलको प्रमुख प्रयोगहरूमध्येको एउटा प्रयोग भनेको विभिन्न ग्राफ तथा चित्र बनाउनु पनि हो । तथ्याङ्क प्रस्तुतिकरणका लागि ग्राफको प्रयोगका बारेमा हामीले पाठ ८ मा अध्ययन गरिसकेका छौं । यहाँ, हामीले एक्सेलको प्रयोगबाट ग्राफको निर्माण कसरी गर्न बारेमा संक्षिप्त चर्चा गर्नेछौं ।

देहायका चरणहरू पूरा गरेपछि हामी इच्छाअनुसारको चार्ट सिर्जना गर्न सक्दछौं ।

क) उदाहरण: एक्सेलको तालिका २२.३ मा संरचित तालिकामा प्रस्तुत गरिएको तथ्याङ्कलाई चार्टमा प्रस्तुत गर्न खोजिएको छ ।

तालिका २२.३: प्रदेशअनुसार धानको उत्पादन

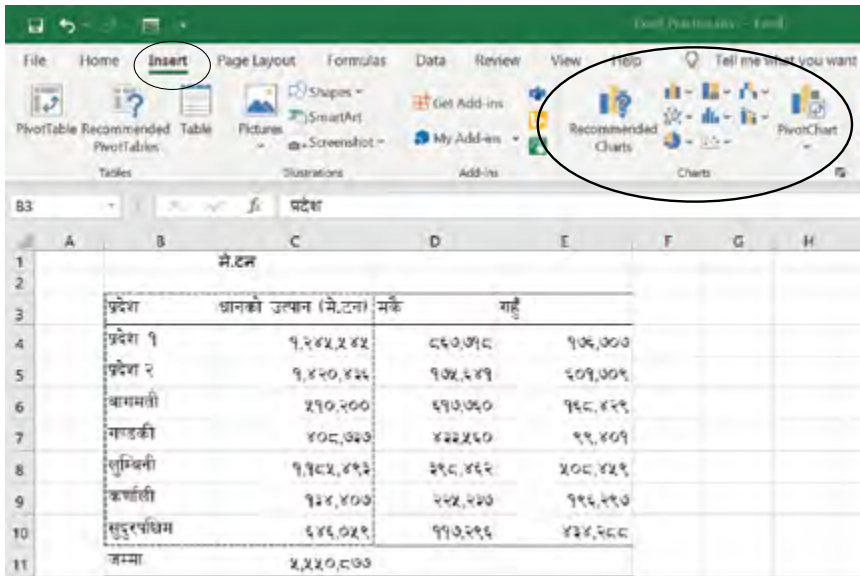
प्रदेश	धानको उत्पादन (मे.टन)
प्रदेश १	१२,४५,५४५
मधेश प्रदेश	१४,२०,४३६
बागमती	५,१०,२००
गण्डकी	४,०८,७३७
लुम्बिनी	११,८५,४९३
कर्णाली	१,३४,४०७
सुदूरपश्चिम	६,४६,०५९

स्रोत: कृषि तथा पशुपन्छी विकास मन्त्रालय, २०७६/७७

ख) जुन तालिकालाई चार्टमा ढाल्ने हो त्यसलाई सेलेक्ट गर्ने । सेलेक्सन गर्दा लहर र महलका शीर्षकलाई समेत गर्ने ।

ग) सेलेक्सन सकिएपछि Insert उपमेनुमा जाने ।

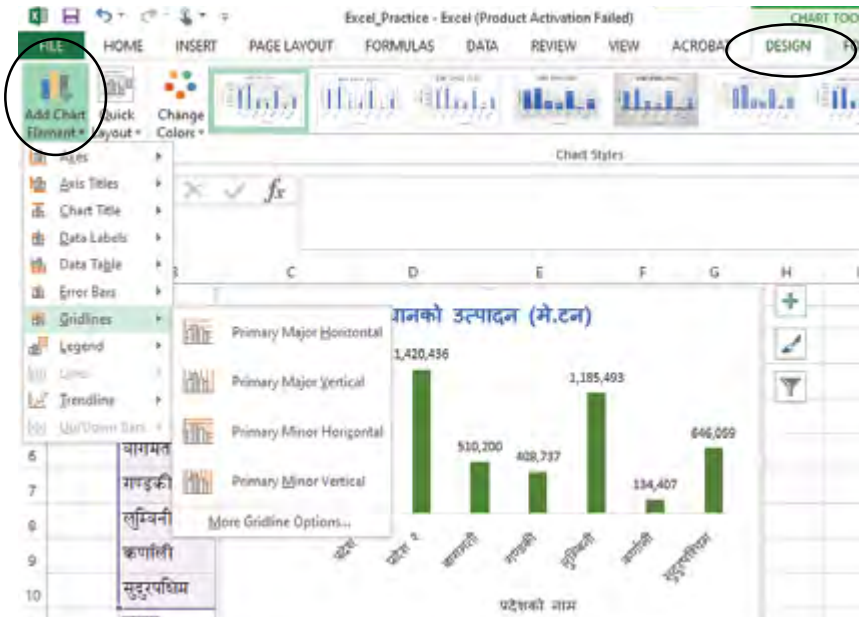
चित्र नं. २४: एक्सेलमा चार्ट उपमेनु रहेको स्थान



घ) त्यहाँ गइसकेपछि चार्ट उपखण्डभित्र रहेका विभिन्न डिजाइनका कुनै एक उपयुक्त आइकनलाई थिच्ने । यसरी थिचिसकेपछि त्यहाँभित्र पनि केही विकल्पहरू देखिन्छन् । तीमध्ये पनि कुनै एकलाई रोज्ने ।

ड) माथि ख देखि घ सम्मका चरण पूरा गरेपछि काम गरिरहेको सिटमा चार्ट सिर्जना हुन्छ।
 च) सिर्जना भइसकेको चार्टलाई आवश्यकतानुसार त्यहाँभित्रका अवयवहरू संशोधन र परिमार्जन गर्न सकिन्छ। नयाँ कुरा थप्न वा भइरहेका कुरा हटाउन पनि सकिन्छ। यसका लागि परिमार्जन गर्न खोजेको अवयवलाई सेलेक्सन गरी Right Click मार्फत् गर्न सकिन्छ। थप परिमार्जन गर्नका लागि चार्टलाई सेलेक्ट गरिसकेपछि Chart Design उपमेनुमा जानुपर्दछ। यस उपमेनु भित्रको Add Chart Element लगायतका अन्य उपमेनुहरू प्रयोग गरेर चार्टलाई आकर्षक बनाउन सकिन्छ। तर चार्टमा धेरै कुराहरू राखेर भद्दा बनाउनु चाहिँ हुँदैन।

चित्र नं. २५: एक्सेलमा चार्टलाई सम्पादन गर्ने तरिका

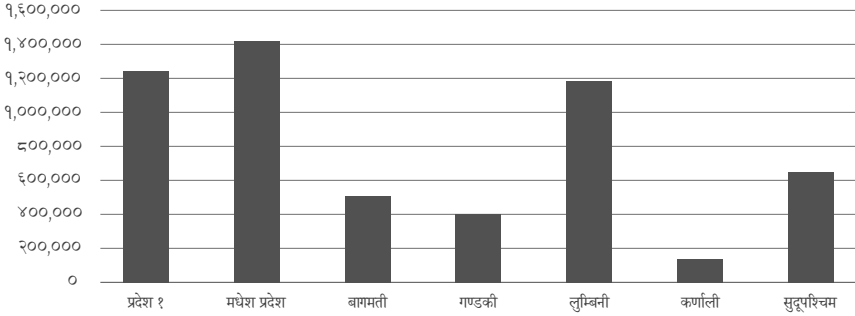


छ) सिर्जित चार्टलाई एक्सेल सिटकै कुनै नयाँ सिटमा, त्यही सिटमा वा माइक्रोसफ्ट वर्डमा वा पावरप्वाइन्ट स्लाइडमा लगेर राख्न सकिन्छ। यसका लागि कपी र पेष्टको प्रयोग गर्नुपर्दछ।

२२.५.१ बार डायग्राम बनाउने विधि

एक्सेलमा बार डायग्राम बनाउनका लागि माथि उल्लेख गरे बमोजिमका चरणहरूकै अनुसरण गर्नुपर्दछ। यसका लागि Insert उपमेनु भित्रको Insert Column or Bar Chart भित्र गएर कुनै एक डिजाइन छान्नुपर्दछ। माथि तालिका २२.२ को तथ्याङ्कको बार डायग्राम चित्र नं. २६ को जस्तो बनाउन सकिन्छ।

चित्र नं. २६: प्रदेशअनुसार धानको उत्पादन
धान



प्रस्तुत चित्र नं. २६ को चार्टमा प्रस्तुत गरिएको तथ्याङ्क के हो भन्नेमा पाठकवर्गलाई निकै द्विविधा उत्पन्न भएको छ । चार्टको शीर्षक धान लेखिएको छ, धानको उत्पादन हो कि धान खेती गरिएको जमिनको क्षेत्रफल हो कि भन्ने कुरा यकिन हुन सक्दैन । त्यस्तै, यो तथ्याङ्क कहिलेको हो भन्नेमा पनि कुनै सूचना दिइएको छैन । तसर्थ, तथ्याङ्कको चार्टमार्फत् प्रस्तुति गर्दा अत्यावश्यक सूचनाहरू उपलब्ध गराइएको हुनुपर्दछ भने नचाहिने चिजहरू नराख्दा पनि हुन्छ । जस्तै: चार्टमा तेर्सो धर्काहरू हटाउँदा पनि केही फरक पर्दछन् । अब यो चार्टलाई थप आकर्षक र सूचनामूलक बनाउन केही सम्पादनका कार्यहरू निम्नानुसार गरौं ।

क) शीर्षक: यसका लागि शीर्षक भएको स्थानमा गएर Double Click गर्ने । त्यसपछि विभिन्न फर्मेटिङका सुविधाहरू प्रयोग गर्ने । उदाहरणका लागि शीर्षकलाई धानको उत्पादन (मे.टन) बनाई अक्षरका रङ्गलाई निलो पारेर बोल्ड पनि गरिएको छ (चित्र नं. १७) ।

ख) तेर्सो धर्का हटाउने: यसका लागि चार्टलाई सेलेक्ट गरेर Chart Design उपमेनुभित्र जाने र Gridlines को Primary Major Horizontal लाई Click गर्ने । यसमा एकपटक क्लिक गर्दा त्यस्ता धर्का हट्दछन् भने अर्को पटक क्लिक गर्दा फेरि थपिन्छन् । अन्य विकल्पहरूमा पनि क्लिक गरेर थप ज्ञान हासिल गर्न सकिन्छ ।

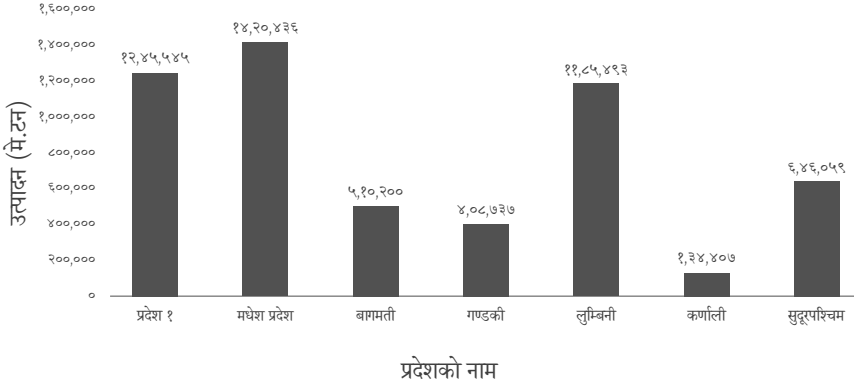
ग) बारमा तथ्याङ्क देखाउने: यसका लागि चार्टलाई सेलेक्ट गरेर Chart Design उपमेनुभित्र जाने र Data Labels को कुनै एक विकल्पलाई Click गर्ने । यसमा एकपटक क्लिक गर्दा त्यस्ता तथ्याङ्क देखिन्छन् भने अर्कोपटक क्लिक गर्दा हट्दछन् । अन्य विकल्पहरूमा पनि क्लिक गरेर थप ज्ञान हासिल गर्न सकिन्छ ।

घ) अक्षको शीर्षक राख्ने: यसका लागि चार्टलाई सेलेक्ट गरेर Chart Design उपमेनुभिन्न जाने र Axis Titles को कुनै एक विकल्पलाई (Horizontal or Vertical Axis) Click गर्ने । त्यसपछि चार्टमा थपिएको Text Box मा गएर अक्षको शीर्षक टाइप गर्ने ।

ङ) बारका रङ्ग परिवर्तन गर्ने: यसका लागि कुनै एउटा बारमा गएर क्लिक गर्ने । यसो गर्दा सबै बारहरू सेलेक्ट हुन्छन् (कुनै विशेष बार मात्र सेलेक्ट गर्नका लागि पुनः त्यही बारमा गएर क्लिक गर्ने) । त्यसपछि Right Click गर्ने । त्यसपछि Format Data Series... क्लिक गर्ने । त्यसपछि रङ्गको बाल्टीजस्तो बटन (Fill and Line) मा क्लिक गर्ने । अब Fill भित्र गएर Solid Fill क्लिक गरेर तलको Color भित्र रहेका विभिन्न रङ्गमध्ये कुनै एक छान्ने ।

यसैगरी अन्य धेरै फर्मेटिङका कार्यहरू गरेर चार्टलाई सूचनामूलक, स्पष्ट र आकर्षक बनाउन सकिन्छ । यहाँ फर्मेटिङका सबै विकल्प वर्णन गरिएको छैन । पाठकवर्गलाई यसमा उपलब्ध विभिन्न विकल्पहरूलाई प्रयोग गरेर थप अभ्यास गर्न आग्रह गरिएको छ । माथि भनिबमोजिम कार्य गर्दा देहायको चित्र नं. १७ बमोजिमको चार्ट तयार भएको छ । यस चित्रको शीर्षक, तथ्याङ्कको स्रोत र अन्य न्यूनतम अवयवहरू उपलब्ध गराइएको हुनाले जो कोहीले चार्टलाई सजिलै बुझ्दछन् । यस चित्रबाट तथ्याङ्कको विभिन्न आयामबाट वर्णन गर्न सकिन्छ । खासगरी बार डायग्राम तुलनात्मक अध्ययनका लागि निकै उपयोगी हुन्छ । जस्तै तलको चित्र नं. २७ मा आर्थिक वर्ष २०७६/७७ मा सबैभन्दा धेरै धानको उत्पादन मधेश प्रदेश मा भएको देखिन्छ भने सबैभन्दा कम कर्णाली प्रदेशमा देखिन्छ । त्यसैगरी प्रदेश १ र लुम्बिनी प्रदेशमा उक्त वर्ष धानको उत्पादन लगभग लगभग उस्तै देखिन्छ । यस चार्टबाट तुलनात्मक अध्ययनका साथसाथै निरपेक्ष अध्ययन पनि गर्न सकिन्छ । जस्तै आर्थिक वर्ष २०७६/७७ मा गण्डकी प्रदेशको धानको उत्पादन ४,०८,७३७ मे.टन रहेको थियो ।

चित्र नं. २७: प्रदेशअनुसार धानको उत्पादन
धानको उत्पादन (मे.टन)



स्रोत: कृषि तथा पशुपन्छी विकास मन्त्रालय, २०७६/७७

एक्सेलले धेरै प्रकारका चार्ट र ग्राफ बनाउन अवश्य मद्दत गर्दछ । तर, प्रश्न उठ्छ कस्तो अवस्था र कस्तो प्रकारको तथ्याङ्कको लागि कुन ढाँचाको चार्ट प्रयोग गर्ने भन्ने सम्बन्धमा पाठ ८ मा संक्षेपमा उल्लेख गरिएको छ ।

थप उदाहरण: एक्सेलमा चार्टको प्रयोग गरेर आउटलायर्स पत्ता लगाउने तरिका

आर्थिक वर्ष २०७६/७७ मा काठमाण्डौँ जिल्लाका नगरपालिकाहरूमा बी.सी.जी. खोप पाउने बच्चाहरूको संख्या तलको तालिका २२.४ मा दिइएको छ । यस तालिकामा प्रस्तुत गरिएका तथ्याङ्कमा Outliers छन् कि छैनन् भनेर थाहा पाउन एक्सेलमा Box Plot को प्रयोग गर्न सकिन्छ ।

नोट: कुनै पनि तथ्याङ्क समूहमा अस्वाभाविक ठुलो वा सानो मानलाई Outliers भनिन्छ । यस्ता Outliers हरूले तथ्याङ्क विश्लेषणमा ठुलो प्रभाव पारेका हुन्छन् । Outliers को उपस्थितिमा गरिएको तथ्याङ्क विश्लेषणले सही सूचना दिन सक्दैनन् । तसर्थ, हामीले तथ्याङ्क विश्लेषण गर्नुपूर्व त्यस्ता Outliers पत्ता लगाएर तिनीहरूको ठाउँमा अंकगणितीय मध्यक राख्ने वा प्रशस्त तथ्याङ्क छन् भने तिनीहरूलाई हटाएर विश्लेषण गर्ने गर्नुपर्दछ ।

तालिका २२.४: काठमाण्डौँ जिल्ला अन्तर्गतका नगरपालिकाहरूमा बी.सी.जी. खोप पाएका बच्चा संख्या

नगरपालिका	बी.सी.जी. खोप पाएका बच्चा संख्या
शंखरापुर नगरपालिका	२०६
कागेश्वरी मनहरा नगरपालिका	८२८

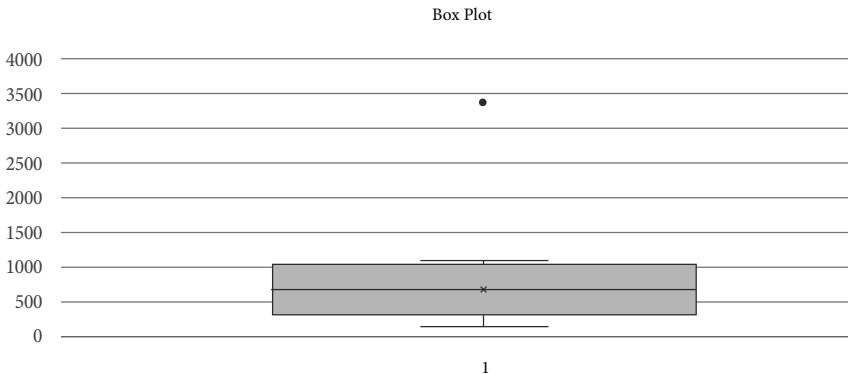
गोकर्णेश्वर नगरपालिका	३,३६६
बुढानिलकण्ठ नगरपालिका	१,००४
टोखा नगरपालिका	१,०९४
तारकेश्वर नगरपालिका	४०८
नागार्जुन नगरपालिका	३३४
कीर्तिपुर नगरपालिका	८२४
चन्द्रागिरी नगरपालिका	५४६
दक्षिणकाली नगरपालिका	१३४

स्रोत: स्वास्थ्य सेवा विभाग, २०७६/७७

एक्सेलको प्रयोगबाट Outliers पत्ता लगाउने विधि

- क) दिइएका तथ्याङ्कहरू सेलेक्ट गर्ने,
 ख) Insert उपमेनुमा गई यस भित्रको Insert Statistics Chart भित्र जाने,
 ग) त्यसपछि Box and Whisker मा क्लिक गर्ने ।
 घ) चित्र नं. २८ बमोजिमको Box Plot तयार हुन्छ ।

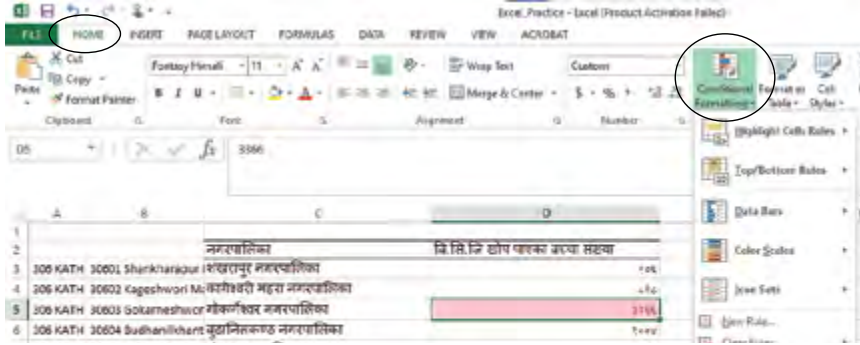
चित्र नं. २८: बक्स प्लट



यस बक्स प्लटमा एउटा बिन्दु निकैमाथि देखिएको छ । यो बिन्दु अन्य आँकडा रहेको आयात (Rectangle) र बोर्डर रेखाभन्दा बाहिर रहेकोले यो Outliers हो । अब दिइएको तथ्याङ्कमा यो बिन्दु कुन हो भनेर हेर्नका लागि Chart Design भित्रको Add Chart Element भित्र Data Labels विकल्पमा गएर कुनै एउटा विकल्प क्लिक गरेपछि उक्त बिन्दुले कुन आँकडा प्रतिनिधित्व गरेको छ भन्ने कुरा थाहा पाउन सकिन्छ । यस उदाहरणमा उक्त आँकडा ३३६६ हो (चित्र २९) । फेरि, सुरूको तालिकामा वा

तथ्याङ्कमा ३,३६६ कहाँ छ वा कुन हो भनेर पत्ता लगाउन हामीले Home भित्रको Conditional Formatting को प्रयोग गर्न सकिन्छ । साना साना आकारका तथ्याङ्कमा यस्ता आँकडा हेरेरै पनि थाहा पाउन सकिन्छ भने ठुला ठुला आकारका तथ्याङ्कमा यस्ता आँकडा पत्ता लगाउन Conditional Formatting को प्रयोग गर्नुपर्दछ ।

चित्र. २९: एक्सेलमा Conditional Formatting गर्ने तरिका



जस्तै: यस उदाहरणमा ३,३६६ कहाँ छ भनेर हेरेरै थाहा पाउन सकिन्छ । यो आँकडा गोकर्णेश्वर नगरपालिकाको हो । Conditional Formatting गर्ने विधि निम्नानुसार छ ।

पहिला दिइएको तथ्याङ्क तालिका सेलेक्ट गर्ने,

Home भित्रको Conditional Formatting भित्र जाने,

Highlight Cells Rule भित्रको Equal to.. विकल्प छान्ने,

Format Cells That Equal to भन्ने ठाउँमा 3366 टाइप गर्ने र with light red fill with dark red text सेलेक्ट गर्ने,

त्यसपछि तालिका २२.५ मा देखाए जस्तै सम्बन्धित Outliers लाई रङ्गिन गराएर देखाउँदछ ।


तालिका २२.५ बी.सी.जी. खोप पाएका बच्चा संख्याको आउटलायस

नगरपालिका	बी.सी.जी. खोप पाएका बच्चा संख्या
शंखरापुर नगरपालिका	२०६
कागेश्वरी मनहरा नगरपालिका	८२८
गोकर्णेश्वर नगरपालिका	३,३६६

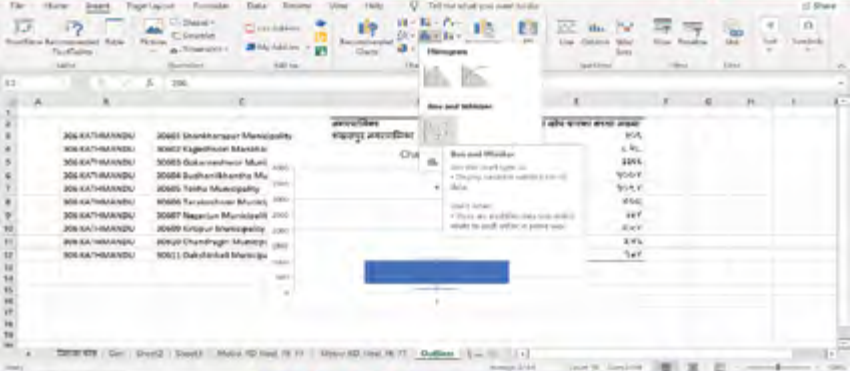
बुढानिलकण्ठ नगरपालिका	१,००४
टोखा नगरपालिका	१,०९४
तारकेश्वर नगरपालिका	४०८
नागार्जुन नगरपालिका	३३४
कीर्तिपुर नगरपालिका	८२४
चन्द्रागिरी नगरपालिका	५४६
दक्षिणकाली नगरपालिका	१३४

स्रोत: स्वास्थ्य सेवा विभाग, २०७६/७७

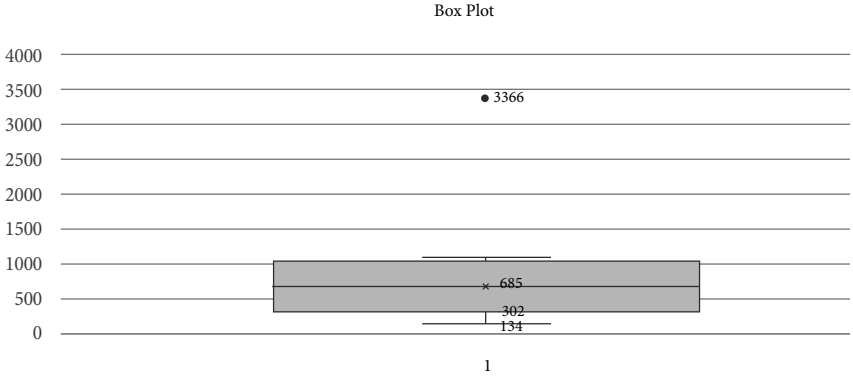
चित्र नं. ३०: एक्सेलको प्रयोग गरी Outliers पत्ता लगाउने विधि



नगरपालिका	वि.सि.ति.को पाएक बट्टा संख्या
KATHMANDU 30601 Shankarapur Municipality	२०५
KATHMANDU 30602 Kageshwori Manahara Municipality	६२८
KATHMANDU 30603 Gokarneshwor Municipality	२२६६
KATHMANDU 30604 Budhanilkantha Municipality	१००४
KATHMANDU 30605 Tokha Municipality	१०९४
KATHMANDU 30606 Tarakeshwar Municipality	४०८
KATHMANDU 30607 Nagarjun Municipality	३३४
KATHMANDU 30609 Kiripur Municipality	८२४
KATHMANDU 30610 Chandragiri Municipality	५४६
KATHMANDU 30611 Dakshinkali Municipality	१३४



नगरपालिका	वि.सि.ति.को पाएक बट्टा संख्या
306-KATHMANDU 30601 Shankarapur Municipality	२०५
306-KATHMANDU 30602 Kageshwori Manahara Municipality	६२८
306-KATHMANDU 30603 Gokarneshwor Municipality	२२६६
306-KATHMANDU 30604 Budhanilkantha Municipality	१००४
306-KATHMANDU 30605 Tokha Municipality	१०९४
306-KATHMANDU 30606 Tarakeshwar Municipality	४०८
306-KATHMANDU 30607 Nagarjun Municipality	३३४
306-KATHMANDU 30609 Kiripur Municipality	८२४
306-KATHMANDU 30610 Chandragiri Municipality	५४६
306-KATHMANDU 30611 Dakshinkali Municipality	१३४



२२.६ अभ्यासका लागि प्रश्नहरू

क) कुनै एक उपयुक्त विकल्पमा ठिक चिन्ह लगाउनुहोस् ।

१. एक्सेलमा रहेको Formula उपमेनुको उपयोग के हो ?

क) तालिका निर्माण

ख) ग्राफ निर्माण

ग) फरम्याटिंग

घ) फरमुलाको प्रयोगद्वारा चाहिएको गणितीय वा तथ्याङ्कीय हिसाबहरू गर्न

२. एक्सेलको सिटमा विभिन्न सेलमा भएको फिल्डहरूलाई सेलेक्सन गरेपछि आन्तरिक रूपमा प्रोसेसिंग गरेर एक्सेलले आफैँ दिने ३ वटा कुराहरू के के हुन् ?

क) Average, Count र Aggregate

ख) Average, Maximum Value र Minimum Value

ग) Average, Count र Sum

घ) Maximum, Minimum र Count

३. एक्सेलमा कुनै पनि Formula प्रयोग गर्दा सम्बन्धित सेलमा सबैभन्दा पहिले नथिची नहुने चिन्ह कुन हो ?

क) + (Plus)

ख) - (Minus)

ग) / (Slash)

घ) = (Equal)

४. चार्टको प्रयोग किन गरिन्छ ?

क) तथ्याङ्क प्रस्तुत गर्न

- ख) तथ्याङ्कको प्रवृत्ति र विशेषता देखाउन
 ग) सहजरूपमा सबैले बुझ्ने गरी तथ्याङ्क प्रस्तुत गर्न
 घ) माथिका सबै

५. पाइ चार्टको प्रयोग कहाँ उपयुक्त हुन्छ ?

- क) समय अन्तरालमा भएको परिवर्तन देखाउन
 ख) बार ग्राफ र रेखाचित्र उपयोगी नहुने सन्दर्भमा
 ग) उतारचढावयुक्त तथ्याङ्क भएको अवस्थामा
 घ) समग्रलाई खण्ड खण्डमा तिनको मानसहित देखाउनु पर्दा

६. चार्ट अध्ययन गर्दा कुन कुरामा ध्यान दिनु जरूरी छ ?

- क) शीर्षक
 ख) संकेत सूची
 ग) सन्दर्भ समय र स्रोत
 घ) माथिका सबै

ख) गुल्मी जिल्लाको चन्द्राकोट गाउँपालिकाका वडा नम्बरहरू Row मा र वडास्तरको जनसंख्या (लिङ्गानुसार) Column मा राख्नुहोस् र त्यसलाई Transpose गर्नुहोस् । यस तथ्याङ्कलाई उपयुक्त विधिको प्रयोग गरेर चार्टमा प्रस्तुत गर्नुहोस् । तथ्याङ्क प्राप्त गर्नका लागि <https://cbs.gov.np/population-2011/> मा जानुहोस् ।

२२.७ सन्दर्भ सामग्री

एक्सेलको थप अध्ययनका लागि पाठक वर्गलाई युट्युबको प्रयोग गर्न सल्लाह दिइएको छ । युट्युबमा गएर How to plot pie chart, line chart अथवा यस्तै अरू कुनै चार्टको खोजी गरेर अध्ययन गर्न सकिन्छ ।

नोट: यस पाठमा हामीले एक्सेलका केही विशेषताहरूको मात्र प्रयोग गरेर अभ्यास गरेका छौं । यति अध्ययन गरिसकेपछि पाठकवर्गले स्वयं Hit and Trial को विधिबाट एक्सेलका विभिन्न विशेषता एवं प्रयोगका बारेमा थप अध्ययन गर्न सक्नुहुनेछ ।

स्टेटासम्बन्धी आधारभूत ज्ञान

२३.१ STATA को परिचय र यसका आधारभूत विशेषताहरू

STATA डाटा व्यवस्थापन र विश्लेषण अर्थात् डाटा विज्ञान (Data Science) को लागि उपयोगी हुने सफ्टवेयर हो । यो सफ्टवेयर आदेश (Command) र मेनु (Menu) दुवै प्रणालीमा आधारित छ । यसले गर्ने तथ्याङ्कीय विश्लेषणका नतिजाहरू निकै नै विश्वसनीय मानिन्छ । यो सफ्टवेयर प्रयोग गर्न निकै सजिलो छ र यसका कमाण्डहरू अनलाइनमा प्रशस्त भेटिन्छन् । साथै यो सफ्टवेयरसँग यसको प्रयोग र विभिन्न तथ्याङ्कीय कार्य गर्ने तरिकाका बारेमा निर्देशिका पनि संलग्न रहेको हुन्छ । यस पाठमा हामीले STATA का सामान्य परिचयात्मक अध्ययन गर्नेछौं र यो पाठ पढिसकेपछि STATA बाट तथ्याङ्क विश्लेषणका धेरै कार्य गर्न सकिन्छ ।

२३.१.१ STATA का सामान्य विशेषताहरू

कम्प्युटरमा स्टेटा खोलिसकेपछि चित्र २३.१ मा देखाइएको जस्तो विन्डो देखिन्छ । चित्रमा देखाइएको बाहेक डाटा ब्राउजर, डाटा इडिटर, ग्राफ, भिउअर, डु फाइलजस्ता विन्डोहरू हुन्छन् । यी विन्डोहरूलाई सम्बन्धित बटनमा थिचेर निकाल्न सकिन्छ ।

चित्र नं.: २३.१ स्टेटाको विन्डा



स्टेटा लेआउटसम्बन्धी केही टिप्सहरू:

फन्ट परिवर्तन गर्न: स्टेटा विन्डोमा Right Click गरेर Font भित्र आवश्यकताअनुसार फन्ट छान्न सकिन्छ ।

स्टेटा विन्डोको रङ्ग परिवर्तन गर्न: स्टेटा विन्डोमा Right Click गरेर Preferences भित्रको Background Color बाट मनपर्ने रङ्ग छान्न सकिन्छ ।

मेमोरी र चर तय गर्न:

- स्टेटाको Command विन्डोमा memory टाइप गरेर इन्टर थिच्ने । नतिजा विन्डोमा मेमोरी सम्बन्धित विवरण देखिन्छ । साधारणतया पछिल्ला भर्सनमा मेमोरी परिवर्तन गरिरहनु पर्दैन ।
- चरहरूको संख्या निर्धारण गर्न Command विन्डोमा set maxvar टाइप गरेर चाहिएको संख्या टाइप गरेर इन्टर थिच्नुपर्दछ । यस्तो संख्या २,०४८ र १,२०,००० को बिचको हुनुपर्दछ ।

डाइरेक्टोरी हेर्न र परिवर्तन गर्न: स्टेटाको Command विन्डोमा dir टाइप गरेर इन्टर थिच्ने । नतिजा विन्डोमा आफूले काम गरिरहेको डाइरेक्टोरी भित्रका फाइलहरू देखिन्छ । यसलाई परिवर्तन गर्न परेमा Command विन्डोमा cd "DM\CBS_official\Agri_census" टाइप गरेर इन्टर थिच्ने । यहाँ स्टेटाको कमाण्ड भनेको cd हो र " " भित्र लेखिएको शब्दहरूले फाइलको लोकेसन बताएको हो ।

नोट: स्टेटाका कमाण्डहरू टाइप गरिसकेपछि तिनीहरूको नतिजा हेर्नका लागि इन्टर थिच्नुपर्दछ । यस पाठमा प्रत्येक कमाण्डपछि इन्टर थिच्ने नभनिए तापनि अभ्यास गर्ने बेला इन्टर थिच्नुपर्दछ भनेर बुझ्नुपर्दछ ।

२३.२ स्टेटामा डाटा लोड गर्ने तरिका, तथ्याङ्कको आयात तथा निर्यात

स्टेटामा डाटा लोड गर्ने विभिन्न तरिका हुन्छन् ।

फाइल मेनुबाट फाइल खोलेर स्टेटाका डाटा लोड गर्न सकिन्छ ।

स्टेटाका फाइलहरू सोभै डबल क्लिकबाट पनि लोड गर्न सकिन्छ ।

अन्य डाटाहरू जस्तै: एक्सेल, Text data in fixed format आदिलाई फाइल मेनुभित्रको Import बाट डाटा लोड गर्न सकिन्छ ।

नोट: खुलिरहेको स्टेटाको मेमोरीमा डाटा हुनसक्दछ । यदि हामीले नयाँ डाटा फाइल लोड गर्ने हो भने मेमोरीमा भएको डाटा हटाउनु पर्दछ । यसका लागि clear कमाण्ड प्रयोग गरिन्छ । स्टेटा प्रोग्राम नै बन्द गर्नु परेको खण्डमा भने exit कमाण्ड प्रयोग गर्नुपर्दछ ।

२३.३ मेनुको परिचय

स्टेटा खास गरी कमाण्डको प्रयोग गरेर प्रयोग गर्ने सफ्टवेयर भए तापनि धेरै कार्यहरू मेनुको प्रयोगबाट पनि गर्न सकिन्छ । मेनुमा उपलब्ध केही प्रमुख सुविधा र तिनीहरूको प्रयोगका बारेमा तालिका २३.१ मा दिइएको छ ।

तालिका २३.१: स्टेटाको मेनुबारको प्रयोग

मेनु	गर्ने काम
File	फाइल खोल्न, सेभ गर्न, लग फाइलमा कमाण्डहरू सेभ गर्न, डाटा फाइलको आयात तथा निर्यात गर्न, प्रिन्ट गर्न, बाहिर निस्कन आदि ।
Edit	कपी पेष्ट गर्न, नतिजा विन्डो सफा गर्न, केही कुरा खोज्न आदि ।
Data	डाटाको वर्णन गर्न, हेर्न, डाटा सम्पादन गर्न, चरको नाम परिमार्जन गर्न, नयाँ चर बनाउन, दुई डाटासेट जोड्न, आदि ।
Graphics	ग्राफ बनाउन र सेभ गर्न ।
Statistics	तथ्याङ्कको विश्लेषण गर्न जस्तै: तालिका निर्माण, मध्यक मध्यिका निकाल्न, प्रतिगमन विश्लेषण आदि ।
User	प्रयोगकर्ताले परिभाषित गरेका कार्य सुचारू गर्न ।
Window	विभिन्न प्रकारका विन्डो खोल्न बन्द गर्न
Help	स्टेटासम्बन्धी विस्तृत अध्ययन गर्न

तालिका २३.२: स्टेटाका आधारभूत केही कमाण्डहरू

कमाण्ड	गर्ने काम
dir	काम गरिरहेको डाइरेक्टोरीमा के के फोल्डर छन् भनेर हेर्न ।
cd	काम गरिरहेको डाइरेक्टोरी परिवर्तन गर्न ।
memory	मेमोरीसम्बन्धी जानकारी प्राप्त गर्न ।
sum	तथ्याङ्कको सारांश जानकारी लिन ।
des	चरको बारेमा जानकारी लिन ।
set more off	नतिजा विन्डोमा लामो नतिजा आइरहने कार्यलाई रोक्न ।
clear	मेमोरीमा भएका डाटा हटाउन ।
save	फाइल सेभ गर्न ।
use	फाइल खोल्न ।
exit	स्टेटाबाट बाहिर निस्कन ।
preserve	स्टेटाको सुरूको डाटालाई मेमोरीमा राखिराख्न प्रयोग गरिन्छ । स्टेटाको प्रयोगकर्ताले काम सुरू गर्नुपूर्व यो कमाण्डको प्रयोग गर्नु राम्रो मानिन्छ ।
restore	preserve गरिएका डाटा पुनः प्राप्त गर्न ।

h	help छोटो रूप हो । केही कुराको बारेमा मद्दत चाहिएमा । जस्तै: set-more off ले के काम गर्दछ भनेर जान्न मन लाग्यो भने h set more off टाइप गरेर इन्टर थिच्नुपर्दछ । यो कमाण्ड स्टेटामा अत्यन्त उपयोगी कमाण्ड हो तसर्थ, माथिका कमाण्डहरूले के काम गर्दछन् भनेर help हेरौं । h dir, h cd, h memory, ... ।
br	browse को छोटो रूप हो । यो कमाण्डको प्रयोगले स्टेटामा डाटा हेर्न सकिन्छ ।

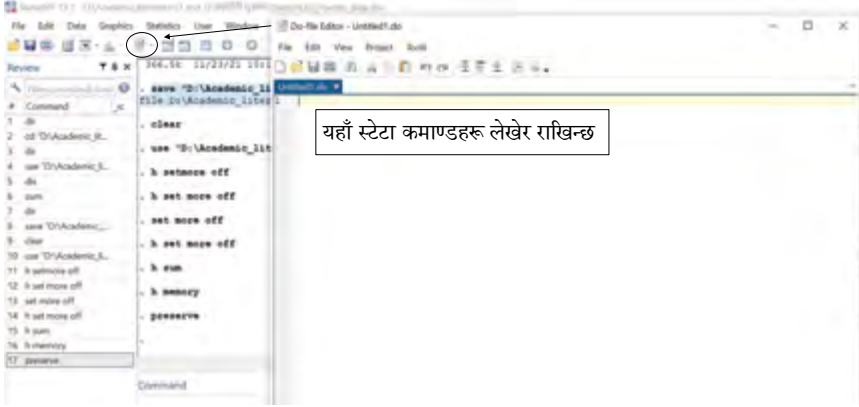
नोट: कतिपय कमाण्डहरू शब्दको पूरा रूप नलेखी संक्षिप्त रूपमा पनि प्रयोग गर्न सकिन्छ ।

जस्तै: h, br, dir, sum,

२३.४ STATA कमाण्डहरू, “Do” फाइल र लग फाइल

स्टेटामा कमाण्डको प्रयोग गरेर विभिन्न कार्यहरू गर्न सकिने बारेमा केही जानकारी प्राप्त भइसकेको छ । यस्ता कमाण्डहरू भविष्यमा प्रयोग गर्ने उद्देश्यले व्यवस्थित रूपमा संग्रह गरी राख्ने फाइललाई Do फाइल भनिन्छ । उदाहरणका लागि नेपाल जीवनस्तर सर्वेक्षण, २०१०/११ को डाटा व्यवस्थापन गरी गरिबी मापन गर्दा प्रयोग गरिएका कमाण्डहरू सिलसिलेवार Do फाइलमा राखियो भनेपछि जो कोहीले कुनै पनि बेला डाटा फाइल र Do फाइल प्रयोग गरेर नतिजा हेर्न वा निकाल्न सक्दछन् । यसरी Do फाइल बनाएर राख्दाको अर्को फाइदा भनेको डाटा व्यवस्थापनका उस्तै उस्तै कार्यहरूलाई सरल र सहज बनाउनु पनि हो । तसर्थ, स्टेटामा कुनै काम गर्दा प्रयोग गरिएका कमाण्डहरूको संग्रह बनाएर Do फाइल सुरक्षित साथ राख्नुपर्दछ । Do फाइल बनाउनका लागि मेनुबार मुनिको टुलबार बाट New do file editor क्लिक गर्न सकिन्छ । यसमा क्लिक गरेपछि एउटा नयाँ विन्डो प्राप्त हुन्छ र त्यसमा हामीले प्रयोग गरेका कमाण्डहरू सिलसिलेवार राख्न सकिन्छ । यो फाइललाई सेभ गरेर सुरक्षित राख्नुपर्दछ । यसो गरेपछि डाटा फाइल सेभ गरिरहनु पर्दैन । पछि यिनै सुरक्षित कमाण्डहरू प्रयोग गरेर तथ्याङ्क व्यवस्थापनका काम गर्न सकिन्छ ।

चित्र नं. २३.२: स्टेटामा Do file editor विन्डोको स्वरूप



स्टेटाका सबै क्रियाकलाप रेकर्ड गरेर राख्ने फाइललाई लग फाइल भनिन्छ । स्टेटा खोलेर लग फाइल बनाउन log using “फाइलको लोकेसन र नाम दिने” भन्ने कमाण्ड प्रयोग गरिन्छ भने log close ले खुलेको लग फाइल बन्द गर्दछ ।

नोट:

- स्टेटाका कमाण्ड case sensitive हुन्छन् । सबै अक्षर अंग्रेजी साना अल्फाबेट प्रयोग गर्नुपर्दछ ।
- स्टेटामा चरका नाम case sensitive हुन्छन् । यिनीहरूले space स्वीकार गर्दैनन् । साना ठुला अक्षर जे भए पनि कमाण्ड लेख्ने बेला दुरूस्त टाइप गर्ने वा चरको सूचीबाट क्लिक गरेर प्रयोग गर्नुपर्दछ ।

उदाहरण: चरको सूचीमा लिङ्गलाई sex ले जनाइएको छ र यसको बारेमा केही जानकारी लिनु पर्‍यो भने, des sex नै लेख्नुपर्दछ । des Sex लेखेर हुँदैन (यहाँ des स्टेटाको कमाण्ड हो भने sex कुनै एउटा चर हो) । des sex कमाण्ड दिनु भनेको स्टेटालाई डाटामा भएको sex भन्ने चरको बारेमा केही जानकारी देऊ भनेको हो । यसको जानकारी नतिजा विन्डोमा हेर्न पाइन्छ ।

- स्टेटालाई क्याल्कुलेटरका रूपमा पनि प्रयोग गर्न सकिन्छ । जस्तै : २ र ९ जोड्न पर्‍यो भने dis 2 + 9 गर्नुपर्दछ । अन्य गणितीय क्रियाका चिन्हहरू तालिका २३.३ मा दिइएको छ ।

अभ्यास: NLSS_2010_11 फोल्डर भित्रको STATA9 फोल्डरमा भएको xh01_s01 फाइल खोलेर v01_02 भन्ने चर हेरौं । यो चर sex को लागि प्रयोग गरिएको छ । अब, des v01_02, des V01_02, v01 02 जस्ता कमाण्डहरू

प्रयोग गरेर हेरौं । प्राप्त नतिजालाई अध्ययन मनन गरौं । स्टेटाले लिन नमानेका गलत कमाण्डहरूको नतिजाका सन्देश राता अक्षरमा प्राप्त हुन्छ भने सही कमाण्डका नतिजा काला अक्षरमा प्राप्त हुन्छ ।

गणितीय क्रिया: स्टेटाले गणितीय क्रियाका लागि निम्नानुसारका चिन्हहरू प्रयोग गर्दछ ।

तालिका २३.३: स्टेटामा प्रयोग हुने गणितीय चिन्ह र प्रयोग

चिन्ह	प्रयोग	चिन्ह	प्रयोग	चिन्ह	प्रयोग
+	जोड	^	घाताङ्क	>	भन्दा ठुलो
-	घटाउ	&	र	<	भन्दा सानो
*	गुणन		अथवा	>=	बराबर वा ठुलो
/	भाग	!	हैन	<=	बराबर वा सानो
==	बराबर	!=	बराबर छैन		

तालिका २३.४: स्टेटाका थप केही क्रियाहरू

क्रियाहरू	के गर्दछ ?	उदाहरण
dis abs(x)	x लाई सकारात्मक बनाउँछ ।	dis abs (-2) = 2
dis exp(x)	e को x घाताङ्क हिसाब गर्दछ ।	dis exp(1) = 2.7182818
dis ln(x)	x को प्राकृतिक लग निकाल्दछ ।	dis ln(10) = 2.3025851
dis log(x)	x को प्राकृतिक लग निकाल्दछ ।	dis log(10) = 2.3025851
dis log10(x)	x को 10 base लग निकाल्दछ ।	dis log10(10) = 1
dis sqrt(x)	x को वर्गमूल निकाल्दछ ।	dis sqrt(4) = 2
dis int(x)	x को दशमलव पछाडिका संख्या हटाउँदछ ।	dis int(3.67) = 3
dis round(x)	x को rounded मान देखाउँदछ ।	dis round(3.67) = 4

२३.५ STATA मा डाटा व्यवस्थापन

स्टेटाको प्रमुख कार्य भनेकै तथ्याङ्क व्यवस्थापन हो । तथ्याङ्क व्यवस्थापनमा अनगिन्ती कार्यहरू पर्दछन् । तीमध्ये प्रमुख केही कार्यका बारेमा यहाँ उल्लेख गरिएको छ । यस खण्डको अध्ययन गर्न उदाहरणसहित अभ्यास गर्ने उद्देश्यले निम्न बमोजिमको स्टाटा फाइलको डाटा प्रयोग गरौं ।

डाटा फाइल: NLSS_2010_11 फोल्डरभित्रको STATA9 फोल्डरमा भएको xh01_s01 फाइल

क) चरका फर्मेट मिलाउने:**कमाण्ड:** form

उदाहरण: *form v01_03%4.2f* यसले चर *v01_03* लाई चार अंकको संख्या बनाउँदछ जसमा दशमलव पछि २ संख्या हुन्छ ।

form v01_03 %06.0f यसले चर *v01_03* लाई शून्यबाट सुरू भई ६ अंकको संख्याको बनाउँदछ जसको दशमलव पछि अंक हुँदैन ।

ख) चरका गुण हेर्ने:**कमाण्ड:** describe

उदाहरण: *des v01_02* (यहाँ, चर *v01_02* ले लिङ्ग जनाएको छ)

variable name: चरको नाम (space हुनु हुँदैन, 32 अक्षरभन्दा कम)

storage type: चरको प्रकार (byte, int, long, float, double, str)

display format: सांख्यिक फर्मेटका रूप (g, f, e, s)

value label: ex: १= पुरुष, २= महिला (*v01_02*)

variable label: चरको नामका बारेमा विस्तृत विवरण (*1.02 sex*)

ग) चरको प्रकार:

स्टेटाले दुई प्रकारका चरहरूलाई डाटाका रूपमा लिन्छ । एउटा सांख्यिक (Numeric) हो भने अर्को गैरसांख्यिक (String) हो । एक प्रकारको चरलाई अर्कोमा परिवर्तन गर्नका लागि *tostring* अथवा *destring* कमाण्ड प्रयोग गर्नुपर्दछ ।

कमाण्ड: *tostring* (सांख्यिक चरलाई गैरसांख्यिक चर बनाउँदछ)

उदाहरण: *tostring v01_03, replace* (परिवर्तन भयो वा भएन भनेर जाँच गर्न *des v01_03* कमाण्ड प्रयोग गरेर हेर्न सकिन्छ । यस कमाण्डमा कमा (,) पछि राखिने शब्द विकल्प हो । सांख्यिक चरलाई गैरसांख्यिक चर बनाइसकेपछि पुरानो चरलाई यथावत राखेर नयाँ गैरसांख्यिक चर बनाउन पनि सकिन्छ । यसका लागि *tostring v01_03, generate* (चरको नयाँ नाम लेख्ने) । यसरी चरको प्रकार परिवर्तन गर्दा पहिलेको चर नै परिवर्तन गर्ने वा नयाँ चर सिर्जना गर्ने भन्ने बारे ध्यान पुऱ्याए मात्र कमाण्डको प्रयोग गर्नुपर्दछ ।

कमाण्ड: *destring* (गैरसांख्यिक चरलाई सांख्यिक चर बनाउँदछ)

उदाहरण: *destring v01_03, replace* (यहाँ चर *v01_03* लाई पहिला गैरसांख्यिक बनाएर पुनः सांख्यिक चरमा परिवर्तन गर्न लागिएको छ ।)

डाटा व्यवस्थापनमा चरको प्रकार परिवर्तन निकै महत्त्वपूर्ण हुन्छ । कतिपय चरहरू गैरसांख्यिक रूपमा राखिएका हुन्छन् तर वास्तवमा तिनीहरू सांख्यिक प्रकृतिका हुन् भने यस किसिमको परिवर्तन अपरिहार्य हुन्छ । कतिपय कमाण्डहरू प्रयोग गर्ने बेला पनि गैरसांख्यिक वा सांख्यिक के हो विशेष ध्यान दिनुपर्दछ ।

घ) चर राख्ने र हटाउने:

कमाण्ड: *keep* (यो कमाण्ड प्रयोग गरेर स्टेटा डाटाको चर सूचीमा चाहिने मात्र चरहरू राख्न सकिन्छ ।) जस्तै: यदि हामीलाई *v01_02* / *v01_03* मात्र चाहिएको हो भने *keep v01_02 v01_03* प्रयोग गर्नुपर्दछ ।

कमाण्ड: *drop* (यो कमाण्डले चरका सूचीबाट नचाहिएका चरहरू हटाउन सकिन्छ) । जस्तै यदि हामीलाई चर *v01_02* चाहिएको छैन भने *drop v01_02* प्रयोग गर्नुपर्दछ ।

स्मरण रहोस्, यस किसिमका *keep*, *drop* जस्ता कमाण्ड प्रयोग गर्नुपूर्व *preserve* भन्ने कमाण्ड प्रयोग गर्नु राम्रो मानिन्छ । किनभने पछि पुनः हटाएका चरहरू चाहिएमा *restore* भन्ने कमाण्ड प्रयोग गर्न सकिन्छ ।

ङ) अन्य केही उपयोगी कमाण्ड र तिनको प्रयोग तलको तालिकामा दिइएको छ ।

तालिका २३.५: डाटा व्यवस्थापनका केही स्टेटा कमाण्डहरू

कमाण्ड	काम	उदाहरण
<i>sort</i>	चरका मानलाई सानोबाट ठुलो वा ठुलोबाट सानो क्रम मिलाउन । यसमा Option "stable" प्रयोग गरियो भने उही मान भएका मानहरूलाई पहिला जुन क्रममा छ त्यही क्रममा राखिदिन्छ । कमाण्ड: <i>sort v01_03, stable</i>	<i>sort v01_03</i>
<i>gsort</i>	चरका मानलाई आवश्यकताअनुसार सानोबाट ठुलो वा ठुलोबाट सानो क्रम मिलाउन । यदि ठुलोबाट सानो क्रम मिलाउनु पन्थोभने चरको अगाडि "-" चिन्ह प्रयोग गर्नुपर्दछ । यदि क्रम मिलाइएका चरलाई नयाँ चरको रूपमा सिर्जना गर्नुपन्थो भने, अप्सन प्रयोग गर्नुपर्दछ । जस्तै: <i>gsort v01_03, gen (age) mfirst</i> यहाँ, <i>mfirst</i> ले <i>missing</i> मानहरूलाई पहिला राख् भन्ने बुझाउँदछ ।	सानोबाट ठुलो: <i>gsort v01_03</i> ठुलोबाट सानो: <i>gsort -v01_03</i>

keep if	<p>कुनै विशेष मान लिनेहरूलाई मात्र राख्ने । जस्तै: महिलाको मात्र अध्ययन गर्नुपन्थो भने यो कमाण्ड प्रयोग गरिन्छ । यो कमाण्ड प्रयोग गर्दा ==, >=, <=, != >, < जस्ता चिन्हहरू प्रयोग गरिन्छ ।</p>	<p><i>keep if v01_02==2</i> यहाँ चर v01_02 लिङ्ग हो र 2 भनेको महिला हो ।</p>
drop if	<p>कुनै विशेष मान लिनेहरूलाई मात्र हटाउने । जस्तै महिला हटाउनु पन्थो भने यो कमाण्ड प्रयोग गरिन्छ । यो कमाण्ड प्रयोग गर्दा पनि ==, >=, <=, != >, < जस्ता चिन्हहरू प्रयोग गरिन्छ ।</p>	<p><i>drop if v01_02==2</i> यहाँ चर v01_02 लिङ्ग हो र 2 भनेको महिला हो ।</p>
generate	<p>नयाँ चर निर्माण गर्दछ । जस्तै: व्यक्तिगत उमेरको डाटाबाट उमेर समूह निर्माण गर्न यो कमाण्डको प्रयोग गर्न सकिन्छ । यसका लागि अर्को कमाण्ड <i>replace</i> पनि प्रयोग गर्नुपर्दछ । gen agegrp = 1 if v01_03>=0& v01_03<=14 यो कमाण्डले १४ वर्ष माथिका व्यक्तिको हकमा खाली राखिदिन्छ, तसर्थ, तिनीहरूलाई <i>replace</i> गर्ने अर्को कमाण्ड चाहिन्छ ।</p> <p>कुनै गणितीय क्रियाबाट पनि नयाँ चर सिर्जना गर्न सकिन्छ । जस्तै उमेरको वर्ग निकालेर नयाँ चर बनाउन gen agesq = v01_03^2</p>	<p>gen agegrp = 1 if v01_03>=0& v01_03<=14</p> <p>replace agegrp = 2 if v01_03>=15 & v01_03<=59</p> <p>replace agegrp = 3 if v01_03>=60 & v01_03<=98</p> <p>gen ageraise = v01_03+5 (प्रत्येक व्यक्तिको उमेर ५ वर्षले बढाइएको)</p>
egen	<p>मध्यक, योगफल जस्ता फङ्सनको प्रयोग गरेर नयाँ चर सिर्जना गर्न । यो कमाण्ड प्रयोग गर्नुपूर्व जुन चरको फङ्सन निकाल्ने हो त्यसलाई <i>sort</i> गर्नुपर्दछ । यहाँ दिइएको उदाहरणमा <i>unid</i> भन्ने चर डाटा सेटमा छैन । यसले <i>unique identity</i> भन्ने बुझाएको छ । यो बनाउन <i>xhpsu</i>, <i>xhnum</i> र <i>xstra</i> प्रयोग भएको छ । यसका लागि तलको नोट हेर्नुहोस् ।</p>	<p>sort xhnum egen hhsiz = count(v01_02), by (unid)</p>

नोट: डाटा सेटमा युनिक आइडीको निकै ठुलो महत्त्व हुन्छ । यो भनेको कुनै पनि डाटा आइटमलाई अरूसँग नजुध्ने गरी परिचय दिने आइडी हो । नेपाल जीवनस्तर सर्वेक्षणको तथ्याङ्कमा यस्तो किसिमको युनिक आइडी उपलब्ध छैन र हामी आफैँले तयार गर्नुपर्दछ । छनोटमा परेका घरपरिवारको युनिक आइडी बनाउन तिनवटा चर प्रयोग गर्न सकिन्छ । यहाँ, *xhpsu*, *xhnum* र *xstra* क्रमशः प्राथमिक नमुना एकाइ (Primary Sampling Unit = PSU), घरपरिवार संख्या र स्ट्राटा हुन् । यसमा घरपरिवार संख्या युनिक हुन सक्दैन किनभने उही संख्या अर्को स्ट्राटामा पनि प्रयोग

भएको हुन्छ । यी कुराहरू हेर्न हामीले list or tab कमाण्डको प्रयोग गर्न सक्दछौं । यदि हामीले माथि उल्लेख गरिएका तिनवटा चरलाई गाँसेर एउटै बनाउन सक्यौं भने त्यो युनिक आइडी हुन्छ । तर त्यसरी गाँस्नका लागि तिनवटै चरहरू गैरसांख्यिक हुनुपर्दछ । des कमाण्डको प्रयोग गरेर सांख्यिक वा गैरसांख्यिक छुट्टयाएपछि यदि सांख्यिक रहेछन् भने *tostring* चरको नाम, *gen*(नयाँ चरको नाम = a or b or c गर्न सकिन्छ) गरेर गैरसांख्यिक बनाउने । त्यसपछि *gen* *unid* = a + b + c गर्नुपर्दछ । स्मरण रहोस्, a + b + c मा यदि a, b, c सांख्यिक भए भने तिनीहरूको योग प्राप्त हुन्छ, युनिक आइडी प्राप्त हुँदैन । तलका कमाण्ड प्रयोग गरेर युनिक आइडी बनाउने अभ्यास गरौं ।

```
sort xhpsu
sort xhnum
sort xstra
tostring xhpsu, gen(a)
tostring xhnum, gen(b)
tostring xstra, gen(c)
gen unid=a+b+c
```

अब, **egen** *hhsz* = *count*(v01_02), *by* (*unid*) यसको बारेमा बुझौं ।

यहाँ, v01_02 भनेको परिवारका सदस्यको लिङ्ग हो । यदि परिवारअनुसार लिङ्ग गणना गर्ने हो भने त्यो परिवारको जम्मा सदस्य संख्या प्राप्त हुन्छ । तसर्थ, परिवार आकार (*hhsz*) पत्ता लगाउन *egen* भन्ने कमाण्डको प्रयोग गरिएको छ । यही विधिबाट यदि हामीलाई व्यक्तिगत आमदानीको डाटा छ भने पारिवारिक आमदानी पत्ता लगाउन सकिन्छ ।

कमाण्ड	काम	उदाहरण
deduplicates drop	एउटै एकाइका दुई वा सो भन्दा बढी अवलो कन भएमा एउटा मात्र कायम गर्दछ ।	deduplicates drop
label variable	चरलाई चिन्न सजिलो हुने गरी वर्णन गर्ने । यसले ८० वटा अक्षर लिन्छ । धेरै शब्द प्रयोग गर्दा “ ” को प्रयोग गर्नुपर्दछ । यस्तो वर्णनलाई हटाउनु परेमा <i>label variable</i> v01_02 मात्र गरे पुग्छ ।	<i>label variable</i> v01_03 “व्यक्तिको उमेर”
rename	चरको नाम परिवर्तन गर्न ।	rename v01_02 sex
recode	दिइएको नियमको परिधिभित्र रहेर चरका मानहरू परिवर्तन गर्न यो कमाण्ड स्टेटामा तथ्याङ्क व्यवस्थापन गर्न निकै उपयोगी कमाण्ड हो । यसलाई विस्तृत बुझ्नका लागि तलको बक्स हेरौं । (स्टेटा कमाण्डलाई <i>syntax</i> , <i>code</i> भन्ने चलन पनि छ ।	recode sex (1=2) (2=1) (महिलालाई पुरुष र पुरुषलाई महिला बनाइयो ।)

syntax: **recode** varlist (rule)..... ,gen (newvar)

उदाहरण: **recode** v01_03 (min/5 = 1) (6/20 = 2) (21/max = 3) (miss = 9), gen (agegrp)

rule	Example (rule)	अर्थ
# = #	3 = 1	मान 3 लाई 1 बनाउन
## = #	1/5 = 4	मान 1 देखि 5 लाई 4 बनाउन
nonmissing = #	nonmiss = 8	अरू nonmissing लाई 8 बनाउन
missing = #	miss = 9	अन्य missings लाई 9 बनाउन

बक्स २३.१: रिकोडका नियम

कमाण्ड	काम	उदाहरण
label define	चरका मानलाई कुनै शब्दसँग सम्बन्धित बनाउने काम गर्दछ। माथि बक्स २३.१ मा सिजर्जा गरिएको नयाँ चर agegrp लाई कसरी लेवल परिभाषित गर्ने भन्ने सँगैको उदाहरणमा दिइएको छ।	label define agegrp “५ वर्ष भन्दा कम” 2 “६ देखि २० वर्ष” 3 “२१ वर्षभन्दा माथि” ९ “उमेर उल्लेख नगरिएका”
label values	माथि परिभाषित गरिएको लेवललाई सम्बन्धित चरमा लागु गर्न। यहाँ, agegrp भन्ने चरमा agegrp भन्ने लेभल लागु गरिएको छ।	label values agegrp agegrp
order	चरहरूलाई चरको सूचीमा मिलाएर राख्न।	order v01_04 v01_03 v01_02
aorder	अल्फावेटअनुसार मिलाउन।	aorder v01_04 v01_03 v01_02
move	चरलाई एक ठाउँबाट अर्को ठाउँमा लैजान	move v01_01 v01_04
merge	दुईवटा डाटासेटलाई जोड्न। यसका लागि युनिक आइडी आवश्यक पर्दछ। युनिक आइडी बनाउने तरिका हामीले माथि वर्णन गरिसकेका छौं। अब, दुईवटा डाटासेट xh00_s00 र xh01_s01 जोड्ने उदाहरण हेरौं। यो कमाण्ड प्रयोग गर्न सुरुमा दुवै फाइलमा unid बनाउने र कुनै एक फाइल खोलेर अर्को जोड्ने।	merge 1M1 unid using xh00_s00 नोट: दुवै फाइल एउटै लोकेसनमा हुनुपर्दछ, अन्यथा लोकेसन दिनुपर्ने हुन्छ।
append	कुनै डाटा फाइलमा अर्को फाइलका थप मानहरू भण्डारण गर्न। यसमा चर थप हुँदैन तर चरका मानहरू थप हुन्छन्।	append using xh00_s00

२३.६ डाटा संक्षेपीकरण

तथ्याङ्क व्यवस्थापनको एउटा महत्त्वपूर्ण काम भनेको डाटाको संक्षेपीकरण गर्नु हो । डाटाको गन्ती गर्नु, सूची बनाउनु, तालिका बनाउनु जस्ता कार्यहरू यसअन्तर्गत पर्दछन् । यसका लागि प्रयोग हुने कमाण्ड र तिनीहरूको प्रयोगका उदाहरण तलको तालिका २३.६ मा दिइएको छ ।

तालिका २३.६: डाटा संक्षेपीकरण गर्ने कमाण्ड

कमाण्ड	काम	उदाहरण
codebook	यसले एक किसिमको तथ्याङ्कको पनि तथ्याङ्क देखाउँदछ ।	codebook v01_01 codebook v01_02
count if	दिइएको अवस्था पालना गर्ने गरी सम्बन्धित चरका मानहरूको संख्या बताउँदछ ।	count if v01_02==1 (पुरुषको संख्या देखाउँदछ)
tabulate	तथ्याङ्कलाई तालिकामा प्रस्तुत गर्दछ । एकतर्फी वा दुईतर्फी तालिका बनाउन प्रयोग गरिन्छ । परिवारमूलि महिला र पुरुष कति कति छन् भनेर हेर्न पन्थो भने: <i>tab v01_04 v01_02</i>	tab v01_02 tab v01_02 v01_04
tabstat	तथ्याङ्कको सारांश सहितको तालिका बनाउन । (चित्र नं. २३.३)	tabstat v01_03, by(v01_02) stat(mean sd min max)
corr	चरहरूबिचको सहसम्बन्ध हेर्न ।	corr v01_16 v01_03
reg	प्रतिगमन विश्लेषण गर्न । दिइएको उदाहरणमा आमाको माथिल्लो शैक्षिक योग्यता र उमेर बिचको प्रतिगमन विश्लेषण गरिएको छ ।	reg v01_16 v01_03

चित्र नं. २३.३: कमाण्ड **tabstat** को नतिजा

```
. tabstat v01_03 , by( v01_02 ) stat(mean sd min max)
```

```
Summary for variables: v01_03
```

```
by categories of: v01_02 (1.02 sex)
```

v01_02	mean	sd	min	max
male	26.79148	20.84801	0	99
female	27.045	19.62326	0	99
Total	26.92707	20.20228	0	99

२३.७ डाटा भिजुअलाइजेसन

स्टेटाको प्रयोगबाट तथ्याङ्क प्रस्तुतिकरणमा प्रयोग हुने विभिन्न चार्टहरू बनाउन सकिन्छ । केही प्रमुख चार्टहरू बनाउने कमाण्डहरूका बारेमा यहाँ उल्लेख गरिएको छ ।

बक्स २३.१: रिकोडका नियम

कमाण्ड	काम	उदाहरण
graph bar	बार डायग्राम बनाउन । यसले बारहरूलाई प्रतिशत मानमा देखाउँदछ ।	graph bar, over (v01_02)
graph hbar (count)	तेर्सो बार डायग्राम बनाउन । यसले बारहरूलाई बारम्बारतामा देखाउँदछ ।	graph hbar (count), over(v01_02)
histogram	हिस्टोग्राम बनाउन	histogram v01_03, width(5)
graph box	बक्स प्लट बनाउन । graph hbox र graph vbox गरी दुई प्रकारका बक्स प्लट बनाउन सकिन्छ । यो ग्राफबाट आउटलायर्स समेत हेर्न सकिन्छ ।	graph box v01_03
graph pie	वृत्तचित्र बनाउन ।	graph pie, over(v01_02)
twoway scatter	स्क्याटर डायग्राम बनाउन	twoway scatter v01_16 v01_03
scatter	स्क्याटर डायग्राम बनाउन	scatter v01_16 v01_03
graph save	बनाएको ग्राफलाई स्टेटा फाइलमा सेभ गर्न । यसका लागि पहिला ग्राफ बनाउने अनि यो कमाण्ड प्रयोग गर्ने	graph save "Sex Distribution"
gr export	बनाएको ग्राफलाई अन्य (.png, .wmf, .emf, .ps) फाइलमा सेभ गर्न । यसका लागि पहिला ग्राफ बनाउने अनि यो कमाण्ड प्रयोग गर्ने	gr export sex_distribution.png
gr combine	स्टेटा फाइलमा सेभ गरिएका ग्राफहरूलाई एउटै चित्रमा देखाउने । दिइएको उदाहरणमा Sex र Marriage दुईवटा स्टेटा फाइल (graph save कमाण्डबाट सेभ गरिएका) का ग्राफ हुन् ।	graph bar, over (v01_02) graph save "Sex" graph bar, over(v01_06) graph save "Marriage" gr combine "Sex" "Marriage"

नोट: माथि तालिकामा दिइएका कमाण्डहरू स्टेटामा ग्राफ बनाउनका लागि आधारभूत ज्ञान प्रदान गर्नका लागि मात्र हो। स्टेटामा ग्राफ बनाउन होस् वा अन्य कुनै काम गर्न, help को सहायता लिन सकिन्छ। उदाहरणका लागि बारग्राफका बारेमा थप जानकारी हासिल गर्न, h graph bar प्रयोग गर्न सकिन्छ।

२३.८ परियोजना कार्य

यस पाठमा चर्चा गरिएका स्टेटा कमाण्डहरूको प्रयोग गरेर उपलब्ध गराइएको NLSS_2010-11 फोल्डर भित्रका दुईवटा डाटा फाइल xh01_s01 / xh44_s20 प्रयोग गरेर निम्नानुसारका कार्य गर्नुहोस्।

क) डाटा फाइल xh01_s01 र xh44_s20 लाई merge गर्नुहोस्।

ख) दुई चरहरू v20_05 (height) र v20_07 (weight) लाई rename गर्नुहोस्।

ग) बच्चाको तौल र उचाइबिचको सहसम्बन्ध निकाल्नुहोस्।

घ) बच्चाको तौल र उचाइबिच प्रतिगमन विश्लेषण गर्नुहोस्।

२३.९ अभ्यासका लागि प्रश्नहरू

१) xh01_s01 मा भएको उमेर चरलाई ५ वर्षे उमेर समूहको नयाँ चर बनाएर तालिकामा प्रस्तुत गर्नुहोस्। साथै, यस तालिकालाई हिस्टोग्राममा पनि देखाउनुहोस्।

२३.१० सन्दर्भ सामग्री

स्टेटा अध्ययन गर्नका लागि स्टेटाको help कमाण्ड वा मेनु बारको help उपमेनु निकै उपयोगी हुन्छ। साथै, आजकल यससम्बन्धी अनलाइनमा थुप्रै सामग्रीहरू पाइन्छन्। कुनै काम गर्न परेको खण्डमा गुगलमा गएर, how to...in STATA ? भनेर खोजी गर्ने हो भने प्रशस्त मार्गदर्शन प्राप्त हुन्छ। गुगल र युट्युबको व्यापक प्रयोग गर्न सकेमा स्टेटाका धेरै कामहरू गर्न सकिन्छ। यसका अतिरिक्त स्टेटाको वेबसाइटमा भएको निम्न सामग्रीहरू उपयोगी छन्।

१. <https://www.STATA.com/links/video-tutorials/>



अनुसूचीहरू

अनुसूची १

वस्तुगत प्रश्न समाधानका लागि सुझाव

पाठ ३: तथ्याङ्कीय साक्षरतासम्बन्धी अवधारणा

च) तलका भनाइहरू ठिक वा बेठिक छुट्याई कारण दिनुहोस् ।

१) विद्यार्थीहरूले तथ्याङ्कसम्बन्धी जानकारी राख्नुपर्दछ ।

उत्तर: ठिक

२) पत्रकारहरू तथ्याङ्कीय साक्षर नभए पनि हुन्छ ।

उत्तर: बेठिक

३) स्थानीय तहमा कार्यरत कर्मचारीहरूलाई डिजिटल साक्षरताको त्यति आवश्यकता छैन ।

उत्तर: बेठिक

४) तथ्याङ्कीय क्रियाकलाप संघीय सरकारको अधिकार क्षेत्रभित्र मात्र पर्ने विषय हुन् ।

उत्तर: बेठिक

५) डिजिटल साक्षरताले मानिसको दक्षता वृद्धि गरेको छ ।

उत्तर: ठिक

पाठ ५: मापन एकाइसम्बन्धी अवधारणा

क) खाली ठाउँ भर्नुहोस् ।

१) परिवारको मासिक आमदानीतहको मापन हो ।

उत्तर : अनुपात

२) चर प्रकारका हुन्छन् ।

उत्तर: दुई

३) गुणात्मक विशेषताको अध्ययन गर्नुपर्दा मापन एकाइ प्रयोग गरिन्छ ।

उत्तर: नाम मात्रको वा क्रमिक तहका मापन

४) मापनका तह हुन्छन् ।

उत्तर: चार

ग) जोडा मिलाउनुहोस् ।

नाम मात्र तहको मापन	→	विद्यार्थीको तौल
क्रमिक तहको मापन	→	फुटबल खेलाडीको जर्सी नम्बर
अन्तराल तहको मापन	→	मानव विकास सूचकाङ्कमा देशहरूको स्थान
अनुपात तहको मापन	→	बिरामीको ज्वरोको तापक्रम

पाठ ६: तथ्याङ्कका स्रोत तथा तथ्याङ्क व्यवस्थापन

क) खाली ठाउँ भर्नुहोस् ।

१) कुनै पनि अध्ययनका लागि सुरुमा संकलन गरेको तथ्याङ्कलाई भनिन्छ ।

उत्तर: प्राथमिक तथ्याङ्क

२) कुनै जनसंख्याको प्रतिनिधित्व हुने उप-समूहलाई भनिन्छ ।

उत्तर: नमुना

३),,, तथ्याङ्कका स्रोतहरू हुन् ।

उत्तर: गणना, नमुना सर्वेक्षण, प्रशासनिक अभिलेख, बिग डाटा

४) तथ्याङ्कलाई विभिन्न उस्तै प्रकृतिका समूहमा विभाजन गर्ने कार्यलाई भनिन्छ ।

उत्तर: तथ्याङ्क वर्गीकरण

५) कुनै पनि तालिकाको लहर (Row) को शीर्षकलाई भनिन्छ ।

उत्तर: स्तब

ख) तलका भनाइहरू ठिक वा बेठिक छुट्याउनुहोस् ।

१) प्राथमिक तथ्याङ्कभन्दा माध्यमिक तथ्याङ्क संकलन गर्न कम खर्चिलो हुन्छ ।

उत्तर: ठिक

२) गणकहरूले तथ्याङ्क संकलन गर्न प्रयोग गर्ने प्रश्नावलीलाई सेड्युल भनिन्छ ।

उत्तर: ठिक

३) तथ्याङ्क विश्लेषणमा आउटलायर्सहरूलाई ध्यान नदिए पनि हुन्छ ।

उत्तर: बेठिक

४) दुई वा दुईभन्दा बढी चर वा विशेषता दर्शाउने तालिकाहरूलाई सामान्य तालिका भनिन्छ ।

उत्तर: बेठिक

५) ठुला ठुला सङ्ख्या भएको तथ्याङ्कलाई बिग डाटा भनिन्छ ।

उत्तर: बेठिक

पाठ ७: तथ्याङ्क र सूचक

क) तथ्याङ्क केका लागि प्रयोग गरिन्छ ?

- १) योजना तर्जुमा
- २) नीति निर्माण
- ३) अध्ययन अनुसन्धान
- ४) माथिका सबै

ख) तथ्याङ्कका प्रयोगकर्ताहरू किन सजग हुनुपर्छ ?

- १) तथ्याङ्कलाई बुझ्न
- २) तथ्याङ्कको गलत प्रयोगबाट हुनसक्ने जोखिमबाट मुक्त हुन
- ३) तथ्याङ्कका सीमाहरू बारे जानकारी प्राप्त गर्न
- ४) माथिका सबै

ग) सार्क राष्ट्रहरूको विगत ५ वर्षको कुल गार्हस्थ्य उत्पादनका आँकडा कस्तो डाटा हो ?

- १) खण्डित डाटा
- २) बिग डाटा
- ३) अविच्छिन्न डाटा
- ४) मेटा डाटा

घ) कुनै सडकमा कुनै वर्षमा गणना गरिएको सडक दुर्घटनाबाट भएको मृत्यु विवरण

यसप्रकार छ :

बिच सडकमा मृत्यु : १२४ जना,

सडकपेटीमा मृत्यु : ४,०४८ जना ।

उल्लिखित आँकडाको आधारमा निम्नानुसारको निचोड निकाल्न सकिन्छ कि सकिँदैन, किन ?

“सडक बिचमा थोरै मृत्यु भएका कारण सडक पेटीभन्दा बिचमा हिँड्नु बढी सुरक्षित देखिन्छ”।

उत्तर: सकिँदैन

ङ) तलको तालिका अध्ययन गरी सोधिएको प्रश्नको जवाफ दिनुहोस् ।

कक्षा ७ मा अध्ययनरत दुई विद्यार्थीको नाम	प्रथम त्रैमासिक परीक्षाको प्राप्ताङ्क प्रतिशत	दोस्रो त्रैमासिक परीक्षाको प्राप्ताङ्क प्रतिशत	अन्तिम परीक्षाको प्राप्ताङ्क प्रतिशत	औसत वार्षिक प्राप्ताङ्क प्रतिशत
सुजन खतिवडा	८०	६०	४०	६०
सजना दर्लामी	५५	६०	६५	६०

“औसत वार्षिक प्राप्ताङ्क प्रतिशत सूचकका आधारमा सुजन र सजनाको अध्ययन प्रगति समान छ”। यो भनाइ कतिको सही छ, यसले तथ्याङ्क विश्लेषणमा सूचक प्रयोग गर्दा के कुरामा ख्याल गर्नुपर्दछ भन्ने कुरा इङ्कित गर्न खोजेको छ, व्याख्या गर्नुहोस् । (उत्तरका लागि संकेत: चित्र नं. ७.१ अनुसार व्याख्या गर्नुहोस्)

उत्तर: भनाइ सही छैन ।

च) जोडा मिलाउनुहोस् ।

आँकडा	→	तथ्याङ्कको व्याख्या
सूचक	→	खण्डित श्रेणी
“WHAT” प्रश्न	→	अविच्छिन्न श्रेणी
“WHY” प्रश्न	→	सूचनाको कच्चा पदार्थ
परिवारको सदस्य संख्या	→	निर्णय प्रक्रिया
परिवारको वार्षिक आमदानी	→	तथ्याङ्कको विश्लेषण

पाठ ८: तथ्याङ्क प्रस्तुतिकरण

क) तथ्याङ्कअनुसार प्रस्तुतिकरणको कुन विधि उत्तम होला ?

१) नेपालको विगत १० वर्षको कुल गार्हस्थ्य उत्पादन वृद्धिदर

उत्तर: समय श्रृङ्खलाको ग्राफ

२) पाँच वर्ष उमेर समूहअनुसार विद्यार्थीहरूको वितरण

उत्तर: हिस्टोग्राम

३) कुनै परिवारले खाना, कपडा, औषधी, मनोरञ्जन र अन्य शीर्षकमा गरेको मासिक खर्च

उत्तर: वृत्तचित्र वा बार डायग्राम

४) कुनै गाउँपालिकाको वडाअनुसार निरपेक्ष गरिब परिवारको संख्या

उत्तर: नक्सा वा बार डायग्राम

- ५) विगत तिन वर्षमा भारत, चीन, अमेरिका र जापानले नेपाललाई दिएको अनुदान र कम रू.

उत्तर: बहु-स्तम्भचित्र

ख) तलका भनाईहरू ठिक वा बेठिक कारण सहित बताउनुहोस् ।

- १) तथ्याङ्कको प्रस्तुतिकरण नगरिकन पनि तथ्याङ्कको विश्लेषण र व्याख्या गर्न सकिन्छ ।

उत्तर: बेठिक

- २) स्तम्भचित्रको प्रयोगले तुलनात्मक अध्ययन गर्न सजिलो हुन्छ ।

उत्तर: ठिक

- ३) वृत्तचित्र ग्राफको उदाहरण हो ।

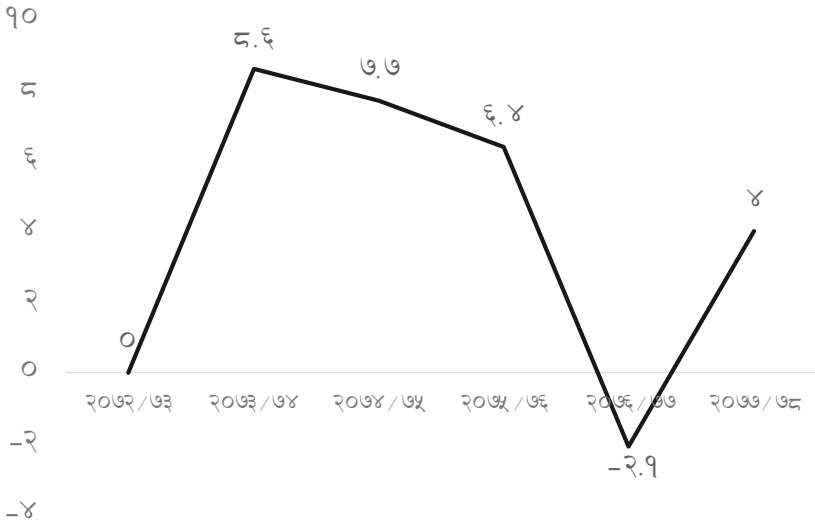
उत्तर: बेठिक

- ४) नीतिनिर्माताहरूका लागि तथ्याङ्क प्रस्तुतिकरणले खासै महत्त्व राख्दैन ।

उत्तर: बेठिक

ङ) तलको चित्र अध्ययन गरेर सोधिएका प्रश्नहरूको जवाफ दिनुहोस् ।

चित्र नं. ८.१२: कुल गार्हस्थ्य उत्पादन वार्षिक वृद्धिदर



स्रोत: केन्द्रीय तथ्याङ्क विभाग, नेपालको राष्ट्रिय लेखा अनुमान, २०७७/७८

१) यो चित्र कुन प्रकारको चित्र हो ?

उत्तर: समय श्रृङ्खलाको ग्राफ

२) चित्रमा के देखाइएको छ ?

उत्तर: नेपालको विगत छ वर्षको कुल गार्हस्थ्य उत्पादन वार्षिक वृद्धिदर

३) कुन आ.व.को वृद्धिदर नकारात्मक छ ?

उत्तर: २०७६/७७

४) सबैभन्दा बढी वृद्धिदर कुन आ.व. मा देखिन्छ ?

उत्तर: २०७३/७४

च) तलको तालिका अध्ययन गरेर सोधिएका प्रश्नको जवाफ दिनुहोस् ।

तालिका ८.५: विगत १० वर्षमा पेट्रोलियम पदार्थको आयात परिमाण

आ.व.	पेट्रोल	डिजेल	मट्टितेल	हवाई इन्धन	हलुका डिजेल	फर्नान्स		र्याँस (मे. टन.)
						(किलोलिटर)	(मे. टन.)	
२०७६/७७	५,१२,१२८	१४,७३,५३६	१८,९२४	१,३७,४२४	०	०	०	४,४९,०६३
२०७५/७६	५,६६,८२७	१७,१४,९१७	२५,००४	२,००,१०८	०	०	०	४,२९,६०९
२०७४/७५	४,८८,६७५	१५,८८,८६९	२२,३३७	१,९७,२२०	०	०	०	३,७०,५६०
२०७३/७४	४,०७,२७०	१३,१९,८७३	१९,६०७	१,६४,८३६	०	३६	३६	३,१२,९२८
२०७२/७३	२,४०,३८६	७,८५,६८५	१४,१९४	८३,८१९	०	७७	७७	२,१४,१९४
२०७१/७२	२,८७,४७३	९,२१,७१४	१९,६५३	१,४१,४०४	०	८८३	८८३	२,५८,२९९
२०७०/७१	२,५३,३८१	८,०८,५६७	१८,४०९	१,२५,६७८	०	२,१७२	२,१७२	२,३२,६६०
२०६९/७०	२,२३,०८७	७,२१,२०३	२४,०६५	१,१५,८९६	२६०	२,४५६	२,४५६	२,०७,०३८
२०६८/६९	२,०२,४६७	६,५३,५६०	४१,६०९	१,०९,९०४	०	४४०	४४०	१,८१,४११
२०६७/६८	१,८८,०८२	६,५२,७६४	४३,३९९	९९,९९०	२२८	१,४३४	१,४३४	१,५९,२८६

स्रोत: नेपाल आयल निगम लिमिटेड^१

1 Nepal Oil Corporation. *Import and Sales of Petroleum Products*. Kathmandu, Nepal; Nepal Oil Corporation. Retrieved on November 27, 2021 from: <http://noc.org.np/import>

१) यो तालिकामा के प्रस्तुत गरिएको छ ?

उत्तर: विगत १० वर्षमा नेपालले गरेको पेट्रोलियम पदार्थको आयात परिमाण

२) पेट्रोल, डिजेल र मट्टितेल सबैभन्दा बढी आयात भएको वर्षहरू कुन कुन हुन् ?

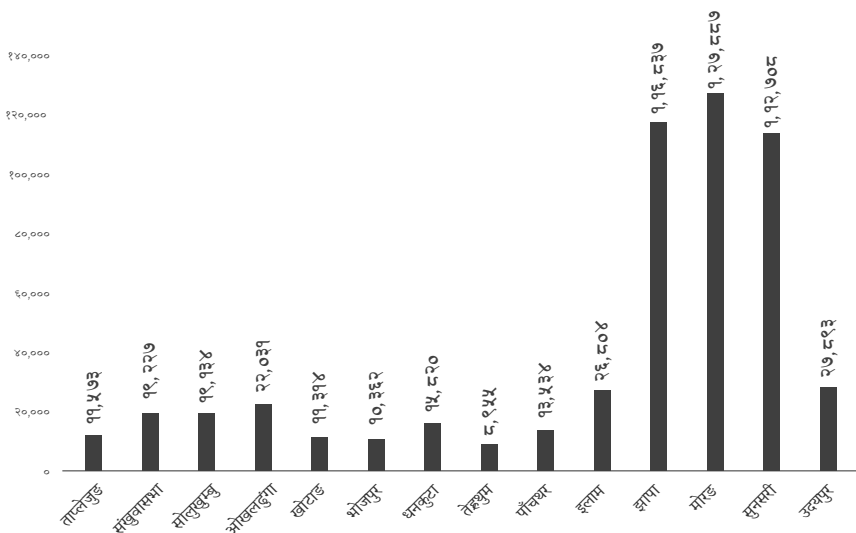
उत्तर: पेट्रोल - २०७५/७६, डिजेल - २०७५/७६, मट्टितेल - २०६७/६८

३) ग्याँसको आयात सबैभन्दा बढी र सबैभन्दा कम कुन कुन वर्षमा भएको छ ?

उत्तर: बढी - २०७५/७६, कम - २०६८/६९

छ) तलको चित्र हेरेर निम्नानुसारका प्रश्नको जवाफ दिनुहोस् ।

चित्र नं. ८.१२: कुल गार्हस्थ्य उत्पादन वार्षिक वृद्धिदर



स्रोत: केन्द्रीय तथ्याङ्क विभाग, आर्थिक गणना, २०१८

१) यो चित्र कुन प्रकारको चित्र हो ?

उत्तर: बार डायग्राम

२) यो चित्रले के देखाउन खोजेको छ ?

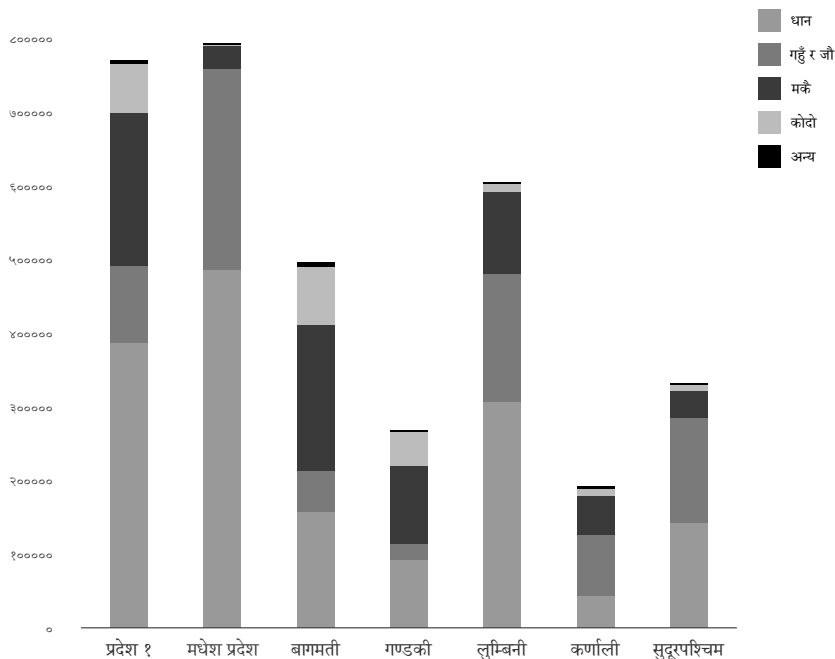
उत्तर: प्रदेश १ अन्तर्गतका जिल्लाहरूमा आर्थिक क्रियाकलापमा संलग्न जनसंख्या विवरण

३) प्रदेश १ मा आर्थिक क्रियाकलापमा संलग्न सबैभन्दा बढी र सबैभन्दा कम जनसंख्या भएको दुई जिल्ला कुन कुन हुन् ?

उत्तर: बढी - मोरङ, कम - तेह्रथुम

ज) तलको चित्र अध्ययन गरेर सोधिएका प्रश्नहरूको जवाफ दिनुहोस् ।

चित्र नं. ८.१४: प्रदेश र अन्नअनुसार अन्न खेती गरिने जग्गाको क्षेत्रफल (हेक्टर)



स्रोत: केन्द्रीय तथ्याङ्क विभाग, राष्ट्रिय कृषिगणना, २०६८^२

१) यो कस्तो प्रकारको चित्र हो ?

उत्तर: उप-विभाजित स्तम्भचित्र

२) यस चित्रमा केको तथ्याङ्क प्रस्तुत गरिएको छ ?

उत्तर: प्रदेश र अन्नअनुसार अन्न खेती गरिने जग्गाको क्षेत्रफल (हेक्टर)

३) सबैभन्दा बढी धान खेती गर्ने प्रदेश कुन हो ?

उत्तर: मधेश प्रदेश

४) कोदोको सबैभन्दा बढी खेती गर्ने प्रदेश कुन हो ?

उत्तर: बागमती

2 Central Bureau of Statistics. (2011). *Province Summary of Agriculture Census 2011*. Kathmandu, Nepal; Central Bureau of Statistics. Retrieved on November 28, 2021 from: <https://cbs.gov.np/province-summary-of-agriculture-census-2011/>

५) मकैको सबैभन्दा कम खेती गर्ने प्रदेश कुन हो ?

उत्तर: मधेश प्रदेश

पाठ ९: तथ्याङ्क विश्लेषणका लागि आधारभूत गणितीय क्रियाकलाप

क) खाली ठाउँ भर्नुहोस्

१) पूर्ण संख्या भनेको सहितको प्राकृतिक संख्या हो ।

उत्तर: शून्य

२) दुई संख्याहरूको तुलना गर्न को प्रयोग गरिन्छ ।

उत्तर: अनुपात

३) जन्मदर, मृत्युदर जस्ता दरहरू प्रायः प्रति मा व्यक्त गरिन्छ ।

उत्तर: १०००

४) सय भागको केही भागलाई भनिन्छ ।

उत्तर: प्रतिशत

५) अंश भन्दा हर ठुलो भएको भिन्नलाई भनिन्छ ।

उत्तर: **Improper fraction**

ख) ठिक र बेठिक छुट्याउनुहोस् ।

१) $\sqrt{१६}$ भिन्नमा व्यक्त गर्न नसकिने संख्या हो ।

उत्तर: बेठिक

२) ८ पनि एउटा भिन्न संख्या हो ।

उत्तर: ठिक

३) $i + ५$ एउटा जटिल संख्या हो ।

उत्तर: ठिक

४) दर र अनुपात एउटै कुरा हुन् ।

उत्तर: बेठिक

५) भिन्नलाई प्रतिशतमा लैजान १०० ले भाग गर्नुपर्दछ ।

उत्तर: बेठिक

पाठ १०: सूचकाङ्क र यसको प्रयोग

ख) तलका भनाइहरूलाई कारणसहित ठिक वा बेठिक छुट्याउनुहोस् ।

१) सूचकाङ्कहरू अर्थतन्त्रका ब्यारोमिटर हुन् ।

उत्तर: ठिक

२) जीविकोपार्जन सूचकाङ्क एक प्रकारको परिमाण सूचकाङ्क हो ।

उत्तर: बेठिक

३) वास्तविक ज्याला पत्ता लगाउन थोक सूचकाङ्क प्रयोग गरिन्छ ।

उत्तर: बेठिक

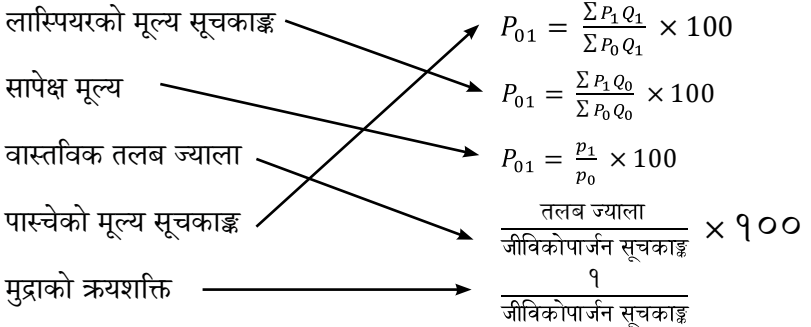
४) पास्चेको सूचकाङ्कमा आधार वर्षको परिमाणलाई भारको रूपमा प्रयोग गरिन्छ ।

उत्तर: बेठिक

५) सूचकाङ्क निर्माणमा तथ्याङ्कको गुणस्तरको कुनै असर हुँदैन ।

उत्तर: बेठिक

ग) जोडा मिलाउनुहोस् ।



पाठ ११: केन्द्रीय प्रवृत्ति मापन

ख) तलका तथ्याङ्कहरूको लागि उपयुक्त औसत छान्नुहोस् ।

१) एक वर्षको मासिक ब्याजदरहरूको औसत

(अंकगणितीय मध्यक, ज्यामितीय मध्यक, हार्मोनिक मध्यक, मध्यिका, बहुलक)

२) विद्यार्थीले परीक्षामा प्राप्त गरेको प्रतिशतको औसत

(अंकगणितीय मध्यक, ज्यामितीय मध्यक, हार्मोनिक मध्यक, मध्यिका, बहुलक)

- ३) केन्द्रीय तथ्याङ्क विभागका कर्मचारीहरूको मासिक औसत तलब
(अंकगणितीय मध्यक, ज्यामितीय मध्यक, हार्मोनिक मध्यक, मध्यिका, बहुलक)
- ४) मन्त्रालयहरूको विद्युत महसुलको औसत मासिक खर्च
(अंकगणितीय मध्यक, ज्यामितीय मध्यक, हार्मोनिक मध्यक, मध्यिका, बहुलक)
- ५) सम्पत्तिका आधारमा देशका परिवारलाई बराबर दुई भाग लगाउने औसत
(अंकगणितीय मध्यक, ज्यामितीय मध्यक, हार्मोनिक मध्यक, मध्यिका, बहुलक)

पाठ १२: विचलनको मापन

क) तलकामध्ये कुनचाहिँ भनाइ गलत छ ?

- १) विचलन मापन स्पष्ट रूपमा परिभाषित भएको हुनुपर्दछ ।
- २) विचलन मापन सबै आँकडाहरू आधारित भएर निकालिएको हुनुपर्दछ ।
- ३) विचलन मापन थप गणितीय क्रियाहरूमा उपयोगी हुनुपर्दछ ।
- ४) विचलन मापन अत्यन्त ठुला वा साना आँकडाले प्रभावित पार्नु हुँदैन ।
- ५) माथिका सबै गलत छन् ।

ख) तलकामध्ये कुनचाहिँ विचलन मापन होइन ?

- १) स्तरीय भिन्नता
- २) भेरियन्स
- ३) मध्यिका
- ४) मध्य विचलन
- ५) माथिका कुनै पनि होइन

ग) खाली ठाउँ भर्नुहोस् ।

- १) विस्तार = - सबैभन्दा सानो आँकडा
उत्तर: सबैभन्दा ठुलो आँकडा

- २) तथ्याङ्क वितरणमा केन्द्रीय मानका बारेमा औसतहरूले बताउँछन् भने कति छरिएर रहेका छन् भनेर ले बताउँछ ।
उत्तर: विचलन मापन

- ३) असमानता अध्ययन गर्नका लागि र को प्रयोग गरिन्छ ।

उत्तर: लरेन्ज बक्ररेखा र जिनीको गुणाङ्क

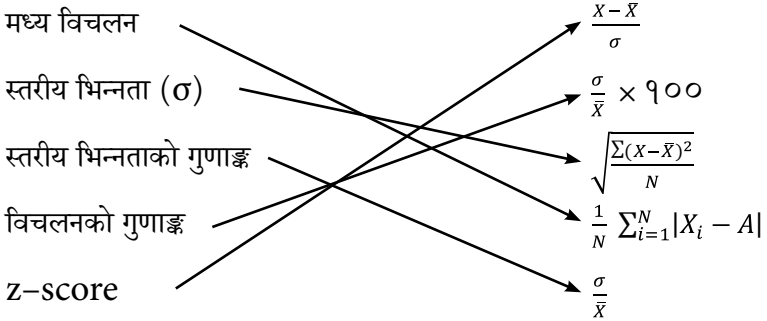
४) दिइएका आँकडाहरूको उद्गम बिन्दु तथा स्केल परिवर्तन गरेर आउने मानलाई भनिन्छ ।

उत्तर: जेड स्कोर

५) भेरियन्स भनेको को वर्ग हो ।

उत्तर: स्तरीय भिन्नता

घ) जोडा मिलाउनुहोस् ।



पाठ १३: सहसम्बन्ध र यसको प्रयोग

क) खाली ठाउँ भर्नुहोस् ।

१) दुईवटा पूर्ण रूपमा असम्बन्धित चरहरूबिच पनि सहसम्बन्धको गुणाङ्क सकारात्मक वा नकारात्मक आयो भने त्यसलाई भनिन्छ ।

उत्तर: स्परियस अथवा अर्थहीन सहसम्बन्ध

२) सहसम्बन्धको गुणाङ्कको वर्गलाई भनिन्छ ।

उत्तर: कोअफिसियन्ट अफ डिटरमिनेशन

३) एउटा चरमा आएको परिवर्तनले अर्को चरमा पनि परिवर्तन आउँछ भने ती चरहरूबिच भएको मानिन्छ ।

उत्तर: सहसम्बन्ध

४) वस्तुको मूल्य बढ्दा उपभोगको वा मागको मात्रा घट्दछ । यस्तो अवस्थामा ती दुई चरहरूबिचको सहसम्बन्ध हुन्छ ।

उत्तर: नकारात्मक

५) सहसम्बन्धले एउटा चरको मान हुँदा अर्को चरको मान पनि हुन्छ भनेर देखाउँछ ।

उत्तर: परिवर्तन, परिवर्तन

ख) तलका भनाइ सही या गलत के छन्, कारणसहित बताउनुहोस् ।

१) r को मान $< .6$ आउनुलाई अर्थहीन सहसम्बन्ध भनिन्छ ।

उत्तर: गलत

२) एउटा चरको मान बढ्दा अर्को चरको मानमा कुनै असर पर्दैन भने ती दुई चरहरूबिच नकारात्मक सहसम्बन्ध छ ।

उत्तर: गलत

३) R^2 ले एउटा चरले अर्को चरलाई कुन हदसम्म असर गरिरहेको छ भनेर सामान्य धारणा दिन्छ ।

उत्तर: सही

४) सहसम्बन्ध विश्लेषणको सहयोगले कारण र असरका बारेमा अध्ययन गर्न सकिन्छ ।

उत्तर: गलत

५) कतिपय अवस्थामा कुनै पनि सम्बन्ध नभएका चरहरूबिच पनि संयोगवश r को मान शून्य नआई अन्य कुनै अंक आउन सक्दछ ।

उत्तर: सही

ग) तलका चरहरूबिच सामान्यतया सकारात्मक, नकारात्मक वा शून्य सहसम्बन्ध के हुन्छ, कारणसहित उल्लेख गर्नुहोस् ।

१) वस्तुको मूल्य र उपभोग

उत्तर: नकारात्मक

२) मौसमको तापक्रम र बाक्लो लुगाको बिक्री

उत्तर: नकारात्मक

३) कृषि क्षेत्रको उत्पादन र कुल गार्हस्थ्य उत्पादन

उत्तर: सकारात्मक

४) सरकारले विकासमा गरेको खर्च र कुल गार्हस्थ्य उत्पादन

उत्तर: सकारात्मक

५) परिवारको आमदानी र खर्च

उत्तर: सकारात्मक

पाठ १४: समय श्रृङ्खला

ख) ठिक वा बेठिक छुट्याउनुहोस् ।

१) विभिन्न समय अवधिमा मिलाएर राखेको तथ्याङ्कलाई क्रस-सेक्सन तथ्याङ्क भनिन्छ ।

उत्तर: बेठिक

२) जनसंख्या वृद्धिदर निकाल्न अंकगणितीय वृद्धिदर प्रयोग गर्नु उपयुक्त हुन्छ ।

उत्तर: बेठिक

३) समय श्रृङ्खलाको प्रवृत्ति ग्राफको सहयोगबाट पनि पत्ता लगाउन सकिन्छ ।

उत्तर: ठिक

ग) खाली ठाउँ भर्नुहोस् ।

१) समयको कुनै एक बिन्दुमा लिइएको वा प्रस्तुत गरिएको तथ्याङ्क तथ्याङ्क हो ।

उत्तर: क्रस-सेक्सन तथ्याङ्क

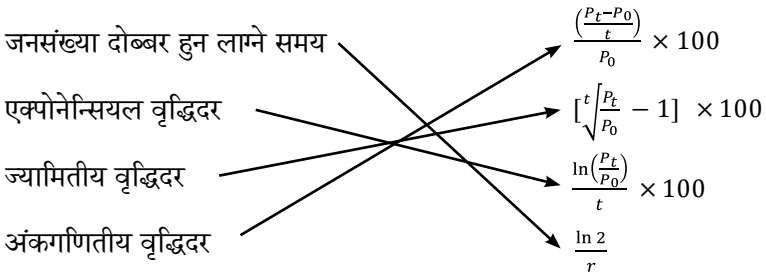
२) कुनै पनि समयमा घटबढ भइरहन सक्ने वा आकस्मिक घटनाका कारण परिवर्तन हुने प्रकृतिको श्रृङ्खलालाई उतारचढाव भनिन्छ ।

उत्तर: अनियमित

३) विभिन्न समयका क्रस-सेक्सन तथ्याङ्क समूहलाई एउटै तथ्याङ्क समूहमा मिलाएर बनाइएको तथ्याङ्कलाई तथ्याङ्क भनिन्छ ।

उत्तर: प्यानल

घ) जोडा मिलाउनुहोस् ।



पाठ १५: नर्मल डिस्ट्रिब्युसन

क) खाली ठाउँ भर्नुहोस् ।

१) नर्मल डिस्ट्रिब्युसनले चरको सम्भाव्यता वितरणको अध्ययन गर्दछ ।

उत्तर: अविच्छिन्न

२) सम्भाव्यताको मान देखि सम्म मात्र हुन्छ ।

उत्तर: ०, १

३) निश्चित भएको घटनाको सम्भाव्यता हुन्छ ।

उत्तर: १

४) असम्भव घटनाको सम्भाव्यता हुन्छ ।

उत्तर: ०

५) स्टान्डर्ड नर्मल पछ्याउने चरको मध्यक र हुन्छ ।

उत्तर: ०, स्तरीय भिन्नता

ग) ठिक बेठिक छुट्याउनुहोस् ।

१) $P(-\infty < Z < 2) = ०.९७५$ हुन्छ ।

उत्तर: बेठिक

२) नर्मल बक्ररेखा र X अक्षबिच $\mu \pm 2\sigma$ ले कुल क्षेत्रफलको करिब ९५.४४% समेट्छ ।

उत्तर: ठिक

३) सम्भाव्यताको मान नकारात्मक पनि हुन सक्दछ ।

उत्तर: बेठिक

४) पाँच वर्षभरिमा कुनै एउटा फेब्रुअरी महिनामा ३१ वटा दिन हुने सम्भावना शून्य हुन्छ ।

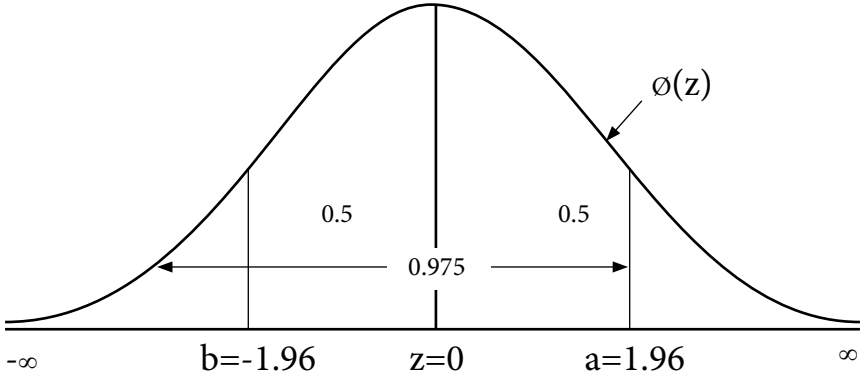
उत्तर: ठिक

५) स्टान्डर्ड नर्मल बक्ररेखा र X अक्षबिचको पूरा क्षेत्रफल १ हुन्छ ।

उत्तर: ठिक

छ) तलको चित्र अध्ययन गरेर सोधिएका प्रश्नहरूको जवाफ दिनुहोस् ।

चित्र १५.५ स्टाण्डर्ड नर्मल बक्ररेखासम्बन्धी अभ्यास



१) यो चित्रमा $P(a < z < \infty)$ कति हुन्छ ?

उत्तर: ०.०२५

२) यो चित्रमा $P(-\infty < z < \infty)$ कति हुन्छ ?

उत्तर: १

३) यो चित्रमा $P(-\infty < z < b)$ कति हुन्छ ?

उत्तर: ०.०२५

४) यो चित्रमा $P(b < z < \infty)$ कति हुन्छ ?

उत्तर: ०.९७५

५) यो चित्रमा $P(z = 1.96)$ कति हुन्छ ?

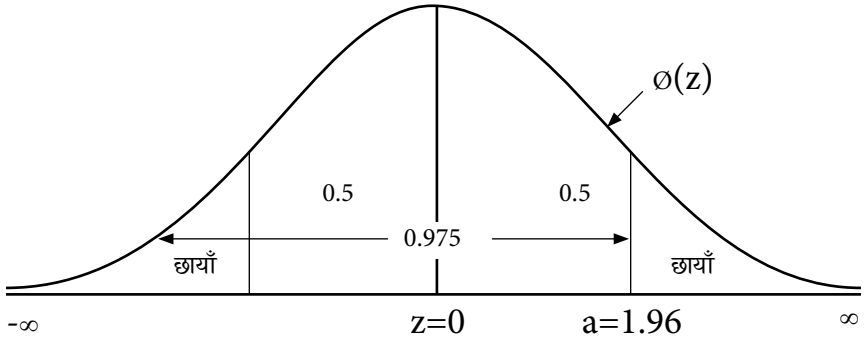
उत्तर: ०

६) यो चित्रमा $P(-\infty < z < 0)$ कति हुन्छ ?

उत्तर: ०.५

७) $P(-\infty < z < b)$ र $P(a < z < \infty)$ जनाउने क्षेत्रफललाई चित्रमा छायाँ पारेर

देखाउनुहोस् । साथै यी दुई क्षेत्रफलहरू बराबर छन् वा छैनन् तर्क प्रस्तुत गर्नुहोस् ।



पाठ १६: अनुमान तथा नमुनासम्बन्धी आधारभूत जानकारी

ग) ठिक बेठिक छुट्याउनुहोस् ।

१) पर्याप्तता भएका सबै आँकलकहरू निष्पक्ष हुन्छन् ।

उत्तर: बेठिक

२) नमुना छनोटको अध्ययनले कम खर्चमा विश्वासिलो नतिजा पाइन्छ ।

उत्तर: ठिक

३) स्नो बल नमुना छनोट एउटा सम्भाव्य नमुना छनोटको उदाहरण हो ।

उत्तर: बेठिक

पाठ १७: परिकल्पना परीक्षण

क) ठिक बेठिक छुट्याउनुहोस् ।

१) शून्य परिकल्पना जनसंख्याका मानका बारेमा गरिएको दाबी वा कथन हो ।

उत्तर: ठिक

२) सही कथनलाई अस्वीकार गर्नुलाई परीक्षणको शक्ति भनिन्छ ।

उत्तर: बेठिक

३) जनसंख्याको स्तरीय भिन्नता थाह छ र यसको मध्यकको बारेमा परिकल्पना

परीक्षण गर्नु छ भने जेड परीक्षण गरिन्छ ।

उत्तर: ठिक

४) क्रिटिकल मान र क्रिटिकल क्षेत्र भनेको एउटै कुरा हो ।

उत्तर: बेठिक

ग) खाली ठाउँ भर्नुहोस्

१) दुईवटा जनसंख्याको भेरियन्स तुलना गर्नको लागि प्रयोग गरिने टेस्ट स्टाटिस्टिक्स हो ।

उत्तर: एफ टेस्ट

२) $(1 - \alpha)$ लाई भनिन्छ ।

उत्तर: विश्वस्तताको तह

३) P (गलत कथनलाई अस्वीकार गर्न नसक्नु) =

उत्तर: β

४) लेख्दा जनसंख्या मानका बारेमा “=” (= वा. वा ;) चिन्हको अनिवार्य प्रयोग गरेर लेख्नुपर्दछ ।

उत्तर: शून्य परिकल्पना

पाठ १८: प्रतिगमन विश्लेषण

ख) खाली ठाउँ भर्नुहोस् ।

१) दुईवटा चरहरूबिचको सहसम्बन्धको वर्गलाई भनिन्छ ।

उत्तर: निर्धारणको गुणाङ्क

२) चरहरूको सम्बन्धलाई हेरेर प्रतिगमन नमुना तय गर्ने कार्यलाई भनिन्छ ।

उत्तर: नमुना स्पेसिफिकेशन

३) चरहरू बिचको सम्बन्धलाई सिधा रेखाद्वारा व्याख्या गर्न सकिन्छ भने त्यस्तो प्रतिगमनलाई भनिन्छ ।

उत्तर: सिधा रेखीय

४) नमुनाका मानहरूबाट दुई चरहरूबिचको सम्बन्ध स्थापित गर्ने गुणाङ्कहरूको अनुमान निकालेर गरिने अध्ययनलाई भनिन्छ ।

उत्तर: प्रतिगमन विश्लेषण

पाठ २२: एक्सेलसम्बन्धी आधारभूत ज्ञान

क) कुनै एक उपयुक्त विकल्पमा ठिक चिन्ह लगाउनुहोस् ।

१) एक्सेलमा रहेको Formula उपमेनुको उपयोग के हो ?

क) तालिका निर्माण

ख) ग्राफ निर्माण

- ग) फरम्याटिंग
घ) फरमुलाको प्रयोगद्वारा चाहिएको गणितीय वा तथ्याङ्कीय हिसाबहरू गर्न
- २) एक्सेलको सिटमा विभिन्न सेलमा भएको फिल्डहरूलाई सेलेक्सन गरेपछि आन्तरिक रूपमा प्रोसेसिंग गरेर एक्सेलले आफैँ दिने ३ वटा कुराहरू के के हुन् ?
क) Average, Count र Aggregate
ख) Average, Maximum Value / Minimum Value
ग) **Average, Count / Sum**
घ) Maximum, Minimum / Count
- ३) एक्सेलमा कुनै पनि Formula प्रयोग गर्दा सम्बन्धित सेलमा सबैभन्दा पहिले नथिची नहुने चिन्ह कुन हो ?
क) + (Plus)
ख) - (Minus)
ग) / (Slash)
घ) = (**Equal**)
- ४) चार्टको प्रयोग किन गरिन्छ ?
क) तथ्याङ्क प्रस्तुत गर्न
ख) तथ्याङ्कको प्रवृत्ति र विशेषता देखाउन
ग) सहज रूपमा सबैले बुझ्ने गरी तथ्याङ्क प्रस्तुत गर्न
घ) **माथिका सबै**
- ५) पाइ चार्टको प्रयोग कहाँ उपयुक्त हुन्छ ?
क) समय अन्तरालमा भएको परिवर्तन देखाउन
ख) बार ग्राफ र रेखाचित्र उपयोगी नहुने सन्दर्भमा
ग) उतारचढावयुक्त तथ्याङ्क भएको अवस्थामा
घ) **समग्रलाई खण्ड खण्डमा तिनको मानसहित देखाउनु पर्दा**
- ६) चार्ट अध्ययन गर्दा कुन कुरामा ध्यान दिनु जरूरी छ ?
क) शीर्षक
ख) संकेत सूची
ग) सन्दर्भ समय र स्रोत
घ) **माथिका सबै**

अनुसूची २
सांख्यिकीय तालिकाहरू

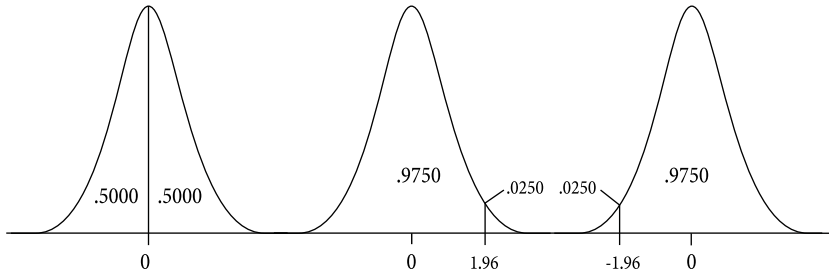
अनुसूची २.१: जेड तालिका

तालिाकाको क्षेत्रफलले $P(-\infty < z < a)$ लाई जनाएको छ ।

Z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5	0.504	0.508	0.512	0.516	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.591	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.648	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.67	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.695	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.719	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.758	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.791	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.834	0.8365	0.8389
1	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.877	0.879	0.881	0.883
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.898	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.937	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545

1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.975	0.9756	0.9761	0.9767
2	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.983	0.9834	0.9838	0.9842	0.9846	0.985	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.989
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.992	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.994	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.996	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.997	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.998	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.999	0.999

Examples: $z.5000 = 0.00$; $z.9750 = +1.96$; $z.0250 = -1.96$



अनुसूची २.२: टी-तालिका (एकतर्फी पुच्छर)

DF	A = 0.1	0.05	0.025	0.01	0.005	0.001	0.0005
1	3.078	6.314	12.706	31.821	63.656	318.289	636.578
2	1.886	2.92	4.303	6.965	9.925	22.328	31.6
3	1.638	2.353	3.182	4.541	5.841	10.214	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.61
5	1.476	2.015	2.571	3.365	4.032	5.894	6.869
6	1.44	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.86	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.25	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.93	4.318
13	1.35	1.771	2.16	2.65	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.14
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.12	2.583	2.921	3.686	4.015
17	1.333	1.74	2.11	2.567	2.898	3.646	3.965
18	1.33	1.734	2.101	2.552	2.878	3.61	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.85
21	1.323	1.721	2.08	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.5	2.807	3.485	3.768
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745

25	1.316	1.708	2.06	2.485	2.787	3.45	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.689
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.66
30	1.31	1.697	2.042	2.457	2.75	3.385	3.646
60	1.296	1.671	2	2.39	2.66	3.232	3.46
120	1.289	1.658	1.98	2.358	2.617	3.16	3.373
1000	1.282	1.646	1.962	2.33	2.581	3.098	3.3



In partnership with



Implementing partner



The Asia Foundation
Improving Lives, Expanding Opportunities