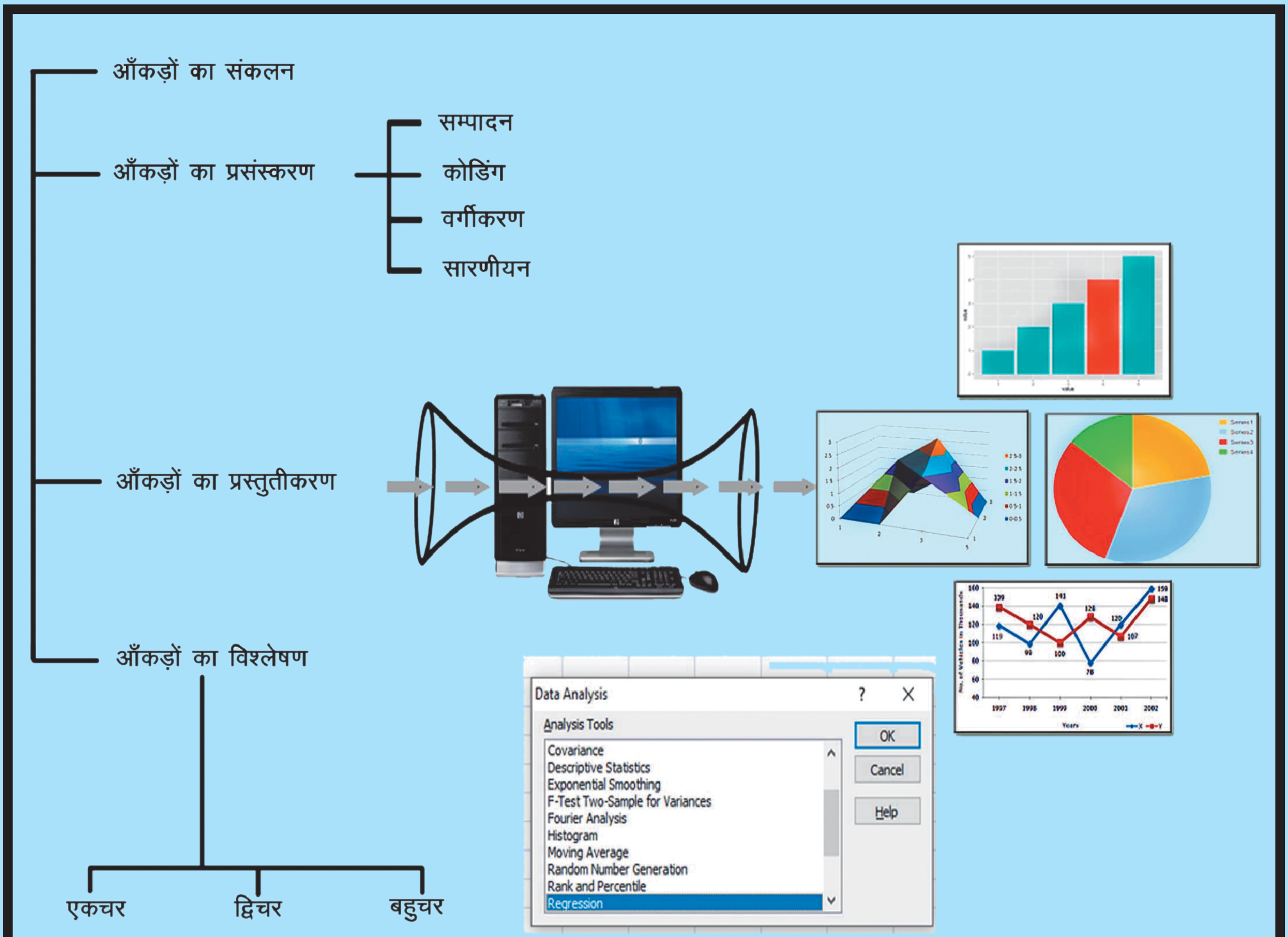
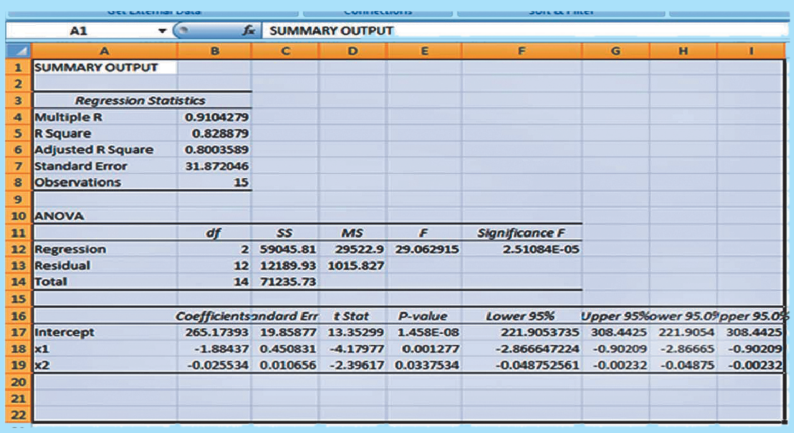
इग्नू
जन-जन का
विश्वविद्यालय

सामाजिक विज्ञान विद्यापीठ
इन्दिरा गांधी राष्ट्रीय मुक्त विश्वविद्यालय

BECS - 184
आँकड़ों का विश्लेषण





A1 SUMMARY OUTPUT

	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
2									
3	*Regression Statistics*								
4	Multiple R	0.9104279							
5	R Square	0.828879							
6	Adjusted R Square	0.8003589							
7	Standard Error	31.872046							
8	Observations	15							
9									
10	ANOVA								
11		*df*	*SS*	*MS*	*F*	*Significance F*			
12	Regression	2	59045.81	29522.9	29.062915	2.51084E-05			
13	Residual	12	12189.93	1015.827					
14	Total	14	71235.73						
15									
16		*Coefficients*	*andard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*pper 95.0%*
17	Intercept	265.17393	19.85877	13.35299	1.458E-08	221.9053735	308.4425	221.9054	308.4425
18	x1	-1.88437	0.450831	-4.17977	0.001277	-2.866647224	-0.90209	-2.86665	-0.90209
19	x2	-0.025534	0.010656	-2.39617	0.0337534	-0.048752561	-0.00232	-0.04875	-0.00232
20									
21									
22									

BECS-184

आँकड़ों का विश्लेषण

सामाजिक विज्ञान विद्यापीठ
इंदिरा गांधी राष्ट्रीय मुक्त विश्वविद्यालय

विशेषज्ञ समिति

पाठ्यक्रम संयोजक – प्रो. नारायण प्रसाद

संपादक – प्रो. जी.के.शुक्ला

खंड 1	**गणितीय एवं सांख्यिकीय संकल्पनाएँ : एक सिंहावलोकन**	**इकाई लेखक**	**इकाई संपादक**	**हिंदी अनुवाद**
इकाई 1	गणितीय संकल्पनाएँ	प्रो. मनोहर लाल, अवकाश प्राप्त आचार्य कंप्यूटर एवं सूचना विज्ञान विद्यापीठ, इग्नू, नई दिल्ली	श्री. अक्षय कुमार सह-आचार्य कंप्यूटर एवं सूचना विज्ञान विभाग, इग्नू, नई दिल्ली	डॉ. जगमोहन राय सह-आचार्य–गणित पीजीडीएवी (सायंकालीन) कॉलेज, दिल्ली विश्वविद्यालय, दिल्ली
इकाई 2	सांख्यिकीय संकल्पनाएँ	डॉ. जे.सी. शर्मा अवकाश प्राप्त सलाहकार विद्युत नवीन एवं नवीकरणीय ऊर्जा मंत्रालय, भारत सरकार, नई दिल्ली	प्रो. नारायण प्रसाद एवं श्री. बी.एस. बागला	डॉ. जगमोहन राय सह-आचार्य–गणित पीजीडीएवी(सायंकालीन) कॉलेज, दिल्ली विश्वविद्यालय, दिल्ली
इकाई 3	सांख्यिकीय सॉफ्टवेयर का परिचय	श्री. अक्षय कुमार सह-आचार्य कंप्यूटर एवं सूचना विज्ञान, विद्यापीठ इग्नू, नई दिल्ली	प्रो. नारायण प्रसाद एवं श्री. बी.एस. बागला	डॉ. जगमोहन राय सह-आचार्य–गणित पीजीडीएवी(सायंकालीन) कॉलेज, दिल्ली विश्वविद्यालय, दिल्ली

खंड 2	आँकड़ों का संग्रहण और प्रस्तुति	इकाई लेखक	इकाई संपादक	हिंदी अनुवाद
इकाई 4	*आँकड़ों का संग्रहण : विधियाँ तथा स्रोत	श्री एस. एस. सूर्यनारायणन अवकाशप्राप्त संयुक्त सलाहकार, योजना आयोग (वर्तमान में नीति आयोग), नई दिल्ली प्रो. नारायण प्रसाद इग्नू, नई दिल्ली	श्री बी.एस. बागला	डॉ. जगमोहन राय सह-आचार्य–गणित पीजीडीएवी(सायंकालीन) कॉलेज, दिल्ली विश्वविद्यालय, दिल्ली
इकाई 5	आँकड़ा संकलन के उपस्कर	डॉ. जय कुमार रंजन सहा. आचार्य–मनोविज्ञान बी.एच.यू, वाराणसी	प्रो. नारायण प्रसाद आचार्य – अर्थशास्त्र इग्नू, नई दिल्ली	डॉ. जगमोहन राय सह-आचार्य–गणित पीजीडीएवी(सायंकालीन) कॉलेज, दिल्ली विश्वविद्यालय, दिल्ली
इकाई 6	आँकड़ों की प्रस्तुति	प्रो. एच.एस. अस्थाना आचार्य – मनोविज्ञान बी.एच.यू. वाराणसी	प्रो. नारायण प्रसाद एवं श्री. अक्षय कुमार	डॉ. जगमोहन राय सह-आचार्य–गणित पीजीडीएवी(सायंकालीन) कॉलेज, दिल्ली विश्वविद्यालय, दिल्ली

खंड 3	परिमाणात्मक आँकड़ों का विश्लेषण			
इकाई 7	एक विचर डेटा विश्लेषण	डॉ. जे.सी. शर्मा अवकाशप्राप्त वरिष्ठ आर्थिक सलाहकार, विद्युत नवीन एवं नवकरणीय ऊर्जा मंत्रालय भारत सरकार, नई दिल्ली एवं श्री अक्षय कुमार सह-आचार्य, कंप्यूटर एवं सूचना विज्ञान विभाग, इग्नू, नई दिल्ली	प्रो. नारायण प्रसाद श्री. बी.एस. बागला सुश्री चेताली अरोड़ा	डॉ. जगमोहन राय सह-आचार्य–गणित पीजीडीएवी(सायंकालीन) कॉलेज, दिल्ली विश्वविद्यालय, दिल्ली
इकाई 8	द्वि-विचर आँकड़ों का विश्लेषण	डॉ. जे.सी. शर्मा अवकाशप्राप्त वरिष्ठ आर्थिक सलाहकार, विद्युत नवीन एवं नवकरणीय ऊर्जा मंत्रालय, भारत सरकार, नई दिल्ली एवं श्री अक्षय कुमार सह-आचार्य, कंप्यूटर एवं सूचना विज्ञान विभाग, इग्नू, नई दिल्ली	प्रो. नारायण प्रसाद श्री. बी.एस. बागला सुश्री चेताली अरोड़ा	श्री.बी.एस. बागला अवकाशप्राप्त सह-आचार्य–अर्थशास्त्र पीजीडीएवी कॉलेज, दिल्ली विश्वविद्यालय, दिल्ली

*यह इकाई इग्नू के MEC-009 अर्थशास्त्र में शोध पद्धतियाँ नामक पाठ्यक्रम की इकाई 6 से अनुकूलित की गई है।

इकाई 9	बहु-विचर आँकड़ों का विश्लेषण	डॉ. जे.सी. शर्मा अवकाशप्राप्त वरिष्ठ आर्थिक सलाहकार, विद्युत नवीन एवं नवीकरणीय ऊर्जा मंत्रालय भारत सरकार, नई दिल्ली एवं श्री अक्षय कुमार सह-आचार्य, कंप्यूटर एवं सूचना विज्ञान विभाग, इग्नू, नई दिल्ली	प्रो. नारायण प्रसाद श्री. बी.एस. बागला सुश्री चेताली अरोड़ा	श्री.बी.एस. बागला अवकाशप्राप्त सह-आचार्य–अर्थशास्त्र पीजीडीएवी कॉलेज, दिल्ली विश्वविद्यालय, दिल्ली
खंड 4	**संयुक्त सूचकांक एवं गुणात्मक आँकड़े**			
इकाई 10	**सामाजिक विज्ञानों में संयुक्त सूचकों की रचना	डॉ. सुनील कुमार मिश्रा सह-अध्येता, मानव विकास संस्थान, इग्नू, नई दिल्ली	प्रो. नारायण प्रसाद	श्री.बी.एस. बागला अवकाशप्राप्त सह-आचार्य–अर्थशास्त्र पीजीडीएवी कॉलेज, दिल्ली विश्वविद्यालय, दिल्ली
इकाई 11	गुणवाची आँकड़ों का विश्लेषण	डॉ. चारु जैन सह-अध्येता, एनसीएईआर, नई दिल्ली	प्रो. नारायण प्रसाद एवं श्री. बी.एस. बागला	श्री.बी.एस.बागला अवकाश प्राप्त सह आचार्य–अर्थशास्त्र पीजीडीएवी कॉलेज, दिल्ली विश्वविद्यालय, दिल्ली

सामग्री निर्माण

श्री तिलक राज
सहायक कुलसचिव (प्रकाशन)
एम.पी.डी.डी., इग्नू, नई दिल्ली

श्री यशपाल
अनुभाग अधिकारी (प्रकाशन)
एम.पी.डी.डी., इग्नू, नई दिल्ली

सुश्री कामिनी डोगरा
आशु लिपिक, इग्नू, नई दिल्ली

जनवरी, 2021

ISBN : 978-93-90773-17-6

इन्दिरा गांधी राष्ट्रीय मुक्त विश्वविद्यालय नई दिल्ली की ओर से कुलसचिव, सामग्री निर्माण एवं वितरण प्रभाग, इग्नू, नई दिल्ली द्वारा मुद्रित एवं प्रकाशित।

लेजर टाइप सेटिंग : टेसा मीडिया एण्ड कम्प्यूटर्स, सी–206, शाहीन बाग, जामिया नगर, नई दिल्ली

मुद्रक : मैसर्स ए–वन ऑफसेट प्रिंटर्स, 5/34, कीर्ति नगर, इंडस्ट्रियल एरिया, नई दिल्ली-110015 द्वारा मुद्रित।

** यह इकाई इग्नू के MEC-109 अर्थशास्त्र में शोध पद्धतियाँ नामक पाठ्यक्रम की इकाई 12 से अनुकूलित की गई है।

विषय-वस्तु

पाठ्यक्रम परिचय : आँकड़ों का विश्लेषण

आज के सूचना युगीन समाज में सूचना या जानकारी ही देश की सामाजिक, आर्थिक, राजनीतिक और सांस्कृतिक विकास की एक प्रमुख प्रचालक शक्ति बन गई है। सूचना का सृजन होता है, विभिन्न स्वरूपों में इसका दृश्यांकन होता है, अन्य सूचना स्वरूपों से उसे संबंधित किया जाता है और फिर आर्थिक एवं निर्णय प्रक्रियाओं में उसका उपयोग होता है।

सूचना सृजन का एक निर्णायक घटक आँकड़े हैं। आँकड़े सूचनाओं का ही अपरिष्कृत स्वरूप हैं, जिन्हें एकत्र, संस्कारित एवं प्रसंसाधित किया जाता है। अतः आंकड़ों का विश्लेषण कर उन्हें अपेक्षाकृत सरल स्वरूप में तालिकाओं और रेखाचित्रों में प्रस्तुत किया जाता है ताकि उनका कहीं अधिक सार्थक रूप से उपयोग हो सके। इसी कारण से आज स्प्रैडशीट्स का प्रयोग करते हुए आँकड़ों के संकलन, प्रस्तुति और विश्लेषण में उपस्करों के ज्ञान की अर्थव्यवस्था के प्रायः सभी क्षेत्रकों में बहुत माँग है। हमारा यह कौशल संवर्धन पाठ्यक्रम आपको 'Excel' स्प्रैडशीट सॉफ्टवेयर का प्रयोग करते हुए उपर्युक्त उपस्करों एवं तकनीकों से सम्पन्न कर आपकी रोज़गार प्राप्ति क्षमता में सुधार करेगा। इस पाठ्यक्रम को चार खंडों में विभाजित किया गया है।

खंड 1 गणितीय एवं सांख्यिकी संकल्पनाएँ : एक सिंहावलोकन : एक आँकड़ा विश्लेषक से इस कार्य में प्रयुक्त विभिन्न गणितीय एवं सांख्यिकी संकल्पनाओं के ज्ञान से संपन्न होने की अपेक्षा की जाती है। फिर, आँकड़ा विश्लेषण में अनेक सांख्यिकी सॉफ्टवेयर भी प्रयोग होते हैं। इसी दृष्टि से इस खंड में आपको सॉफ्टवेयर प्रयोग की एक झलक के साथ उपयुक्त गणितीय एवं सांख्यिकी संकल्पनाओं से परिचित कराया जा रहा है। इस खंड में तीन इकाइयाँ हैं :

इकाई 1 समुच्चय सिद्धांत, संबंध एवं फलन, तर्कशास्त्र एवं प्रमाण निरूपण की प्रविधियों से संबंधित है।

इकाई 2 प्रारंभिक सांख्यिकीय संकल्पनाओं, यथा, सांख्यिकों/आँकड़ों, चर और उसके प्रकार भेद, आँकड़ों के स्वरूप एवं स्रोत और मापन के पैमानों के विवरण के माध्यम से सांख्यिकीय ज्ञान के आधार का निर्माण करती है। इसमें एक-आयामी आँकड़ों के विश्लेषण की प्रतिशत या अनुपात, केंद्रीय प्रवृत्ति के मापकों, विसर्जन/विसरण तथा प्रावस्था के मापकों आदि पर भी विचार किया गया है।

इकाई 3 आपको एक्सेल शीट में आँकड़े भरना और उनका संपादन करना एक सॉफ्टवेयर अनुप्रयोग से अन्य में आँकड़ों का आयात-निर्यात करना सिखाएगी। यहीं आपको आँकड़ों के आधार पर सरल प्राक्कलन और कुछ अन्य प्रक्रियाएँ करना भी सिखाया जाएगा। यहीं आप विभिन्न सांख्यिकीय विश्लेषण विधियाँ भी सीख पाएंगे।

खंड 2 आँकड़ा संकलन और प्रस्तुति विषयक : यह खंड आँकड़े एकत्र करने और उनकी प्रस्तुति से संबंधित है। इसमें भी तीन इकाइयाँ हैं।

इकाई 4 में समष्टि एवं प्रतिदर्श स्तर पर आँकड़ा संकलन, प्रकल्प के नियोजन एवं संगठन रचना विषयक विभिन्न चरणों पर चर्चा करते हुए आपका परिचय अधिकारिक द्वितीय आंकड़ा स्रोतों से भी कराया जाएगा।

इकाई 5 में परिमाणवाची एवं गुणवाची शोध हेतु आँकड़ा संकलन हेतु प्रयुक्त विभिन्न उपस्करों पर चर्चा की गई है।

इकाई 6 में आँकड़ों के वर्गों एवं उपवर्गों में संयोजन, तालिका निर्माण के सामान्य नियमों तथा विभिन्न प्रकार के दंड चित्रों, आयत चित्रों, बारंबारता/प्रायिकता बहुभुजों और प्रायिकता वक्रों आदि से भी आपका परिचय कराया जाएगा।

खंड 3 परिमाणात्मक आँकड़ों के विश्लेषण से संबंधित है। इस खंड में भी 3 ही इकाइयाँ हैं।

इकाई 7 में निष्कर्षात्मक विश्लेषण पर बल दिया गया है। इस प्रकार के विश्लेषण के संदर्भ में प्रसामान्य आवंटन के गुणधर्मों, अवधारणा परीक्षण की संकल्पना तथा इस प्रक्रिया में संभावित त्रुटियों के प्रकार भेदों, विश्वस्ति अंतराल आदि पर चर्चा की गई है।

इकाई 8 प्रतीपगमन विधि से द्वि-विचर आँकड़ों के विश्लेषण से संबंधित है।

इकाई 9 में बहु-विचर विश्लेषण में प्रयुक्त विभिन्न तकनीकों का प्रयोग समझाते हुए इस प्रकार के विश्लेषण के अनुप्रयोग एवं परिणामों की व्याख्या संबंधी दिशा-निर्देश प्रदान किए गए हैं।

खंड 4 संयुक्त सूचकांकों और गुणात्मक आँकड़ों के विवेचन से संबंधित है। इसमें केवल दो इकाइयाँ हैं।

इकाई 10 संयुक्त सूचकांकों के निर्माण की विधियाँ समझाते हुए पाठक को इन सूचकाकों से प्राप्त परिणामों की व्याख्या करना सिखाएगी।

इकाई 11 पाठ्यक्रम की इस अंतिम इकाई में गुणवाची शोध प्रकल्प की रचना और गुणवाची आँकड़ों के विश्लेषण की तकनीकों से आपका परिचय कराया जाएगा।

खंड 1
गणितीय एवं सांख्यिकीय संकल्पनाएँ: एक सिंहावलोकन

इकाई 1 गणितीय संकल्पनाएँ

संरचना

1.0 उद्देश्य

इस इकाई के अध्ययन के पश्चात्, आप निम्नलिखित से भली-भाँति परिचित हो सकेंगे:

- समुच्चय, समुच्चयों की संकेतन पद्धति, समुच्चयों में संबंध तथा समुच्चयों पर संक्रियाओं एवं उनके गुणधर्मों से;
- संबंध तथा फलन की संकल्पनाओं तथा उनके गुणधर्मों से;
- विभिन्न संख्या समुच्चयों, उनके बीच संबंधों तथा उनके गुणधर्मों से;
- प्रतीकात्मक तर्क-शास्त्र तथा तार्किक संयोजनों की संकल्पनाओं से; तथा
- उपपत्ति की विभिन्न विधियों से।

1.1 विषय प्रवेश

हम आदिकाल से संख्याओं के रूप में गणित का उपयोग अपनी संपत्ति का हिसाब रखने के लिए, उसमें बढ़ोत्तरी करने के लिए तथा आबंटन के लिए करते चले आ रहे हैं। औद्योगिक क्रांति के पश्चात् हमारे उत्पादन, आबंटन, उपभोग, विनिमय, व्यापार की पद्धतियाँ अत्यंत जटिल होती चली गईं। इससे हम गणित के और जटिल उपकरणों तथा सूत्रों को अपनाने को प्रेरित/बाध्य हुए।

अर्थशास्त्र की समस्याओं को हल करने में गणित किस प्रकार उपयोगी है?

आइये, हम अर्थशास्त्र तथा गणित की कुछ मूलभूत संकल्पनाओं को पुनः स्मरण करें :

अर्थशास्त्र

- पसंद, समझौताकारी समन्वयन (trade off), प्रोत्साहन, विनिमय, सूचना, आबंटन, कमी / अभाव
- बाज़ार (उत्पाद बाज़ार, श्रम बाज़ार, पूँजी बाज़ार), वस्तुएँ (उपभोग एवं पूंजीगत), सेवाएं
- सहसंबंध, कारण-कार्य संबंध
- अर्थव्यवस्था तथा आर्थिक पद्धति
- ये संरचनाएं मूलभूत आर्थिक प्रश्नों के सूत्रीकरण में उपयोगी सिद्ध होती हैं, जैसे– किस वस्तु का उत्पादन किया जाए?
- यह उत्पादन कैसे किया जाए?
- किसके लिए उत्पादन किया जाए?

गणित

- **संख्याएँ** : उदाहरण के लिए, 3, −12, 8.7, 2/3, $\sqrt{2}$, 557 और 7 – 3i इत्यादि जहाँ 3 एक प्राकृतिक संख्या है, −12 एक पूर्णांक है, 8.7 और 2/3 दोनों **परिमेय (rational) संख्याएँ** हैं तथा $\sqrt{2}$ और $\sqrt[5]{7}$ दोनों **अपरिमेय (irrational) संख्याएँ** हैं। ये सभी संख्याएँ 3, −12, 8.7, 2/3, $\sqrt{2}$, $\sqrt[5]{7}$ वास्तविक (real) संख्याओं की श्रेणी में आती हैं। संख्या 7 – 3i एक सम्मिश्र (complex) संख्या है, यह एक वास्तविक संख्या नहीं है। इन संख्याओं को जिस वर्ग क्रम में लिखा गया है, कोई भी संख्या पिछले किसी वर्ग में सम्मिलित नहीं है। उदाहरण के लिए, $\sqrt{2}$ एक अपरिमेय संख्या है परंतु यह एक परिमेय संख्या नहीं है, यह एक पूर्णांक भी नहीं है और एक प्राकृतिक संख्या भी नहीं है।
- **अंकगणितीय संक्रियाएँ** : +, −, ×, ÷, घात, जैसे कि 5^3 इत्यादि अंकगणित की विभिन्न संक्रियाओं के उदाहरण हैं।
- **अंकगणितीय संबंध** : उदाहरण के लिए, =, ≠, ≤ तथा ≥
- **चर, अचर, समीकरण एवं हल** : उदाहरण के लिए, $x^2 - 16 = 0$ एक समीकरण है जिसमें x एक चर है 16 तथा 0 और दो अचर हैं। इस समीकरण के दो हल हैं, $x = 4$ तथा $x = -4$
- **समुच्चय** : (कुत्ता, बिल्ली, भैंस, घोड़ा, भेड़ बकरी) एक समुच्चय का एक उदाहरण है। इस समुच्चय में कुछ पालतू पशुओं के नाम सम्मिलित किए गए हैं। {4, − 4}समीकरण $x^2 - 16 = 0$ के दोनों हलों का समुच्चय है और {0, 1, 2, 3, 4, 5, 6, 7, 8, 9} उन 10 अंकों का समुच्चय है जिनका प्रयोग दशमलव प्रणाली के अंतर्गत संख्याओं को निरूपित करने के लिए किया जाता है। ध्यान रहे कि किसी समुच्चय को व्यक्त करने के लिए हम धनुकोष्ठकों '{' तथा '}' का इस्तेमाल करते हैं जैसे कि यहाँ उदाहरणों में किया गया है।

- **अंतराल**
  - अंतराल [3, 7.5] ऐसी सभी वास्तविक संख्याओं का समुच्चय है जो 3 और 7.5 के बीच स्थित हैं और इसमें दोनों संख्याएँ 3 तथा 7.5 भी सम्मिलित हैं। ऐसे अंतराल को एक संवृत (closed) अंतराल कहते हैं।
  - अंतराल] 3, 7.5] जिसे (3, 7.5) से भी व्यक्त किया जाता है, ऐसी सभी वास्तविक संख्याओं का समुच्चय है जो 3 और 7.5 के बीच स्थित हैं। इसमें 7.5 सम्मिलित है परंतु 3 नहीं। ऐसे अंतराल को एक अर्द्ध-संवृत (semi-closed) या अर्द्ध-विवृत (semi-open) अंतराल कहते हैं।
  - अंतराल [3, 7.5[जिसे (3, 7.5) से भी व्यक्त किया जाता है। इसमें 3 तो सम्मिलित किया जाता है परन्तु 7.5 शामिल नहीं है। ऐसे अन्तराल को अर्द्ध-संवृत या अर्द्ध-विरल अंतराल कहते हैं।
  - अंतराल] 3, 7.5[, जिसे (3, 7.5]) से भी व्यक्त किया जाता है सभी वास्तविक संख्याओं का समुच्चय है जो 3 और 7.5 के बीच स्थित हैं, परंतु इसमें 3 और 7.5 में से कोई भी संख्या सम्मिलित नहीं होती। ऐसे अंतराल को एक विवृत अंतराल कहते हैं।
- **किसी संख्या का निरपेक्ष मान** : किसी वास्तविक संख्या x का निरपेक्ष मान $|x| = \begin{cases} -x, & x < 0 \\ x, & x \geq 0 \end{cases}$ होता है। उदाहरण के लिए, $|5| = 5$ तथा $|-5| = 5$ है।
- ***प्रतीक, चर, अचर तथा प्राचल***

  व्यापक रूप में एक **प्रतीक**, एक संकेत चिन्ह होता है, जिसका प्रयोग वस्तुतः किसी अन्य वस्तु/गणितीय इकाई को व्यक्त करने के लिए किया जाता है।

  एक **चर** एक ऐसा प्रतीक होता है जिसे मापा जा सकता है और जो परिवर्तनशील हो सकता है। चरों को व्यक्त करने के लिए हम सामान्यतः अंग्रेज़ी वर्णमाला के अंतिम वर्णों जैसे कि p, q, r, s, x, y, z इत्यादि का प्रयोग करते हैं।

  गुणांक विशुद्ध रूप से देखा जाए तो एक प्रतीक न होकर एक चिन्ह होता है जो स्वयं को ही प्रकट करता है, जैसे कि चिन्ह 5.67 स्वयं को प्रकट करता है किसी अन्य वस्तु को नहीं।

  एक **अचर**, मान लीजिए c, एक प्रकार का सामान्यीकृत अचर होता है। इसे यदि किसी गणितीय व्यंजक (जैसे कि $3x^2 + 2x + c$) या किसी समीकरण (जैसे कि $3x^2 + 2x + c = 0$) में प्रयोग किया जाए तो c को पूरे/पूरी तर्क/उपपत्ति के दौरान अथवा पूरी हल–प्रक्रिया में, एक अचर माना जाता है। गणित में प्राचल सामान्यतः a, b, c या लैटिन भाषा के प्रारंभिक अक्षरों द्वारा व्यक्त किये जाते हैं।

गणितीय संकल्पनाओं का उपयोग प्रश्नों/समस्याओं के गणितीय निरूपण/ सूत्रीकरण के लिए उन्हें गणितीय संपदा के प्रयोग से हल करने की प्रक्रिया में तथा पुनः प्राप्त हल का वर्णन दी हुई समस्या के हल के रूप में करने के लिए किया जाता है।

अर्थशास्त्र में उपजी समस्याओं को गणित की सहायता से हल करने में सामान्यतः इस पांच चरणों वाली प्रक्रिया को अपनाया जाता है

i) अर्थशास्त्र के क्षेत्र की समस्या को अर्थशास्त्र की संकल्पनाओं के पदों में परिभाषित कीजिए,

ii) इस प्रकार परिभाषित समस्या का गणितीय सूत्रीकरण कीजिए, अर्थात् उसे गणितीय संकल्पनाओं के पदों में व्यक्त कीजिए,

iii) गणितीय संपदाओं जैसे कि विभिन्न विधियाँ, एलगोरिथ्मस्/प्रक्रियाओं, नियमों, सिद्धांतों, रचनाओं के सप्रयोग से इस गणितीय समस्या का हल ज्ञात कीजिए अथवा तार्किक रूप से निष्कर्ष ज्ञात करें। इस प्रकार परिणाम एक गणितीय इकाई के रूप में होगा, जैसे कि एक संख्या, एक व्यंजक, एक सारिणी, एक आलेख या एक तार्किक कथन इत्यादि।

iv) अंतिम परिणाम को, जो कि एक गणितीय इकाई है, पुनः अर्थशास्त्र की संकल्पनाओं के पदों में व्यक्त कीजिए।

v) जाँच कीजिए कि पिछले चरण में प्राप्त परिणाम यथोचित तथा स्वीकार्य है।

ऊपर दिया चरण (ii) ***गणितीय प्रतिरूपण*** (mathematical modeling) कहलाता है। किसी समस्या का हल उतना ही अच्छा होगा जितना अच्छा उस समस्या का गणितीय प्रतिरूपण होगा।

गणित का प्रयोग करने का महत्त्व नीचे दिए तथ्यों में निहित है :

a) यदि किसी समस्या का गणितीय प्रतिरूप ध्यानपूर्वक तथा यथोचित रूप से सूत्रबद्ध किया जाए, तो ऊपर दी गई पांच चरणों वाली प्रक्रिया से प्राप्त हल के अच्छा होने की अपेक्षा होती है,

b) चरण (iii) में गणित की समस्याओं को हल करने के लिए गणित में भरपूर ख़ज़ाना है, तथा

c) गणित, प्राकृतिक विज्ञान की विशेषकर तथा कुछ सीमा तक सामाजिक विज्ञान की भी, समस्याओं को हल करने के लिए दुनिया की सबसे अधिक परखी–आजमाई हुई, सक्षम तथा सफल प्रणाली है।

1.2 समुच्चय-सिद्धांत

समुच्चय की संकल्पना इतनी आधारभूत संकल्पना है कि आज के समय में गणित के किसी भी क्षेत्र का परिचय बिना समुच्चय सिद्धांतों की चर्चा किए नहीं किया जा सकता है। परंतु गणित जैसे महत्त्वपूर्ण शैक्षणिक विषय की यह मूलभूत संकल्पना, स्वयं अपरिभाषित है। वास्तव में, समुच्चय की इस मूलभूत संकल्पना को पूर्णतयः परिभाषित करने के सभी प्रयास असफल रहे हैं तथा इसकी परिभाषा में प्रयोग किए जाने वाले शब्दों तथा संकल्पनाओं की भी परिभाषा/व्याख्या करने की आवश्यकता पड़ती है तथा हम एक कभी न समाप्त होने वाले पाश में उलझते चले जाते हैं।

परंतु समुच्चय गणित की एक महत्त्वपूर्ण संकल्पना है जो संख्याओं, ज्यामितीय आकृतियों इत्यादि की संकल्पना से एकदम अलग है। अतः हम इसकी एक ऐसी परिभाषा को लेकर चलते हैं जिससे हमारा काम भी चल जाए और हम इस संकल्पना के मूलभाव से भली-भांति परिचित भी हो जाएं। यह परिभाषा इस प्रकार है :

समुच्चय भिन्न वस्तुओं का एक सुनिश्चित संग्रह है।

इस परिभाषा के संदर्भ में थोड़ा स्पष्टीकरण अनिवार्य है। मान लीजिए S एक समुच्चय है। अतः वह भिन्न वस्तुओं का एक सुनिश्चित संग्रह होगा। संज्ञा 'संग्रह' के साथ लगाए गए 'विशेषण' सुनिश्चित का अर्थ है कि यदि हम कोई भी वस्तु, मान लीजिए x लें तो हम कम से कम सैद्धांतिक रूप में यह बता सकें कि x, S में है अथवा नहीं। आइये, हम नीचे दिए वस्तुओं के दो संग्रहों C_1 तथा C_2 पर विचार करें।

C_1 = वर्ष 2017 में इग्नू के स्नातक पाठ्यक्रमों में दाखिला लेने वाले सभी विद्यार्थियों का संग्रह

C_2 = वर्ष 2017 में इग्नू के स्नातक पाठ्यक्रमों में दाखिला लेने वाले कुछ विद्यार्थियों का संग्रह

यहाँ C_1 एक सुनिश्चित संग्रह है परंतु C_2 एक सुनिश्चित संग्रह नहीं है। अतः, C_1 एक समुच्चय है जबकि C_2 एक समुच्चय नहीं है।

1.2.1 समुच्चय की संकल्पना तथा उनका निरूपण

हम समुच्चयों के निरूपण की संकेतन विधियों की व्याख्या कुछ उदाहरणों एवं टिप्पणियों द्वारा करेंगे।

i) *लैटिन* वर्णमाला में पाँच (V) स्वर वर्ण होते हैं : a, e, i, o और u। इस तथ्य को इस प्रकार व्यक्त किया जाता है

$$V = \{a, e, i, o, u\}$$

a, e, i, o, u में से प्रत्येक समुच्चय V का एक अवयव कहलाता है।

अवयवों/समुच्चयों इत्यादि का संकेतन : तथ्य e समुच्चय V का एक अवयव है को $\boldsymbol{e \in V}$ द्वारा व्यक्त किया जाता है। इसी प्रकार तथ्य b समुच्चय V का एक अवयव नहीं है, को $\boldsymbol{b \notin V}$ द्वारा व्यक्त किया जाता है।

सामान्यतः, एक समुच्चय को अंग्रेज़ी वर्णमाला के बड़े अक्षरों द्वारा अथवा एक बड़े अक्षर से प्रारंभ करके अक्षरों और अंकों के एक अनुक्रम से व्यक्त किया जाता है। इस उदाहरण में हमने स्वर वर्णों के समुच्चय को V से व्यक्त किया है। इसी प्रकार किसी छोटे अक्षर से अथवा एक छोटे अक्षर से प्रारंभ करके, अक्षरों और अंकों के एक अनुक्रम से व्यक्त किया जाता है। समुच्चयों तथा उनके अवयवों को निरूपित करने की यह व्यवस्था इतनी दृढ़ नहीं है कि इसमें कोई परिवर्तन न किया जा सके। ये निरूपण अन्य प्रकार से भी किया जा सकता है। परंतु किसी समुच्चय या अवयव के लिए प्रयोग किए जाने वाले नाम में कम से कम एक अक्षर अवश्य होना चाहिए, इसे केवल अंकों से अनुक्रम से व्यक्त नहीं किया जाता।

ii) किसी समुच्चय के अवयवों को जब हम धनु कोष्ठकों में लिखते हैं, तो उनके क्रम का कोई महत्त्व नहीं होता। उदाहरण के लिए, समुच्चय $V = \{a, e, i, o, u\}$ को $V = \{o, a, i, u, e\}$ के रूप में भी लिखा जा सकता है।

iii) किसी समुच्चय में अवयवों की पुनरावृत्ति की अनुमति नहीं है अर्थात् सामान्यतः किसी भी अवयव को दोबारा नहीं लिखा जाता। उदाहरण के लिए, वर्णित

समुच्चय V को $V = \{a, e, i, o, e, u\}$ के रूप में लिखना उचित नहीं है क्योंकि इसमें अवयव e को दो बार लिखा गया है।

iv) यदि किसी समुच्चय में अवयवों की संख्या अनंत हो, तो समुच्चय के प्रत्येक अवयव का उल्लेख करना संभव नहीं होता। उदाहरण के लिए, ऐसे सभी पूर्णांकों के समुच्चय पर विचार कीजिए जो किसी पूर्णांक का वर्ग हो। इस समुच्चय के अवयव हैं : 0, 1, 4, 9, 16 इत्यादि। हम देख सकते हैं कि इस समुच्चय के प्रत्येक अवयव को लिखना संभव नहीं है। ऐसी स्थिति में, हम समुच्चय को व्यक्त करने के लिए हम अपनी संकेतन विधि को निम्न रूप से संशोधित करके इस प्रकार लिखते हैं :

$$Sq\_int = \{0, 1, 4, 9, \ldots\}$$

ध्यान दें कि यहाँ हमने समुच्चय के पहले कुछ अवयव लिखकर आगे तीन बिंदु, ..., लगाकर छोड़ दिया है। समुच्चयों के निरूपण की विभिन्न विधियाँ हैं।

समुच्चय निरूपण संकेत

सूची या तालिका स्वरूप : एक सूची या तालिका स्वरूप में समुच्चय के सभी घटकों को सूचीबद्ध किया जाता है, अर्द्ध विराम चिन्हों द्वारा घटकों को पृथक–पृथक रखा जाता है और पूरी सूची को कोष्ठक { }के बीच में ही लिखा जाता है। उदाहरण के लिए, सभी पूर्णांकों का समुच्चय इस प्रकार दर्शाया जा सकता है $Z = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$

समुच्चय निर्माण स्वरूप : set*Sq_in_less_10001* पर विचार कीजिए। इसके प्रत्येक अवयव के दो गुणधर्म हैं : (a) वह एक पूर्णांक का वर्ग होना चाहिए (b) वह 10001 से छोटा होना चाहिए।

समुच्चय-निर्माण संकेतन में, इन गुणधर्मों का प्रयोग करके, इस समुच्चय को इस प्रकार व्यक्त किया जाता है : $Sq\_int\_less\_10001 = \{x: x = y^2$, y एक पूर्णांक है, तथा $x < 10001$ है} इसे इस प्रकार पढ़ा जाता है समुच्चय *Sq_int_less_10001* का अवयव x होगा, जहाँ (जहाँ को ':' या '|' से व्यक्त किया जाता है) x शर्त 'किसी पूर्णांक y के लिए $x = y^2$ हो' को संतुष्ट करना चाहिए तथा $x < 10001$ होना चाहिए।

किसी समुच्चय को समुच्चय-निर्माण विधि से व्यक्त करने के विभिन्न वैकल्पिक तरीके हैं। जैसे कि समुच्चय को निम्न वैकल्पिक रूपों में लिखा जा सकता है :

समुच्चय $Decimal\_digits = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ को इस प्रकार व्यक्त किया जा सकता है।

$Decimal\_digits = \{x: x \in N$ तथा $x < 10\}$ या

$Decimal\_digits = \{x \in N: x < 10\}$ इत्यादि

किसी समुच्चय को समुच्चय-निर्माण विधि द्वारा सुविधा से व्यक्त किया जा सकता है यदि इसके अवयवों द्वारा संतुष्ट किए जाने वाले गुणधर्मों का वर्णन सरलता से किया

जा सके, फिर चाहे समुच्चय के अवयवों की संख्या कितनी भी बड़ी अथवा अनंत ही क्यों न हो।

|S| द्वारा अभिव्यक्त किसी समुच्चय S के अवयवों की संख्या

यदि x एक संख्या है तो सामान्यतः, |x| का प्रयोग, x के निरपेक्ष मान को दर्शाने के लिए किया जाता है। परंतु, यदि यही संकेतन किसी समुच्चय, मान लीजिए X, के साथ प्रयोग किया जाए, तो X के अवयवों की संख्या को व्यक्त करता है। अर्थात् |X|, समुच्चय X के अवयवों की संख्या को दर्शाता है। उदाहरण के लिए, यदि $X= \{1, 2, 3\}$, है, तो $|X| = 3$ होगा। इसी प्रकार समुच्चय $= \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ के लिए |Decimal_digits|= 10 होगा।

अनंत समुच्चय : अनंत संख्या को चिन्ह ∞ से व्यक्त किया जाता है। ∞ की परिभाषा के अनुसार $\infty = \infty - 1 = \infty - 2 = \infty - 3 = \ldots$ होता है।

यदि किसी समुच्चय के अवयवों की संख्या अनंत हो तो हम ऐसे समुच्चय को एक अपरिमित समुच्चय कहते हैं। समुच्चयों $Sq\_int = \{0, 1, 4, 9, \ldots\}$ तथा $Z = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ दोनों ही अपरिमित समुच्चय हैं। ऐसा समुच्चय जिसके अवयवों की संख्या सीमित हो (अर्थात् अनंत न हो) एक परिमित समुच्चय कहलाता है।

ध्यान दें कि $|Z|= \infty$ है। इसी प्रकार $|Sq\_int|= \infty$ है, परंतु $|Sq\_int\_less\_10001| \neq \infty$ है।

1.2.2 समुच्चयों में संबंध

हम 'समुच्चयों के बीच संबंध' तथा 'एक समुच्चय में संबंध' की संकल्पनाओं की औपचारिक परिभाषा बाद में देंगे। अभी के लिए हम अपनी चर्चा को केवल दो दिए हुए समुच्चयों, मान लीजिए X और Y के बीच कुछ विशिष्ट संबंधों जैसे कि 'का एक उपसमुच्चय है', 'के समान है' तथा ''एक समुचित उपसमुच्चय'' तक सीमित रखते हैं।

एक समुच्चय का उपसमुच्चय

हम जानते हैं कि समुच्चय वस्तुओं के विशिष्ट संग्रह होते हैं। यह भी संभव है कि किसी समुच्चय का प्रत्येक अवयव किसी अन्य समुच्चय का भी एक अवयव हो। यदि ऐसा है तो हम पहले समुच्चय को दूसरे समुच्चय का उपसमुच्चय कहते हैं।

औपचारिक रूप से, दो समुच्चयों X और Y के लिए, यदि ऐसा हो कि जब भी x, समुच्चय X का अवयव हो तो X समुच्चय Y का भी अवयव हो, तो X, Y का उपसमुच्चय कहलाता है। इस तथ्य को हम संकेतन $X \subseteq Y$ से व्यक्त करते हैं।

उदाहरण के लिए, यदि *Alphabet* अंग्रेज़ी वर्णमान के सभी अक्षरों a, b, c …, w, x, y, z का समुच्चय है अर्थात् Alphabet={a,b, c …, w, x, y, z}, तथा V= {a, e, i, o, u} है, तो $V \subseteq$ Alphabet होगा। यह कहने कि बजाए कि *X, Y* का उपसमुच्चय है, हम यह भी कह सकते हैं कि समुच्चय *X*, समुच्चय *Y* में सम्मिलित है। चिन्ह '$\subseteq$' को 'सम्मिलित होने की असमिका' भी कहते हैं।

हम चिन्ह $\notin$ का प्रयोग 'में स्थित नहीं है' को दर्शाने के लिए करेंगे।

समुचित उपसमुच्चय : मान लीजिए x, y का उपसमुच्चय है अर्थात् जब भी $x \in X$ तो $x \in Y$ है। यदि साथ ही यह भी सत्य हो कि कम से कम एक Y ऐसा है कि $y \in Y$ परंतु $y \notin X$ तो X, Y का समुचित उपसमुच्चय कहलाता है। इसे हम संकेतन $X \subset Y$ से व्यक्त करते हैं। इसे X, Y का एक समुचित उपसमुच्चय है या 'X, Y में समुचित रूप से सम्मिलित है' पढ़ा जाता है।

टिप्पणी : उपसमुच्चय के लिए प्रयोग किए जाने वाले संकेतन का पूर्णतयः मानकीकरण नहीं है। कुछ लोग '$\subset$' का प्रयोग एक उपसमुच्चय के लिए करते हैं, चाहे वह एक उचित उपसमुच्चय हो अथवा नहीं। परंतु हम अपनी चर्चा में चिन्ह '$\subseteq$' का प्रयोग उपसमुच्चय के लिए तथा '$\subset$' का प्रयोग समुचित उपसमुच्चय के लिए करेंगे।

उदाहरण के लिए, यदि

Alphabet = {a, b, c …, w, x, y, z} और V = {a, e, i, o, u} है

तो हम इसे

$$V \subset Alphabet$$

द्वारा व्यक्त करेंगे क्योंकि $b \in Alphabet$ परंतु $b \notin V$

चिन्ह $\not\subset$: यदि X, Y का उपसमुच्चय नहीं है, तो इस तथ्य को $\mathbf{X \not\subset Y}$ से व्यक्त किया जाता है।

दो समुच्चयों की समानता

दो समुच्चय समान कहलाते हैं यदि $X \subseteq Y$ तथा $Y \subseteq X$ हो। दो समुच्चयों X और Y की समानता का संकेतन $X=Y$ है।

उदाहरण के लिए : समुच्चय *Decimal-digits* = {0, 1, 2, 3, 4, 5, 6, 7, 8, 9} तथा $X = \{x \in N: x < 10\}$ समान समुच्चय है अर्थात् *Decimal-digits* = X है।

दो समुच्चयों X और Y की असमानता, $X \neq Y$ द्वारा दर्शित : दो समुच्चय X और Y समान नहीं होंगे (या असमान होंगे) यदि निम्नलिखित में से कम से कम एक सत्य हो:

i) एक ऐसा x हो जो X में हो परंतु Y में न हो, अर्थात् $x \in X$ परंतु $x \notin Y$ हो, अथवा

ii) एक ऐसा x हो जो Y में हो परंतु X में न हो, अर्थात् $x \in Y$ परंतु $x \notin X$ हो।

टिप्पणी : गणित तथा तर्कशास्त्र में वाक्यांश *A* या *B* में शब्द 'या' का अर्थ होता है : *A* या या *B* या *A* और *B* दोनों।

बोध प्रश्न 1

1) इनमें से कौन से समूह समुच्चय हैं? अपने उत्तर का औचित्य बताइए!
   i) वर्ष के M अक्षर से प्रारंभ होने वाले महीने का समूह
   ii) वर्ष के Z अक्षर से प्रारंभ होने वाले महीनों का समूह
   iii) विश्व के 10 सबसे गुणी जीवित लेखकों का समूह
   iv) विश्व के 11 सबसे अधिक प्रख्यात फुटबाल खिलाड़ियों की टीम
   v) $\{1, 2, 2, 3\}$
   vi) $\{1, 2, \{2,3\}, 3\}$.
2) निम्नलिखित संख्याओं के प्रत्येक समुच्चय (i) को सूची/तालिका रूप में लिखें, (ii) की गणनीयता बताएँ
   क) 180 के धनात्मक अभाज्य गुणन खंड
   ख) द्विघात समीकरण $x^2 - 3x - 10 = 0$ के हल।
3) इनमें से कौन–कौन समुच्चय हैं? यदि कोई समुच्चय है तो
   i) इसे समुच्चय निर्माण विधि से लिखें
   ii) उसकी गणनीयता बताएँ
   क) $\{6, 9, 12, 15, 18, 21, 24, 27\}$
   ख) $\{2, 4, 8, 16 \ldots\}$
   ग) $\{\{1\}, \{1, 2\}, \{1, 2, 3\} \ldots \{1, 2, 3, 4 \ldots 10\}\}$.

1.2.3 विशेष समुच्चय : परिभाषाएँ तथा गुणधर्म

सार्वत्रिक समुच्चय (universal set)

किसी वास्तविक जीवन से जुड़ी किसी समस्या का हल ज्ञात करने में गणित का प्रयोग करने के लिए हमें किसी विशेष संदर्भ में चर्चा हेतु किसी आधारभूत प्रांत से संबंधित वस्तुओं के समुच्चयों पर विचार करना पड़ता है। वह आधारभूत समुच्चय जिसके अवयवों एवं उपसमुच्चयों पर किसी विशेष संदर्भ में, हमें विचार करना पड़ता है, सार्वत्रिक समुच्चय कहलाता है तथा सामान्यतः इसे U से व्यक्त किया जाता है।

रिक्त समुच्चय (The Empty or Null Set)

एक ऐसा समुच्चय जिसमें कोई भी अवयव न हो, रिक्त समुच्चय कहलाता है। यह अजीब लग सकता है परंतु ऐसे समुच्चयों का अस्तित्व एक वास्तविकता है। उदाहरण के लिए, संग्रह

$$S = \{x : x \text{ सम पूर्णांक है, और } x^2 = 9\}$$

पर विचार कीजिए। S एक सुनिश्चित संग्रह है, अतः, एक समुच्चय है। परंतु ऐसा कोई भी सम पूर्णांक नहीं है जिसका वर्ग 9 के बराबर हो अर्थात् इस समुच्चय में कोई भी अवयव नहीं है। यह एक रिक्त समुच्चय है। स्वाभाविक रूप से एक रिक्त समुच्चय के अवयवों की संख्या 0 होगी। **एक रिक्त समुच्चय को Ø अथवा { } द्वारा व्यक्त किया जाता है।**

टिप्पणी : Ø प्रत्येक समुच्चय X का एक उपसमुच्चय होता है अर्थात् प्रत्येक समुच्चय X के लिए $\emptyset \subseteq X$ होता है।

1.2.4 समुच्चयों के रूप में अंतराल : संकेतन

अंतरालों के संकेतन के बारे में ऊपर उल्लेख किया जा चुका है। यह कहा जा सकता है कि प्रत्येक अंतराल – चाहे वह संवृत हो, विवृत या अर्द्ध संवृत/विवृत – वास्तविक संख्याओं के समुच्चय का एक उपसमुच्चय होता है, एक ऐसा उपसमुच्चय है जिसमें यह गुणधर्म होता है कि यदि x और y अंतराल I के दो अवयव हैं तथा $x \leq y$ है तो x और y के मध्य स्थित प्रत्येक वास्तविक संख्या भी उसका एक अवयव होगी।

परंतु, एक अंतराल को व्यक्त करने के लिए धनुकोष्ठकों '{' तथा '}' के स्थान पर वर्ग कोष्ठकों '[' तथा ']' और/या लघु कोष्ठकों '(' तथा ')' का प्रयोग किया जाता है। जैसा कि नीचे दर्शाया गया है।

किन्हीं भी वास्तविक संख्याओं a तथा b के लिए, जहाँ $\boldsymbol{a < b}$ हो

- अंतराल [a, b] a तथा b के बीच की सभी वास्तविक संख्याओं को दर्शाता है। इसमें a और b भी सम्मिलित हैं। ऐसे अंतराल को एवं *संवृत अंतराल* कहते हैं।
- अंतराल] a, b] जिसका अन्य पाठ्यांतर ***(a, b]*** भी है, a तथा b के बीच की वास्तविक संख्याओं का ऐसा समुच्चय है जिसमें 'b' तो शामिल है किंतु a नहीं है। ऐसे अंतराल को अर्द्ध-संवृत्त या अर्द्ध-विवृत अंतराल कहते हैं।
- अंतराल [a, b[जिसका पाठ्यांतर (***[a, b)***) है, a तथा b के बीच की उन वास्तविक संख्या का समुच्चय है जिसमें a तो शामिल है किंतु b नहीं। ऐसे अंतराल को भी अर्द्ध-संवृत या अर्द्ध-विवृत कहा जाता है।
- अंतराल]a, b[– जिसका पाठ्यांतर ***(a, b)*** भी है a और b के बीच की उन वास्तविक संख्याओं का समुच्चय है जिसमें a और b दोनों ही संख्याएँ स्वयं शामिल नहीं हैं। ऐसे अंतराल को अनावृत्त या विवृत अंतराल कहा जाता है।

1.2.5 संख्याओं के समुच्चय (मानक संकेतनों के साथ)

अब हम संक्षेप में संख्याओं के उन समुच्चयों पर चर्चा करेंगे जिनका प्रयोग बार–बार किया जाता है तथा उनके नामों/संकेतनों का भी उल्लेख करेंगे।

$\mathbf{N} = \{1, 2, 3\ldots\}$ को प्राकृतिक संख्याओं का समुच्चय कहते हैं।

$\mathbf{W} = \{0. 1, 2, 3 \ldots\}$ को पूर्ण संख्याओं का समुच्चय कहते हैं।

एक बात पर विशेष ध्यान दें : काफी समय से '0' शून्य को भी एक प्राकृतिक संख्या माना जाने लगा है। अतः, अब $N=W$ और पूर्ण संख्या समुच्चय को अलग से दिखाना आवश्यक नहीं रहा है।

$\mathbf{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3\ldots\}$ को पूर्णांकों का समुच्चय कहते हैं

$\mathbf{Q} = \{x | x=p/q$ के साथ $q \neq 0$, और $p, q \in \mathbf{Z}\}$ को परिमेय संख्याओं का समुच्चय कहते हैं।

टिप्पणी : यह सरलतापूर्वक देखा जा सकता है कि $\emptyset \subset \mathbf{N} \subset \mathbf{W} \subset \mathbf{Z} \subset \mathbf{Q}$ होता है।

इन समुच्चय–असमिकाओं में Z⊂Q को छोड़कर शेष सभी सुस्पष्ट हैं। Z⊂Q की सत्यता जानने के लिए ध्यान दें कि यदि $p \in \boldsymbol{Z}$ **है** तो $p = (p/1)$ लिखा जा सकता है, अतः $p = (p/1) \in Q$ होगा। साथ ही, $3/4 \in Q$ है तथा $3/4 \notin Z$ है।

R, वास्तविक संख्याओं का समुच्चय

वास्तविक संख्याओं के समुच्चय को **R** से व्यक्त किया जाता है। इसे इस प्रकार परिभाषित किया जा सकता है : $\mathbf{R} = \{\mathbf{x} | \text{ x एक वास्तविक संख्या है}\}$

यह सरलता से देखा जा सकता है एक वर्ग जिसकी भुजा की लंबाई 1 इकाई हो, उसके विकर्ण की लंबाई $\sqrt{2}$ होती है। अतः यदि एक सरल रेखा पर 0 से प्रारंभ करके हम रेखा के दाईं ओर $\sqrt{2}$ इकाई की दूरी तय करें, तो हमें वास्तविक संख्या $\sqrt{2}$ के संगत एक बिंदु प्राप्त होगा। परंतु, यह सिद्ध किया जा चुका है ऐसी कोई परिमेय संख्या p/q नहीं है जिसका मान $\sqrt{2}$ के बराबर हो। ऐसी वास्तविक संख्या जो एक परिमेय संख्या न हो, एक अपरिमेय संख्या कहलाती है। यह सिद्ध किया जा चुका है कि अपरिमेय वास्तविक संख्याओं की संख्या, परिमेय वास्तविक संख्याओं की संख्या से कहीं अधिक है। **अपरिमेय संख्याओं के समुच्चय को R–Q या R ~ Q से व्यक्त किया जा सकता है।**

टिप्पणी : ऊपर की गई चर्चा के आधार पर हम संख्याओं के विभिन्न समुच्चयों के बीच सम्मिलित होने के संबंध अर्थात् '⊂' का विस्तार इस प्रकार कर सकते हैं :

$$\emptyset \subset N \subset W \subset Z \subset Q \subset R$$

इन संबंधों की सत्यता को सरलता से परखा जा सकता है।

सम्मिश्र संख्याओं का समुच्चय C

$x^2 + 1 = 0$ या $x^2 = -1$ प्रकार के समीकरणों का हल ज्ञात करने के संदर्भ में हमें वास्तविक संख्याओं के समुच्चय के आगे भी विस्तार करने की आवश्यकता पड़ती है।

हम जानते हैं कि प्रत्येक वास्तविक संख्या का वर्ग, एक धनात्मक वास्तविक संख्या होती है। अतः, इस समीकरण का हल एक वास्तविक संख्या नहीं हो सकती।

सम्मिश्र संख्याओं के समुच्चयों का विकास $x^2+1 = 0$ प्रकार (जिनका हल वास्तविक संख्याओं में उपलब्ध नहीं है) के समीकरणों को हल करने के प्रयास के संदर्भ में किया गया।

एक ऐसी संख्या की कल्पना की गई जिसका वर्ग –1 के बराबर हो और इसे i से व्यक्त किया गया अर्थात् $C^2 = -1$ माना गया। इसके आधार पर C को इस प्रकार परिभाषित किया गया :

$$C = \{z: z = x + i\, y: \; x, y \in R\}$$

(टिप्पणी : $x + i\, y$ को $x + y\, i$ भी लिखा जा सकता है)

असमिका (inequality) $R \subset C$ की सत्यता को इस प्रकार देखा जा सकता है : मान लीजिए $x \in R$ है। अतः, $x = x + i\, 0 \in C$ होगा। साथ ही, ऐसे कई $1 + i\, 2$ जैसी

सम्मिश्र संख्याएँ हैं जहाँ $1 + i \notin R$. इसी प्रकार, जैसे कि ऊपर चर्चा किया जा सकता है कि $1 \notin R$ है परंतु $i= 0 +1\,i \in C$ है।

अतः, उप-समुच्चयों की सर्वसमिका (inequality) को आगे बढ़ाते हैं तथा प्राप्त करते हैं:

$$\emptyset \subset N \subset W \subset Z \subset Q \subset R \subset C$$

टिप्पणी : संकेतक Ø,N,W, Z, Q,R, C मानक संकेतन हैं तथा इनका प्रयोग करते हुए इन्हें पुनः परिभाषित नहीं करना पड़ता। परंतु, अन्य नाम/संकेतन जैसे कि Sq_int इत्यादि मानक नहीं हैं। हम इसको किसी अन्य नाम जैसे कि Sq_int से भी व्यक्त कर सकते हैं। समुच्चयों के ऐसे संकेतनों/नामों के लिए, प्रत्येक नए संदर्भ में, समुच्चय की परिभाषा पुनः देने की आवश्यकता पड़ती है।

बोध प्रश्न 2

1) समुच्चयों के युग्म X तथा Y के लिए, बताइए कि X = Y है अथवा नहीं। अपने उत्तर की व्याख्या भी कीजिए

   i) $X = \{2, 4, 6, 8, 10\}$ तथा $Y = \{x$: x एक सम धनात्मक पूर्णांक है, तथा $x \leq 10$ है$\}$

   ii) $X = \{x$: 10 का गुणज है$\}$, तथा $Y = \{10, 15, 20, 25, 30 \ldots\}$है

2) नीचे दिए समुच्चयों X और Y के लिए $X \subseteq Y, Y \subseteq X, X \subset Y, Y \subset X$ या $X \not\subset Y$ में से कौन-सा संबंध सही है :

   i) X=$\{x$: x इग्नू का एक विद्यार्थी है जिसका एक विषय अर्थशास्त्र है$\}$,
   Y=$\{x$: x इग्नू का एक विद्यार्थी है$\}$

   ii) X= $\{x$: x तल में एक त्रिभुज है$\}$,
   Y= $\{x$: x तल में एक चतुर्भुज है$\}$

3) $X = \{1, \{2, 3\}, 3, 4\}$समुच्चय के लिए नीचे दिए कथनों में से कौन–से कथन सत्य हैं :

   i) $2 \in X$

   ii) $\emptyset \in X$

   iii) $\emptyset \not\subset X$

   iv) $\{2,3\} \in X$

   v) $\{2,3\} \subset X$

1.2.6 समुच्चयों पर संक्रियाएँ (Set Operations)

हम जानते हैं कि दी हुई संख्याओं पर संख्या–संक्रियाओं के प्रयोग से हम नई संख्याएं प्राप्त करते हैं। उदाहरण के लिए, एक दी हुई संख्या, मान लीजिए 2 के लिए हम उसका योगात्मक व्युत्क्रम (–2) प्राप्त करते हैं। योगात्मक व्युत्क्रम, गुणात्मक व्युत्क्रम तथा वर्गमूल, ये भी किसी संख्या पर एक एकांगी संक्रिया होती है। ये संक्रियाएँ एकांगी संक्रियाएँ कहलाती हैं क्योंकि इन संक्रियाओं को एक संख्या पर लगाया जाता है और इससे हमें एक अन्य (सामान्यतः नयी) संख्या प्राप्त होती है। उदाहरण के लिए,

यदि हम संख्या 2 पर गुणात्मक व्युत्क्रम संक्रिया का प्रयोग करें तो हम $2^{-1} = (1/2)$ प्राप्त करते हैं। इसी प्रकार, दो दी हुई संख्याओं, जैसे कि 2 और –5, पर विभिन्न युग्मक संक्रियाएँ जैसे कि योग (+), व्यवकलन (minus) (–), गुणा (×), भाग (÷) तथा घात लगाकर, हम क्रमशः 3 (= 2– 5), 7 (= 2– (–5)), –10 (= 2× (– 5)), –2/5 (= 2/ (– 5)) तथा 2^{-5} (= 2 की घात– 5) प्राप्त करते हैं। योग, व्यवकलन, गुणा, भाग तथा घात में से प्रत्येक एक युग्मक संक्रिया है क्योंकि इनमें से प्रत्येक को लगाने के लिए दो संख्याओं की आवश्यकता होती है और इन संक्रियाओं के फलस्वरूप हमें एक अन्य संख्या प्राप्त होती है।

इसी प्रकार समुच्चयों के लिए भी एकांगी और युग्मक संक्रियाएँ परिभाषित की गयी हैं जिससे पर्याप्त (आवश्यक) संख्या में समुच्चय लेकर उन पर ये संक्रियाएँ लगाकर परिणाम (जो कि एक समुच्चय ही होगा) प्राप्त किया जा सके।

हम केवल दो प्रकार की समुच्चय संक्रियाओं पर विचार करेंगेः (i) एकांगी (unary) *समुच्चय संक्रियाएँ,* तथा (ii) *युग्मक (binary) समुच्चय संक्रियाएँ*। पहले हम युग्मक संक्रियाओं पर चर्चा करते हैं। इसके पश्चात् हम केवल एक एकांगी संक्रिया की चर्चा करेंगे जो कि एक युग्मक संक्रिया की एक विशेष स्थिति है। दो दिए हुए समुच्चयों X और Y के लिए, हम जिन युग्मक समुच्चय संक्रियाओं की चर्चा करेंगे, वे हैं : (i) समुच्चयों का सम्मिलन (union) जिसे चिन्ह '∪' से व्यक्त किया जाता है जैसे कि $X \cup Y$ में, (ii) समुच्चयों का उभयनिष्ठ (intersection) जिसे कि चिन्ह '∩' से व्यक्त किया जाता है जैसे कि $X \cap Y$ में, (iii) दो समुच्चयों का अंतर जिसे '– ' या '~' से व्यक्त किया जाता है जैसे कि X – Y या X~Y में, (iv) दो समुच्चयों का सममित अंतर जिसे Δ से व्यक्त किया जाता है जैसे कि X Δ Y में। **अंत में, हम एकांगी संक्रिया एक समुच्चय का पूरक, जिसे ' , ' से व्यक्त किया जाता है, जैसे कि *X'* में।**

नोटः चिन्ह '∪' समुच्चयों का सम्मिलन व्यक्त करता है जबकि

चिन्ह 'U' सार्वत्रिक समुच्चय को निरूपित करता है।

प्रत्येक संक्रिया पर चर्चा हम इस रूप में करेंगे : पहले संक्रिया की परिभाषा दी जाएगी, जिसे एक उदाहरण द्वारा स्पष्ट किया जाएगा तथा अंततः संक्रिया के कुछ गुणधर्मों का उल्लेख किया जाएगा।

नीचे दी गई चर्चा में, विशेषकर गुणधर्मों में *X, Y* और *Z* कोई भी स्वछंद/यादृच्छिक समुच्चय हैं।

क) *दो समुच्चयों X और Y का सम्मिलिन X∪ Y को इस प्रकार परिभाषित करते हैं।*

$$X \cup Y = \{x: x \in X \textbf{ या } x \in Y\}$$

उदाहरण के लिए, यदि X= {1, 2, 3, 4} तथा Y = {2, 3, 5, 8, 9}है,

तो $X \cup Y = \{1, 2, 3, 4, 5, 8, 9\}$ होगा।

ध्यान दें कि जो अवयव X और Y दोनों में उपस्थित हों, तो उन्हें $X \cup Y$ में केवल एक बार लिखा जाता है क्योंकि एक समुच्चय का वर्णन करते हुए उसके अवयवों को दोहराया नहीं जाता।

समुच्चयों के सम्मिलन के गुणधर्म

किन्हीं भी समुच्चयों *X, Y* और *Z* के लिए निम्नलिखित सर्वसमिकाएं (identies) सत्य हैं:

i) $X \cup Y = Y \cup X$ (समुच्चयों के सम्मिलन का क्रम–विनिमय (commutative) नियम)

ii) $(X \cup Y) \cup Z = X \cup (Y \cup Z$ का साहचर्य नियम)

iii) $X \cup \emptyset = X$ (Ø (तत्समक नियमः रिक्त समुच्चय सम्मिलन की संक्रिया के लिए तत्समक है)

iv) $X \cup X = X$ (सम्मिलन के संदर्भ में वर्गसम नियम)

v) $U \cup X = U$

ख) दो समुच्चयों *X* और *Y* का उभयनिष्ठ $X \cap Y$ इस प्रकार परिभाषित किया जाता है:

$$X \cap Y = \{x: x \in X \text{ तथा } x \in Y\}$$

उदाहरण के लिए, यदि X= {1, 2, 3, 4} और Y = {2, 5, 8, 3, 9} है तो $X \cap Y = \{2, 3\}$ होगा।

समुच्चयों के उभयनिष्ठ के कुछ गुणधर्म

किन्हीं भी समुच्चयों *X, Y* तथा *Z* के लिए निम्नलिखित सर्वसमिकाएं सत्य हैं :

i) $X \cap Y = Y \cap X$ (समुच्चयों के उभयनिष्ठ का क्रम-विनिमय नियम)

ii) $(X \cap Y) \cap Z = X \cap(Y \cap Z)$ (समुच्चयों के उभयनिष्ठ ∩ का साहचर्य (associative) नियम)

iii) $X \cap U = X$ (तत्समक नियम : सार्वत्रिक समुच्चय U, उभयनिष्ठ के लिए तत्समक है)

iv) $X \cap X = X$ (उभयनिष्ठ के संदर्भ में वर्गसम नियम '∩')

v) $\emptyset \cap X = \emptyset$

ग) दो समुच्चयों *X* और *Y* का अंतर *X – Y* इस प्रकार परिभाषित किया जाता है :

$$X - Y = \{x: x \in X \text{ तथा } x \notin Y\}$$

उदाहरण के लिए

यदि X= {1, 2, 3, 4} तथा Y = {2, 5, 8, 3, 9} है,

ते $X - Y = \{1, 4\}$ तथा $Y - X = \{5, 8, 9\}$ होगा

समुच्चयों के अंतर के कुछ गुणधर्म

i) अंतर संकारक (Difference Operator) क्रम विनिमय नियम का पालन नहीं करता अर्थात् यह आवश्यक नहीं है कि प्रत्येक X और Y के लिए Y, $X - Y \neq Y - X$ हो। ऊपर दिए गए उदाहरण में हम देख सकते हैं कि $X - Y \neq Y - X$ है।

एक अन्य सरल उदाहरण :

मान लीजिए $X = \{1, 2, 3, 4\}$ तथा $Y = \emptyset$ है, तो $X - Y = X \neq \emptyset = Y - X$ है

ii) अंतर संकारक साहचर्य नियम का पालन नहीं करता अर्थात् यह भी संभव है कि $(X - Y) - Z \neq X - (Y - Z)$ हो।

माना $X = \{1, 2, 3, 4\}$, $Y = \{2, 3, 5, 8, 9\}$ तथा $Z = \{1, 5, 7, 8\}$ है

तो $X - Y = \{1, 4\}$, होगा। अतः, $(X - Y) - Z = \{4\}$ होगा

दूसरी ओर

$Y - Z = \{2, 3, 9\}$ होगा तथा $X - (Y - Z) = \{1, 4\}$ होगा।

अतः, $(X - Y) - Z = \{4\} \neq \{1, 4\} = X - (Y - Z)$

एक और सरल उदाहरण :

माना $X = \{1, 2, 3, 4\}$, $Y = \emptyset$ तथा $Z = \{1\}$ है तो $(X - Y) - Z = \{1, 2, 3, 4\} - \{1\} = \{2, 3, 4\}$ होगा।

$(Y - Z) = \emptyset$ है, अतः $X - (Y - Z) = \{1, 2, 3, 4\} - \emptyset = \{1, 2, 3, 4\}$ होगा।

iii) अंतर संकारक के लिए द्विपक्षीय/दोतरफा तत्समक का अस्तित्व नहीं है। परंतु $\emptyset$ एक तरफ से तत्समक का काम करता है क्योंकि $X - \emptyset = X$ है परंतु $\emptyset - X = \emptyset$ है, X नहीं।

iv) अंतर संकारक वर्गसम नियम का पालन नहीं करता क्योंकि $X - X = \emptyset$ होता है।

घ) दो समुच्चयों X और Y के लिए, उनका 'सममित अंतर' एक अन्य युग्मक संक्रिया है जिसे $X \Delta Y$ से व्यक्त किया जाता है तथा

$\mathbf{X \Delta Y} = (X - Y) \cup (Y - X)$ के रूप में परिभाषित किया जाता है।

$\mathbf{X \Delta Y}$ को परिभाषित करने का एक वैकल्पिक तरीका $\mathbf{X \Delta Y} = (X \cup Y) - (X \cap Y)$ है।

यदि $X = \{1, 2, 3, 4\}$ तथा $Y = \{2, 4, 5, 7\}$ है,

तो $X \Delta Y = (X \cup Y) - (X \cap Y) = \{1, 2, 3, 4, 5\} - \{2, 4\} = \{1, 3, 5, 7\}$ होगा।

युग्मक संक्रिया Δ से संबंधित कुछ महत्त्वपूर्ण गुणधर्म इस प्रकार हैं :

किन्हीं भी समुच्चयों X, Y और Z के लिए

i) $X \Delta Y = Y \Delta X$, (Δ क्रम विनिमय नियम का पालन करता है)

ii) $(X \Delta Y) \Delta Z = X \Delta (Y \Delta Z)$, (Δसहचर्य नियम का पालन करता है)

iii) $X \Delta \varnothing = \varnothing \Delta X = X$ (तत्समक नियम : Ø, Δ के लिए तत्समक है)

iv) $X \Delta X' = X' \Delta X = U$ जहाँ ∪ सार्वत्रिक समुच्चय है क्योंकि $X \cup X' = U$ तथा $X \cap X' = \varnothing$ है)

ङ) किसी समुच्चय X के लिए, एकांगी संक्रिया 'समुच्चय के पूरक' को X' से व्यक्त किया जाता है तथा X' U–X के रूप में परिभाषित किया जाता है जहाँ U सार्वत्रिक समुच्चय है।

पूरक का एक महत्त्वपूर्ण गुणधर्म यह है कि $(X')' = X$ होता है।

नीचे समुच्चयों की संक्रियाओं के कुछ ऐसे गुणधर्मों का उल्लेख कर रहे हैं जिनमें दो या दो से अधिक संक्रियाएँ एक साथ सम्मिलित हैं।

i) $Z \cup (X \cap Y) = (X \cap Y) \cup Z = (X \cup Z) \cap (Y \cup Z)$

ii) $(Y \cup Z) \cap X = X \cap Y \cup Z) = (X \cap Y) \cup (X \cap Z)$

iii) $(X \cup Y)' = X' \cup Y'$ (डी-मॉरगन का नियम)

iv) $(X \cup Y)' = X' \cap Y'$ (डी-मॉरगन का नियम)

1.3 संबंध एवं फलन

उपभाग 1.2.2 में की गई चर्चा तथा ऊपर की गई टिप्पणियों के माध्यम से हमें संबंध की संकल्पना की अनौपचारिक जानकारी प्राप्त हो चुकी है। आगे होने वाली चर्चा के लिए यह जानना पर्याप्त है कि एक संबंध सामान्यतः दो (जिनकी संख्या एक अथवा अधिक भी हो सकती है) मान लेता है तथा *'सत्य' या 'असत्य' परिणाम देता है।*

गणित तथा अन्य शैक्षणिक विषयों में, अनेक औपचारिक संकल्पनाएं नित्य प्रति होने वाली वास्तविकताओं, घटनाओं इत्यादि में निहित होती हैं जिन्हें हम जाने-अनजाने देखते हैं। इस भाग में हम एक ऐसी ही संकल्पना 'संबंध' को औपचारिक रूप से परिभाषित करेंगे। हम 'पिता, 'पुत्री, 'मित्र' इत्यादि संबंधों के बारे में जानते हैं। इनका गणित में क्या महत्त्व है? महत्त्व यह है कि इनमें से प्रत्येक एक समुच्चय के बारे में व्यक्त किया जाता है जैसे कि अभी हम आगे की जाने वाली चर्चा में देखेंगे। इसके लिए पहले हमें 'क्रमित समुच्चय' (ordered) तथा 'क्रमित (ordered) युग्म' की संकल्पनाओं को जानने की आवश्यकता है तथा 'कार्तीजन गुणन' की संकल्पना की भी।

क्रमित युग्म तथा क्रमित समुच्चय : कार्तीजन ज्यामिति से हम जानते हैं कि बिंदु (1, 2), बिंदु (2, 1) से अलग होता है, अर्थात्, $(1, 2) \neq (2,1)$ होता है। अतः किसी बिंदु के निर्देशांकों को निरूपित करने वाले संख्या-युग्म में संख्याएं किस क्रम में लिखी जाएं, यह महत्त्वपूर्ण होता है। दूसरी ओर, सामान्य समुच्चय को लिखते हुए उसके अवयव किस क्रम में लिखे जाएं, इसका कोई महत्त्व नहीं होता; जैसे कि समुच्चय (1, 2) तथा (2,1) समान होते हैं।

अतः, एक क्रमित समुच्चय वह समुच्चय होता है जिसमें आने वाले अवयवों के क्रम का ध्यान रखना आवश्यक हो। यदि a, b, c …एक क्रमित समुच्चय के अवयव हैं, तो इस क्रमित समुच्चय को ***(a, b, c, …)*** के रूप में लिखा जाता है।

समुच्चयों का कार्तीय गुणन : दो अरिक्त समुच्चयों X और Y का कार्तीयजन गुणन, जिसे X × Y से व्यक्त किया जाता है, ऐसे क्रमित युग्मों का समुच्चय हैं जिनमें पहला घटक X से तथा दूसरा घटक Y से लिया गया हो, अर्थात्

$$X \times Y = \{(x, y): x \in X \text{ तथा } y \in Y\}.$$

यदि X और Y में से कम से कम एक समुच्चय (रिक्त समुच्चय) $\emptyset$ है, तो $X \times Y$ भी $\emptyset$ होगा।

टिप्पणी : ध्यान दें कि $X \times Y$ का प्रत्येक अवयव (x, y) एक क्रमित समुच्चय (युग्म) है, परंतु $X \times Y$ स्वयं एक क्रमित समुच्चय होना अनिवार्य नहीं है।

टिप्पणी : $|X \times Y| = |X| . |Y|$ अर्थात् X × Y के अवयवों की संख्या |X| तथा |Y| के अवयवों की संख्या के गुणनफल के बराबर होती है।

अब हम गणित की महत्त्वपूर्ण संकल्पना 'संबंध' को परिभाषित करने का प्रयास करते हैं। किसी संबंध की सहज धारणा को गणित की औपचारिक संकल्पना 'संबंध' में परिवर्तित किया जा सकता है।

किसी अरिक्त समुच्चय X से किसी अरिक्त समुच्चय Y तक एक (गणितीय) संबंध $X \times Y$ का एक उपसमुच्चय होता है। विलोमतः, $X \times Y$ का प्रत्येक उपसमुच्चय हमें समुच्चय X से Y तक एक संबंध प्रदान करता है।

एक उदाहरण : हम जानते हैं कि किसी निश्चित समय पर प्रत्येक व्यक्ति की ऊँचाई स्थिर होती है। अतः, HEIGHT एक गणितीय संबंध के रूप में इस प्रकार व्यक्त किया जा सकता है :

मान लीजिए X सभी व्यक्तियों का एक समुच्चय है तथा $Y = \{x: x \in R$ और $x > 0\}$ है। तो संबंध HEIGHT, X × Y का उपसमुच्चय

HEIGHT = {(x∈ y; y∈ y. यहाँ पर y इंचों में x की ऊँचाई व्यक्त करती है।

हम पहले से ही अंकगणितीय संबंधों जैसे कि $=, \neq, <, \leq$ तथा $\geq$ इत्यादि तथा समुच्चयों के मध्य संबंधों जैसे कि $\subseteq, =, \subset, \not\subset, \neq$ से अच्छी तरह परिचित हैं। हम इन ऊपर पर दी गई संबंध की औपचारिक परिभाषा के अनुसार पुनः विचार कर सकते हैं।

उदाहरण के लिए, आइये, हम प्राकृतिक संख्याओं के समुच्चय N पर संबंध '<' वे संदर्भ में विचार करें। संबंध की परिभाषा के अनुसार, इसे N × N के नीचे दिए उपसमुच्चय के रूप में व्यक्त किया जा सकता है :

Less-than-on-N $= \{ (1,2), (1,3), (2,3), (1,4), (2,4), (3,4), (1,5), (2,5), (3,5), (4,5), \ldots\}$

या $\quad$ $'<' = \{(x, y): x, y \in N,$ तथा $x < y\}.$

टिप्पणी : ध्यान रहे कि दाएं पक्ष और बाएं पक्ष के बीच लगा चिन्ह '=' यहाँ यह दर्शाता है कि बाएं पक्ष को दाएं पक्ष के रूप में परिभाषित किया जा सकता है।

परिभाषाएँ : एक अवयव (element) का प्रतिबिंब, एक संबंध का प्रांत (Domain), परिसर (Range) तथा सहप्रांत (Codomain)

एक समुच्चय X से एक समुच्चय Y तक एक संबंध R के लिए तथा R के किसी भी क्रमित-युग्म (a,b) के लिए अवयव b, अवयव a का प्रतिबिंब कहलाता है।

एक समुच्चय X से एक समुच्चय Y तक एक संबंध R के क्रमित-युग्मों के सभी प्रथम घटकों का समुच्चय संबंध R का प्रांत कहलाता है।

एक समुच्चय X से एक समुच्चय Y तक एक संबंध R के क्रमित-युग्मों के सभी द्वितीय घटकों का समुच्चय R का परिसर कहलाता है। पूरा समुच्चय Y, संबंध R का सहप्रांत कहलाता है। ध्यान दें कि किसी भी संबंध R के लिए R का परिसर $\boldsymbol{R \subseteq R}$ का सहप्रांत होता है।

उदाहरण के लिए, संबंध दिल्ली से हवाई अंतर = {(चेन्नई, 1759), (मुंबई, 1153), (कोलकाता, 1305), (हैदराबाद, 1259), (बैंगलुरू, 1744)} पर विचार कीजिए।

इस संबंध में चेन्नई का प्रतिबिंब 1759, हैदराबाद का 1259 है, ... इत्यादि।

इस संबंध का प्रांत समुच्चय (चेन्नई, मुंबई, कोलकाता, हैदराबाद, बैंगलुरू) है तथा परिसर समुच्चय {1759, 1153, 1305, 1259, 1744}है।

फलनः

वर्तमान रूप में फलन की संकल्पना प्रयोग संभवतः 17वीं शताब्दी के अंत में पहली बार गॉटफ्रीड लैबनिट्ज़ ने जोहान बरनौली को लिखे एक पत्र में किया। परंतु, फलन की एक विशेष प्रकार के संबंध के रूप में कल्पना कहीं 20वीं शताब्दी के प्रारंभ में की गई।

पहले हम, पाठकों का परिचय फलन की संकल्पना से स्वतंत्र रूप से करवायेंगे, तत्पश्चात् यह दर्शाएंगे कि फलन को एक विशेष प्रकार के संबंध के रूप में माना जा सकता है।

फलन : मान लीजिए X और Y दो अरिक्त समुच्चय हैं। X से Y तक एक फलन f, जिसे $f: X \rightarrow Y$ के रूप में व्यक्त किया जाता है, एक ऐसा नियम होता है जो X के प्रत्येक अवयव x से y के एक अद्वितीय अवयव, मान लीजिए y, को संबद्ध करता है।

X को फलन f का प्रांत तथा Y को f का सहप्रांत कहते हैं। साथ ही, यदि $f(x) = y$ है, तो y को x का प्रतिबिंब तथा x को y का पूर्व प्रतिबिंब कहते हैं।

फलन को 'प्रतिचित्र' तथा 'प्रतिचित्रण' नामों से भी जाना जाता है।

यदि Y= R हो, तो फलन $f: X \rightarrow R$ को एक ***वास्तविक–मान*** फलन कहा जाता है। यदि Y= C हो तो $f: X \rightarrow C$ फलन को एक सम्मिश्र मान फलन कहा जाता है। इसी प्रकार यदि हम सह–प्रांत Y को Z तथा Q लें, तो हमें क्रमशः ***पूर्णांक–मान*** फलन तथा परिमेय–मान फलन प्राप्त होते हैं।

बोध प्रश्न 3

1) दो समुच्चय दिए गए हैं : $X = \{1, \{2, 3\}, \{4, 5\}, 6\}$ तथा $Y = \{1,\{2, 4\}, \{5. 6, 7\}\}$इनके मान आंकलित करें :

   (i) $X \cup Y$, (ii) $X \cap Y$, (iii) $X - Y$ (iv) $X \Delta Y$

2) मान लीजिए $X = \{2, 3, 4, 5, 6, 8\}$। मान लें R इस X पर इस प्रकार परिभाषित एक संबंध है :

   $\{(x, y): x, y \in X, y, x$ द्वारा विभाज्य है तथा $x \neq y$ $\}$तो

   i) R को सूची रूप में लिखिए

   ii) R का प्रांत आंकलित करें

   iii) R का क्षेत्र ज्ञात करें।

3) इस प्रकार परिभाषित संबंध R के प्रांत एवं क्षेत्र आंकलित करें :

   $R = \{(x, x + 5): x \in \{0, 1, 2, 3, 4, 5\}, y \in \{5, 6, 7,$ and $x^2 < y\}$.

1.4 तर्कशास्त्र

इस विषय पर हम अपनी चर्चा इन दो प्रश्नों से प्रारंभ करेंगे : तर्कशास्त्र क्या है? तर्कशास्त्र का अध्ययन क्यों आवश्यक है?

हम इन प्रश्नों के संक्षिप्त उत्तरों से प्रारंभ करते हुए हम अपनी चर्चा को आवश्यकतानुसार आगे बढ़ाएंगे।

हम प्रारंभ में ही बताना चाहते हैं कि तर्कशास्त्र तर्कजा/तर्कण से संबंधित है, या ऐसा कहें कि ठीक तरीके से तर्क करने से संबंधित है। यह हमें बताता है सिखाता है कि हम ठीक प्रकार से तर्क करें जिससे दिए हुए तथ्यों के आधार पर

i) हम सदैव (एक ही) सही निष्कर्ष पर पहुँचें, तथा

ii) दिए हुए तथ्यों से निष्कर्ष तक पहुँचने की प्रक्रिया, गणितीय परिकलन की तरह पर्याप्त रूप से ठोस/संतोषजनक होनी चाहिए, जिससे संबंधित पक्षों में विवाद शेष रहने की संभावना न्यूनतम हो जाए/न रह जाए।

वास्तव में, तर्कशास्त्र दिए हुए तथ्यों तथा धारणाओं/सूचनाओं से निष्कर्ष तक पहुंचने की विधियाँ उपलब्ध करवाता है तथा उनकी जाँच करता है। परंतु, किसी भी शैक्षणिक क्षेत्र की प्रारंभिक धारणाएँ क्या होंगी, यह उस शैक्षणिक क्षेत्र की विषयवस्तु से सम्बंधित होती है, फिर चाहे वह भौतिकशास्त्र हो, अर्थशास्त्र अथवा राजनीतिशास्त्र; तर्कशास्त्र की नहीं। *इस कथन पर बल देने के लिए, हम यह दोहरा रहे हैं कि तर्कशास्त्र का उपयोग केवल दी हुई धारणाओं से निष्कर्ष निकालने तक है।*

उदाहरण के लिए,

i) अर्थशास्त्र के मूलभूत सिद्धांत संकल्पनाएँ जैसे कि अभाव, पसंद/चुनाव/विकल्प, अदला–बदली/समन्वयन, प्रोत्साहन, विनिमय, सूचना, बंटन इत्यादि तथा अर्थशास्त्र के मौलिक नियम जैसे कि 'घटते हुए प्रतिफल का नियम' इत्यादि क्या होंगे यह अर्थशास्त्री तय करेंगे, तर्कशास्त्री नहीं।

ii) यह केवल गणितज्ञ तय करेंगे कि (क) ज्यामिती की मूलभूत धारणाएँ यूक्लिड अभिगृहीताओं पर आधारित होनी चाहिए और (ख) संख्याओं की मूलभूत मान्यताएँ पिएनो की अभिगृहीताओं (axioms) पर आधारित होनी चाहिए।

तर्कशास्त्र के मूलभूत तत्व हैं : (i) तार्किक अचर/स्थिरांक अर्थात्, सत्य तथा असत्य जिन्हें क्रमशः 0 तथा 1 भी लिखा जाता है, (ii) कथन, (iii) तर्क बल्कि तर्क रूप– तार्किक संयोजक तथा निष्कर्ष के नियम।

अब हम तर्क की मूलभूत संकल्पना के बारे में जानते हैं। वाद–प्रतिवाद या तर्क–वितर्क (argument) को तर्कशास्त्र की केंद्रीय संकल्पना माना जा सकता है, जैसे कि अणु, भौतिकशास्त्र के संदर्भ में एक केंद्रीय संकल्पना है।

तर्कशास्त्र में 'तर्क' या 'वाद' की संकल्पना, रोज़मर्रा की भाषा में बहस की धारणा से अलग है। तर्कशास्त्र में तर्क (n+1) कथनों, मान लीजिए S_1, S_2 … S_n, $S_{n+}1$ का एक अनुक्रम है, जहाँ $n \in N$ है (कुछ स्थितियों में n+1 भी हो सकता है) जिनमें से कथन S_{n+1} अन्य कथनों से निष्कर्ष के रूप में प्राप्त होता है। इस प्रकार के तर्क को सामान्यतः निम्न प्रकार से ऊर्ध्वाधर रूप में व्यक्त किया जाता है :

$$\begin{array}{c} S_1, \\ S_2, \\ \ldots \\ S_n \\ \hline S_{n+1} \end{array}$$

अथवा, (क्षैतिज रूप में) जैसे कि

$$\frac{S_1, S_1 \ldots\ S_n}{S_{n+1}}$$

ये निरूपण यह दर्शाते हैं कि कथन S_{n+1} (जिसे एक क्षैतिज रेखा के नीचे लिखा जाता है) निष्कर्ष के रूप में दिए हुए/माने गए कथनों S_1, S_2 … S_n (जो सभी क्षैतिज रेखा के ऊपर लिखे जाते हैं) से प्राप्त होता है। निष्कर्ष का अर्थ है कि हम यह मानते हैं कि दिए हुए सभी कथन S_1, S_2 … S_n सत्य हैं तथा तर्क के माध्यम से हम प्राप्त करते हैं कि S_{n+1} भी सत्य है। निष्कर्ष की गुणवत्ता/विश्वसनीयता माने गए (दिए हुए) कथनों के अतिरिक्त निष्कर्ष ज्ञात करने के लिए प्रयोग की गई विधि/पद्धति पर भी निर्भर करती है। कुछ सुपरिचित विधियाँ हैं : (i) निगमन (deduction), (ii) आगमन (induction), तथा (iii) अपकर्षण (abduction)। हम इन विधियों पर आगे चर्चा करेंगे। पहले, हम 'कथन' की संकल्पना की चर्चा करते हैं। कथन, तर्क का एक महत्वपूर्ण घटक है।

कथन : कथन एक कथात्मक/घोषणात्मक वाक्य होता है अर्थात् एक ऐसा वाक्य जिसके लिए यह कहा जा सके कि वह 'सत्य' है या 'असत्य'। उदाहरण के लिए, नीचे दिए तीन वाक्यों में से प्रत्येक एक कथन है : (i) चीनी मीठी होती है, (ii) सूर्य पश्चिम से उगता है, (iii) पीटर एक स्नातक है। कथन (i) सत्य है, कथन (ii) असत्य है। इसी प्रकार कथन (iii) के लिए भी 'सत्य' और 'असत्य' में से केवल एक ही मान

संभव होगा, चाहे हमें वह मालूम न हो। 'सत्य' और 'असत्य' में से प्रत्येक मान जो कि किसी कथन के साथ संबद्ध होता है, कथन का सत्यमान कहलाता है।

एक वाक्य तथा एक कथन में अंतर स्पष्ट करने के लिए आइये, हम नीचे दिए कुछ वाक्यों पर विचार करें जिनमें से कोई भी कथन नहीं है :

– भारत का प्रथम प्रधानमंत्री कौन था? *(प्रश्नवाचक वाक्य)*

– कृपया, मुझे वह किताब दें। *(आदेशात्मक वाक्य)*

– तुम्हें नियमित रूप से व्यायाम करना चाहिए। (*आदेशात्मक कदाचित कर्त्तव्यवाची)*

– हुर्रे! हम ट्रॉफी जीत गए। *(विस्मयादि बोधक वाक्य)*

हम देख सकते हैं कि इन वाक्यों के लिए 'सत्य' अथवा 'असत्य' नहीं कहा जा सकता है।

प्रतीकात्मक भाषा तथा तार्किक संयोजक

प्रतीकात्मक भाषा क्यों?

क) प्रतीक बड़े कथनों का छोटा रूप होता है; उदाहरण के लिए, कथन 'चेरापूँजी' में जो कि भारत के उत्तर–पूर्व में एक कस्बा है, 1901–2001 की अवधि में, दुनिया भर में सर्वाधिक वर्षा हुई' को हम प्रतीक '***CMRain***' या **CR** या **P** केवल से ही व्यक्त कर सकते हैं।

ख) एक तार्किक संदर्भ में हम, सामान्यतः, किसी कथन के तार्किक मान 'सत्य' या 'असत्य' में रुचि रखते हैं, उसकी विषयवस्तु या संरचना में नहीं। प्रतीक सामान्य भाषा के वाक्यों की अपेक्षा किसी कथन के अन्य पक्षों को एक ओर करते हुए, अमूर्त रूप से केवल उसके तार्किक पक्ष को बेहतर रूप से उजागर करते हैं।

ग) प्रतीकात्मक निरूपण में, एक अकेला प्रतीक किसी कथन के अनेक समान रूपों को निरूपित कर सकता है। उदाहरण के लिए, 'कालिदास ने मेघदूत लिखा', 'मेघदूत कालिदास के द्वारा लिखा गया', 'मेघदूत कालिदास की कृतियों में से एक है' इत्यादि अलग-अलग वाक्य होते हुए भी वास्तव में समान कथन हैं। शायद कोई कंप्यूटर प्रणाली यह न पहचान सके। लेकिन, प्रतीकात्मक रूप से यदि हम 'रचना (कालिदास, मेघदूत) सत्य है' के रूप में लिखें तो इसमें ये सभी कथन समाहित हो जाते हैं।

इस प्रकार, कथनों के प्रतीकात्मक निरूपण के और कई लाभ हो सकते हैं परंतु इस पर और अधिक विस्तार से चर्चा हम इस इकाई में नहीं कर पाएंगे।

प्रतीकात्मक भाषा : तर्कशास्त्र के लिए प्रतीकात्मक भाषा में निम्न प्रकार के प्रतीकों की आवश्यकता होती है :

i) **कथनों के प्रतीकात्मक नामों** *P, Q, R, CMRain* तथा *EF* इत्यादि की, जैसे कि ऊपर प्रयोग किए गए हैं। केवल आण्विक (atomic) कथनों को ही प्रतीकात्मक नाम दिए जाते हैं। यहाँ आण्विक कथनों से हमारा तात्पर्य ऐसे कथनों

से है जिनमें कोई भी *(i) तार्किक संयोजक तथा (ii) परिमाणक न हों,* जिनकी चर्चा हम नीचे दिए बिंदु (iv) में अभी करेंगे।

ii) **तार्किक चरों की** : तर्कशास्त्र में एक प्रतीकात्मक नाम एक तार्किक चर होता है जो किसी भी समय पर दो संभव तार्किक स्थिर मानों T और F में से (जिसकी चर्चा हम नीचे दिए बिंदु (iii) में करेंगे) केवल एक ही मान ले सकता है।

iii) **प्रतीकात्मक अचरों की** : सत्य-मान 'सत्य' और 'असत्य' प्रतीकात्मक तर्कशास्त्र के दो अचर हैं। इन मानों को 'T' और 'F' या क्रमशः '1' और '0; से व्यक्त किया जाता है।

iv) **तार्किक संयोजक, परिमाणक, प्रतीकात्मक व्यंजक, सुपरिभाषित सूत्र (जिसे हम संक्षेप में *wff* लिखते हैं)** इत्यादि। हम इनमें से प्रत्येक की विस्तृत चर्चा करेंगे।

v) **कथन रूपों (प्रतीकात्मक प्राचलों) की** : ये सामान्यतः कथनों से संबंधित सर्वसमिकाएं इत्यादि होते हैं, जिन्हें किसी भी लिपि के अक्षरों द्वारा प्रतीकात्मक रूप से व्यक्त किया जा सकता है। उदाहरण के लिए, डी-मॉरगन नियम को :~ $(\Phi \vee \Psi) \equiv \sim \Phi \wedge \Psi$ के रूप में लिखा जा सकता है **जहाँ** Φ और Ψ ग्रीक अक्षर हैं तथा इन्हें इस सूत्र में किसी भी ***wff*** से प्रतिस्थापित किया जा सकता है।

आइये, अब हम बिंदुओं (iv) और (v) पर विस्तार से चर्चा करें।

तार्किक संयोजक :

पाँच तार्किक संयोजक जिनका प्रयोग बार-बार जाता है, नीचे अपने नामों के साथ दिए गए हैं :

'$\sim$' (निषेधक), (Negation) $\wedge$ (और) (Conjunction, and, $\vee$ (या) (disjunction or), $\rightarrow$ (अंतर्भाव या सप्रतिबंध) (conditional or implication), $\leftrightarrow$ (द्वि-अंतर्भाव या द्वि-सप्रतिबंध या यदि और केवल यदि)

इनमें से प्रत्येक समुच्चय *{T, F}* तक एक फलन होता है। इनकी परिभाषाएँ नीचे दी गई हैं :

i) ' $\sim$ ' निषेधन जो कि एक एकांगी तार्किक सकारक है, *{T, F}* से *{T, F}* तक एक फलन है और इस इस प्रकार से परिभाषित किया जाता है :

$$\sim(T) = F \text{ तथा } \sim(F) = T$$

परिमाणक

उपभाग 1.1.2 में परिमाणकों और तार्किक संकारकों के अंतर्गत हम पहले ही परिमाणकों से आपको परिचित करा चुके हैं यथा :

i) $(\forall x)$ – अर्थात सभी ऐसे x के लिए कि $x \in X'$,

ii) $(\exists x)$ – ऐसे x का अस्तित्व है कि $x \in X'$

iii) $(\nexists\, x \in X)$ – अर्थात किसी ऐसे x के लिए नहीं कि $x \in X'$ अथवा ऐसा कोई x नहीं है कि $x \in X'$ आदि

सत्यता तालिका : सत्यता तालिका एक तार्किक या एक से अधिक फलनों/संकारकों के लिए एक ऐसी तालिका को कहते हैं जिसमें फलन के प्रांत के प्रत्येक अवयव के संगत प्राप्त होने वाले परिणाम को सूचीबद्ध किया जाता है।

उदाहरण के लिए, निषेधन संकारक की सत्यता तालिका है।

A	~A
T	F
F	T

शेष चारों तार्किक संयोजक युग्मक संकारक हैं। ये सभी *{T, F} x {T, F}* से *{T, F}* तक फलन हैं तथा इन्हें इस प्रकार परिभाषित किया जाता है :

ii) फलन '**∧**' (और) को इस प्रकार परिभाषित किया जाता है : $(T\wedge T) = T$, $(T\wedge F) = F$, $(F\wedge T) = F$ और $(F\wedge F) = F$

iii) फलन '∨ '(या) को इस प्रकार परिभाषित किया जाता हैः $(T\vee T) = T$, $(T\vee F) = T$, $(F\vee T) = T$ और $(F \vee F) = F$,

iv) फलन '→' (अंतर्भाव/सप्रतिबंध) को इस प्रकार परिभाषित किया जाता है : $(T\rightarrow T) = T$, $(T\rightarrow F) = F$, $(F\rightarrow T) = T$ और $(F \rightarrow F) = T$,

v) फलन '**↔**' (द्वि-सप्रतिबंध/द्वि-अंतर्भाव) को इस प्रकार परिभाषित किया जाता है : $(T\leftrightarrow T) = T$, $(T\leftrightarrow F) = F$, $(F\leftrightarrow T) = F$ और $(F\leftrightarrow F) = T$.

चार युग्मक तार्किक संकारकों की सत्यता तालिकाएं हमने नीचे एक ही सारणी में सूचीबद्ध की हैं :

A	B	$A \wedge B$	$\mathbf{A \vee B}$	$A \rightarrow B$	$\mathbf{A \leftrightarrow B}$
T	T	*T*	**T**	*T*	**T**
T	F	*F*	**T**	*F*	**F**
F	T	*F*	**T**	*T*	**F**
F	F	*F*	**F**	*T*	**T**

टिप्पणियाँ : विभिन्न तार्किक संकारकों के लिए प्रयुक्त होने वाले प्रतीक इस अर्थ में मानक/अद्वितीय नहीं हैं कि अभ्यास में इनके लिए अन्य प्रतीक भी प्रयोग किए जाते हैं। उदाहरण के लिए, (i) निषेधन के लिए ~ के अतिरिक्त प्रतीक ¬ तथा ! भी प्रयोग किए जाते हैं। इसी प्रकार (ii) और के लिए ∧ के अतिरिक्त, प्रतीक '· ' तथा & भी प्रयोग में लाए जाते हैं, तथा (iii) या के लिए ∀ के अतिरिक्त प्रतीकों + तथा // का भी प्रयोग किया जाता है। इसी प्रकार अन्य संयोजकों/संकारकों के लिए भी एक से अधिक प्रतीक प्रयोग किए जाते हैं।

सुपरिभाषित सूत्र (संक्षेप में *wff* अथवा कथन)

किसी प्राकृतिक भाषा की तरह ही, प्रतीकात्मक/सांकेतिक भाषा में भी ऐसे नियम उपलब्ध हैं जिनसे निर्धारित किया जा सके कि कोई दिया हुआ सांकेतिक व्यंतक निरूपण के सही रूप में लिखा गया है अथवा नहीं। उदाहरण के लिए, युग्मक तार्किक

संकारकों के अंतःप्रत्यय संकेतन में, $P\rightarrow (S \vee U)$ एक सही निरूपण परंतु $\rightarrow P\ (S \vee U),\ P \rightarrow\vee\ S\ U$ तथा $\rightarrow\vee\ P\ \ SU$ इत्यादि सही निरूपण नहीं है।

सही रूप में लिखे गए प्रत्येक औपचारिक व्यंजक को हम सुपरिभाषित सूत्र (***wff***) कहते हैं। हम इस विषय पर और अधिक चर्चा नहीं करेंगे।

किसी सुपरिभाषित सूत्र की व्याख्या तथा मूल्यांकन : जैसा कि उल्लेख किया गया कि एक तार्किक प्रतीक एक सूत्र पर दो संभव सत्यमानों T और F में से केवल एक ही मान लेता है।

यदि हमारे पास एक तार्किक पद E है जिसमें दो तार्किक प्रतीक हों तो इन प्रतीकों को दिए जा सकने वाले चार संभव मूल्यमान होंगे। इन चारों संभावित मूल्यमान निर्दिष्टियों को E के प्रतीक चिन्हों की व्याख्या कहा जाता है और फिर तो हम उक्त व्याख्या के अनुरूप E का तार्किक मूल्यमान ज्ञात कर सकते हैं। उदाहरण के लिए, *wff* E में

$$E : (\sim ((A \vee (\sim B)) \rightarrow (A \leftrightarrow B))$$

दो चिन्ह A और B हैं। आइये अब क्रमशः A तथा B को हम उनके उपयुक्त सत्यमान निर्दिष्ट करें। दूसरे शब्दों में, हम एक व्याख्या I मानकर A को F तथा B को T मान दे सकते हैं। इस I को {~ A, B}भी लिख सकते हैं। अतः पद E का सत्यमान हमें सहज ही A तथा B के स्थानों पर क्रमशः F तथा T लिखने से मिल जाएगा। अतः हम पाएंगे :

$$(\tilde{}\,((F\ \ (\tilde{}\,T)) \rightarrow (F \leftrightarrow T))$$

उपर्युक्त सत्यमान तालिका का प्रयोग कर हम उत्तरोत्तर रूप से पाएंगे :

$$\sim (\ (F \vee F) \rightarrow F\ )$$

$$\sim (\ F \rightarrow F\ )$$

$$\sim (\ T\ )$$

$$F$$

अतः जो व्याख्या A को F तथा B को T मान निर्दिष्ट करती है उसके अनुसार पदबंध E का सत्यमान 'F' होगा।

तार्किक रूप से समान सूत्र (या तार्किक सर्वसमिकाएं) : दो सूत्र G1 तथा G2 तार्किक रूप से समान कहलाते हैं यदि सूत्र की प्रत्येक व्याख्या के लिए (अर्थात् सूत्रों G1 तथा G2 में उपस्थित प्रत्येक आण्विक कथन से समान सत्यमान संबद्ध करने पर) G1 तथा G2 का समान सत्यमान प्राप्त हो। दूसरे शब्दों में, दो सूत्र G1 तथा G2 तार्किक रूप से समान होंगे यदि एक जैसे सत्यमान नियत करने पर, G1 सत्य होगा यदि G2 सत्य है तथा G1 असत्य होगा यदि G2 असत्य है।

उदाहरण : हम यह जाँच कर सकते हैं कि सूत्र G_1: $\sim (A \rightarrow B)$ तथा सूत्र G_2: $A \wedge \sim B$ समान हैं। इसके लिए हम G1 और G2 की नीचे दी सत्यमान तालिका का परीक्षण करते हैं :

A	B	$\sim B$	$(A \to B)$	$\sim(A \to B)$	$A \wedge \sim B$
T	T	F	T	F	F
T	F	T	F	T	T
F	T	F	T	F	F
F	F	T	T	F	F

हम पाते हैं कि अंतिम दो स्तरों के संगत मान समान हैं।

कथात्मक तर्क के कुछ समान सूत्रों की सूची :

इस सूची में E, G और H एक कथन रूप या एक व्यापक *wff* को निरूपित करते हैं। यह आवश्यक नहीं है कि E,G,H आण्विक कथन ही हो ये कोई *wff* भी हो सकते हैं।

(1.1) $E \leftrightarrow G = (E \to G) \wedge (G \to E)$

(1.2) $E \to G = \sim E \vee G$

(1.3)(a) $E \vee G = G \vee E$ ($\vee$ का क्रम विनिमेय नियम)

(1.3)(b) $E \wedge G = G \wedge E$ ($\wedge$ का क्रम विनिमेय नियम)

(1.4)(a) $(E \vee G) \vee H = E \vee (G \vee H)$ ($\vee$ का साहचर्य नियम)

(1.4)(b) $(E \wedge G) \wedge H = E \wedge (G \wedge H)$ ($\wedge$ का साहचर्य नियम)

(1.5)(a) $(E \vee (G \wedge H)) = (E \vee G) \wedge (E \vee H)$ ($\vee$ का $\wedge$ पर बंटन नियम)

(1.5)(b) $(E \wedge (G \vee H)) = (E \wedge G) \vee (E \wedge H)$ ($\wedge$का $\vee$ पर बंटन नियम)

(1.6)(a) $E \vee$ असत्य $= E$ (असत्य, $\vee$ का तत्समक है)

(1.6)(b) $E \wedge$ सत्य $= E$ (सत्य, $\wedge$ का तत्समक है)

(1.7)(a) $E \vee$ सत्य $=$ सत्य

(1.7)(b) $E \wedge$ असत्य $=$ असत्य

(1.8)(a) $E \vee \sim E =$ सत्य (बहिष्कृत मध्य का नियम)

(1.8)(b) $E \wedge \sim E =$ असत्य

(1.9)(a) $E \wedge E = E$ ($\wedge$ एक वर्गसम संकारक है)

(1.9)(b) $E \vee E = E$ ($\vee$ एक वर्गसम संकारक है)

(1.10) $\sim(\sim E) = E$

(1.11)(a) $\sim(E \vee G) = \sim E \wedge \sim G$ (डी मॉरगन का नियम)

(1.11)(b) $\sim(E \wedge G) = \sim E \vee \sim G$ (डी मॉरगन का नियम)

1.5 उपपत्ति की विधियाँ

इस उपभाग में हम गणित में उपपत्ति की कुछ महत्त्वपूर्ण विधियों का अध्ययन करेंगे :

i) प्रत्यक्ष विधि तथा प्रत्युदाहरण (counter examples)

ii) विराधोक्ति विधि (contradiction) तथा प्रतिधनात्मक विधि (contra positive)

iii) मामलों (cases) द्वारा प्रमाण

iv) अस्तित्व प्रमाण

v) गणितीय आगमन द्वारा प्रमाण

vi) समतुल्यता का प्रमाण

सामान्यतः, अलग-अलग प्रकार की समस्याओं के लिए, अलग-अलग विधियाँ बेहतर हो सकती हैं तो कुछ प्रकार की समस्याओं के लिए एक से अधिक विधियाँ समान रूप से उपयोगी हो सकती हैं।

जिन समस्याओं में उपपत्ति की आवश्यकता होती है, उनमें स्पष्ट रूप से दो भाग होते हैं जो कि इस प्रकार हैं (i) दिए हुए तथ्यों/अभिगृहीतों (axioms) इत्यादि का समुच्चय; इन सबको एक साथ मिलाकर, हम परिकल्पना (Hypotheses) कहते हैं, तथा (ii) वह जो हमें सिद्ध करना है। इसे हम निष्कर्ष कहते हैं। इनके अतिरिक्त भी, परोक्ष रूप से, अनेक परिभाषाएं, विधियाँ, नीतियाँ, पहले सिद्ध किए गए प्रमेय इत्यादि भी इस प्रकार की उपपत्तियों में प्रयोग होते हैं। इन्हें हम 'प्रदत्त' मानते हैं।

उदाहरण के लिए, मान लीजिए कि हमें सिद्ध करना है, यदि m और n विषम पूर्णांक हैं, तो उनका गुणनफल $m.n$ भी विषम होगा।

इस प्रश्न में दी हुई स्पष्ट मान्यताएँ हैं : (i) m तथा n दोनों पूर्णांक हैं, तथा (ii) और दोनों विषम हैं। हमें सिद्ध करना है : उनका गुणनफल $m.n$ विषम है। परोक्ष रूप से हम यह भी मानते हैं कि हमें कई उपयुक्त परिभाषाएँ, दिए हुए तत्वों के उपयुक्त गुणधर्म, निष्कर्ष निकालने के वैध नियम तथा उपपत्ति की विभिन्न विधियाँ भी दी हुई हैं। इनमें से हम उन नियमों तथा विधियों को चुन सकते हैं जो हमें इस प्रश्न को कुशलता से तथा संतोषजनक रूप से हल कर सकते हैं। उपपत्ति की विभिन्न विधियों पर विस्तृत चर्चा इस पाठ्यक्रम के विषय क्षेत्र से बाहर है।

1.6 सार-संक्षेप

इस पाठ्यक्रम की इस पहली इकाई में हमने पाठकों को गणित की कुछ मूलभूत संरचनाओं से अवगत करवाया। आगे आने वाली इकाइयों में ये संकल्पनाएँ अत्यंत उपयोगी सिद्ध होंगी। समुच्चयों की संकल्पना गणित की सर्वाधिक मूलभूत संकल्पना है तथा अनेक वास्तविक समस्याओं को हल करने में इसका उपयोग किया जाता है। इस इकाई में हमने समुच्चयों की संकल्पना तथा उनके निरूपण, संकेतन के साथ-साथ समुच्चयों पर विभिन्न संक्रियाओं तथा कुछ महत्त्वपूर्ण संख्या समुच्चयों का भी अध्ययन किया है। हमने गणित में संबंधों तथा फलनों की संकल्पनाओं तथा तर्कशास्त्र तथा उपपत्ति की विधियों का भी अध्ययन किया। पाठकों को चाहिए कि तार्किक संयोजकों की सत्यमान तालिकाओं का अध्ययन ध्यानपूर्वक करें क्योंकि ये तार्किक व्यंजकों के अध्ययन में अत्यंत उपयोगी होती हैं।

1.7 संदर्भ ग्रंथादि

1) K. Sydsaeter and P. Hammond, *Mathematics for Economic Analysis*, Fourth Edition, Pearson, Educational Asia: Delhi, 2012.

2) NCERT *Chapters 1, NCERT Math class XI,* NCERT publication 2018.

3) NCERT *Chapters 1& 2, NCERT Math class IX,* NCERT publication 2018.

4) IGNOU *Unit 2/ Block 1& Unit 1 of Block 2, MCSE-003: Artificial Intelligence and Knowledge Management,* IGNOU, (2008).

5) R. Johnsonbaugh, *(2009). Discrete Mathematics (Seventh Edition)* Pearson Education Asia.

6) Thomas J. Mckay (1989). *Modern Formal Logic* MacMillan Publishing Company.

7) Copi I., Cohen C. McMahon K. (2014). *Introduction to Logic (Fourteenth Edition)* Pearson.

1.8 बोध प्रश्नों के उत्तर अथवा संकेत

बोध प्रश्न 1

1) (i), (ii) तथा (vi) में से प्रत्येक एक सुनिश्चित संग्रह है, अतः एक समुच्चय है। दूसरी ओर संग्रह (iii) और (iv) सुनिश्चित न होने के कारण समुच्चय नहीं है। (v) भी एक समुच्चय नहीं है, क्योंकि समुच्चयों में अवयवों को दोहराने की अनुमति नहीं है।

2) क) $\{2, 3, 5,\}$ अवयवों की संख्या=3
   ख) $\{5, -2\}$ अवयवों की संख्या = 2

3) क) यह एक समुच्चय है, $\{x: x \in N, 3, x$ को विभाजित करता है, तथा $6 \leq x \leq 27\}$ अवयवों की संख्या=8

   ख) यह एक समुच्चय है, $\{x: x= 2^n, n \in N\}$,
   अवयवों की संख्या=∞

   ग) यह एक समुच्चय है, $\{X_k: X_k= \{1, \ldots, i_k\}, i_k \in N\}$,
   अवयवों की संख्या=10

बोध प्रश्न 2

1) i) $X=Y$, (ii) $X \neq Y$, क्योंकि $15 \in Y$, तथा $15 \notin X$

2) i) $X \subset Y$ (ii) $X \not\subset Y$ तथा $Y \not\subset X$

3) i) असत्य (ii) असत्य (iii) असत्य (iv) सत्य (v) असत्य

बोध प्रश्न 3

1) a) $X \cup Y = \{1, \{2, 3\}, \{4, 5\}, 6, \{2, 4\}, \{5. 6, 7\}\}$; $X \cap Y = \{1\}$; $(X - Y) = \{\{2, 3\}, \{4, 5\}, 6\}$ $X \Delta Y =\{\{2, 3\}, \{4, 5\}, 6, \{2, 4\}, \{5. 6, 7\}\}$

2) $R = \{ (2,4), (2,6), (2,8), (3, 6), (4, 8)\}$, R का प्रांत = $\{2, 3, 4\}$ तथा R का परिसर = $\{4, 6, 8\}$

3) R का प्रांत= $\{0, 1, 2\}$ तथा R का परिसर = $\{5, 6, 7\}$

इकाई 2 सांख्यिकीय संकल्पनाएँ

संरचना

2.0 उद्देश्य

इस इकाई के अध्ययन के पश्चात्, आप निम्नलिखित से भली-भाँति से परिचित हो जाएंगे :

- डेटा, चर, चरों के वर्गीकरण तथा माप इत्यादि संकल्पनाओं से;

- किसी बंटन में केंद्रीय प्रवृत्ति तथा परिवर्तनशीलता की संकल्पनाओं की रूपरेखा से;
- केंद्रीय प्रवृत्ति तथा परिवर्तनशीलता के मापों के परिकलन की विधियों से;
- किसी स्थिति-विशेष में सबसे अधिक उपयोगी मापक चुनने की विधि से;
- स्थिति के मापक की संकल्पना से;
- अनुसंधानमूलक डेटा विश्लेषण की प्रक्रिया से; तथा
- स्प्रेडशीट पैकेज की सहायता से प्रतिशत, माध्य, बहुलक, माध्यिका, वैषम्य (skewness) तथा ककुदता (kurtosis) इत्यादि ज्ञात करने की विधि से।

2.1 विषय प्रवेश

डेटा विश्लेषण की प्रक्रिया में, अनेक गणितीय तथा सांख्यिकीय संकल्पनाओं का प्रयोग होता है। इकाई 1 में संख्याओं, समुच्चयों, संबंधों, फलनों इत्यादि अनेक गणितीय संकल्पनाओं की व्याख्या की गई है। डेटा विश्लेषण करने से पहले हमें यह जानने की आवश्यकता होगी कि डेटा से हमारा क्या तात्पर्य है? इसके विभिन्न प्रकार क्या है? डेटा के स्रोत क्या हैं? डेटा विश्लेषण के दौरान हमें किस प्रकार के चरों की आवश्यकता पड़ सकती है? इनके माप के विभिन्न पैमाने क्या हैं? हम डेटा की विभिन्न विशेषताओं को किस प्रकार सारबद्ध कर सकते हैं? इस इकाई में हम इन सभी विषयों पर चर्चा करेंगे। साथ ही, चरों की विशेषताओं का वर्णन करने या उन्हें सारबद्ध करने के प्रयास में हमें केंद्रीय प्रवृत्ति, बंटन, परिक्षेपण तथा स्थिति इत्यादि की आवश्यकता पड़ती है। ये सभी विधियाँ, वर्णनात्मक सांख्यिकी के अंतर्गत वर्गीकृत की जा सकती हैं। ये सभी विधियाँ मूलभूत अथवा आधारभूत हैं तथा इस इकाई में हम इनका अध्ययन करेंगे। आइये, हम डेटा/सांख्यिकी की संकल्पनाओं की व्याख्या से प्रारंभ करें।

2.2 कुछ प्राथमिक संकल्पनाएँ

2.2.1 आँकड़े/डेटा

शब्द सांख्यिकी का प्रयोग एकवचन तथा बहुवचन में अनेक अर्थों में प्रयोग किया जाता है। अतः, इसे दो अलग-अलग प्रकार से परिभाषित किया जा सकता है। जब शब्द 'Statistics' का प्रयोग एकवचन में किया जाए तो इससे हमारा तात्पर्य होता है सांख्यिकीय विधियाँ। बहुवचन में इसका प्रयोग 'डेटा/आँकड़ों' के लिए किया जाता है। इस संदर्भ में, यह एक विज्ञान की एक ऐसी शाखा के रूप में परिभाषित की जाती है जो सांख्यिकीय डेटा के संग्रहण, उसके वर्गीकरण, विश्लेषण तथा विवेचन से संबंधित है। दूसरे शब्दों में, सांख्यिकी, डेटा को वैज्ञानिक विधियों द्वारा संग्रहण करने, उसे व्यवस्थित करने, सारबद्ध करने, उसके प्रस्तुतीकरण और विश्लेषण के साथ-साथ उससे वैध निष्कर्ष ज्ञात करने तथा विश्लेषण के आधार पर उचित/तर्कसंगत निर्णय लेने से संबद्ध है। यह संख्यात्मक डेटा के पद्धतिबद्ध रूप से संग्रहण तथा विवेचन से संबद्ध है। डेटा का यही पद्धतिबद्ध संग्रहण, इसे अन्य प्रकार की सूचना से अलग करता है।

2.2.2 चर तथा इसके प्रकार

यहाँ पद **'चर'** को परिभाषित करना उचित होगा। चर गुणात्मक अथवा परिमाणात्मक वर्गों में बाँटे जा सकते हैं।

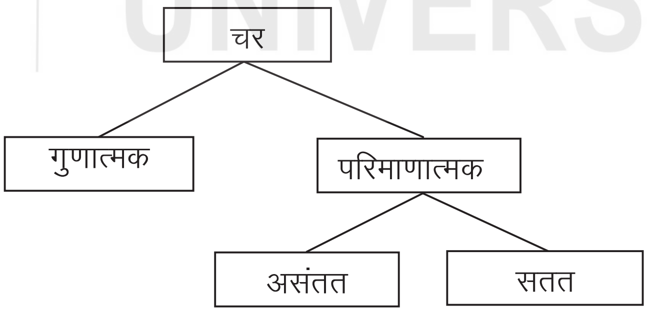
गुणात्मक चर वे चर होते हैं जिन्हें कुछ विशेषताओं अथवा गुणों के आधार पर विभिन्न वर्गों में रखा जा सके। उदाहरण के लिए, यदि व्यक्तियों को लिंग (पुरुष, स्त्री इत्यादि) के आधार पर बांटा जाए तो चर 'लिंग' गुणात्मक होगा।

परिमाणात्मक चर संख्यात्मक होते हैं तथा उन्हें क्रमबद्ध या श्रेणीबद्ध किया जा सकता है। उदाहरण के लिए, चर 'आयु' संख्यात्मक है तथा लोगों को उनकी आयु के अनुसार क्रमबद्ध किया जा सकता है। ऊँचाई, भार, शरीर का तापमान इत्यादि परिमाणात्मक चरों के अन्य उदाहरण हैं। परिमाणात्मक चर पुनः दो अन्य वर्गों में विभाजित किए जा सकते हैं : असंतत चर तथा संतत चर।

असंतत चर वे चर होते हैं जिन्हें 0, 1, 2, 3 इत्यादि मान दिए जा सकें। इन्हें गणनीय कहा जाता है। किसी परिवार में बच्चों की संख्या, एक कक्षा में विद्यार्थियों की संख्या तथा किसी स्विचबोर्ड कर्मी द्वारा किसी मास के प्रत्येक दिन प्राप्त होने वाली फोन कॉल की संख्या, असंतत चरों के कुछ उदाहरण हैं। *असंतत चर वे मान लेते हैं जिन्हें गिना जा सकता है।*

सतत् चर, तुलनात्मक रूप से, वे चर होते हैं जो दो मानों के मध्य के अंतराल में असंख्य मान ले सकते हैं। उदाहरण के लिए, तापमान एक सतत् चर है क्योंकि यह चर दो दिए हुए तापमानों के बीच में असंख्य मान ले सकता है। सतत चर, दो दिए हुए विशिष्ट मानों के बीच में अनंत मान ले सकता है। इन्हें नाप कर प्राप्त किया जा सकता है। इनमें बहुधा भिन्नें तथा दशमलव वाली संख्याएँ सम्मिलित होती हैं। चरों को नामित मान, क्रमवाची या अंतराल के रूप में भी वर्गीकृत किया जा सकता है।

चरों के वर्गीकरण को निम्नलिखित रूप से सारबद्ध किया जा सकता है :



गुणात्मक या परिमाणात्मक चरों के रूप में वर्गीकृत किए जाने के अतिरिक्त, चरों को अन्य आधारों पर भी वर्गीकृत किया जा सकता है। जैसे कि उनकी गणना की विधि के आधार पर अथवा माप की विधि के आधार पर।

2.2.3 डेटा के स्रोत

डेटा गुणात्मक या परिमाणात्मक चरों के मानों का एक समुच्चय होता हैं डेटा के विशिष्ट खंड सूचना के विशिष्ट खंड होते हैं। यद्यपि डेटा की संकल्पना प्रायः वैज्ञानिक अनुसंधान से संबद्ध होती है, डेटा का संग्रहण बहुत-से संगठनों तथा संस्थाओं द्वारा किया जाता है जिनमें व्यापारिक संस्थाएँ (उदाहरणार्थ, विक्रय डेटा,

राजस्व, लाभ, स्टॉक-कीमतें इत्यादि), सरकारें (उदाहरणार्थ, अपराध दर, बेरोज़गारी दर, साक्षरता दर इत्यादि) तथा गैर-सरकारी संगठन सम्मिलित हैं।

प्राथमिक तथा द्वितीयक स्रोत : डेटा का संग्रह एक प्राथमिक स्रोत द्वारा किया जा सकता है (जिसमें अनुसंधानकर्ता डेटा एकत्रित करने वाला प्रथम व्यक्ति होता है) या द्वितीयक स्रोत द्वारा (जिसमें अनुसंधानकर्ता अन्य स्रोतों द्वारा पहले से संग्रहित डेटा का उपयोग करता है, जैसे कि किसी वैज्ञानिक पत्रिका द्वारा प्रसारित किया गया डेटा)।

2.2.4 डेटा के प्रकार

डेटा तीन प्रकार के हो सकते हैं– काल श्रेणी (Time Series), समस्तरीय (Cross-Section) तथा संयोजित (Pooled)।

1) **काल श्रेणी डेटा** : यह एक चर के द्वारा भिन्न-भिन्न समय पर लिए गए मानों से संबंधित प्रेक्षणों का एक समुच्चय होता है। ऐसा डेटा नियमित समय-अंतरालों पर एकत्रित किया जा सकता है। जैसे कि प्रतिदिन (उदाहरणार्थ, उपभोक्ता कीमत सूचकांक इत्यादि), त्रैमासिक रूप से (उदाहरणार्थ, जी.डी.पी. इत्यादि) अथवा वार्षिक रूप से (उदाहरणार्थ, सरकारी बजट, इत्यादि)।

2) **समस्तरीय डेटा** : समस्तरीय डेटा वह डेटा होता है जो कि एक या एक से अधिक चरों के लिए एक ही समय पर एकत्रित किया जाता है। जैसे कि रजिस्ट्रार द्वारा जनगणना के माध्यम से एकत्र किया गया डेटा।

3) **संयोजित डेटा** : संयोजित डेटा में, समय काल डेटा तथा समस्तरीय डेटा दोनों के तत्व समाहित होते हैं। उदाहरण के लिए, किसी समयांतराल के लिए मान लीजिये 2000 से 2013 के लिए भारत के विभिन्न राज्यों के लिए बचत, निवेश तथा जी.डी.पी. से संबंधित डेटा।

4) **समस्तरीय, देशांतरीय या सूक्ष्म समस्तरीय डेटा (Panel, Longitudinal or Micro Panel Data):** यह एक विशेष प्रकार का संयोजित डेटा होता है जिसमें किसी समय अंतराल पर एक ही वर्ग (मान लीजिए एक परिवार या फर्म) का सर्वेक्षण किया जाता है।

2.2.5 माप के पैमाने

- माप के पहले स्तर को माप का नामित/सांकेतिक स्तर कहते हैं। कॉलेज प्रशिक्षकों का एक प्रतिदर्श जिसे पढ़ाये जाने वाले विषय (जैसे कि अंग्रेज़ी, इतिहास, मनोविज्ञान या गणित इत्यादि) के अनुसार वर्गीकृत किया गया हो, सांकेतिक स्तर मापक का एक उदाहरण है। इस स्तर पर डेटा को क्रमबद्ध अथवा श्रेणीबद्ध नहीं किया जा सकता। सांकेतिक स्तर मापन डेटा को परस्पर अनन्य या गैर-अतिव्यापी सर्वग्राही वर्गों में बांटता है जिन पर कोई क्रम या श्रेणी नहीं लगाई जा सकती।

- माप का दूसरा स्तर क्रमवाचक स्तर कहलाता है। इस स्तर पर मापा गया डेटा वर्गों में बांटा जा सकता है तथा ये वर्ग क्रमबद्ध या श्रेणीबद्ध किए जा सकते हैं। उदाहरण के लिए, विद्यार्थियों द्वारा किए गए मूल्यांकन के आधार पर आमंत्रित वक्ताओं को श्रेष्ठ, औसत या खराब श्रेणियों में रखा जा सकता है। ध्यान रहे कि क्रमवाचक स्तर पर माप के लिए दो इकाइयों में माप के अंतर के लिए कोई

स्पष्ट विधि नहीं होती। अतः और अधिक विश्लेषण के लिए इन वर्गों को एक संख्यात्मक पैमाने में परिवर्तित किया जाना चाहिए।

- मापन का तीसरा स्तर, अंतराल स्तर कहलाता हैं इस स्तर में और क्रमवाचक स्तर में मुख्य अंतर यह है कि इस स्तर में दो इकाइयों में अंतर की एक स्पष्ट विधि होती है। उदाहरण के लिए, अनेक मानक मनोवैज्ञानिक परीक्षणों में अंतराल के पैमाने पर मापे गए मान प्राप्त होते हैं। आई.क्यू (IQ) एक ऐसे चर का उदाहरण है। 109 के आई.क्यू. तथा 110 के आई.क्यू. में 1 बिंदु का स्पष्ट तथा अर्थपूर्ण अंतर है। तापमान, अंतराल माप का एक अन्य उदाहरण है क्योंकि 72ºF और 73ºF में 1ºF का स्पष्ट तथा अर्थपूर्ण अंतर है। अंतराल पैमाने में एक गुणधर्म की कमी होती है : इसमें शुद्ध शून्य नहीं होता। उदाहरण के लिए, आई. क्यू. परीक्षण उन व्यक्तियों का माप नहीं ज्ञात कर सकते जिनकी बुद्धिमता शून्य हो। इसी प्रकार 0ºF का अर्थ यह नहीं होता कि ताप/ऊष्मा बिल्कुल भी नहीं है।

- माप का अंतिम स्तर अनुपात स्तर कहलाता है। ऐसे पैमाने जो ऊँचाई, भार, क्षेत्रफल तथा प्राप्त फोन कॉल्स का माप देते हैं, अनुपातिक पैमानों के उदाहरण कहलाते हैं। अनुपातिक पैमानों में इकाइयों में स्पष्ट अंतर होता है (1 इंच, 1 कि.ग्रा., इत्यादि) तथा एक शुद्ध शून्य भी होता है। इसके अतिरिक्त, अनुपातिक पैमाना, विभिन्न मानों में एक शुद्ध अनुपात भी दर्शाता है। उदाहरण के लिए, यदि एक व्यक्ति 200 पाउंड का भार उठा सकता है तथा एक अन्य व्यक्ति 100 पाउंड का भार उठा सकता है, तो उनके बीच में 2 : 1 का अनुपात होता है। दूसरे शब्दों में, पहला व्यक्ति, दूसरे व्यक्ति से दो गुना भार उठा सकता है। माप के अनुपात स्तर में अंतराल माप के सभी गुणधर्म होते हैं तथा एक शुद्ध शून्य भी होता है। इसके अतिरिक्त, यदि एक ही चर के मान समष्टि की दो अलग इकाइयों के लिए ज्ञात किए जाएँ, तो उनमें एक शुद्ध अनुपात भी होता है।

चरों की परिभाषा तथा वर्गीकरण महत्त्वपूर्ण है क्योंकि इनके आधार पर हम निर्धारित करते हैं एकल विचर, द्वि-विचर तथा बहु-विचर डेटा विश्लेषण में कौन-सी विधियों का प्रयोग किया जाएगा।

2.2.6 डेटा का वर्गीकरण

डेटा संग्रह करने के पश्चात्, हमारे सामने समस्या होती है उसे एक ऐसे रूप में व्यवस्थित करने की, जिससे हम कुछ निष्कर्ष निकाल सकें। कुछ समानताओं के आधार पर व्यवस्थित करने को वर्गीकरण कहते हैं। वर्गीकरण में डेटा को वर्गों में बांट दिया जाता है जिससे हमें डेटा के बारे में व्यापक अंतर्दृष्टि उपलब्ध होती है। ऐसा करने से असंसाधित डेटा पर काम करने से होने वाली परेशानियाँ तथा जटिलताएँ कम हो जाती हैं। वर्गीकरण डेटा को घनीभूत रूप में प्रस्तुत करता है जिससे डेटा का विश्लेषण किफायती रूप से सटीकता के साथ किया जा सकता है। इसकी समानताओं और विविधताओं की व्याख्या की जा सकती है तथा तुलना सुविधाजनक हो जाती है। यह कहा जा सकता है कि वर्गीकरण एक ऐसा आवश्यक और महत्त्वपूर्ण कदम है जो अन्वेषक को परिणामों को एक अर्थपूर्ण तथा सुविधाजनक रूप में सारबद्ध करने में सहायता करता है। इकाई 6 में पाठकों का परिचय वर्गीकरण की कला तथा शिल्प से करवाया गया है। वहाँ हमने सारणियों, संतत तथा असंतत बारंबरता आबंटनों, समूहीकृत बारंबारता आबंटनों के बारे में चर्चा की है। इस इकाई में, हम

आपका परिचय एक दिए हुए डेटा समुच्चय की आवश्यक विशेषताओं को संक्षेप में प्रस्तुत करने की विधियों से करवाएंगे।

2.3 वर्णनात्मक सांख्यिकी

इस विश्व में अधिकतर प्रेक्षण परिवर्तनशील होते हैं, विशेषकर इंसानी व्यवहार से संबद्ध प्रेक्षण। अंतः ऐसे डेटा के आधार पर बुद्धिमत्तापूर्ण निर्णय लेने के लिए यह आवश्यक हो जाता है कि इन प्रेक्षणों को विधिपूर्वक रूप से व्यक्त किया जाए, अर्थात्, प्रेक्षणों को एक एकल अनुमान के रूप में व्यक्त करने की आवश्यकता होगी जो कि प्रेक्षणों को सारबद्ध कर सकें। वर्णनात्मक सांख्यिकी, सांख्यिकी की वह शाखा है जो प्राप्त डेटा के वर्णन से संबद्ध है। इन वर्णकों के आधार पर इस डेटा की विशेषताओं के संगत समष्टि का एक विशिष्ट समूह परिभाषित किया जा सकता है। वर्णनात्मक सांख्यिकी में वर्गीकरण, सारणीकरण, डेटा का आरेखीय तथा चित्रात्मक रूप में प्रस्तुतीकरण, केंद्रीय प्रवृत्ति तथा परिवर्तनशीलता के माप इत्यादि सम्मिलित होते हैं। ये माप अनुसंधानकर्ता को डेटा की प्रवृत्ति जानने में सहायता करते हैं जिससे संबद्ध घटना के वर्णन में और अधिक सरलता होती है।

सांख्यिकी की शाखा जो कि डेटा के वर्णन तथा उसे सारबद्ध करने से संबद्ध है, वर्णनात्मक सांख्यिकी कहलाती है। इसमें आलेखों, चार्टों तथा सारणियों की रचना तथा विभिन्न वर्णनात्मक मापों जैसे कि औसतों, प्रसरण और प्रतिशतक इत्यादि का परिकलन सम्मिलित होते हैं। अतः यह स्पष्ट है कि वर्णनात्मक सांख्यिकी केवल संख्याओं के सारणीकरण तथा उनके आलेखीय चित्रण तक ही सीमित नहीं है, इसका विषय-क्षेत्र इससे कहीं अधिक है।

वर्णनात्मक सांख्यिकी का उपयोग किसी अध्ययन के लिए डेटा की मूलभूत विशेषताओं का वर्णन करने के लिए किया जाता है। वर्णनात्मक सांख्यिकी में हम केवल यह बताते हैं कि दिया हुआ डेटा क्या दर्शाता है। वर्णनात्मक सांख्यिकी तक हम डेटा के बड़ी मात्रा को बुद्धिमतापूर्वक सरलीकृत करते हैं। वर्णनात्मक सांख्यिकी अत्यधिक बड़े डेटा को सरल तथा संक्षिप्त रूप में परिवर्तित कर देती है। उदाहरण के लिए, बैटिंग औसत केवल एक अंक है जिसका प्रयोग यह दर्शाने के लिए किया जा सकता है कि बेसबॉल में एक बल्लेबाज का प्रदर्शन कैसा है? यह संख्या/अंक बल्लेबाज द्वारा लगाई गई हिट्स को उन बॉलों की संख्या से विभाजित करके निकाली जाती है तो बल्ले पर आयीं और इसे तीन दशमलव अंकों तक ज्ञात किया जाता है। एक बल्लेबाज, जिसका हिट रेट .333 है, बल्ले पर आने वाली प्रत्येक तीन बॉलों में से एक को हिट करता है। एक बल्लेबाज जिसका हिट रेट औसत .250 है चार में से एक बॉल को हिट करता है। यह एक अकेली संख्या असंतत घटनाओं की एक बड़ी संख्या को निरूपित करती है। या ग्रेड बिंदु एवरेज (GPA/जीपीए) पर विचार कीजिए, जो अनेक विषयों के आधार पर विद्यार्थी के व्यापक प्रदर्शन को अभिव्यक्त करती है। वर्णनात्मक सांख्यिकी का प्रयोग बहुध निम्नलिखित की जाँच के लिए किया जाता है :

- **केंद्रीय प्रवृत्ति** अर्थात् डेटा (स्थिति) कहाँ स्थित है यदि उसे माध्य, माध्यिका और बहुलक के माध्यम से नापा जाए।

- **डेटा का परिक्षेपण** (परिवर्तनशीलता) अर्थात् डेटा का फैलाव कितना है यदि उसे प्रसरण तथा उसके वर्गमूल, मानक विचलन के माध्यम से नापा जाए।

- **डेटा का वैषम्य** (की असममिति) अर्थात् यदि विषम सूचकांक के माध्यम से नापा जाए तो पैमाने के उच्च या निम्न छोर पर डेटा किस प्रकार संकेद्रित है।
- **डेटा की ककुदता (Kurtosis)** (शिखरता) अर्थात् यदि ककुदता सूचकांक द्वारा नापा जाए तो डेटा एक बिंदु के आस-पास किस प्रकार केंद्रित है।

बोध प्रश्न 1

1) 'एकल विचर डेटा विश्लेषण' की व्याख्या कीजिए।

....

....

....

....

2) वर्णनात्मक सांख्यिकी को परिभाषित कीजिए तथा डेटा विश्लेषण में इसके मौलिक योगदान पर टिप्पणी कीजिए।

....

....

....

....

3) पद 'परिमाणात्मक चर' को परिभाषित कीजिए तथा इसके माप के विभिन्न स्तरों का वर्णन कीजिए।

....

....

....

....

2.4 परिमाणात्मक डेटा : प्रतिशत तथा केंद्रीय प्रवृत्ति के माप

2.4.1 प्रतिशत के प्रयोग

प्रतिशत किसी सूचना को पूर्ण के अनुपात में व्यक्त करता है तथा यह अधिकतया प्रयोग किया जाने वाला प्राचल है। इसको समझना काफी सरल है। उदाहरण के लिए, यह कहने की अपेक्षा कि एक परीक्षा में 400 में से 280 विद्यार्थी उत्तीर्ण हुए हैं, यह समझना अधिक सरल है कि परीक्षा में भाग लेने वाले कुल विद्यार्थियों में से 70% उत्तीर्ण हुए हैं। प्रतिशत तुलनात्मक संबंध को दर्शाने का एक अच्छा और सरल तरीका है – चाहे यह तुलना प्रतिक्रिया देने वालों के मध्य की जा रही हो अथवा प्रतिक्रियाओं के मध्य। प्रतिशत किसी वर्गीकृत बारंबारता बंटन को दर्शाने के लिए भी उपयोगी है। यद्यपि प्रतिशत का परिकलन अत्यंत सरल है, तो भी हमें नीचे दी गई सावधानियाँ बरतनी चाहिएं जिससे त्रुटियाँ होने की संभावना न रहे :

- ***सही आधार का प्रयोग करना :*** आधार वह संख्या होती है जिसका प्रतिशत ज्ञात किया जाना है। सही आधार का प्रयोग करना तथा जिस आधार का प्रयोग किया जा रहा है उसे स्पष्ट रूप से इंगित करना आवश्यक होता है। ऊपर दिए उदाहरण में 70% का अर्थ कुल विद्यार्थियों का 70% या कि उन विद्यार्थियों का 70% जिनसे कोई प्रश्न पूछा गया था? इसे और स्पष्ट करने के लिए मान लीजिये कि कोई प्रश्न पूछे जाने पर, 10 व्यक्ति कोई प्रतिक्रिया नहीं देते, 70 'हाँ' में उत्तर देते हैं तथा 20 'नहीं' में। यदि हम आधार या विभाजक के तौर पर 100 का प्रयोग करते हैं, तो हम कहेंगे कि 70% लोगों ने 'हाँ' में उत्तर दिया; परंतु यदि आधार 90 लें अर्थात् उन लोगों की संख्या लें जिन्होंने अपनी प्रतिक्रिया दी, तो हम पाएंगे कि 78% लोगों की प्रतिक्रिया 'हाँ' में थी क्योंकि हम नहीं जानते कि जिन लोगों ने अपनी प्रतिक्रिया नहीं दी, उनका उत्तर 'हाँ' में होता या 'नहीं' में।

- ***प्रतिशतों का योग :*** प्रतिशतों का योग केवल तभी किया जाता है जब वर्ग/श्रेणियाँ परस्पर अनन्य हों अर्थात् अतिव्याप (overlap) न हों। बहु विकल्पी प्रश्नों में यह स्थिति नहीं होती क्योंकि वहाँ प्रतिक्रिया देने वाला एक से अधिक विकल्प चुन सकता है। उदाहरण के लिए, यदि कुछ किसानों/लोगों से गेहूँ के उत्पादन में प्रयोग होने वाले उर्वरकों के बारे में सूचना प्राप्त करने के स्रोतों के बारे में पूछ जाए तो प्रतिक्रिया देने वाला व्यक्ति एक या एक से अधिक उत्तर दे सकता है जैसे कि उर्वरक विक्रेता, साथी किसान, समाचार-पत्र में विज्ञापन इत्यादि। ये उत्तर/प्रतिक्रियाएँ परस्पर अनन्य नहीं हैं, अतः इनके प्रतिशतों को नहीं जोड़ा जा सकता।

- ***प्रतिशतों का औसत :*** प्रतिशतों को जोड़कर उनका औसत ज्ञात करने से होने वाली त्रुटि से बचना चाहिए। इससे गलत उत्तर प्राप्त होगा जैसा कि नीचे दी तालिका में देखा जा सकता है :

ज़िला	पूर्ण की गई परियोजनाएँ	प्रतिशत
A	31,000	74
B	8,000	60
C	12,000	65
D	26,000	75
E	11,000	50
F	28,000	72
कुल परियोजनाएँ 165000	**1,16,000**	**औसत 66 (त्रुटिपूर्ण)**

इस तालिका में, एक अनुसंधानकर्त्ता ने छह जिलों के प्रतिशतों का गलत/त्रुटिपूर्ण औसत ज्ञात किया है। जबकि सही प्रतिशत ज्ञात करने के लिए पूर्ण की गई परियोजनाओं की संख्या (116000) को कुल स्वीकृत परियोजनाओं (165000) से विभाजित किया जाना चाहिए था जिससे हमें पूरी की गई परियोजनाओं का सही औसत 70.3% प्राप्त होता।

2.4.2 केंद्रीय प्रवृत्ति के माप

जब हम संख्यात्मक डेटा पर काम करते हैं, तो यह स्वाभाविक प्रतीत होता है कि अधिकतर डेटा समुच्चयों में प्रेक्षित मानों में किसी आंतरिक मान के आस-पास समूहीकृत होने की प्रवृत्ति पाई जाती है; जिनमें से कुछ केंद्रीय मान उस डेटा की विशिष्टताएँ दर्शाते हैं। इस प्रवृत्ति को केंद्रीय प्रवृत्ति कहते हैं। डेटा के किसी दिए हुए समुच्चय के लिए, हम स्थिति के किस माप का प्रयोग करें यह इस पर निर्भर करता है कि हम मध्य से क्या समझते हैं; विभिन्न परिभाषाओं से विभिन्न माप उत्पन्न होते हैं। हम सामान्यतः प्रयोग में आने वाले कुछ मापों पर विचार करेंगे। ये माप हैं माध्य, माध्यिका तथा बहुलक। इन मानों को ज्ञात करने के सूत्र इस पर निर्भर करते हैं कि डेटा वर्गीकृत हैं अथवा अवर्गीकृत।

माध्य, जिसे समांतर माध्य भी कहते हैं, **प्रेक्षणों/डेटा के मानों** के योग को प्रेक्षणों/मानों की कुल संख्या से भाग करके ज्ञात किया जाता है। उदाहरण के लिए, 3, 2, 6, 5 और 4 का माध्य, 3 + 2 + 6 + 5 + 4 के मान को 5 से भाग करके प्राप्त होता है। अतः, इस डेटा का माध्य 20/5 = 4 है। यदि डेटा के मानों को X से निरूपित किया जाए, तो गणित की भाषा में माध्य, मानों के योग को, मानों की कुल संख्या से भाग करके प्राप्त होता है। चिन्ह $\bar{X}$ प्रेक्षणों/प्रतिदर्श के माध्य को निरूपित करता है।

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \cdots + X_n}{n} = \frac{\sum X}{n}$$

जहाँ n प्रतिदर्श के प्रेक्षणों की कुल संख्या को निरूपित करता है।

एक समष्टि के माध्य के लिए ग्रीक अक्षर μ (म्यू) का प्रयोग किया जाता है :

$$\mu = \frac{X_1 + X_2 + X_3 + \cdots + X_N}{N} = \frac{\sum X}{N}$$

जहाँ N समष्टि में प्रेक्षणों की कुल संख्या को निरूपित करता है।

सांख्यिकी में ग्रीक अक्षरों का प्रयोग प्राचलों को व्यक्त करने के लिए तथा रोमन अक्षरों का प्रयोग विचरों (variates) को व्यक्त करने के लिए किया जाता है।

किसी वर्गीकृत डेटा का माध्य ज्ञात करने के विभिन्न चरणों को नीचे दी तालिका में सारबद्ध किया गया है :

वर्गीकृत डेटा का माध्य ज्ञात करना

<table>
<tr><td>चरण 1</td><td colspan="4">एक तालिका बनाइए जैसे कि नीचे दर्शाया गया है</td></tr>
<tr><td></td><td>A
वर्ग</td><td>B
बारंबारता f</td><td>C
मध्य बिंदु X_m</td><td>D
$f.X_m$</td></tr>
<tr><td>चरण 2</td><td colspan="4">प्रत्येक वर्ग का मध्य बिंदु ज्ञात कीजिए तथा उसे स्तंभ C में लिखिये।</td></tr>
<tr><td>चरण 3</td><td colspan="4">प्रत्येक वर्ग के मध्य बिंदु को उस वर्ग की बारंबारता से गुणा करिये तथा प्राप्त गुणनफल को स्तंभ D में लिखिये।</td></tr>
<tr><td>चरण 4</td><td colspan="4">स्तंभ D की सभी प्रविष्टियों का योग ज्ञात कीजिए।</td></tr>
<tr><td>चरण 5</td><td colspan="4">स्तंभ D से प्राप्त योग को स्तंभ B में दी गई बारंबारताओं के योग से विभाजित कीजिये। माध्य ज्ञात करने के लिए सूत्र है :</td></tr>
</table>

	$\bar{X} = \frac{\sum f.X_m}{n}$ **टिप्पणी** : संकेत $\sum f.X_m$ का अर्थ है प्रत्येक वर्ग के लिए बारंबारता (f) और मध्यबिंदु (X_m) के गुणनफल का योग।

साधारण माध्य ज्ञात करने के लिए, हम मान लेते हैं कि आबंटन में सभी मानों का महत्त्व समान है। परंतु, व्यवहारिक जीवन में ऐसा होना आवश्यक नहीं है। यदि कुछ प्रविष्टियों का महत्त्व अन्य प्रविष्टियों से अधिक हो, तो साधारण माध्य आबंटन का सही रूप से प्रतिनिधित्व नहीं करता।

ऐसी स्थितियों में हमें विभिन्न प्रविष्टियों को उपयुक्त महत्त्व देना पड़ता है। उदाहरण के लिए, यदि हम व्यक्तियों के एक समूह की जीवनयापन की लागत में होने वाले परिवर्तन का अनुमान लगाना चाहें तो उनके द्वारा उपभोग की जाने वाली वस्तुओं की कीमत का साधारण माध्य से काम नहीं चलेगा क्योंकि सभी वस्तुओं का महत्त्व समान नहीं होता जैसे कि चावल, गेहूँ और दालों का महत्त्व चाय, मिष्ठान इत्यादि से अधिक होगा। बल्कि इस स्थिति में भारित समांतर माध्य अधिक उपयुक्त होगा क्योंकि इसमें प्रत्येक वस्तु/प्रविष्टि को उपयुक्त भार/महत्त्व के साथ प्रयोग में लाया जाता है। इस स्थिति में, माध्य ज्ञात करते समय प्रत्येक घटना/प्रविष्टि को एक विशिष्ट मान से गुणा किया जाता है, जिसे प्रविष्टि का भार कहा जाता है तथा इस प्रकार प्राप्त गुणनफलों के योग को, भारों के कुल योग से विभाजित किया जाता है।

यदि $x_1, x_2 \ldots x_n$ चर x के मान हों तथा $w_1, w_2 \ldots w_n$ इन मानों के संगत भार, तो भारित माध्य होगा : A.M. = $\overline{x_w} = \frac{w_1x_1+w_2x_2+\cdots+w_nx_n}{w_1+w_2+\cdots+w_n} = \frac{\sum w_1x_1}{\sum w_1}$

किसी डेटा की **माध्यिका** (median) डेटा प्रविष्टियों के ठीक मध्य वाली प्रविष्टि होती जबकि सभी प्रविष्टियों को परिमाण के क्रम में व्यवस्थित किया जाए। उदाहरण के लिए, यदि 11 व्यक्तियों का भार किया जाए और उन्हें भार के बढ़ते या घटते हुए क्रम व्यवस्थित किया जाए, तो छठी प्रविष्टि का मान माध्यिका होता है क्योंकि प्रविष्टि के नीचे और ऊपर दोनों ओर पांच-पांच प्रविष्टियाँ हैं। माध्यिका पारिवारिक आय का प्रयोग अक्सर माध्य से बेहतर औसत के रूप में किया जाता है क्योंकि माध्यिका ठीक मध्य वाली आय को निरूपित करती है। ठीक 50% परिवारों की आय माध्यिका से ऊपर होगी और ठीक 50% परिवारों की आय माध्यिका से नीचे। यदि प्रविष्टियों की संख्या सम हो, तो माध्यिका मध्य की दो प्रविष्टियों का औसत ज्ञात करके प्राप्त होती है। इसका सूत्र है :

माध्यिका (Median) = Md= $\left(\frac{n+1}{2}\right)$**वीं प्रविष्टि**

वर्गीकृत डेटा में डेटा की प्रविष्टियों के साथ बारंबारताएँ भी संबद्ध होती हैं। वर्गीकरण एक संतत बारंबारता आबंटन के रूप में भी हो सकता है और असंतत बारंबारता आबंटन के रूप में भी। आबंटन किसी भी प्रकार का हो, हमें प्रविष्टियों की कुल संख्या ज्ञात करने के लिए, संचयी बारंबारताएँ ज्ञात करनी पड़ती हैं। किसी भी वर्ग की संचयी बारंबारता, उस वर्ग की बारंबारता तथा उससे पीछे वाले सभी वर्गों की बारंबारताओं के योग के बराबर होती है। अर्थात् संचयी बारंबारता, क्रमिक रूप से बारंबारताओं का योग करके ज्ञात की जाती है जिससे अंतिम संचयी बारंबारता,

प्रविष्टियों की कुल संख्या के बराबर होती है। एक वर्गीकृत डेटा की माध्यिका का परिकलन निम्नलिखित चरणों का प्रयोग करके किया जा सकता है :

चरण 1 : संचयी बारंबारताएँ ज्ञात कीजिए।

चरण 2 : $\left(\frac{N}{2}\right)$ ज्ञात कीजिए।

चरण 3 : संचयी बारंबारता का वह मान ढूंढिये जो $\left(\frac{N}{2}\right)$ के बराबर या अधिक छोटा हो। इस बारंबारता के संगत वर्ग अंतराल माध्यिका वर्ग कहलाता है।

$$\text{माध्यिका} = l + \frac{\frac{N}{2} - m}{f} \times c$$

जहाँ l = माध्यिका वर्ग की निचली सीमा

m = माध्यिका वर्ग से पिछले वर्ग की संचयी बारंबारता

c = माध्यिका वर्ग की चौड़ाई

f = माध्यिका वर्ग की बारंबारता

N = कुल बारंबारता

टिप्पणी : यदि वर्ग अंतराल समावेशित रूप में दिए हों तो पहले उन्हें अनन्य रूप में परिवर्तित करें तथा तब उनकी निचली सीमा लें।

***बहुलक* डेटा** में सबसे अधिक बार आने वाले मान को कहते हैं। क्रॉक्स्टन तथा कॉउडन के अनुसार *''किसी दिए हुए बंटन के लिए बहुलक उस बिंदु का मान होता है जिसके आस-पास प्रविष्टियाँ सबसे सघन रूप में संकेंद्रित होती हैं। इसे मानों की एक श्रेणी का सबसे विशिष्ट मान कहा जा सकता है।''* यह दर्शाता है कि बारंबारता का संकेंद्रण एक दिए हुए मान के आस-पास है। अतः, जहाँ हम सबसे अधिक संकेंद्रण वाले बिंदु के बारे में जानना चाहते हैं, वहाँ बहुलक को वरीयता दी जाती है। अर्थात् यह एक स्थिति/अवस्था का माप है। यदि किसी प्रतिदर्श में 25 वर्ष की आयु वाले लोग सबसे अधिक हों तो उस प्रतिदर्श का बहुलक 25 होगा। बहुलक केंद्रीय प्रवृत्ति का सबसे कम प्रयोग किया जाने वाला माप है, विशेषकर जब डेटा समुच्चय बड़ा हो। तो भी बहुलक किसी डेटा समुच्चय के वर्णन में महत्त्वपूर्ण होता है, विशेष तौर पर उस स्थिति में जब एक से अधिक मान बार-बार आ रहे हों।

ऐसी स्थिति में, डेटा को द्वि बहुलक या बहु बहुलकी डेटा/बंटन कहते हैं, यह इस पर निर्भर करता है कि डेटा में दो मान बार-बार आते हैं अथवा दो से अधिक मान। एक संतत डेटा में, हम पहले सबसे बड़ी बारंबारता ज्ञात करते हैं। इस बारंबारता के संगत वर्ग को बहुलक वर्ग कहते हैं। इसके पश्चात्, नीचे दिए सूत्र से बहुलक का मान ज्ञात किया जाता है :

$$\textbf{बहुलक} = M_0 = l + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times c$$

जहाँ l = बहुलक वर्ग की निम्न सीमा

$\Delta_1 = f_1 - f_0$

$\Delta_2 = f_1 - f_2$

f_1 =बहुलक वर्ग की बारंबारता

f_0 = बहुलक वर्ग से पिछले वाले वर्ग की बारंबारता

f_2 = बहुलक वर्ग से आगे वाले वर्ग की बारंबारता

$$\text{बहुलक} = l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times c$$

टिप्पणियाँ :

1) यदि $(2f_1 - f_0 - f_2)$ का मान शून्य हो जाए, तो बहुलक का मान नीचे दिए सूत्र के द्वारा निरपेक्ष मानों के प्रयोग से ज्ञात किया जाता है।

2) $M_0 = l + \frac{(f_1 - f_0)}{|f_1 - f_0| + |f_1 - f_2|} \times c$

3) यदि बहुलक पहले वर्ग अंतराल में स्थित हो तो f_0 को 0 लिया जाता है।

2.4.3 औसतों के बीच आनुभविक संबंध

एक सममित (symmetrical) आबंटन में तीन साधारण औसतों में संबंध **माध्य = माध्यिका = बहुलक** होता है। किसी मध्यम रूप से असममित आबंटन में औसतों के बीच संबंध **बहुलक = 3 माध्यिका–2 माध्य** होता है। यह संबंध प्रो. कार्ल पीयरसन द्वारा दिया गया था।

2.4.4 कौन-सा माप चुना जाए?

अनुसंधानकर्त्ताओं तथा सांख्यिकीविदों के लिए यह जानना अत्यंत आवश्यक होता है कि कब कौन-से औसत का प्रयोग किया जाना चाहिए। नीचे केंद्रीय प्रवृत्ति के विभिन्न औसतों के गुणधर्म तथा प्रयोग दिए जा रहे हैं तथा किसी दी हुई स्थिति में उनके उपयोग पर भी प्रकाश डाला गया है :

माध्य

- माध्य डेटा के सभी मानों का प्रयोग करके ज्ञात किया जाता है।
- यदि प्रतिदर्श एक ही समष्टि से प्राप्त किए जाएं तथा तीनों औसत इन्हीं प्रतिदर्शों से ज्ञात किए जाएं, तो माध्यिका तथा बहुलक की अपेक्षा माध्य कम परिवर्तित होता है।
- माध्य का उपयोग अन्य मान ज्ञात करने के लिए भी किया जाता है जैसे कि प्रसरण।
- किसी डेटा समुच्चय के लिए माध्य अद्वितीय होता है तथा यह आवश्यक नहीं है कि इसका मान, डेटा के किसी मान के बराबर ही हो।
- माध्य, विशेषकर समांतर माध्य पर बीजगणितीय संक्रियाएँ संभव होती हैं जैसे कि दो या दो से अधिक श्रेणियों का संयुक्त माध्य ज्ञात करना।
- विवृतांत वर्गों (open ended class) वाले बारंबारता आबंटनों के लिए माध्य ज्ञात नहीं किया जा सकता।
- डेटा में कुछ अत्यधिक बड़े और अत्यधिक छोटे मानों वाली प्रविष्टियों की उपस्थिति से, जिन्हें बाहरी कारक कहते हैं, माध्य प्रभावित होता है तथा ऐसी स्थितियों में यह उपयुक्त औसत नहीं होता।

माध्यिका

- माध्यिका का प्रयोग एक डेटा समुच्चय का केंद्रीय या मध्य मान ज्ञात करने के लिए किया जाता है।
- माध्यिका का प्रयोग तब किया जाता है जब यह जानना आवश्यक हो कि डेटा के अधिक मान आबंटन के ऊपरी अर्द्ध भाग में स्थित हैं अथवा निचले अर्द्ध भाग में।
- माध्यिका प्रयोग विवृतांत वर्ग अंतराल वाले आबंटनों में किया जा सकता है।
- अत्यधिक बड़े या अत्यधिक छोटे मानों वाली प्रविष्टियों की उपस्थिति में माध्य की तुलना में माध्यिका कम प्रभावित होती है।

बहुलक

- बहुलक का प्रयोग डेटा की सबसे अधिक बारंबारता वाली विशिष्ट प्रविष्टि को ज्ञात करने के लिए किया जाता है।
- बहुलक का परिकलन सभी औसतों में सबसे सरल है।
- बहुलक का प्रयोग तब भी किया जा सकता है जब डेटा संख्यावाचक न होकर, जातिवाचक हो जैसे कि धार्मिक वरीयता, लिंग या राजनैतिक संबंधन।
- बहुलक का अद्वितीय होना आवश्यक नहीं है। किसी डेटा समुच्चय के लिए एक से अधिक बहुलक भी हो सकते हैं और यह भी हो सकता है कि कोई बहुलक हो ही न।

गतिविधि 1 : केंद्रीय प्रवृत्ति का माप

अर्थशास्त्र में 21 विद्यार्थियों द्वारा प्राप्तांक इस प्रकार हैं — 14, 9, 29, 18, 20, 14, 15, 29, 25, 18, 39, 49, 33, 20, 30, 35, 25, 25, 27, 29, 37 (अधिकतम अंक 50 है)। प्रत्येक विद्यार्थी के प्रतिशत अंक ज्ञात कीजिए तथा एक स्प्रैडशीट पैकेज के प्रयोग से इस डेटा के लिए माध्य, माध्यिका तथा बहुलक का मान ज्ञात कीजिए।

यह कैसे करें?

- स्प्रैडशीट के कोष्ठकों (सैल) A2 से A22 में अंक लिखें। A1 में आप शीर्षक 'Marks in Eco' प्रविष्टि कर सकते हैं।
- प्रत्येक विद्यार्थी के लिए अंकों का प्रतिशत ज्ञात करने के लिए हमें इस विषय के अधिकतम अंकों की आवश्यकता होगी (जो कि इस उदाहरण में 50 हैं), प्रतिशत ज्ञात करने के लिए हमें केवल प्राप्तांकों को अधिकतम अंकों से विभाजित करने की तथा सैल को प्रतिशत (percentage) फारमेट में फारमेट करने की आवश्यकता है। इस उदाहरण में, अधिकतम अंक एक सेल में संग्रहित हैं जो कि विभाजन के लिए प्रयोग में लाए जा सकते हैं। यह 'absolute addressing' को दर्शाने के लिए किया गया है। अतः सैल D2 में "Max Marks" तथा E2 में 50 दर्ज करें। साथ ही, सैल B1 में "%age" दर्ज करें।
- आप सैल A2 से सैल A22 में प्रविष्ट किए गए अंकों का प्रतिशत क्रमशः सैल B2 से B22 सैल में परिकलित करेंगे।
  - पहले सैल B2 में सूत्र ***=(A2/$E$2)*** को प्रविष्ट करें तथा Enter बटन को क्लिक करें। सैल B2 में मान 0.28 दिखेगा तथा Formula Bar में प्रविष्ट किया गया सूत्र दिखेगा (नीचे दिए गए चित्र को देखें) कृपया ध्यान दें कि सैल

A2 में प्रथम विद्यार्थी के अंक दर्ज हैं, इसे अधिकतम अंकों से विभाजित किया गया है तथा प्राप्त परिणाम सैल E2 में दृष्टिगोचर होता है। परंतु सैल E2 में सूत्र *\$E\$2* लिखा गया है। क्यों? ऐसा यह सुनिश्चित करने के लिए किया गया है कि जब आप सैल B2 से इस सूत्र को B2 से B22 तक कॉपी करें, तो यह सैल वही अर्थात् E2 ही रहे। इसे एबसोल्यूट एड्रैसिंग (Absolute Addressing) कहते हैं।

B2 f_x =(A2/$E$2)

	A	B	C	D	E
1	Marks in Eco	%age			
2	14	0.28		Max MARKS	50
3	9				
4	29				

○ सैल B2 से इस सूत्र को सैल B2 से सैल B22 तक कॉपी करने के लिए मॉउस के प्वाइंटर को सैल B2 के नीचे की ओर दाईं ओर वाले बिंदु पर लेकर जाएं। ऐसा करने पर यह बिंदु एक + के चिन्ह में बदल जाएगा। अब माउस का बायां बटन दबाएं तथा इसे माउस को खींचकर सैल B22 तक लाएं जैसा कि नीचे चित्र में दर्शाया गया है :

B2 f_x =(A2/$E$2)

	A	B	C	D	E
1	Marks in Eco	%age			
2	14	0.28		Max MARKS	50
3	9				
4	29				
5	18				
6	20				
7	14				
8	15				
9	29				
10	25				
11	18				
12	39				
13	49				
14	33				
15	20				
16	30				
17	35				
18	25				
19	25				
20	27				
21	29				
22	37				
23					

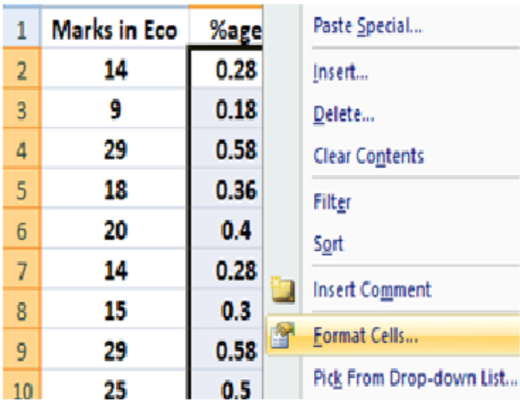
B2 f_x =(A2/$E$2)

	A	B	C	D	E
1	Marks in Eco	%age			
2	14	0.28		Max MARKS	50
3	9	0.18			
4	29	0.58			
5	18	0.36			
6	20	0.4			
7	14	0.28			
8	15	0.3			
9	29	0.58			
10	25	0.5			
11	18	0.36			
12	39	0.78			
13	49	0.98			
14	33	0.66			
15	20	0.4			
16	30	0.6			
17	35	0.7			
18	25	0.5			
19	25	0.5			
20	27	0.54			
21	29	0.58			
22	37	0.74			
23					

○ अब माउस का बटन छोड़ें और स्क्रीन में परिवर्तन हो जाएगा। ऊपर दिए गए चित्र में सबसे दाईं ओर का कॉलम देखिये। विभिन्न सैलों के लिए आप सूत्र की जांच कर सकते हैं। उदाहरण के लिए, सैल B10 हो जाएगा ***=(A10/\$E\$2)***

○ आप देख सकते हैं कि इस कॉलम के सैल B2 से सैल B22 तक का सारा डेटा दशमलव में व्यक्त है। इसे प्रतिशत में परिवर्तित करने के लिए, कॉलम का सैल B2 से सैल B22 तक का भाग सलैक्ट कीजिए और माउस का दायां बटन क्लिक कीजिए। ऐसा करने पर विभिन्न

विकल्पों वाला एक मैन्यू खुला जाएगा। इन विकल्पों में से विकल्प 'Format Cells...' चुनिए जैसा कि नीचे चित्र में दर्शाया गया है :



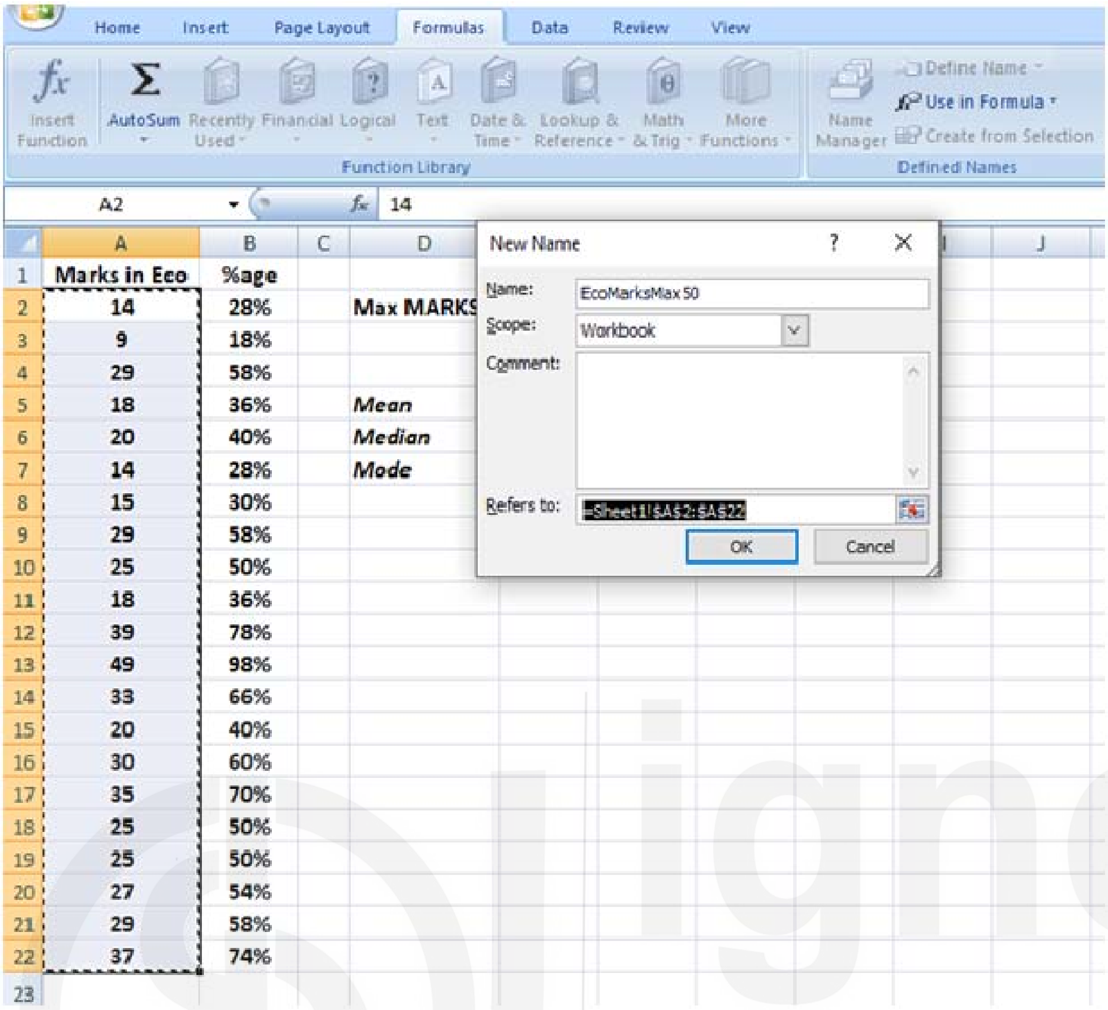
- ○ सैल्स् को फॉरमैट करने का एक डॉयलाग बॉक्स खुल जाएगा जिसमें अनेक विकल्प होंगे। इन विकल्पों में से विकल्प "percentage" चुनिये। 'OK' बटन पर क्लिक कीजिए। आप दशमलव के बाद के अंकों को 0 भी चुन सकते हैं।

- माध्य, माध्यिका और बहुलक ज्ञात करने के लिए, पहले उस सैल रेंज के लिए एक नाम तय कीजिए जिसमें विद्यार्थियों द्वारा प्राप्तांक प्रविष्ट किए गए हैं। यह रेंज (A2 से A22) है। यह नीचे दिए चरणों के माध्यम से किया जा सकता है :

  - ○ सैल A2 से सैल A22 तक की रेंज को चुनिये। ऐसा करने के लिए माउस को सैल A2 पर ले जाइये। माउस के बाएं बटन पर क्लिक कीजिए तथा बटन को दबाए रखते हुए माउस को नीचे सैल A22 तक ले जाइये (ड्रैग कीजिए)। ऐसा करने पर यह रेंज हाईलाइट हो जाएगी (चित्र 2.4.4.1 देखिये)।

  - ○ टैब 'Formulas' को चुनिये और इस प्रकार खुलने वाले रिक्त पर उपलब्ध विकल्पों में से विकल्प 'Define Name' चुनिए। ऐसा करने पर एक New Name के नाम से डायलॉग बॉक्स खुल जाएगा। इस डायलॉग बॉक्स में एक नाम EcoMarksMax50 प्रविष्ट कीजिए (चित्र 2.4.4.1 देखिये)।

  - ○ 'New Name' डायलॉग बॉक्स में OK बटन पर क्लिक कीजिए। अब आपने रेंज A2..A22 के लिए एक नाम EcoMarksMax50 निश्चित कर दिया है। यह रेंज 'New Name' डॉयलॉग बाक्स में *Refers to =Sheet1!$A$2:$A$22*in के रूप में दर्शायी जा रही है क्योंकि यह absolute addressing के नियम का पालन करती है।

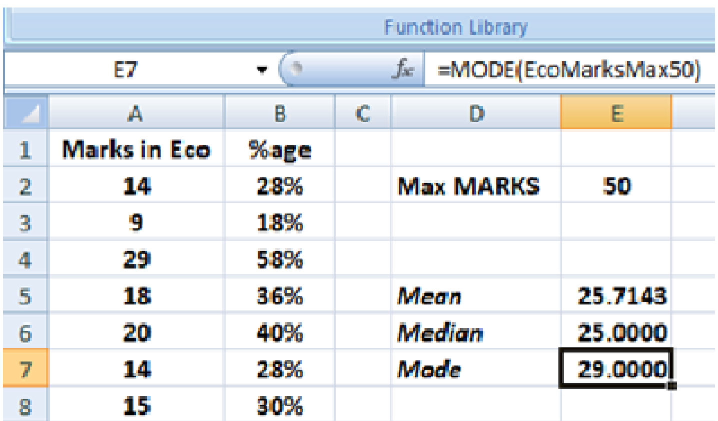


चित्र 2.4.4.1 : किसी रेंज का नाम परिभाषित करना।

अब आप लेबल - *Max MARKS, Mean, Median* तथा *Mode* प्रविष्ट कर सकते हैं जैसे ऊपर दिए चित्र में दर्शाया गया है। हमने maximum marks value (50) को सेल E2 में प्रविष्ट/दर्ज किया है।

अब माध्य के लिए सूत्र **AVERAGE(EcoMarksMax50)**in सैल ***E5*** में; माध्यिका के लिए सूत्र **=MEDIAN(EcoMarksMax50)**in सैल ***E6*** में; तथा बहुलक के लिए सूत्र **Mode(EcoMarksMax50)**in सैल *E7* में प्रविष्ट/दर्ज कीजिए।

आपको कंप्यूटर स्क्रीन पर निम्नलिखित परिकलित मान दिखेंगे +



बोध प्रश्न 2

1) पद 'प्रतिशत' को परिभाषित कीजिए तथा इसके परिकलन तथा प्रयोग में बरती जाने वाली सावधानियों को सूचीबद्ध कीजिए।

....

....

....

....

....

2) किसी एक विचर डेटा श्रेणी की केंद्रीय प्रवृत्ति के माप कौन-कौन से होते है?

....

....

....

....

....

3) उन विशिष्ट स्थितियों की व्याख्या कीजिए जिनमें केंद्रीय प्रवृत्ति के किसी विशिष्ट माप का प्रयोग किया जा सकता है।

....

....

....

....

....

2.5 परिमाणात्मक डेटा – विचलन के माप

केंद्रीय प्रवृत्ति में माप किसी आबंटन के केंद्र की स्थिति ज्ञात करने में सहायक होते हैं, परंतु उनसे यह पता नहीं चलता कि आबंटन में विभिन्न प्रविष्टियाँ केंद्र के दोनों ओर किस प्रकार फैली हुई हैं। किसी बारंबारता आबंटन की इस विशेषता को सामान्यतः विचलन कहते हैं। एक श्रेणी में सभी प्रविष्टियाँ समान नहीं होतीं। मानों में अंतर अथवा विविधता होती है। इस विविधता का परिमाण विचलन के विभिन्न मापों द्वारा ज्ञात किया जाता है। कम विचलन प्रविष्टियों में उच्च एकरूपता को दर्शाता है जबकि अधिक विचलन निम्न स्तर की एकरूपता की ओर इशारा करता है। विचलन का एक आदर्श माप कठोरता से परिभाषित होना चाहिए, सभी प्रविष्टियों पर आधारित होना चाहिए तथा समझने तथा परिकलित करने में सरल होना चाहिए। साथ ही, यह अत्यधिक छोटे या बड़े मानों से प्रभावित नहीं होना चाहिए और इस पर बीजगणितीय संक्रियाएँ संभव होनी चाहिए।

नीचे विचलन के विभिन्न निरपेक्ष तथा तुलनात्मक माप सूचीबद्ध किए गए हैं :

निरपेक्ष माप (ये प्रेक्षणों की इकाइयों के पदों में परिवर्तन के माप को दर्शाते हैं)	तुलनात्मक माप (ये इकाइयों से स्वतंत्र होते हैं)
परिसर	परिसर का गुणांक
माध्य विचलन	माध्य विचलन का गुणांक
मानक विचलन	प्रसरण का गुणांक

2.5.1 परिसर तथा परिसर का गुणांक

परिसर विचलन का सरलतम संभव माप है तथा इसे चर के सबसे बड़े और सबसे छोटे मान के अंतर के रूप में परिभाषित किया जाता है। प्रतीकात्मक रूप में इसे परिसर (Range) $= L - S$ के रूप में व्यक्त किया जाता है जहाँ L = सबसे बड़ा मान तथा S = सबसे छोटा मान है। व्यक्तिगत प्रेक्षणों तथा असंतत श्रेणियों में L तथा S को आसानी से ज्ञात किया जा सकता है। संतत श्रेणियों में या तो L को सबसे ऊँचे वर्ग की उपरिसीमा तथा S को सबसे निचले वर्ग की निचली सीमा लिया जाता है या फिर L को सबसे ऊंचे वर्ग का मध्य-बिंदु तथा S को सबसे निचले वर्ग का मध्य-बिंदु लिया जाता है।

परिसर का गुणांक :

परिसर का गुणांक $= \frac{L-S}{L+S}$ होता है।

2.5.2 माध्य विचलन तथा माध्य विचलन का गुणांक

किसी श्रेणी का मानक विचलन, श्रेणी की केंद्रीय प्रवृत्ति के किसी भी माप जैसे कि माध्य, माध्यिका या बहुलक से श्रेणी की विभिन्न प्रतिविष्टियों के विचलन का समांतर माध्य होता है। इसमें सभी विचलन धनात्मक लिए जाते हैं अर्थात् उनके चिन्हों को छोड़ दिया जाता है। क्लार्क और शेकाडे के अनुसार, ''औसत विचलन, एक आबंटन की प्रविष्टियों के माध्य या माध्यिका से बिखराव का औसत परिमाण होता है, जिसमें विचलनों के चिन्हों (+ या –) पर ध्यान नहीं दिया जाता''। हम सामान्यतः माध्य विचलन को तीन औसतों अर्थात् माध्य, माध्यिका या बहुलक में से किसी की भी सहायता से ज्ञात कर सकते हैं परंतु क्योंकि बहुलक कई बार स्पष्ट रूप से परिभाषित नहीं होता, इसलिए माध्य विचलन का परिकलन माध्य या माध्यिका से ही किया जाता है। माध्य और माध्यिका में भी माध्यिका के प्रयोग को अधिक वरीयता दी जाती है। परंतु सामान्य अभ्यास में, माध्य के अधिक प्रचलित अनुप्रयोगों के कारण, माध्य विचलन अधिकतर माध्य से ही ज्ञात किया जाता है। माध्य विचलन को सामान्यतः एम.डी. (MD) से व्यक्त किया जाता है।

व्यक्तिगत प्रेक्षणों की एक श्रेणी के लिए माध्य विचलन ज्ञात करने की विधि के विभिन्न चरण नीचे दिए हैं :

1) श्रेणी का कोई औसत माध्य, माध्यिका या बहुलक ज्ञात कीजिए।

2) इस औसत से श्रेणी की विभिन्न प्रविष्टियों का विचलन उनके चिन्हें (+ या –) को छोड़ते हुए ज्ञात कीजिए तथा इन विचलनों को $|D|$ से व्यक्त कीजिए।

3) इन सभी विचलनों का योग ज्ञात कीजिए अर्थात् $\sum|D|$ ज्ञात कीजिए।

4) इस योग को प्रविष्टियों की संख्या से विभाजित कीजिए। संकेतात्मक रूप में इसे इस प्रकार व्यक्त किया जा सकता है :

माध्य विचलन (M.D.) $= \frac{\sum|D|}{n}$

वर्गीकृत डेटा की संतत या असंतत श्रेणी के लिए, प्रत्येक मान को उसके संगत बारंबारता से गुणा कीजिए तथा योग $\sum f|D|$ ज्ञात करके, उसे n से भाग कीजिए। इस प्रकार हमें MD का मान प्राप्त हो जाएगा।

माध्य विचलन का गुणांक : केंद्रीय प्रवृत्ति के किसी भी माप से ज्ञात किया गया माध्य विचलन, एक निरपेक्ष माप है। विभिन्न श्रेणियों की माध्य परिवर्तनशीलता की तुलना करने के लिए, एक तुलनात्मक माध्य विचलन की आवश्यकता होती है। माध्य विचलन को, उसमें प्रयोग किए गए औसत के मान से विभाजित करके तुलनात्मक/सापेक्ष माध्य विचलन ज्ञात किया जाता है। अर्थात्

माध्य विचलन का गुणांक = माध्य विचलन/माध्य या माध्यिका या बहुलक

यदि परिणाम प्रतिशत में चाहिए तो माध्य विचलन का गुणांक = माध्य विचलन/माध्य या माध्यिका या बहुलक × 100

2.5.3 मानक विचलन तथा प्रसरण

कार्ल पीयरसन् ने सन् 1893 में मानक विचलन की संकल्पना दी। यह परिक्षेपण (dispersion) का सबसे महत्त्वपूर्ण माप है तथा इसका प्रयोग सांख्यिकी के अनेक सूत्रों में किया जाता है। मानक विचलन को वर्गित विचलनों के योग का वर्गमूल भी कहते हैं। इसका कारण यह है कि यह समांतर माध्य से विचलन के वर्गों के योग के माध्य के बराबर होता है। यह परिशुद्ध परिणाम देता है। समष्टि के एक प्राचल के रूप में, मानक विचलन को ग्रीक अक्षर σ (सिगमा) से व्यक्त किया जाता है। मानक विचलन के वर्ग को प्रसरण कहते हैं।

मानक विचलन का परिकलन करने के लिए आवश्यक चरण नीचे दिए हैं :

1. श्रेणी का वास्तविक माध्य $(\bar{X})$ ज्ञात कीजिए।
2. प्रत्येक प्रविष्टि का माध्य से विचलन $(x = X - \bar{X})$ ज्ञात कीजिए।
3. इन विचलनों का वर्ग लें तथा इस प्रकार प्राप्त वर्गों का योग $\sum x^2$ ज्ञात करें।
4. इस योग $(\sum x^2)$ को प्रेक्षणों की कुल संख्या (n) से भाग करके $\left[\frac{\sum x^2}{n}\right]$ प्राप्त करें।

   $\left[\frac{\sum x^2}{n}\right]$ का वर्गमूल मानक विचलन है।

अतः, $\sigma = \sqrt{\left[\frac{\sum x^2}{n}\right]}$ या $\sqrt{\frac{\sum(x-\bar{x})^2}{n}}$

वर्गीकृत डेटा के लिए सबसे पहले एक तालिका बनाएं जैसे कि नीचे दर्शाया गया है तथा उसके अनुरूप सभी कॉलमों को भरें :

	A	B	C	D	E	F
	वर्ग	बरंबारता (f)	मध्य-बिंदु X_m	$d = X - X_m$ कल्पित माध्य	f.d	$f.d^2$

भिन्नात्मक संख्याओं से विचलन ज्ञात करना अत्यंत कठिन काम है। अतः, समय और श्रम बचाने के लिए हम शार्टकट विधि का उपयोग कर सकते हैं जिसमें विचलन किसी कल्पित माध्य से लिए जा सकते हैं। इसके लिए सूत्र होगा :

$$\sigma = \sqrt{\frac{\sum d^2}{N} - \left(\frac{\sum d}{N}\right)^2}$$

यहाँ d = X–A तथा A कल्पित माध्य है। अर्थात् आबंटन में से किसी भी प्रविष्टि/मान को कल्पित माध्य के रूप में चुना जा सकता है। वर्गीकृत डेटा में, प्रत्येक d को संगत बारंबारता से गुणा करने की आवश्यकता होगी। अर्थात्

$$\sigma = \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2}$$

प्रतिदर्श प्रसरण तथा मानक विचलन (Sample Variance and Standard Deviation)

किसी प्रतिदर्श का प्रसरण ज्ञात करते हुए, सामान्यतः उपरोक्त सूत्र का प्रयोग नहीं किया जाता क्योंकि अधिकतर स्थिति में सांख्यिक (statistic) ज्ञात करने के पीछे हमारा उद्देश्य, संगत प्राचल ज्ञात करना होता है। उदाहरण के लिए, प्रतिदर्श माध्य का प्रयोग समष्टि माध्य के अनुमान के रूप में किया जाता है। उपरोक्त सूत्र समष्टि प्रसरण का सर्वोत्तम अनुमान नहीं देता क्योंकि जब समष्टि बड़ी है और प्रतिदर्श छोटा (जो कि साधारणतया 30 से छोटा होता है), इस सूत्र से ज्ञात प्रसरण सामान्यतः समष्टि प्रसरण को कम आंकता है। इसलिए, n से भाग करने की बजाय, प्रतिदर्श का प्रसरण ज्ञात करने के लिए हम n – 1 से भाग करते हैं जिससे हम थोड़ा बड़ा मान तथा समष्टि प्रसरण का एक निष्पक्ष अनुमान प्राप्त होता है।

प्रतिदर्श प्रसरण को s^2 से व्यक्त किया जाता है तथा इसका सूत्र है

$$s^2 = \frac{\sum(X - \bar{X})^2}{n-1}$$

जहाँ,

$\bar{X}$ = प्रतिदर्श माध्य तथा

n = प्रतिदर्श का आकार है।

मानक विचलन तथा प्रसरण के बारे में कुछ बिंदुओं पर ध्यान देना आवश्यक है :

- प्रथम, प्रसरण के वर्गमूल से हमें मानक विचलन प्राप्त होता है तथा मानक विचलन का वर्ग लेने पर हमें प्रसरण प्राप्त होता हैं।
- दूसरे, प्रसरण प्रत्येक मान के माध्य से दूरी के वर्ग का वास्तविक औसत है। अतः, यदि विभिन्न मान माध्य के पास होंगे तो प्रसरण छोटा होगा। इसके विपरीत, यदि प्रविष्टियों के मान माध्य से दूर होंगे तो, प्रसरण बड़ा होगा।
- तीसरे, प्रसरण ज्ञात करने के लिए प्रविष्टियों और माध्य के बीच की दूरी न लेकर इन दूरियों का वर्ग लेने के पीछे मुख्य कारण यह है कि प्रविष्टियों के मानों के माध्य से वास्तविक विचलन का मान सदा शून्य होता है। जब हम इन विचलनों का वर्ग ले लेते हैं तो ऋणात्मक चिन्ह समाप्त हो जाते हैं।
- चौथे, वर्गमूल लेना क्यों आवश्यक है? वर्गमूल लेने का कारण यह है कि क्योंकि हमने दूरियों/विचलनों का वर्ग ले लिया था, इस प्रकार प्राप्त संख्याओं की इकाइयों, मूल डेटा की इकाइयों के वर्ग हैं। प्रसरण का वर्गमूल लेने से मानक विचलन उन्हीं इकाइयों में आ जाता है जिसमें मूल डेटा था।

- अंततः, जब हम वर्गमूल ज्ञात करते हैं, तो हमेशा इसका धनात्मक मान ही प्रयोग किया जाता है क्योंकि किसी डेटा समुच्चय के प्रसरण तथा मानक विचलन कभी ऋणात्मक नहीं हो सकते।

2.5.4 प्रसरण का गुणांक

मानक विचलन परिक्षेपण (dispersion) का निरपेक्ष मान है। यह इन्हीं इकाइयों के पदों में व्यक्त होता है जिसमें हमारे प्रेक्षण/संख्याएँ संग्रहित तथा व्यक्त होती हैं। विद्यार्थियों की लंबाई/ऊंचाई से मानक विचलन की तुलना विद्यार्थियों के भार के मानक विचलन से नहीं की जा सकती क्योंकि दोनों अलग-अलग इकाइयों में व्यक्त किए जाते हैं; लंबाई सेंटीमीटर में तथा भार किलोग्राम में व्यक्त किया जाता है। अतः यदि हमारा उद्देश्य तुलना करना है तो मानक विचलन को परिक्षेपण के एक तुलनात्मक माप में परिवर्तित करना होगा। यह तुलनात्मक माप प्रसरण गुणांक कहलाता है। प्रसरण गुणांक मानक विचलन को माध्य से भाग तथा 100 से गुणा करके प्राप्त होता हैं प्रतीकात्मक रूप में, इसका सूत्र है :

$$\text{प्रसरण गुणांक (C.V.)} = \frac{\sigma}{\bar{x}} \times 100$$

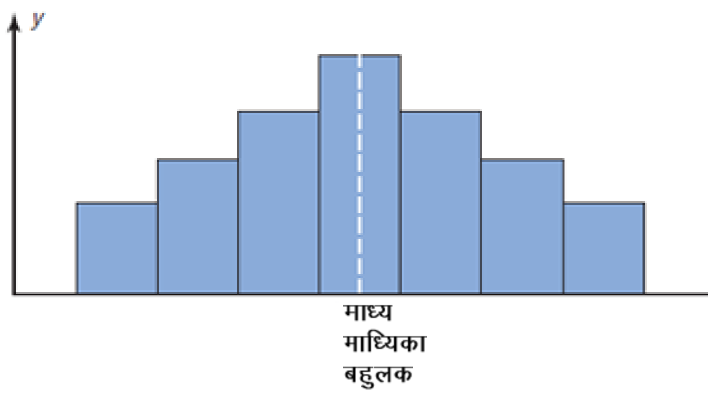
यदि हम दो या दो से अधिक श्रेणियों की परिवर्तनशीलता की तुलना करना चाहें तो हम प्रसरण गुणांक (C.V.) का प्रयोग कर सकते हैं। डेटा की जिस श्रेणी के C.V. का मान बड़ा होगा वह श्रेणी अधिक परिवर्तनशील, कम स्थायी, कम एकरूप, कम संगत या कम समरूप/समांगी होगी। यदि किसी श्रेणी का C.V. छोटा होगा तो वह श्रेणी कम परिवर्तनशील, अधिक स्थायी, अधिक एकरूप, अधिक संगत या अधिक समरूप/समांगी होगी।

2.5.5 बंटनों के आकार (Distribution Shapes)

वैषम्य तथा ककुदता

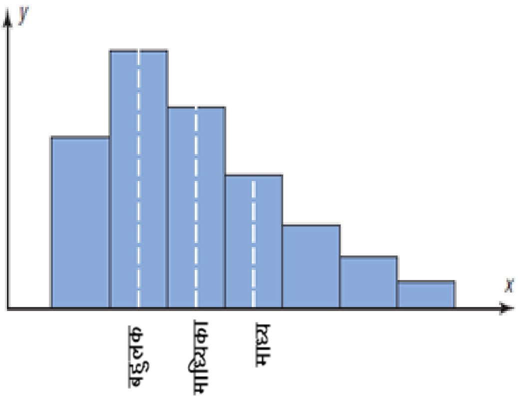
वैषम्य का अर्थ है 'सममिति की कमी'। हम वैषम्य का अध्ययन इसलिए करते हैं कि दिए हुए डेटा की सहायता से जो वक्र हम प्राप्त करते हैं, हमें उनके आकार का अनुमान हो जाए। यदि किसी *आबंटन* में माध्य = माध्यिका = बहुलक हो, तो वह *आबंटन* एक **सममित आबंटन** कहलाता है। यदि किसी आबंटन में माध्य, माध्यिका तथा बहुलक समान न हो तो वह आबंटन सममित नहीं होता। उसे विषम बंटन कहते हैं तथा ऐसा बंटन धनात्मक रूप से विषम अथवा ऋणात्मक रूप से विषम हो सकता है।

सममित बंटन (Symmetrical distribution)

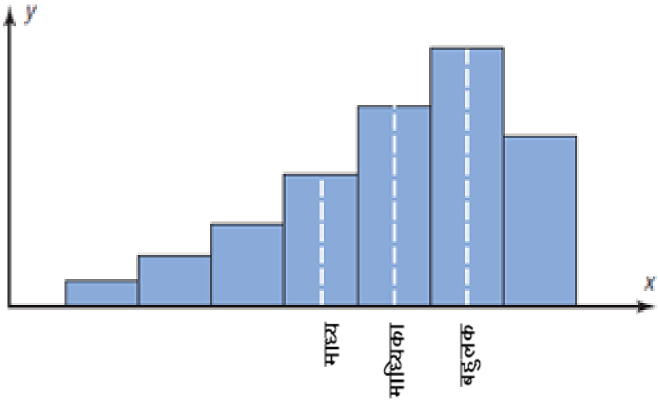


ऊपर दिए चित्र से यह स्पष्ट है कि एक सममित बंटन में माध्य, माध्यिका तथा बहुलक समान/संपाती होते हैं तथा इस केंद्रीय बिंदु के दोनों ओर बारंबारताओं का फैलाव एक जैसा होता है।

विषम बंटन



(क) धनात्मक रूप से विषम या दाईं ओर विषम



(ख) ऋणात्मक रूप से विषम या बाईं ओर विषम

ऊपर दिए चित्रों से यह स्पष्ट होता है कि धनात्मक एवं ऋणात्मक विषम बंटनों में, बहुलक का मान सबसे अधिक तथा माध्य का मान सबसे कम होता है, माध्यिका इन दोनों के मध्य में स्थित होता है। एक धनात्मक रूप से विषम बंटन में, बारंबारताओं का फैलाव बाईं ओर की अपेक्षा, दाईं ओर के मानों के अधिक बड़े परिसर में होता है। इसी प्रकार, एक ऋणात्मक रूप से विषम बंटन में, बारंबारताओं का फैलाव दाईं ओर की अपेक्षा, बाईं ओर के मानों के अधिक बड़े परिसर में होता है।

वैषम्य के माप :

कार्ल पीयरसन् का वैषम्य गुणांक तथा बाऊले का वैषम्य गुणांक, वैषम्य के दो महत्त्वपूर्ण माप हैं।

कार्ल पीयरसन् के अनुसार, *वैषम्य का निरपेक्ष मान = माध्य – बहुलक* होता है। परंतु यह माप दो या दो से अधिक बंटनों में वैषम्य की तुलना करने के लिए उपयुक्त नहीं है क्योंकि अलग-अलग श्रेणियों में माप की इकाइयाँ अलग-अलग हो सकती हैं। इस दुविधा से बचने के लिए हम कार्ल पीयरसन द्वारा दिए गए तुलनात्मक माप वैषम्य गुणांक का प्रयोग करते हैं जिसका सूत्र इस प्रकार है :

कार्ल पीयरसन् का वैषम्य गुणांक = माध्य – बहुलक / मानक विचलन

यदि बहुलक स्पष्ट रूप से परिभाषित न हो, तो यह गुणांक निम्न सूत्र द्वारा ज्ञात किया जा सकता है :

वैषम्य का गुणाक = 3(माध्य–माध्यिका) / मानक विचलन

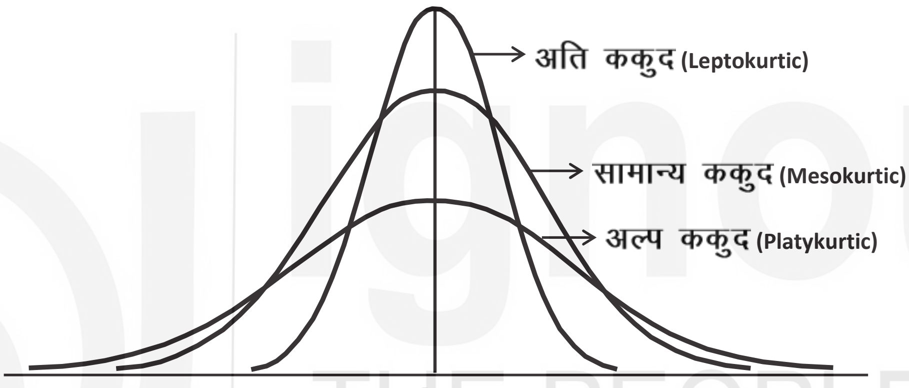
कार्ल पीयरसन् द्वारा वैषम्य के माप की विधि में पूरी श्रेणी का उपयोग करना पड़ता है। प्रो. बाऊले ने वैषम्य गुणांक को ज्ञात करने की एक विधि सुझाई जो चतुर्थकों की तुलनात्मक/सापेक्ष स्थितियों पर आधारित है। एक सममित बंटन में चतुर्थक, माध्यिका के मान से समान दूरी पर होते हैं, अर्थात् माध्यिका – $Q_1 = Q_3$ – माध्यिका होता है।

इस आधार पर बाऊले ने निम्नलिखित सूत्र दिया :

बाऊले का वैषम्य गुणांक (sk) = $Q_3 + Q_1 \; - 2$ माध्यिका $\;/ Q_3 - \; Q_1$

ककुदता

ककुदता आबंटन के आकार का एक अन्य मापक है। वैषम्य ने तो किसी केंद्रीय प्रवृत्ति के मापक के गिर्द सममिति के अभाव को मापा था। ककुदता किसी आबंटन की प्रसामान्य आबंटन के सापेक्ष 'शिखरता' या शिखरता के अभाव का मापन करती है। दूसरे शब्दों में, ककुदता वह प्राचल है जो किसी बारंबारता आबंटन के वक्र के शिखर की उच्चता या उसमें उच्चता के अभाव की कोटि को दर्शाता है। विभिन्न बारंबारता वक्रों को शिखरों के आकार के अनुसार तीन वर्गों में बांटा जा सकता है। ये तीन आकार क्रमशः अति ककुद, सामान्य तथा अल्प ककुद होते हैं। हम सामान्य ककुद वक्र द्वारा प्रसामान्य आबंटन को अभिव्यक्त करते हैं। निम्न चित्र में तीनों प्रकार की ककुदता वाले वक्र दिखाए गए हैं।



कार्ल पीयरसन ने किसी आबंटन की ककुदता की कोटि को इस प्रकार बताया है: $\eta = \beta_2 - 3$ । इस पद में $\beta_2 = \dfrac{\frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)^4}{\left[\frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)^2\right]^2}$, जहाँ N समष्टि का आकार है (प्रतिदर्श के लिए हम उसके आकार को 'n' द्वारा दर्शाते हैं); μ समष्टि के माध्यम को दिखाता है (प्रतिदर्श माध्य $\bar{X}$ द्वारा दिखाया जाता है)। उपर्युक्त सूत्र के आधार पर तीन प्रकार के वक्रों के आकार स्पष्ट होंगे जिनकी परिभाषा इस प्रकार है :

1. सामान्य ककुद वक्र : ऐसा वक्र जो न तो अतिशिखरता दिखाता है और न ही चपटापन, जैसा कि प्रसामान्य आबंटन में होता है। ऐसे वक्र में $\beta_2 = 3$ । अतः, $\eta = 0$ ।

2. अति ककुद वक्र : यह वक्र प्रसामान्य से अधिक उच्च शिखरता दर्शाता है– यानी प्रसामान्य की अपेक्षा अधिक ऊँचाई तक जाता है। ऐसे वक्र में $\beta_2 > 3$ और इसीलिए $\eta > 0$ ।

3. अल्प ककुद वक्र : यह वक्र प्रसामान्य वक्र से कुछ चपटी होती है। इस प्रकार के वक्र में $\beta_2 < 3$ और स्वभावित रूप से $\eta < 0$ ।

बोध प्रश्न 3

1) पदों 'मानक विचलन' तथा 'प्रसरण' की व्याख्या कीजिए तथा उनके परस्पर संबंध तथा उपयोगों पर प्रकाश डालिये।

....

....

....

....

2) परिक्षेपण का कौन-सा माप सबसे अधिक विश्वसनीयता सुनिश्चित करता है तथा क्यों?

....

....

....

....

3) किसी ऋणात्मक रूप से विषम बंटन के लिए माध्य, बहुलक, माध्यिका तथा चतुर्थकों में क्या संबंध है?

....

....

....

....

2.6 परिमाणात्मक डेटा : अवस्थिति के माप

केंद्रीय प्रवृत्ति के मापों तथा परिवर्तनशीलता के मापों के अतिरिक्त, हमें स्थिति के मापों की आवश्यकता पड़ती है। इन मापों में स्टैंडर्ड स्कोर, शतमक, दशमक तथा चतुर्थक सम्मिलित हैं। इनका प्रयोग किसी डेटा सेट में अन्य मानों के सापेक्ष, किसी डेटा मान की स्थिति नियत करने के लिए किया जाता है। उदाहरण के लिए, यदि कोई मान 80वें शतमक पर स्थित है तो इसका अर्थ है कि 80% प्रविष्टियों का मान बंटन में उससे नीचे है तथा 20% ऊपर। माध्यिका वह मान है जो 50वें शतमक के संगत होता है, क्योंकि आधे मान इससे नीचे होते हैं तथा आधे इसके ऊपर। यह भाग स्थिति के इन मापों पर केंद्रित है।

2.6.1 मानक मान / Z- मान (Standard Scores)

मान लीजिये एक विद्यार्थी को संगीत की परीक्षा में 90 अंक मिलें तथा अंग्रेज़ी की परीक्षा में 45। इन अंकों की सीधी तुलना असंभव है, क्योंकि यह हो सकता है कि दोनों परीक्षाएँ समतुल्य न हों। उनमें प्रश्नों की संख्या अलग हो सकती है, प्रत्येक प्रश्न का मान अलग हो सकता है, इत्यादि। फिर भी, यदि हमारे पास दोनों परीक्षाओं के लिए एक जैसा सापेक्ष मानक हो तो यह तुलना की जा सकती है। यह तुलना माध्य तथा मानक विचलन का प्रयोग करती है जिसे मानक मान/अंक या z मान

कहते हैं। एक मानक मान या z मान हमें यह बताता है कि किसी दिए हुए मानों के विशिष्ट बंटन में कोई डेटा मान/प्रविष्टि माध्य से कितने मानक विचलन गुणा ऊपर या नीचे है। यदि एक मानक मान शून्य है, तो डेटा का मान माध्य के बराबर होगा। यह ध्यान देने योग्य है कि यदि z मान धनात्मक है, तो यह मान माध्य से ऊपर होगा; यदि z मान 0 है तो मान माध्य के बराबर होगा और यदि z मान ऋणात्मक है तो मान माध्य से नीचे होगा।

किसी मान के लिए z मान उस मान में से माध्य को घटाकर, प्राप्त परिणाम को मानक विचलन से विभाजित करके प्राप्त किया जाता है। एक मानक मान के लिए प्रतीक z का उपयोग किया जाता है। इसका सूत्र है :

$$z = (\text{मान} - \text{माध्य}) / \text{ मानक विचलन}$$

प्रतिदर्शों के लिए, यह सूत्र है

$$\mathrm{z} = \frac{X - \bar{X}}{s}$$

समष्टि के लिए, यह सूत्र है

$$z = \frac{X - \mu}{\sigma}$$

z मान यह निरूपित करता है कि कोई मान माध्य से कितने मानक विचलन गुणा ऊपर या नीचे है।

उदाहरण :

एक विद्यार्थी ने कैलकुलस की एक परीक्षा में, जिसका माध्य 50 तथा मानक विचलन 10 है, 65 अंक प्राप्त किये। उसने इतिहास की एक परीक्षा में, जिसका माध्य 25 तथा मानक विचलन 5 है, 30 अंक प्राप्त किए। इस विद्यार्थी की दोनों परीक्षाओं में सापेक्ष स्थितियों की तुलना कीजिए।

हल :

पहले हम दोनों परीक्षाओं में विद्यार्थी द्वारा प्राप्तांकों के z मान ज्ञात करते हैं। कैलकुलस के लिए z मान है

$$z = \frac{X - \bar{X}}{s} = \frac{65 - 50}{10} = 1.5$$

इतिहास के लिए z मान है

$$z = \frac{30 - 25}{5} = 1.0$$

क्योंकि कैलकुलस का z मान बड़ा है, विद्यार्थी की कैलकुलस में सापेक्ष स्थिति, इतिहास की सापेक्ष स्थिति से ऊपर है।

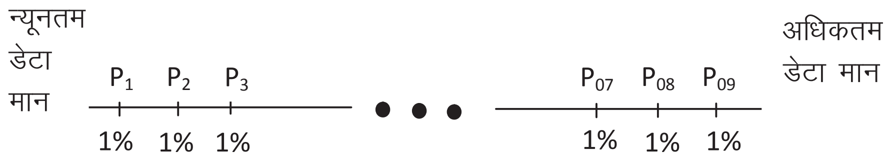
2.6.2 शतमक

शतमक शिक्षा तथा स्वास्थ्य-संबंधी क्षेत्रों में किसी व्यक्ति की एक ग्रुप में स्थिति को इंगित करने वाली स्थिति माप है। शतमक किसी डेटा समतुल्य को 100 बराबर वर्गों में बांटते हैं।

विभिन्न शतमक प्रतीकों

$$P_1, P_2, P_3, \ldots\ldots, P_{99}$$

द्वारा व्यक्त किए जाते हैं तथा बंटन को 100 वर्गों में विभाजित करते हैं।



शतमक तथा प्रतिशत समान नहीं होते। अर्थात्, यदि एक विद्यार्थी एक परीक्षा के 100 प्रश्नों में से 72 प्रश्नों के ठीक उत्तर देता है, तो उसे 72 प्रतिशत अंक प्राप्त होते हैं। इससे पूरी कक्षा में उसकी स्थिति का अनुमान नहीं लगाया जा सकता। हो सकता है उसने कक्षा में अधिकतम अंक प्राप्त किए हों, या न्यूनतम या न्यूनतम तथा अधिकतम के कहीं बीच में। दूसरी ओर, यदि उसके द्वारा प्राप्त 72% अंक 64वें शतमक के संगत हैं तो इसका अर्थ है कि उसने कक्षा के 64% विद्यार्थियों से बेहतर प्रदर्शन किया।

किसी डेटा के शतमक ज्ञात करने के लिए अनेक विधियाँ उपलब्ध हैं। इन विधियों का उपयोग किसी डेटा मान के संगत अनुमानित शतमक ज्ञात करने के लिए अथवा किसी दिए हुए शतमक के संगत एक डेटा मान ज्ञात करने के लिए किया जा सकता है। यदि डेटा समुच्चय बड़ा हो (100 या उससे अधिक प्रविष्टियों वाला), तो इन विधियों से बेहतर परिणाम प्राप्त होते हैं।

$$\text{शतमक} = \frac{(X \text{ से नीचे वाले मानों की संख्या}) + 0.5)}{\text{कुल मानों की संख्या}} \times 100$$

उदाहरण,

एक शिक्षक 10 विद्यार्थियों को 20 अंकों वाली एक परीक्षा देता है। विद्यार्थियों द्वारा प्राप्तांक नीचे दर्शाए गए हैं। स्कोर 12 के संगत शतमक श्रेणी ज्ञात कीजिए।

18, 15, 12, 6, 8, 2, 3, 5, 20, 10

हम इस डेटा को छोटे से बड़े क्रम में व्यवस्थित करते हैं।

$$\text{शतमक} = \frac{(X \text{ से नीचे वाले मानों की संख्या}) + 0.5)}{\text{कुल मानों की संख्या}} \times 100\%$$

अब नीचे दिए सूत्र का उपयोग करते हैं

$$\text{शतमक} = \frac{(X \text{ से नीचे वाले मानों की संख्या}) + 0.5)}{\text{कुल मानों की संख्या}} \times 100$$

क्योंकि 12 के स्कोर से नीचे 6 प्रविष्टियाँ हैं, हम पाते हैं कि

$$\text{शतमक} = \frac{6+0.5}{10} .100\% = 65\text{वाँ शतमक}$$

अतः, एक विद्यार्थी ने, जिसका मान 12 था, कक्षा के 65% विद्यार्थियों से बेहतर प्रदर्शन किया है।

इसी प्रकार, हम एक दिए हुए शतमक के संगत डेटा मान ज्ञात कर सकते हैं। इसकी विधि नीचे दी गई है।

एक दिए हुए शतमक के संगत एक डेटा मान ज्ञात करना :

चरण 1 : डेटा को छोटे से बड़े क्रम में व्यवस्थित कीजिए।

चरण 2 : सूत्र $c = \frac{n.p}{100}$ में विभिन्न चरों के मान रखिये, जहाँ

n = मानों की कुल संख्या, तथा
p = शतमक है

चरण 3 A : यदि c एक पूर्णांक न हों, तो c से बड़ा निकटतम पूर्णांक ले लें। सबसे छोटे मान से गिनना प्रारंभ करें तथा इस पूर्णांक के संगत संख्या चुन लें।

चरण 3B : यदि c एक पूर्णांक है तो सबसे छोटे मान से प्रारंभ करते हुए वह मान चुनें जो cवें तथा $(c + 1)$ वें मानों के ठीक मध्य में हो।

ऊपर दिए उदाहरण में 25वें शतमक के संगत मान ज्ञात कीजिए।

चरण 1 : हम डेटा को बढ़ते हुए क्रम में व्यवस्थित करते हैं।

2, 3, 5, 6, 8, 10, 12, 15, 18, 20

चरण 2 : अब सूत्र $c = \frac{n.p}{100}$ का मान परिकलित करते हैं,

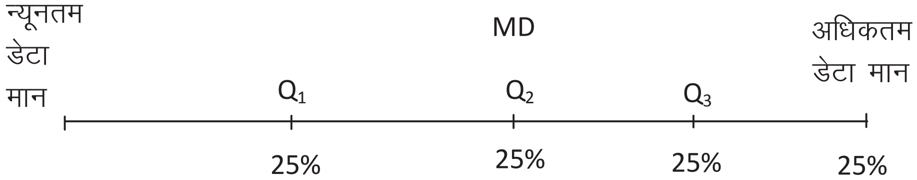
जहाँ
n = मानों की कुल संख्या, तथा
p = शतमक

अतः, $c = \frac{10 \,.\, 25}{100} = 2.5$ है।

चरण 3 : यदि c एक पूर्णांक नहीं है, अतः हम इससे बड़ा निकटतम पूर्णांक लेते हैं। इस प्रकार, हमें $c = 3$ प्राप्त होता है। अतः हम सबसे छोटे मान से प्रारंभ करके तीसरा मान चुनते हैं जो कि 5 है। अतः, मान 5, 25वें शतमक के संगत वांछित मान है।

2.6.3 चतुर्थक तथा दशमक

चतुर्थक किसी बंटन चार वर्गों में विभाजित करते हैं जो कि Q_1, Q_2, Q_3 द्वारा दर्शाए जाते हैं। ध्यान दें कि Q_1 , 25वें शतमक के बराबर है, Q_2, 50वें शतमक या माध्यिका के तथा Q_3, 75वें शतमक के जैसा कि नीचे दिए चित्र में दर्शाया गया है।



चतुर्थक, ऊपर शतमक निकालने के सूत्र का प्रयोग करके ज्ञात किए जा सकते हैं। Q_1 ज्ञात करने के लिए $p = 25$, Q_2 ज्ञात करने के लिए $p = 50$ तथा Q_3 ज्ञात करने के लिए $p = 75$ का प्रयोग करें।

फिर भी, चतुर्थक ज्ञात करने की एक सरल विधि नीचे दी है :

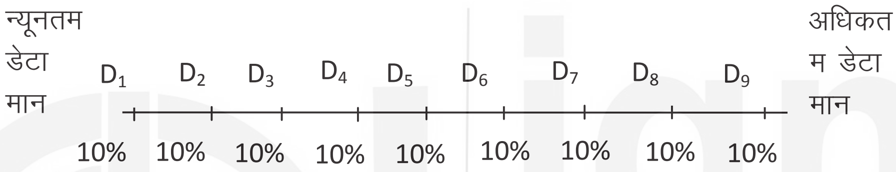
चरण 1 : डेटा को सबसे छोटे मान से लेकर सबसे बड़े मान तक बढ़ते हुए क्रम में व्यवस्थित कर लें।

चरण 2 : डेटा मानों की माध्यिका ज्ञात करें। इससे हमें Q_2 का मान मिल जाएगा।

चरण 3 : Q_2 से नीचे वाले मानों की माध्यिका ज्ञात करें। यह Q_1 का मान होगा।

चरण 4 : Q_2 से ऊपर वाले मानों की माध्यिका ज्ञात करें। यह Q_3 का मान होगा।

दशमक बंटन को 10 वर्गों में विभाजित करते हैं जैसा कि नीचे दिए रेखाचित्र में दर्शाया गया है। इन्हें D_1, D_2 इत्यादि से व्यक्त किया जाता है।



ध्यान दें कि D_1, P_{10} के बराबर है, D_2, P_{20} के, इत्यादि। दशमक भी शतमक ज्ञात करने के सूत्र से सरलता से परिकलित किए जा सकते हैं। शतमकों, दशमकों तथा चतुर्थकों के मध्य संबंधों को इस प्रकार व्यक्त किया जा सकता है :

दशमक $D_1, D_2, D_3. . ..D_9$ द्वारा व्यक्त किए जाते हैं तथा ये क्रमशः $P_{10}, P_{20}, P_{30}, . . . , P_{90}$ के बराबर होते हैं। चतुर्थक Q_1, Q_2, Q_3 द्वारा व्यक्त किए जाते हैं तथा ये P_{25}, P_{50}, P_{75} के बराबर होते हैं। माध्यिका तथा P_{50}, Q_2 और D_5 का मान बराबर होता है।

गतिविधि 2 : डेटा का परिक्षेपण (Dispersion of Data)

गतिविधि 1 में दिए गए डेटा का वैषम्य, ककुदता, प्रसरण, मानक विचलन तथा चतुर्थक ज्ञात कीजिए।

इसे कैसे करें?

अब तक आप यह जान चुके होंगे कि किसी स्प्रैडशीट में सूत्र किस प्रकार प्रविष्ट किया जाता है। चित्र 2.6.1 में गतिविधि 2 के लिए प्राप्त किए गए अंतिम परिणाम दर्शाए गए हैं।

E13 =STDEV(MarksInEcoOutOf50)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Marks in Eco	%age					Z-Score	Percentile	Quartile	Marks			Sorted Marks			
2	14	28%		Max MARKS	50		-1.2043	5%	25%	18	P_{25}		9	Minimum		
3	9	18%					-1.7183	0%	50%	25	P_{50}		14			
4	29	58%					0.3378	60%	75%	30	P_{75}		14			
5	18	36%		*Mean*	25.7143		-0.7931	20%					15			
6	20	40%		*Median*	25.0000		-0.5875	30%		Five Point Summary			18			
7	14	28%		*Mode*	29.0000		-1.2043	5%		*Minimum*	9		18	Quartile 1 at 25%		
8	15	30%		Skewness	0.4429		-1.1015	15%		*Q1*	18		20			
9	29	58%		Kurtosis	0.1916		0.3378	60%		*Median*	25		20			
10	25	50%					-0.0734	40%		*Q3*	30		25			
11	18	36%					-0.7931	20%		*Maximum*	49		25			
12	39	78%					1.3659	95%					25	Median (Quartile 2) at 50%		
13	49	98%		Std Dev	9.7270		2.3939	100%					27			
14	33	66%		Variance	94.6143		0.7490	80%					29			
15	20	40%					-0.5875	30%					29			
16	30	60%					0.4406	75%					29			
17	35	70%					0.9546	85%					30	Quartile 3 at 75%		
18	25	50%					-0.0734	40%					33			
19	25	50%					-0.0734	40%					35			
20	27	54%					0.1322	55%					37			
21	29	58%					0.3378	60%					39			
22	37	74%					1.1602	90%					49	Maximum		

चित्र 2.6.1 : गतिविधि 2 के परिणाम

चित्र 2.6.2(क) तथा (ख) विभिन्न सूत्र दर्शाते हैं जिनकी सहायता से चित्र 2.6.1 प्राप्त किया गया है।

I1 Quartile

	D	E	F	G	H
1				Z-Score	Percentile
2	Max MARI 50			=STANDARDIZE(A2,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A2)
3				=STANDARDIZE(A3,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A3)
4				=STANDARDIZE(A4,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A4)
5	*Mean*	=AVERAGE(MarksInEcoOutOf50)		=STANDARDIZE(A5,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A5)
6	*Median*	=MEDIAN(MarksInEcoOutOf50)		=STANDARDIZE(A6,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A6)
7	*Mode*	=MODE(MarksInEcoOutOf50)		=STANDARDIZE(A7,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A7)
8	Skewness	=SKEW(MarksInEcoOutOf50)		=STANDARDIZE(A8,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A8)
9	Kurtosis	=KURT(MarksInEcoOutOf50)		=STANDARDIZE(A9,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A9)
10				=STANDARDIZE(A10,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A10)
11				=STANDARDIZE(A11,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A11)
12				=STANDARDIZE(A12,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A12)
13	Std Dev	=STDEV(MarksInEcoOutOf50)		=STANDARDIZE(A13,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A13)
14	Variance	=VAR(MarksInEcoOutOf50)		=STANDARDIZE(A14,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A14)
15				=STANDARDIZE(A15,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A15)
16				=STANDARDIZE(A16,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A16)
17				=STANDARDIZE(A17,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A17)
18				=STANDARDIZE(A18,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A18)
19				=STANDARDIZE(A19,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A19)
20				=STANDARDIZE(A20,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A20)
21				=STANDARDIZE(A21,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A21)
22				=STANDARDIZE(A22,$E$5,$E$13)	=PERCENTRANK(MarksInEcoOutOf50,A22)

चित्र 2.6.2(क) : कॉलम E, G तथा H के सूत्र

	I	J	K
1	Quartile	Marks	
2	*0.25*	*=PERCENTILE(MarksInEcoOutOf50,I2)*	P_{25}
3	*0.5*	*=PERCENTILE(MarksInEcoOutOf50,I3)*	P_{50}
4	*0.75*	*=PERCENTILE(MarksInEcoOutOf50,I4)*	P_{75}
5			
6		**Five Point Summary**	
7		*Minimum*	=MIN(MarksInEcoOutOf50)
8		*Q1*	=QUARTILE(MarksInEcoOutOf50,1)
9		*Median*	=QUARTILE(MarksInEcoOutOf50,2)
10		*Q3*	=QUARTILE(MarksInEcoOutOf50,3)
11		*Maximum*	=MAX(MarksInEcoOutOf50)
12			

चित्र 2.6.2(ख) : कॉलम J तथा K के सूत्र

कृपया ध्यान दें कि चित्र 2.6.2(क) तथा (ख) में परिसर/रेंज A2..A22 का नाम *Marks In Economics (out of 50)* है।

कृपया ध्यान दें कि चित्र 2.6.1 में अंक कॉलम M में क्रमबद्ध किए गए हैं जिससे विभिन्न चतुर्थक दर्शाए जा सकते हैं।

बोध प्रश्न 4

1) पद 'अवस्थिति के माप' की व्याख्या कीजिए।

2) अवस्थिति के कौन-से माप का उपयोग व्यापक रूप से किया जाता है और क्यों?

3) अवस्थिति के विभिन्न मापों में अंतर्संबंधों की व्याख्या कीजिए।

2.7 सार-संक्षेप

इस इकाई में डेटा को सारबद्ध करने वाली विभिन्न सांख्यिकीय संकल्पनाओं की चर्चा की गई है। इनमें केंद्रीय प्रवृत्ति के माप, परिवर्तनशीलता या परिक्षेपण के माप तथा स्थिति के माप सम्मिलित हैं। केंद्रीय प्रवृत्ति के सबसे अधिक प्रयोग में आने वाले माप माध्य, माध्यिका तथा बहुलक हैं। परिवर्तनशीलता के सबसे अधिक प्रयोग किए जाने वाले माप परिसर, प्रसरण तथा मानक विचलन हैं। अवस्थिति के सबसे अधिक प्रचलित माप शतमक, चतुर्थक तथा दशमक हैं। इस इकाई में, इस पर भी प्रकाश डाला गया कि डेटा मान, बंटन के केंद्रीय मान के आस-पास किस प्रकार बंटे होते हैं। प्रसरण गुणांक का उपयोग मानक विचलन को माध्य के संबंध में/सापेक्ष परिभाषित करने में किया जाता है। इन विधियों को सामान्यतः पारंपरिक सांख्यिकीय विधियाँ कहा जाता है तथा इनका प्रयोग मुख्यतः डेटा की प्रकृति के बारे में लगाए जाने वाले अनुमानों की पुष्टि करने के लिए किया जाता है। केंद्रीय प्रवृत्ति के मापों तथा परिक्षेपण का उपयोग वास्तविक जीवन से जुड़ी व्यावहारिक समस्याओं को हल करने में व्यापक रूप से किया जाता है।

इस इकाई में, हमने स्प्रैडशीट पैकेज (Excel), के प्रयोग से प्रतिशत, माध्य, बहुलक, माध्यिका, वैषम्य, ककुदता इत्यादि ज्ञात करना भी सीखा है।

2.8 संदर्भ ग्रंथादि

1) Blaikie, N. (2003). *Analyzing quantitative data: From description to explanation*. Sage.

2) Bohrnstedt, G. W., & Knoke, D. (1994). *Statistics for social data analysis*.

3) Bryman, A., & Cramer, D. (1994). *Quantitative data analysis for social scientists* (rev. Taylor & Frances/Routledge.

4) Cramer, D. (2003). *Advanced quantitative data analysis*. McGraw-Hill International.

5) Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating quantitative data*. Prentice Hall.

6) Suen, H. K., &Ary, D. (2014). *Analyzing quantitative behavioral observation data*. Psychology Press.

2.9 बोध प्रश्नों के उत्तर अथवा संकेत

बोध प्रश्न 1

1) भाग 2.2 देखें।

2) भाग 2.3 देखें।

3) भाग 2.3 देखें।

बोध प्रश्न 2

1) उपभाग 2.4.1 देखें।
2) उपभाग 2.4.2 देखें।
3) उपभाग 2.4.4 देखें।

बोध प्रश्न 3

1) उपभाग 2.5.3 देखें।
2) उपभाग 2.5.4 देखें।
3) उपभाग 2.5.5 देखें।

बोध प्रश्न 4

1) उपभाग 2.6 देखें।
2) उपभाग 2.6.1 तथा 2.6.2 देखें।
3) उपभाग 2.6.3 देखें।

इकाई 3 सांख्यिकीय सॉफ्टवेयर का परिचय

संरचना

3.1 विषय प्रवेश

सूचना क्रांति के इस युग में, सूचना आर्थिक विकास के लिए एक प्रमुख संचालक शक्ति है। सूचना का सृजन होता है, विभिन्न दृश्य (Visual Forms) रूपों में प्रस्तुत की जाती है, सूचना के अन्य रूपों के साथ संबद्ध की जाती है तथा आर्थिक तथा अन्य निर्णय लेने संबंधी गतिविधियों में प्रयोग की जाती है।

सूचना किस प्रकार उत्पन्न की जाती है? सूचना की उत्पत्ति में आँकड़े प्रमुख घटक होते हैं। आँकड़ों के विश्लेषण की आवश्यकता होती है और साथ ही उन्हें तालिकाओं तथा आलेखों के सरल रूप में प्रस्तुत करने की आवश्यकता होती है, जिससे उनका उपयोग और अधिक सार्थक रूप से किया जा सके। अतः, हम यह कह सकते हैं कि आँकड़े सूचना का असंसाधित रूप है, जिन्हें संग्रहित करना पड़ता है तथा प्रस्तुत करने के योग्य सूचना बनाने के लिए संसाधित करना पड़ता है।

सांख्यिकीय सॉफ्टवेयर एक ऐसा साधन है जो असंसाधित आँकड़ों को सूचना में परिवर्तित करने में सहायक होता है। इस इकाई में हम पाठकों का परिचय उन मूलभूत विशेषताओं से करवाएंगे जो आँकड़ा विश्लेषण के लिए प्रयोग किए जाने के लिए किसी भी सांख्यिकीय सॉफ्टवेयर में अवश्य होनी चाहिए। क्योंकि हम सॉफ्टवेयर के प्रयोग पर चर्चा करने वाले हैं, हमारे लिये यह आवश्यक हो जाता है कि हम एक उदाहरण लें जिससे किसी भी सांख्यिकीय सॉफ्टवेयर में शामिल की जाने वाली संकल्पनाओं की ठीक से व्याख्या की जा सके।

इस इकाई में हमने इस उद्देश्य की पूर्ति हेतु MS-Excel का चुनाव किया है अर्थात् हम MS-Excel के माध्यम से सांख्यिकीय सॉफ्टवेयर का उपयोग करने का व्यावहारिक ज्ञान प्राप्त करेंगे। परंतु, पाठक इन संकल्पनाओं को सीखने के लिए किसी भी ओपन सोर्स अथवा मालिकाना सॉफ्टवेयर का उपयोग कर सकते हैं। कृपया ध्यान

दें कि हम इस इकाई में Excel के सभी संकल्पनाओं/विशेषताओं पर प्रकाश नहीं डाल पाएंगे बल्कि आँकड़ा संसाधन के लिए आवश्यक केवल कुछ सीमित उपकरणों की ही चर्चा कर पाएंगे।

3.2 सांख्यिकीय सॉफ्टवेयर की आवश्यकता

सांख्यिकीय सॉफ्टवेयर एक ऐसा उपकरण है जिसके द्वारा हम आँकड़ों का सृजन तथा संसाधन करके उन्हें एक सार्थक अर्थपूर्ण रूप में प्रस्तुत कर सकते हैं। किसी भी सांख्यिकी सॉफ्टवेयर में नीचे दी गई मूलभूत कार्य क्षमताएँ/सुविधाएँ सामान्यतः उपलब्ध होती हैं :

- एक ऐसा इंटरफेस/अंतराफलक जिससे आँकड़ों की प्रविष्टि तथा संपादन सरलतापूर्वक तरीके से उपयुक्त शीर्षकों के साथ की जा सके।
- एक सॉफ्टवेयर एप्लीकेशन से दूसरी सॉफ्टवेयर एप्लीकेशन के बीच आँकड़ों के लेन-देन की सरल विधि।
- आँकड़े परिभाषित करने तथा उनके सत्यापन की सुविधा।
- आँकड़ों की गणना के परिकलन की तथा विश्लेषण के लिए आँकड़ों पर sorting तथा filtering जैसी संक्रियाएँ लगाने की सुविधा।
- विभिन्न प्रकार के चार्ट बनाने की सुविधा।
- विभिन्न प्रकार के सांख्यिकीय विश्लेषण करने के लिए उपयुक्त फलनों की सुविधा।

आँकड़ा विश्लेषण के लिए बहुत-से सांख्यिकीय सॉफ्टवेयर उपलब्ध हैं। इनमें (कंप्यूटर की) ऐसी भाषाएँ भी उपलब्ध हैं जिन्हें विशेष तौर पर आँकड़ा विश्लेषण के लिए ही बनाया गया है – उदाहरण के लिए, पाइथन। पाइथन एक ऑब्जेक्ट ओरिएंटिड भाषा है जिसका उपयोग विश्लेषण के लिए, आँकड़े तैयार करने के लिए, आँकड़ा प्रत्यक्षीकरण के लिए तथा आँकड़ा विश्लेषण के लिए किया जा सकता है।

आँकड़ा विश्लेषण के लिए कई ओपन सोर्स सॉफ्टवेयर उपलब्ध हैं। R मुफ़्त में उपलब्ध, एक ऐसा सांख्यिकीय परिवेश है जो सांख्यिकीय विश्लेषण के लिए एक अत्यंत लोकप्रिय साधन/उपकरण बन चुका है। आज PSPP, Greti, MicroOsiris, GNU Octave, Dmelt तथा ऐसे ही कई अन्य ओपन सोर्स सॉफ्टवेयर उपलब्ध हैं जो सांख्यिकीय विश्लेषण के लिए उपयोगी हैं।

इसी प्रकार कई मालिकाना हक वाले सॉफ्टवेयर पैकेज उपलब्ध हैं जो कि सांख्यिकीय विश्लेषण के लिए अत्यंत उपयोगी हैं। जिनमें SPSS, EViews, SAS, MATLAB, MS Excel इत्यादि के अतिरिक्त कई अन्य ऐसे ही सॉफ्टवेयर हैं।

इस इकाई में हम अपनी चर्चा, मुख्यतः MS Excel के आधार पर करेंगे परंतु पाठक Libreoffice Calc के माध्यम से भी इन सुविधाओं (Functionalities) को सीख सकते हैं।

3.3 आँकड़ों का प्रयोग

आँकड़ा विश्लेषण करने के लिए पहला कार्य आँकड़ा सृजन, संपादन/सत्यापन होता है। हम इसकी व्याख्या नीचे दिए भागों में विभिन्न गतिविधियों के द्वारा करेंगे।

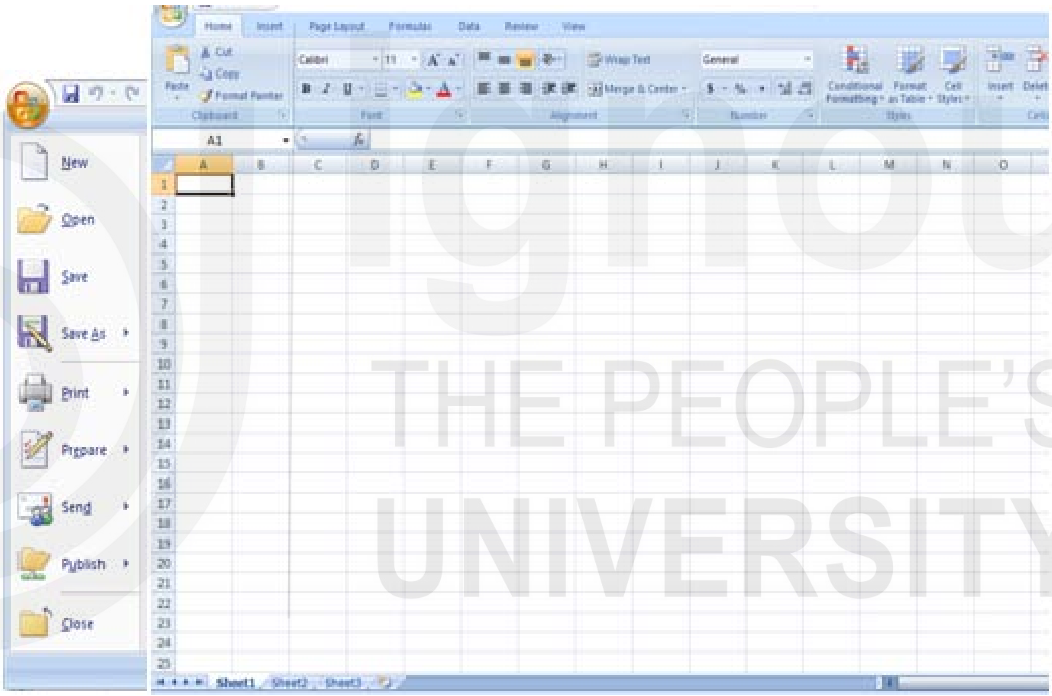
3.3.1 आँकड़ा प्रविष्टि

नीचे दी हुई गतिविधि के माध्यम से हम आँकड़ा प्रविष्टि की सरल प्रक्रिया को समझेंगे।

गतिविधि 1 : 20 विद्यार्थियों द्वारा अर्थशास्त्र विषय में प्राप्त अंकों की प्रविष्टि एक Excel वर्कशीट में करें : 3, 30, 29, 17, 25, 58, 23, 48, 16, 25, 25, 29, 21, 23, 19, 22, 20, 29, 37, 32

यह गतिविधि किस प्रकार करेंगे?

इस गतिविधि को करने के लिए, हम पहले Excel सॉफ्टवेयर को खोलते हैं, स्क्रीन की बायीं ओर के ऊपरी कोने में Office बटन को चुनते हैं तथा तत्पश्चात् विभिन्न विकल्पों में से New विकल्प को चुनते हैं जैसा कि चित्र 3.1(क) में दिखाया गया है। इस प्रकार, हमें एक नई रिक्त वर्कशीट प्राप्त होगी जैसी कि चित्र 3.1(ख) में दिखाई गई है।



(क) नया विकल्प **(ख) एक रिक्त स्प्रेडशीट**

चित्र 3.1 : एक नई वर्कशीट

कृपया ध्यान दें कि Excel में एक पंक्ति तथा एक स्तंभ के काटने से उत्पन्न क्षेत्र को एक सेल या कोष्ठक प्राप्त होता है। यह भी ध्यान दें कि चित्र 3.1(ख) में चयनित सेल A1 है और यह स्तंभ (कॉलम) के नाम के ठीक ऊपर दर्शाया गया है।

अब हम आँकड़ा प्रविष्टि के लिए तैयार हैं। हम स्तंभ A में प्रविष्टियाँ कर सकते हैं। यह बेहतर होगा कि हम सेल A1 के लेबल में अंक (Marks) लिख दें। मान लीजिए कि अधिकतम अंक 50 हैं। यदि हम इस सूचना को स्तंभ B1 तथा C1 में अंकित करना चाहे तो कर सकते हैं जैसे कि दर्शाया गया है (चित्र 3.2 देखें)। हम लेबलों को प्रत्येक सेल के मध्य में कर सकते हैं तथा उन्हें गाढ़े अक्षरों में अंकित कर सकते हैं। पाठक इसे स्वयं करें।

	A	B	C	D
1	Marks	Max Marks	50	
2	3			
3	30			
4	29			
5	17			
6	25			
7	58			
8	23			
9	48			
10	16			
11	25			
12	25			
13	29			
14	21			
15	23			
16	19			
17	22			
18	20			
19	29			
20	37			
21	32			
22				

चित्र 3.2 : एक स्प्रेडशीट में अंकों को अंकित करना

अब हम इस आँकड़ा समूह को सुरक्षित कर सकते हैं। इसके लिए office बटन के मेन्यू विकल्पों में से विकल्प save को चुनना होगा (कृपया चित्र 3.1(क) देखें)। ऐसा करने पर एक डायलॉग बाक्स खुलेगा, जिसमें हम अपनी इस नई फाइल को एक नाम देकर यदि save बटन पर क्लिक करेंगे तो यह फाइल दिए हुए नाम से चुने हुए स्थान पर सुरक्षित (save) हो जाएगी।

ध्यान दें कि सेल A7 में, प्रविष्टि 58 है जो कि अधिकतम अंकों से बड़ी है। क्या पाठक इस गलती को Excel के माध्यम से पकड़ सकते हैं? अगले उपभाग में हम इस प्रक्रिया की व्याख्या करेंगे।

3.3.2 आँकड़ों का सत्यापन करना

हमने ऊपर आँकड़ों की प्रविष्टि करना तथा उसे एक फाइल इकाई 3 – Activty.xls, के नाम से सुरक्षित किया। अब हम यह जाँच करना चाहते हैं कि जो प्रविष्टियाँ हमने कीं वे सब 50 के बराबर या उससे कम हैं। इसके लिए *डेटा टैब* में जाकर, रिबन में 'डेटा टूलस' पर जाते हैं तथा उसमें से आँकड़ों की वैधता (Data validation) चुन सकते हैं, हालांकि हम अभी ही सशर्त स्वरूपण (Conditional Formatting) द्वारा आँकड़ा की वैधता जाँचने की एक सरल विधि के बारे में जानेंगे।

सशर्त स्वरूपण (Conditional Formatting)

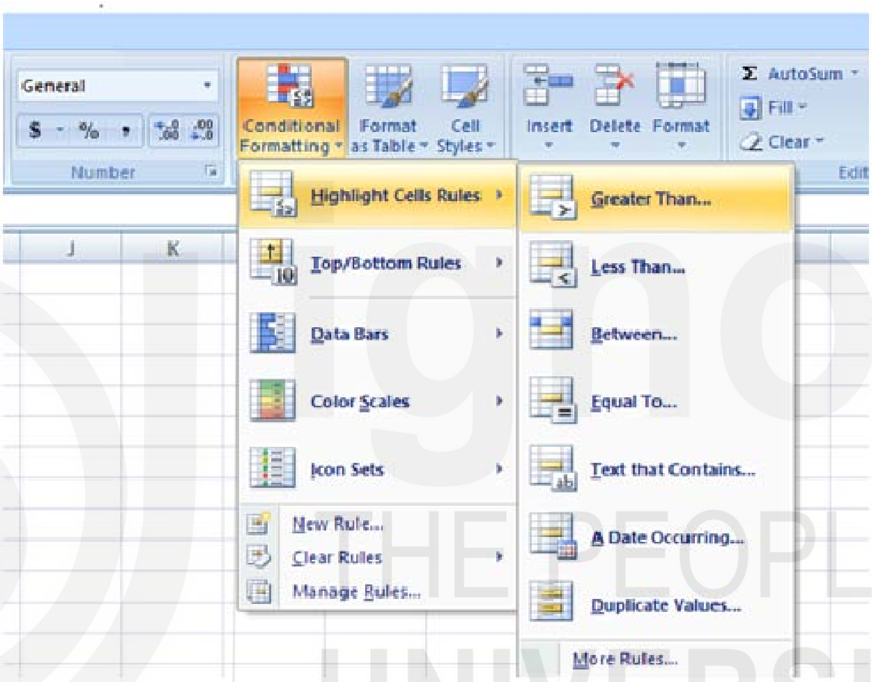
सशर्त स्वरूपण हमें कुछ शर्तों के अनुरूप, जैसे कि किसी मान से बड़े या छोटे मानों वाली प्रविष्टियों को चुनना इत्यादि, आँकड़ों का स्वरूपण करने में हमारी सहायता करता है। यह उन कोष्ठकों को दर्शाने की अत्यंत उपयोगी विधि है जिनके आँकड़े समुच्चय की अन्य प्रविष्टियों से 'अलग' हों।

गतिविधि 2 : गतिविधि 1 में गलत प्रविष्टियों (अधिकतम अंकों से बड़ी प्रविष्टियाँ) को दर्शाने के लिए सशर्त स्वरूपण का प्रयोग कीजिए।

यह गतिविधि कैसे करें?

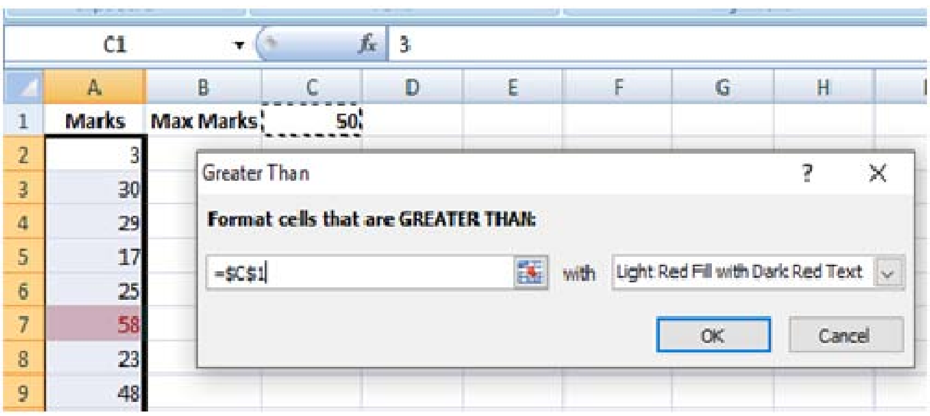
हमारा उद्देश्य यह है कि यदि किसी कोष्ठक में की गई प्रविष्टि का मान अधिकतम अंक मान, जिसे कि सेल C1 में दर्शाया गया है, से अधिक हो तो लाल रंग में प्रदर्शित हो जाए :

सबसे पहले **Home** टैब में जाकर उस सेल रेंज (परिसर) का चयन कीजिए जिसमें कथित नियम प्रारूप लागू करना है। इस उदाहरण में यह रेंज सेल A2 से A21 तक है। इसके पश्चात् **Styles** रिबन में जाकर **Conditional Formatting** चुनें। ऐसा करने पर आपको विकल्पों की एक सूची दिखेगी। इसमें से **Highlight Cell Rules** विकल्प चुनिये। अब विभिन्न शर्तों की एक सूची सामने आ जाएगी, इसमें से **Greater Than** चुनिये (चित्र 3.3 देखें)।



चित्र 3.3 : Creater than rule ('से अधिक है' नियम) के अनुसार सेट को स्पष्ट करना

ऐसा करने पर एक डायलॉग बॉक्स (संवाद कोष्ठ) खुल जाएगा जैसा कि चित्र 3.4 में दर्शाया गया है।



चित्र 3.4 : खुला हुआ डायलॉग बाक्स

कृपया ध्यान दें कि **'Format Cells that are GREATER THAN'** लेबल के नीचे खुले टेक्स्ट बॉक्स (Text Box) में सेल C1 जिसमें अधिकतम अंक लिखे गये हैं' स्वतः चयनित हो गया है। कृपया इस पर भी ध्यान दें कि सेल C1 के चुने जाने को ***=$C$1*** द्वारा दर्शाया गया है। यह MS Excel का एक संपूर्ण/निरपेक्ष संबोधन शैली में एक सूत्र है। इस इकाई में आगे हम इस पर और विस्तार से चर्चा करेंगे। कृपया यह भी ध्यान से देखें कि इसमें वांछित सेल को स्पष्ट हाईलाइट करने के लिए प्रारूप 'Light Red Fill with Dark Red Text' का प्रयोग किया गया है। यदि हम चाहें तो drop down list से अन्य विकल्प चुनकर इसे बदल सकते हैं। इस सशर्त स्वरूपण का परिणाम तुरंत सेल C7 में नज़र आ जाएगा जिसमें 50 से अधिक का मान (गलती से) प्रविष्ट हो गया है। डायलॉग बॉक्स में OK पर क्लिक करें इससे A7 हाईलाइट हो जाएगी। अब इस प्रविष्टि को ठीक किया जा सकता है।

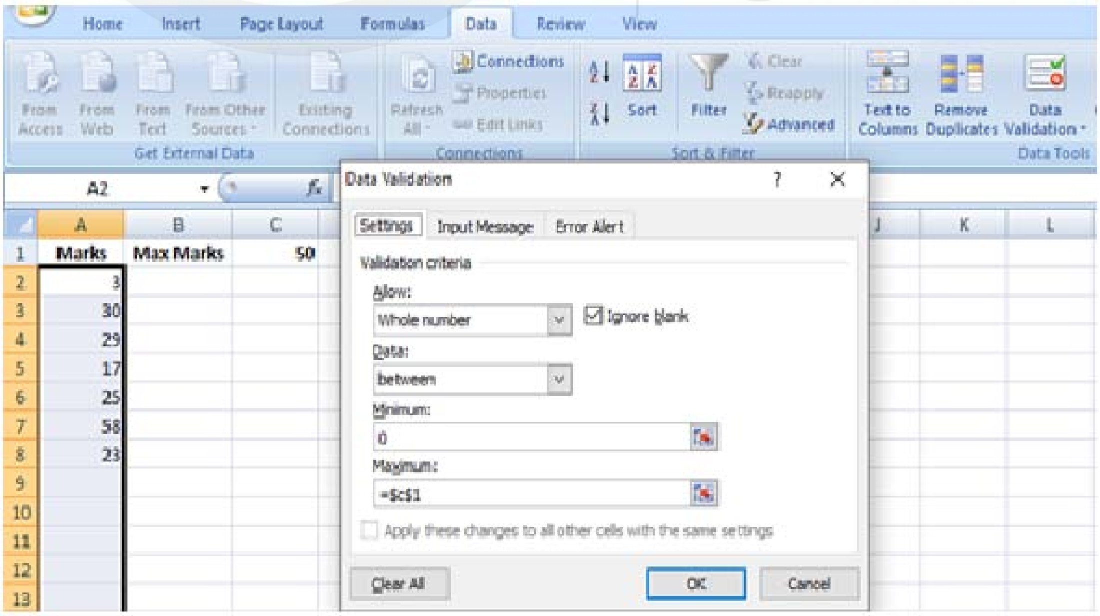
आँकड़ों का वैधीकरण (Data Validation)

आँकड़ों के वैधीकरण का एक तरीका यह भी है कि एक ऐसा आँकड़ा फॉरमेट (प्रारूप) या शर्त चुनी जाए कि (किसी नियम का उल्लंघन करने वाले) गलत आँकड़े की प्रविष्टि करने की अनुमति ही न हो।

गतिविधि 3 : सेल A2 से A21 में आँकड़ा वैधीकरण के लिए एक नियम (उत्पन्न/का सृजन (Create) कीजिए) बनाइए। यह नियम यह जाँच करने में सक्षम होना चाहिए कि की जाने वाली प्रविष्टियों का मान 0 से लेकर 50 (सेल C1 में दिए अधिकतम अंक) तक हो। साथ ही, यह मान लें कि सेल A2 से लेकर सेल A8 तक डेटा में प्रविष्टियाँ की जा चुकी हैं (इसका अर्थ है कि प्रविष्ट किए गए आँकड़ों में एक गलत प्रविष्टि 58 भी सम्मिलित है)।

इस गतिविधि को कैसे करें?

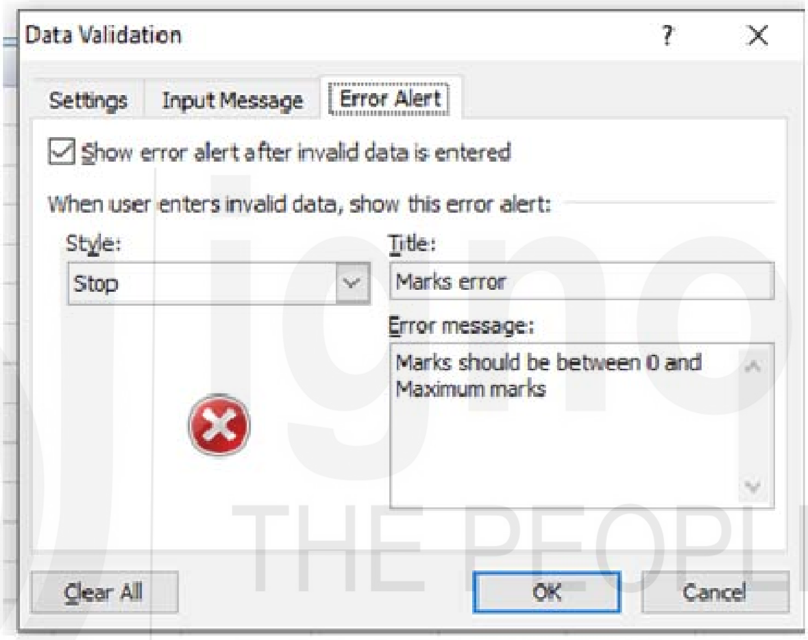
पहले हम सेल A2 से सेल 21 तक निम्न चरणों का प्रयोग करते हुए एक आँकड़ा वैधीकरण नियम का सृजन करते हैं :



चित्र 3.5(क) : **Data Validation-Settings Tab**

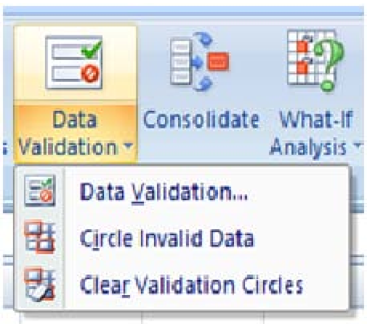
उस सेल रेंज का चयन कीजिए जिनके लिए वांछित वैधता नियम सृजित किया जाना है जो कि यहाँ A2 से A21 है, Data टैब पर जाइये तथा Data Tools रिबन में से Data Validation चुनिए। ऐसा करने पर एक डेटा वैधीकरण डायलॉग बाक्स खुल जाएगा। अब 'Allow' सूची में 'Whole Number' चुनिए, 'Data' सूची में between, 'Minimum' टेक्स्ट बॉक्स में *0* लिखिये तथा 'Maximum' टेक्स्ट बॉक्स में ***=$C$1*** जैसे कि नीचे चित्र 3.5(क) में दर्शाया गया है :

अब 'Data Validation' विंडो में ***Error Alert*** टेब को चुनिए तथा ***'Title'*** टेक्स्ट बॉक्स तथा **'Error Message :'** टेक्स्ट क्षेत्र में उपयुक्त टेक्स्ट/संदेश लिखिये। जैसा कि चित्र 3.5(ख) में दर्शाया गया है। यदि सेल C2 से सेल C21 में मानों की प्रविष्टि करते हुए कोई प्रविष्टि गलत हुई तो ये संदेश स्वतः प्रदर्शित हो जाएंगे।

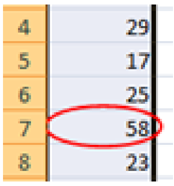


चित्र 3.5(ख) : **Data Validation Error Alert Tab**

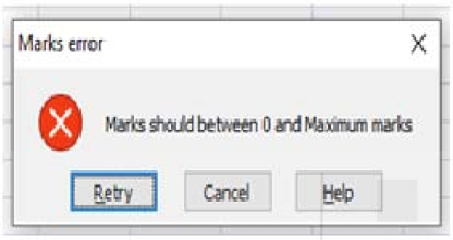
अब OK 'बटन' पर क्लिक कीजिए। इस समय ऐसा प्रतीत होगा कि अभी कुछ भी नहीं हुआ है। अभी मान 58 जो कि (अधिकतम अंक 50 से अधिक होने के कारण) एक गलत प्रविष्टि है वह भी हाईलाइट नहीं होगी। दिए हुए आँकड़ों में गलती को हाईलाइट करने के लिए हमें ***Data Validation*** बटन में टेक्स्ट वाले भाग में क्लिक करना होगा। ऐसा करने पर हमें नीचे दर्शाए गए विकल्प दिखेंगे :



इन विकल्पों में से विकल्प 'Circle Invalid Data' विकल्प चुनिये। हम देखेंगे कि गलत प्रविष्टि वाला सेल एक **वृत्त** द्वारा हाईलाइट हो गया है।



पाठक कृपया ध्यान दें, जैसा कि चित्र 3.5(क) में दर्शाया गया है, अभी तक केवल सेल C8 तक ही प्रविष्टियाँ की गई हैं। जबकि वैधता (Validation) की जाँच के लिए सेल A2 से सेल A21 तक के हिस्से को चुना गया था। अतः, सेल A9 से सेल A21 तक के भाग के लिए, जो कि अभी खाली छोड़ा गया है, डेटा वैलीडेशन वैध रहेगा। अब यदि हम सेल A9 में 52 तथा सेल A10 में –2 लिखें तो हमें निम्न त्रुटि चेतावनी संदेश दिखेगा :



इस त्रुटि चेतावनी संदेश को ध्यान से देखें। इसमें शीर्षक (Title) तथा त्रुटि संदेश (Error Message) वही है जो हमने 'Title' तथा 'Error Message:' टेक्स्ट बॉक्स में लिखे थे (चित्र 3.5ख देखें)।

अतः, आँकड़ा वैधीकरण केवल गलत प्रविष्टियों वाले सेलों को पहचानने में ही हमारी सहायता नहीं करता बल्कि हमें त्रुटिपूर्ण आँकड़ों की प्रविष्टि की अनुमति भी नहीं देता।

बोध प्रश्न 1

1) 5 उपयोगी सांख्यिकीय सॉफ्टवेयर पैकेजों के नाम लिखिये तथा उनकी पाँच महत्त्वपूर्ण विशेषताएँ बताइये।

....

....

....

....

2) आँकड़ा वैधता (Data Validation) के प्रयोग के क्या लाभ हैं?

....

....

....

3) किसी भी सांख्यिकीय सॉफ्टवेयर का प्रयोग करके निम्नलिखित आँकड़ों का सृजन कीजिए :

एक संगठन के 10 कर्मचारियों की मासिक आय इस प्रकार है :
25000, 33000, 43500, 12000, 55000, 104000, 60000, 29000, 31507, 39201

निम्नलिखित वैधता कीजिए:

न्यूनतम आय 15000 तथा अधिकतम आय 100000 होनी चाहिए। ऐसी सभी प्रविष्टियों को सर्कल कीजिए जो वैधीकरण के इस मापदंड को संतुष्ट नहीं करतीं। ऐसी सभी प्रविष्टियों को जो 50000 से अधिक हों प्रारूपिता (Format) करें। ये प्रविष्टियाँ जिन *सैलों* में हों उन्हें लाल रंग से भरिये।

....

....

....

....

....

3.4 सूत्रों तथा फलनों का प्रयोग

अधिकतर स्प्रेडशीट्स तथा सांख्यिकीय **पैकेजों** में सूत्रों तथा फलनों द्वारा परिकलन की सुविधा उपलब्ध होती है। सूत्रों तथा फलनों के महत्त्व को दर्शाने के लिए इस भाग में हम उदाहरण के रूप में MS Excel का प्रयोग करेंगे। अधिकतम सांख्यिकीय सॉफ्टवेयर में सूत्रों तथा फलनों का तर्क लगभग एक जैसा ही होगा यद्यपि उनके प्रयोग की विधियों में कुछ अंतर हो सकते हैं।

क्योंकि प्रत्येक सांख्यिकी सॉफ्टवेयर में फलनों का एक व्यापक भंडार होता है, अतः हम सभी फलनों की व्याख्या करने की बजाय, एक उदाहरण के माध्यम से इनमें से कुछ फलनों का प्रयोग करने की विधि का वर्णन करेंगे। शेष फलनों के बारे में जानकारी, विद्यार्थी जिस भी सॉफ्टवेयर का वे प्रयोग कर रहे हों उस सॉफ्टवेयर की Help में जाकर प्राप्त कर सकते हैं।

गतिविधि 4 : विद्यार्थियों के एक समूह द्वारा प्राप्तांकों की जानकारी नीचे दी गई है। प्रत्येक विद्यार्थी द्वारा प्राप्त अंकों का प्रतिशत तथा ग्रेड ज्ञात कीजिए।

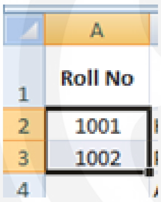
रोल नंबर	विद्यार्थी का नाम	प्राप्तांक (200 में)
1001	कमल जैन	57
1002	प्रेमलता सिंह	91
1003	आद्रियन	155
1004	सपना झा	150
1005	आरिफ़ खान	152
1006	अर्जुनन आर.	136
1007	स्वप्न पटनायक	173
1008	देवेंद्र मिश्रा	154
1009	उमेश चंद्र	83
1010	नवीन कुमार	101

ग्रेड नीचे दिए नियम के अनुसार ज्ञात किये जायेंगे :

ग्रेड ज्ञात करने का नियम	
>=75%	A
>=60% परंतु <75%	B
>= 40% परंतु <60%	C
< 40%	D

इस गतिविधि को कैसे करें?

आँकड़ों की प्रविष्टि : पहले चरण में दिये हुए आँकड़ों को हम एक वर्कशीट में दर्ज (प्रविष्ट) करें– पहले शीर्षक लिखें तथा Fonts रिबन में बटन B को दबाकर उन्हें मोटे (Bold) अक्षरों में परिवर्तित करें। Alignment विकल्पों का प्रयोग करके उन्हें संरेखित भी किया जा सकता है। अब आप नाम तथा अंकों को प्रविष्ट करेंगे। ध्यान रहें कि नाम इत्यादि लिखते हुए CAPS LOCL बटन ON रहे जिससे नामों को बड़े अक्षरों में लिखा जा सके। कृपया ध्यान दें कि सभी रोल नंबर 1001 से 1010 तक एक श्रेणी (series) में हैं। किसी वर्कशीट में संख्याओं की किसी श्रेणी की प्रविष्टि की एक सरल विधि है। पहले सेल A2 में 1001 दर्ज करें तथा सेल A3 में 1002। अब, सेल A2 तथा A3 दोनों का चयन करें : इसके लिए मॉउस के दायीं प्वाइंट को पहले तो सेल A2 पर ले जाएं। अब मॉउस के बाएं बटन को इस सेल पर दबाएं तथा इसे दबाते हुए ही मॉउस को नीचे की ओर ले जाएँ जिससे सेल A3 सेलैक्ट हो जाए। जैसा कि नीचे चित्र में दर्शाया गया है। अब मॉउस के बाएं बटन को मुक्त कर दें।



ध्यान दें कि इस समय A2 तथा A3 दोनों सेलों के चारों ओर एक बॉक्स बन गया है और इस बॉक्स के दाईं ओर के नीचे वाले कोने पर एक छोटा क्रॉस है। जब आप मॉउस को इस क्रॉस पर लेकर जाएँगे तो यह क्रॉस एक + के चिन्ह में परिवर्तित हो जाएगा। इस + पर मॉउस ले जाकर उसके बाएँ बटन को दबाएँ तथा उसे दबाए रखें, प्वाइंटर को नीचे की ओर सेल A11 तक ले जाएँ तथा बटन को मुक्त कर दें। अब सभी सेल श्रेणी 1001 से 1010 तक की संख्याओं से भर जाएँगे। चित्र 3.6 में यह अन्य शीर्षकों तथा अधिकतम अंकों (Maximum Marks) तथा प्राप्तांकों (Marks Obtained) के साथ दर्शाया गया है।

A2 fx 1001

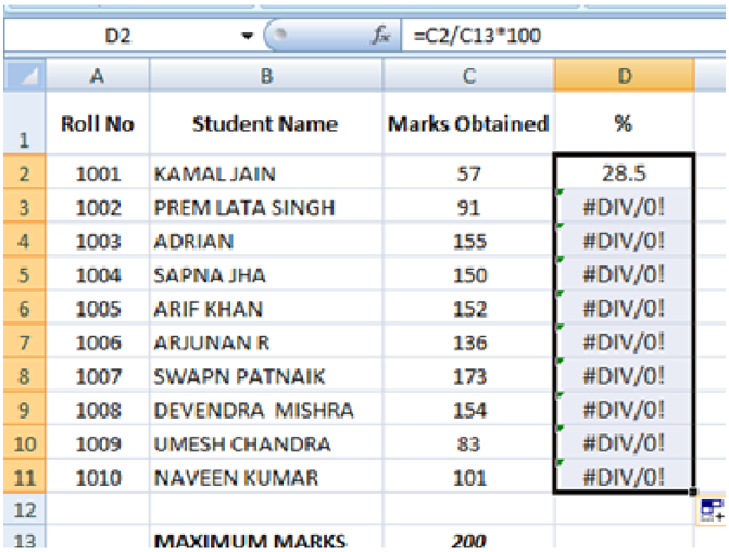
	A	B	C	D	E
1	Roll No	Student Name	Marks Obtained	%	Grade
2	1001	KAMAL JAIN	57		
3	1002	PREM LATA SINGH	91		
4	1003	ADRIAN	155		
5	1004	SAPNA JHA	150		
6	1005	ARIF KHAN	152		
7	1006	ARJUNAN R	136		
8	1007	SWAPN PATNAIK	173		
9	1008	DEVENDRA MISHRA	154		
10	1009	UMESH CHANDRA	83		
11	1010	NAVEEN KUMAR	101		
12					
13		MAXIMUM MARKS	*200*		

चित्र 3.6 : गतिविधि 4 का डेटा

किसी सेल में सूत्र बनाना (दर्ज करना) : एक बार वर्कशीट में आँकड़ों को प्रविष्ट करने के पश्चात् हमें प्रत्येक विद्यार्थी द्वारा प्राप्त प्रतिशत अंकों का परिकलन करने की तथा इस प्रतिशत के आधार विद्यार्थी का ग्रेड ज्ञात करने की आवश्यकता है। इसके लिए हमें प्रतिशत ज्ञात करने का एक सूत्र बनाने की आवश्यकता है। यह सूत्र प्राप्तांक/अधिकतम अंक *100 है। पहले विद्यार्थी के प्राप्तांक सेल C2 में तथा अधिकतम अंक (Maximum Marks) सेल C3 में दर्शाए गए हैं। अतः, परिकलित प्रतिशत सेल D2 में दर्शाई जानी चाहिए। इसके लिए हमें सेल D2 में सूत्र C2/C13*100 को प्रविष्ट करना पड़ेगा। परंतु, यह सूत्र Excel द्वारा एक टेक्स्ट इनपुट के रूप में लिया जा सकता है। अतः, टेक्स्ट तथा सूत्र में अंतर स्पष्ट करने के लिए सभी सूत्र '=' चिन्ह से प्रारंभ किये जाते हैं। अतः, हमें सेल D2 में सूत्र ***=C2/C13*100*** दर्ज करना पड़ेगा। क्योंकि हमें सेल C2 से C11 तक, सभी के लिए प्रतिशत ज्ञात करना है। अतः उक्त सूत्र को क्रमशः सेल D2 से D11 में प्रविष्ट करना होगा है। दूसरे शब्दों में, इन सेलों (कोष्ठिकाओं) में नीचे दिए सूत्र प्रविष्ट/दर्ज करने पड़ेंगे :

सैल	**सूत्र**
D2	=***C2***/C13*100
D3	=***C3***/C13*100
D4	=***C4***/C13*100
....	...
D11	=***C11***/C13*100

ध्यान दें कि ऊपर दी हुई तालिका की कोष्ठिकाओं (Cells) D2 से D11 में दिए गए सूत्र हाईलाइट किए गए हैं। ये सूत्र एक सरल अनुक्रम में हैं। क्या हमें ये सभी सूत्र एक-एक करके प्रविष्ट करने की आवश्यकता है? यदि ऐसा है तो इसमें वास्तव में, बहुत वक्त लगेगा। परंतु ऐसा करने की आवश्यकता नहीं है। Excel में यह सुविधा उपलब्ध है कि जैसे-जैसे हम नीचे की ओर जाएँ एक सूत्र को एक सेल से दूसरे में कॉपी किया जा सके। अतः, सेल D2 में सूत्र प्रविष्ट करने के पश्चात् हमें केवल सूत्र को सेल D3 से लेकर सेल D11 तक कॉपी करना है। अतः, पहले सूत्र सेल D2 में दर्ज करें, फिर सेल D2 का चयन करें जैसा कि नीचे दिखाया गया है तथा मॉउस के प्वांइटर को सेल के दाईं ओर के निचले कोने में ले जाएं जिससे वहाँ का क्रॉस '+' के चिन्ह में बदल जाएगा (चित्र देखे)। अब मॉउस का बायाँ बटन दबाइये तथा बटन को दबाते हुए मॉउस को नीचे D11 तक ले जाइये। अब इस बटन को छोड़ दीजिए। ऐसा करने पर निम्नलिखित स्क्रीन आ जाएगी (चित्र 3.8)।



D2 | f_x =C2/C13*100

	A	B	C	D
1	Roll No	Student Name	Marks Obtained	%
2	1001	KAMAL JAIN	57	28.5
3	1002	PREM LATA SINGH	91	#DIV/0!
4	1003	ADRIAN	155	#DIV/0!
5	1004	SAPNA JHA	150	#DIV/0!
6	1005	ARIF KHAN	152	#DIV/0!
7	1006	ARJUNAN R	136	#DIV/0!
8	1007	SWAPN PATNAIK	173	#DIV/0!
9	1008	DEVENDRA MISHRA	154	#DIV/0!
10	1009	UMESH CHANDRA	83	#DIV/0!
11	1010	NAVEEN KUMAR	101	#DIV/0!
12				
13		MAXIMUM MARKS	200	

चित्र 3.8 : सूत्र कॉपी करना

कृपया ध्यान दें कि सूत्र को सेल D3 से D11 में कॉपी करने पर, ये सभी सेल प्रतिशत दिखाने की बजाय, शून्य से भाग करने पर उत्पन्न त्रुटि दर्शा रहे हैं। क्यों? सूत्र को कॉपी होने वाली किसी समस्या के कारण? यह जानने के लिए सेल D3 से सेल D11 में दिए सूत्रों की जाँच कीजिये। आप इन कोष्ठिकाओं (Cells) में निम्नलिखित सूत्र पाएँगे :

सेल / कोष्ठिका	अपेक्षित सूत्र	सूत्र जो कॉपी किया
D3	=C3/C13*100	=C3/*C14**100
D4	=C4/C13*100	=C4/*C15**100
D5	=C5/C13*100	=C5/*C16**100
…	…	…
D11	=C11/C13*100	=C11/*C22**100

इसका अर्थ है सेल C2 की तरह सेल C13 में भी सूत्र में वृद्धि हो गई है। क्योंकि सेल C4 से सेल C22 में कोई भी मान नहीं है और हम इन कोष्ठिकाओं द्वारा भाग कर रहे हैं, इसलिए हमें 'शून्य द्वारा भाग से उत्पन्न त्रुटि' प्राप्त होती है। इस समस्या को कैसे हल किया जाए? इसके लिए हम एक संकल्पना का प्रयोग करना पड़ेगा जिसे Absolute Addressing कहते हैं जो Excel को यह बताता है कि सूत्र में एक सेल (यहाँ C13) स्थिर रहता है और कॉपी करते समय उसमें परिवर्तन नहीं होना चाहिए। Absolute Addressing के लिए हमें कॉलम से पहले चिन्ह $ लगाना पड़ेगा जैसे कि C$13 (इसे Mixed Addressing) कहते हैं। यदि आप कॉलम में कॉपी कर रहे हैं और कॉलम के Address में कोई परिवर्तन नहीं चाहते या पंक्ति से पहले चिन्ह $ लगाते है। जैसे कि C$13 (इसे भी Mixed Addressing कहते हैं) यदि आप पंक्ति में कॉपी करना चहते हैं या पंक्ति तथा कॉलम दोनों से पहले जैसे कि $C$13 (इसे Absolute Addressing कहते है।) यदि आप यह चाहते हैं कि कॉपी करते समय सेल के Address में न तो पंक्ति और न ही कॉलम में कोई परिवर्तन हो)। इस उदाहरण में हम Absolute Addressing का प्रयोग करते हैं।

अतः सेल D2 में सूत्र ***=C2/C13*100*** के स्थान पर सूत्र ***=C2/$C$13*100*** रखते हैं) और इस सूत्र को सेल D3 से D11 तक कॉपी करते हैं। आप देख सकते हैं कि अब हमें यथोचित परिणाम प्राप्त होगा। चित्र 3.9 में ऐसा ही कॉपी किया हुआ एक सूत्र तथा प्रतिशत का पूरा परिकलन दर्शाया गया है।

D8   fx   =C8/$C$13*100

	A	B	C	D
1	Roll No	Student Name	Marks Obtained	%
2	1001	KAMAL JAIN	57	28.5
3	1002	PREM LATA SINGH	91	45.5
4	1003	ADRIAN	155	77.5
5	1004	SAPNA JHA	150	75
6	1005	ARIF KHAN	152	76
7	1006	ARJUNAN R	136	68
8	1007	SWAPN PATNAIK	173	86.5
9	1008	DEVENDRA MISHRA	154	77
10	1009	UMESH CHANDRA	83	41.5
11	1010	NAVEEN KUMAR	101	50.5
12				
13		MAXIMUM MARKS	200	

चित्र 3.9 : Absolute Addressing द्वारा कॉपी करना

कृपया ऊपर दिए चित्र के सेल D8 में दिए सूत्र पर ध्यान दें।

ग्रेड का परिकलन

ग्रेड ज्ञात करने के लिए हमें ग्रेड परिकलन नियम की आवश्यकता पड़ेगी जो कि गतिविधि 4 में दिया हुआ है। हम इस गतिविधि को IF फलन के प्रयोग द्वारा अथवा VLOOKUP फलन द्वारा कर सकते हैं।

IF फलन द्वारा ग्रेड परिकलन :

Excel में IF फलन का विन्यास (सिंटैक्स) इस प्रकार है,

If (Logical Condition जो कि चेक की जानी है, यदि यह Condition सत्य हो तो क्या करना है, यदि यह Condition असत्य हो तो क्या करना है)

नीचे दी हुई तालिका सेल E2 में ग्रेड के परिकलन के सूत्र का वर्णन किया गया है। कृपया ध्यान दें कि इस तालिका में यह माना गया है कि सेल D2 में प्रतिशत अंक दिए हैं।

Logical Condition	**IF (condition,…,…)**	When condition is TRUE	When condition is FALSE
>=75%	***=IF (D2>=75,…, …)***	=(D2>=75,***"A"***, …) यदि किसी विद्यार्थी के प्रतिशत अंक 75% से ऊपर हैं तो उसे A ग्रेड मिलेगा। कृपया ध्यान दें कि	यदि किसी विद्यार्थी के अंक 75% से कम है। =IF (D2>=75, "A", एक अन्य *IF condition* लिखिये)

		A को " " में रखा गया है क्योंकि यह एक अक्षर है।	
>=60% लेकिन <75% है	=IF (D2>=75,"A", ***IF(D2>=60, …, …)*** कृपया ध्यान दें कि दूसरी IF condition, पहली IF condition के False वाली condition के रूप में लिखी हुई है। अर्थात् दूसरी IF condition केवल तभी परिकलित की जएगी जब अंक 75% से कम होंगें।	= IF (D2>=75,"A", IF (D2>=60, ***"B",*** …) कृपया ध्यान दें कि B ग्रेड के लिए % अंक < 25% है। इसीलिए हम यह जाँच करनी पड़ती है कि IF = D2>=60. अतः अंक >-60 हैं परंतु < 75% है।	अंक 60% से कम हैं =IF (D2>=75,"A", IF(D2>=60, "B", एक अन्य *IF condition* लिखिये
>= 40% लेकिन <60% है	=IF (D2>=75,"A", IF(D2>=60, "B", ***IF (D2>=40, …, …)*** IF (D2>=75,"A", IF(D2>=60, "B", ***IF (D2>=40, …, …)*** कृपया ध्यान दें कि दूसरी IF condition के FALSE वाले भाग में उत्पन्न हुई है। इसका अर्थ है कि तीसरे IF परिकलन केवल तभी किया जाएगा यदि अंक 60% से कम होंगे अर्थात् दोनों बाहरी IF FALSE हों।	=IF (D2>=75,"A", IF(D2>=60, "B", IF (D2>=40, ***"C",*** …) IF (D2>=75,"A", IF(D2>=60, "B", ***IF (D2>=40, …, …)*** कृपया ध्यान दें कि तीसरी IF condition के FALSE वाले भाग में उत्पन्न हुई है। इस का अर्थ है कि तीसरे IF परिकलन केवल तभी किया जाएगा यदि अंक 60% से कम होंगे अर्थात् दोनों बाहरी IF FALSE हों।	अंक निश्चित रूप से 40% से कम हैं =IF (D2>=75,"A", IF(D2>=60, "B", IF (D2>=40, "C", *ग्रेड D होगा*)
< 40%	=IF (D2>=75,"A", IF(D2>=60, "B", IF (D2>=40, "C", ***"D")))*** अब आप सेल E2 के सूत्र को सेल E3 से E11 में कॉपी किया जा सकता है। चित्र 3.10 विद्यार्थियों के ग्रेड दर्शाता है।		

अब सेल E2 के सूत्र को सेल E3 के E11 में कॉपी किया जा सकता है। चित्र 3.10 विद्यार्थियों के ग्रेड दर्शाता है।

E7 f_x =IF(D7>=75,"A",IF(D7>=60,"B",IF(D7>=40,"C","D")))

	A	B	C	D	E	F	G	H
1	Roll No	Student Name	Marks Obtained	%	Grade			
2	1001	KAMAL JAIN	57	28.5	D			
3	1002	PREM LATA SINGH	91	45.5	C			
4	1003	ADRIAN	155	77.5	A			
5	1004	SAPNA JHA	150	75	A			
6	1005	ARIF KHAN	152	76	A			
7	1006	ARJUNAN R	136	68	B			
8	1007	SWAPN PATNAIK	173	86.5	A			
9	1008	DEVENDRA MISHRA	154	77	A			
10	1009	UMESH CHANDRA	83	41.5	C			
11	1010	NAVEEN KUMAR	101	50.5	C			
12								
13		MAXIMUM MARKS	*200*					

चित्र 3.10 : IF फलन द्वारा ग्रेड ज्ञात करना

पाठकों को चाहिए कि Excel फलनों का प्रयोग करना प्रारंभ करें, जैसा कि गतिविधि 4 में दर्शाया गया है। यदि आवश्यकता हो तो एक फलन के अंदर और फलन भी लगाए जा सकते हैं। एक अन्य महत्त्वपूर्ण फलन जिसका उपयोग किसी तालिका के मानों को खोजने के लिए तथा उनके ग्रेड इत्यादि ज्ञात करने के लिए किया जा सकता है, वह है VLOOKUP फलन। आइये, अब हम गतिविधि 4 को VLOOKUP फलन की सहायता से करें।

VLOOKUP फलन द्वारा ग्रेड परिकलन :

Excel में VLOOKUP फलन का विन्यास (Syntax) इस प्रकार है :

=VLOOKUP [मान जिसे खोजना है, वह परिसर जिसमें खोजा जाने वाला मान तथा प्राप्त किया जाने वाला मान स्थित होंगे, प्राप्त किए जाने वाले मानों के लिए निर्दिष्ट परिसर के संगत स्तंभ संख्या, सन्निकट मिलान (TRUE) अथवा सटीक मिलान (FALSE) का वर्णन]।

आइये, अब हम गतिविधि 4 को VLOOKUP फलन द्वारा करें। VLOOKUP फलन के लिए हमें पहले ग्रेड की सीमाएँ निश्चित करने वाले (Cut Off) अंकों की अनुक्रमित तालिका बनानी पड़ती है। इसके लिए निम्न तालिका बनाई गई है :

ग्रेड ज्ञात करने के लिए LOOKUP टेबल (तालिका)		
	प्रतिशत	**ग्रेड**
< 40%	0.0	D
>= 40% लेकिन <60%	40.0	C
>=60% लेकिन <75%	60.0	B
>=75%	75.0	A

कृपया ध्यान दें कि हम इस तालिका के केवल छायांकित (Shaded) भाग का ही उपयोग करेंगे। यह तालिका आप वर्कशीट में तैयार कर सकते हैं। अब आप सेल **E2** में सूत्र को इस प्रकार प्रविष्ट कर सकते हैं :

=VLOOKUP(D2,*$H$4:$I$7*,2,TRUE)

सूत्र की व्याख्या : सेल D2 में वह प्रतिशत अंक हैं जिसे परिसर H4 से 17 जो कि तालिका का छायांकित भाग है, के प्रतिशत स्तंभ में खोजना है। क्योंकि सभी कोष्ठिकाओं (सेलों) में एक ही परिसर के लिए सूत्र कॉपी किया जाना है, इसलिए, Absolute Addressing का प्रयोग किया गया है। इसीलिए इसे ***$H$4:$I$7*** के रूप में लिखा गया है। मान 2 यह दर्शाता है कि D2 के मान का प्रतिशत मान (Percentage Value) के साथ मिलान करने के पश्चात् परिणाम को दर्ज करने के लिए, दूसरे कॉलम/स्तंभ का अर्थात् ग्रेड कॉलम का उपयोग किया जाएगा। क्योंकि यह पूरा सही नहीं हो सकता, अतः, चौथे पैरामीटर/प्राचल के तौर पर TRUE का चयन किया गया है। सेल E2 में सूत्र लिखने के पश्चात्, हम सूत्र को सेल C3 से C11 तक कॉपी कर सकते हैं। चित्र 3.11 इस प्रकार प्राप्त परिणाम को दर्शाता है। कृपया सेल E8 के सूत्र पर ध्यान दें।

E8 fx =VLOOKUP(D8,$H$4:$I$7,2,TRUE)

	A	B	C	D	E	F	G	H	I
1	Roll No	Student Name	Marks Obtained	%	Grade				
2	1001	KAMAL JAIN	57	28.5	D		Lookup Table for Grade		
3	1002	PREM LATA SINGH	91	45.5	C			Percentage	Grade
4	1003	ADRIAN	155	77.5	A		< 40%	0.0	D
5	1004	SAPNA JHA	150	75	A		>= 40% but <60%	40.0	C
6	1005	ARIF KHAN	152	76	A		>=60% but <75%	60.0	B
7	1006	ARJUNAN R	136	68	B		>=75%	75.0	A
8	1007	SWAPN PATNAIK	173	86.5	A				
9	1008	DEVENDRA MISHRA	154	77	A				
10	1009	UMESH CHANDRA	83	41.5	C				
11	1010	NAVEEN KUMAR	101	50.5	C				
12									
13		MAXIMUM MARKS	200						
14									

चित्र 3.11 : VLOOKUP फलन द्वारा ग्रेड का परिकलन

आप चित्र 3.10 तथा 3.11 की तुलना करके यह देख सकते हैं कि दोनों सूत्रों द्वारा प्राप्त परिणाम समान हैं।

गतिविधि 5 : कार्य 1 के डेटा के लिए अधिकतम, न्यूनतम तथा औसत अंक ज्ञात कीजिए। साथ ही, ग्रेड्स के लिए बारंबारता आबंटन ज्ञात कीजिए।

गतिविधि कैसे करें?

हमें केवल अधिकतम, न्यूनतम तथा औसत ज्ञात करने के लिए सूत्र लिखना होगा। ये सूत्र हैं :

अधिकतम अंक	=MAX(C2:C11)	अधिकतम ज्ञात करने के लिए
औसत अंक	=AVERAGE(C2:C11)	औसत ज्ञात करने के लिए
न्यूनतम अंक	=MIN(C2:C11)	न्यूनतम ज्ञात करने के लिए

अब यह कैसे ज्ञात करें कि कौन-से सूत्र का उपयोग करना है? इसके लिए विद्यार्थी उस सॉफ्टवेयर की सहायता लें जिसका वह प्रयोग कर रहे हैं। प्रत्येक सांख्यिकीय सॉफ्टवेयर पैकेज में उपयोगकर्ता की सहायता के लिए समृद्ध Help Documentation (प्रलेखन) उपलब्ध होता है। बारंबारता आबंटन के लिए, आप नीचे दिए सूत्रों का प्रयोग कर सकते हैं (कोष्ठिकाओं (Cells) के संदर्भ देखने के लिए चित्र 3.12 देखें)।

ग्रेड का बारंबारता आबंटन		**टिप्पणी**
A	=COUNTIF($E$2:$E$11, G7)	जिस परिसर के सेलों के ग्रेड ज्ञात करने हैं उनके सैल्स की संख्या की गणना करनी है यदि परिसर में सैल्स विशिष्ट ग्रेड G7 या G8 या G9 या G10 के बराबर हो।
B	=COUNTIF($E$2:$E$11, G8)	
C	=COUNTIF($E$2:$E$11, G9)	
D	=COUNTIF($E$2:$E$11, G10)	
कुल बारंबारता	=SUM(H7:H10)	सभी बारंबारताओं का योग, विद्यार्थियों की कुल संख्या के बराबर होना चाहिए।

H7 =COUNTIF($E$2:$E$11, G7)

	A	B	C	D	E	F	G	H
1	Roll No	Student Name	Marks Obtained	%	Grade			
2	1001	KAMAL JAIN	57	28.5	D		Minimum Score	173.0
3	1002	PREM LATA SINGH	91	45.5	C		Mean Score	125.2
4	1003	ADRIAN	155	77.5	A		Maximum Score	57.0
5	1004	SAPNA JHA	150	75	A			
6	1005	ARIF KHAN	152	76	A		Frequency Distribution of Grade	
7	1006	ARJUNAN R	136	68	B		A	5
8	1007	SWAPN PATNAIK	173	86.5	A		B	1
9	1008	DEVENDRA MISHRA	154	77	A		C	3
10	1009	UMESH CHANDRA	83	41.5	C		D	1
11	1010	NAVEEN KUMAR	101	50.5	C		Total Frequency	10
12								
13		MAXIMUM MARKS	200					

चित्र 3.12 : गतिविधि 5 के परिणाम

बोध प्रश्न 2

1) भाग 3.4 में प्रयुक्त सभी फलनों का उद्देश्य तथा विन्यास (Syntax वाक्य-रचना) लिखिये। प्रत्येक का एक उदाहरण भी दीजिए।

....

....

....

2) परिशुद्ध पताभिगमन (Absolute Addressing) सूत्रों तथा फलनों के क्या लाभ हैं?

....

....

....

3) 5 अलग विषयों में 10 विद्यार्थियों द्वारा प्राप्तांकों के लिए एक स्प्रैडशीट बनाइए। मान लीजिए प्रत्येक विषय के अधिकतम अंक 50 हैं। प्रत्येक विद्यार्थी के प्रत्येक विषय में प्राप्त अंकों का प्रतिशत ज्ञात कीजिए। साथ ही, प्रत्येक विद्यार्थी के लिए कुल प्राप्तांक तथा कुल प्रतिशत अंत ज्ञात कीजिए। कोई भी विद्यार्थी किसी विषय में उत्तीर्ण (Pass) घोषित किया जाएगा यदि वह उस विषय में कम-से-कम 40% अंक प्राप्त करें। कोई भी छात्र सफल (Successful) घोषित किया जाएगा जब वह प्रत्येक विषय में उत्तीर्ण हो। सफल विद्यार्थियों की सूची ज्ञात कीजिए। यदि आवश्यकता हो, तो उपयुक्त मान्यताएँ ली जा सकती हैं।

....

....

....

....

....

3.5 चार्ट बनाना

चार्ट आँकड़ों का आलेखीय/दृश्यात्मक निरूपण करते हैं। प्रत्येक सांख्यिकीय सॉफ्टवेयर में चार्टों का एक बड़ा संग्रह उपलब्ध होता है। सांख्यिकीय विश्लेषण के संदर्भ में चार्टों की विस्तृत चर्चा हम इस पाठ्यक्रम की इकाई 6 में करेंगे। इस भाग में हम एक सरल चार्ट बनाने की प्रक्रिया का केवल एक संक्षिप्त परिचय देंगे।

गतिविधि 6 : चित्र 3.12 में दिए हुए ग्रेड्स के बारंबारता आबंटन के लिए एक कॉलम चार्ट बनाइए। आलेख में अक्षों के उपयुक्त शीर्षक भी होने चाहिए।

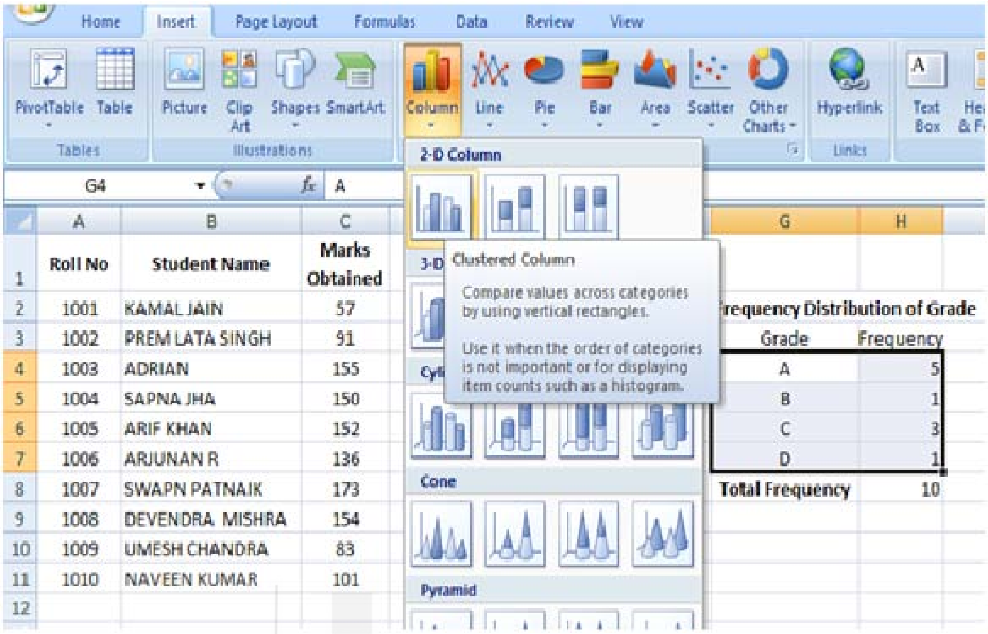
यह गतिविधि किस प्रकार करें?

अक्षों के शीर्षकों के साथ आलेख बनाने के लिए हमने ग्रेड आबंटन डेटा को इस प्रकार संशोधित किया कि सभी कॉलम उपयुक्त शीर्षकों के साथ हों। यह संशोधित डेटा नीचे दी गई तालिका में देखा जा सकता है :

ग्रेड बारंबारता आबंटन	
ग्रेड	**बारंबारता (Frequency)**
A	5
B	1
C	3
D	1
कुल बारंबारता	10

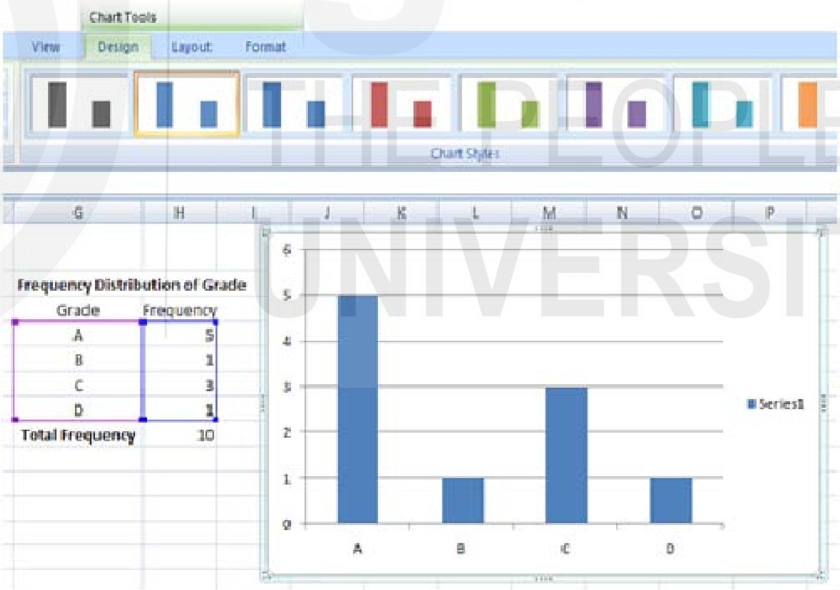
वर्कशीट के छायांकित (shaded) भाग का चयन कीजिए तथा ग्राफ बनाने के लिए **Insert** tab में जाकर Create graph का चयन कीजिए तथा **Charts** रिबन में से **Column** का चयन कीजिए। ऐसा करने पर कॉलम चार्ट के विभिन्न प्रकारों की एक

सूची खुल जाएगी। विकल्प ***@-D Column*** में से ***Clustered Column*** चुनिए, जैसा कि चित्र 3.13 में दर्शाया गया है।



चित्र 3.13 : एक कॉलम चार्ट बनाना

आप देखेंगे कि वर्कशीट में एक कॉलम चार्ट सन्निविष्ट हो जाएगा। परंतु इस चार्ट में कोई शीर्षक नहीं है (चित्र 3.14 देखें)।

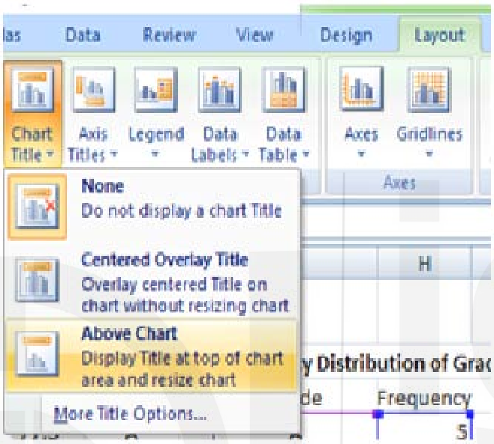


चित्र 3.14 : ग्रेड डेटा के लिए कॉलम चार्ट

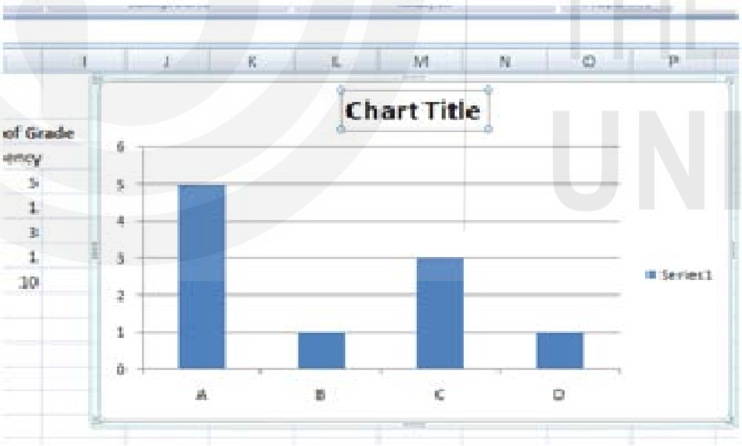
ध्यान दें कि इस चित्र के ऊपरी भाग में Chart tools के साथ तीन और टैब भी दिख रहे हैं : ***डिज़ाइन, ले-आउट*** तथा ***फॉरमेट***। आप विभिन्न विकल्पों में से कोई भी डिज़ाइन चुन सकते हैं। इस गतिविधि के लिए हम अपना ध्यान केवल चार्ट का शीर्षक तथा चार्ट में प्रयोग होने वाले अक्षों (Axes) के शीर्षक उपलब्ध करवाने पर ही केंद्रित करें। साथ ही, चार्ट में दिखाए गए संकेत (series1) भी हटा दिए जाने चाहिए। ये सभी गतिविधियाँ Layout टैब का प्रयोग करके की जा सकती हैं। Layout टैब का चयन करें तथा निम्नलिखित काम करें :

Labels रिबन में chart title चुनें। ऐसा करने पर विभिन्न विकल्प खुल जाएँगे जैसे कि चित्र 3.15 (क) में दर्शाए गए हैं। इनमें से Above chart option चुनें। जब आप यह विकल्प चुनेंगे, चार्ट के ऊपरी हिस्से में chart title आ जाएगा (चित्र 3.15ख देखें)। एक उपयुक्त चार्ट शीर्षक देने के लिए जैसे कि Frequency Distribution of Grades जो कि Cell G2 में है, = लिखें तथा सेल G2 को चुनें। अब हमें Formal bar में ***=Charts!$G$2*** दिखाई देगा (चित्र 3.15ग देखें)।

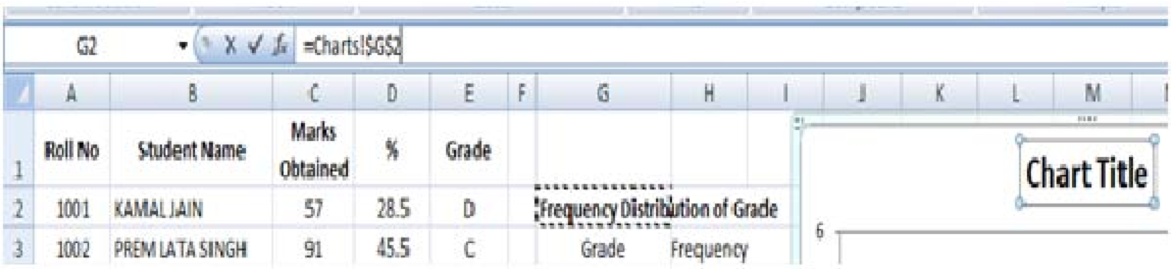
यहाँ charts वर्कशीट का नाम है। यदि आपने वर्कशीट को कोई नाम नहीं दिया है, तो यह sheet1 अथवा sheet2 हो सकता है। ***between*** in *Data:* list, ***0*** in *Minimum:* text box and ***=$C$1*** in the *Maximum:* जैसा कि चित्र 3.15क में दिखाया गया है।



(क) चार्ट शीर्षक विकल्प का चुनाव/चयन



(ख) क में की गई कार्रवाई का परिणाम



(ग) वर्कशीट सेल में चार्ट शीर्षक का चुनाव/चयन

चित्र 3.15 : चार्ट शीर्ष का चयन

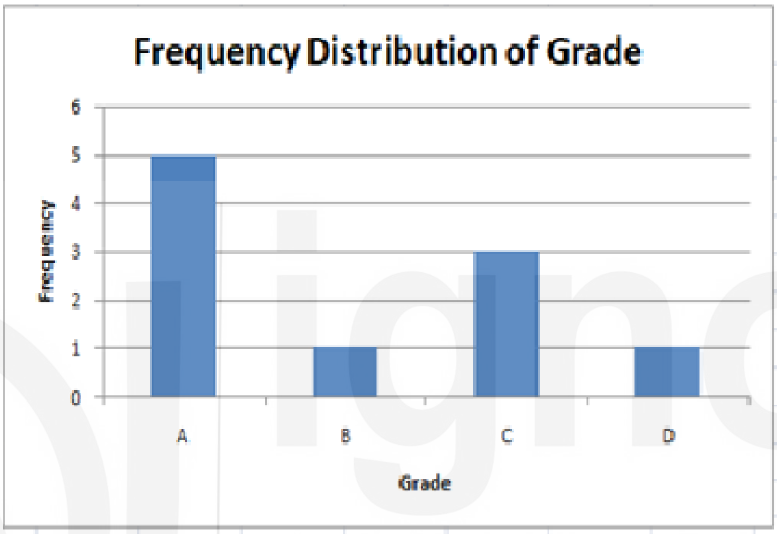
जैसे ही आप Enter बटन को क्लिक करेंगे चार्ट का शीर्षक Frequency Distribution of Grade हो जाएगा (चित्र 3.16 देखें)।

अब, आप इसी प्रकार अक्ष शीर्षक चुन सकते हैं। इसके लिए आपको नीचे दिए विकल्प चुनने होंगे :

Axis Title **→** ***Primary Horizontal Axis Title*** **→** ***Title Below Axis*** **→** (सूत्र पट्टी में = लिखें, सेल G3 को चुनें तथा Enter बटन को क्लिक करें)

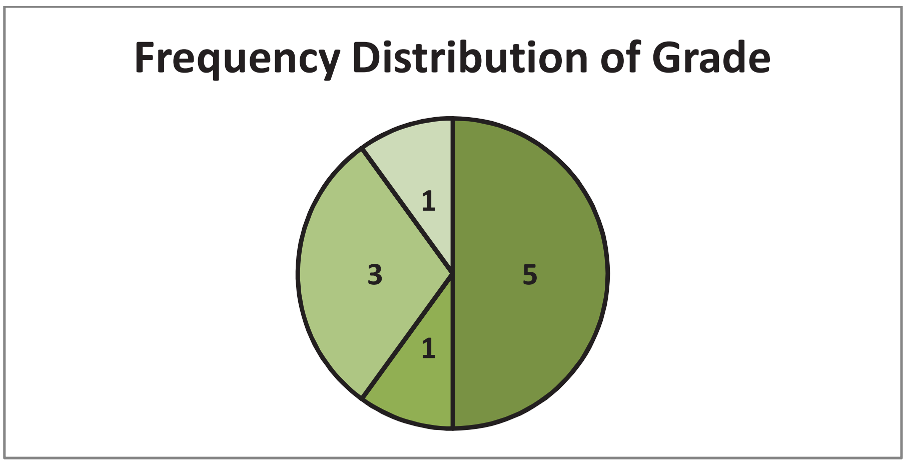
Axis Title **→** ***Primary Vertical Axis Title*** **→** ***Rotated Title*** **→** (put = in formula bar and select cell ***H3*** and press Enter key)

अब ग्राफ/चार्ट में संकेत ***Series1*** को चुनें, मॉउस के दायें बटन को क्लिक करें तथा विकल्प Delete का चयन करें।



चित्र 3.16 : गतिविधि 6 का कॉलम चार्ट

आप इसी डेटा के लिए Pie Chart भी बना सकते हैं। इसके लिए आपको केवल Charts रिबन में जाकर Pie चुनना होगा ओर उसके पश्चात् pie chart का प्रकार चुनना होगा। हमने यहाँ 2-D पाई चार्ट (pie chart) चुना है। डेटा लेबल दर्शाने के लिए विकल्प ***Layout→ Data Labels→Center*** चुनें। अंततः हमें चित्र 3.17 में दर्शाया गया पाई-चार्ट प्राप्त होगा।



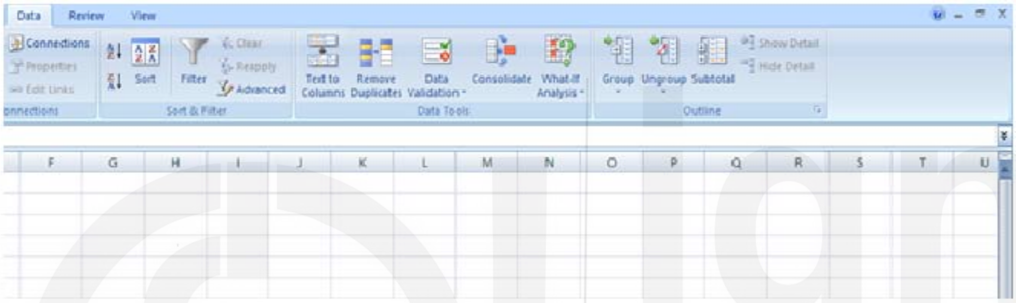
चित्र 3.17 : गतिविधि 6 के लिए पाई चार्ट

इसी प्रकार, आप सांख्यिकीय सॉफ्टवेयर के प्रयोग से आप अन्य चार्ट भी बना सकते हैं।

3.6 आँकड़ा विश्लेषण टैब को सक्रिय करना

एक महत्त्वपूर्ण गतिविधि, जिसे हम किसी सांख्यिकीय सॉफ्टवेयर के माध्यम से कर सकते हैं, वह है किसी आँकड़ा समूह के लिए विभिन्न सांख्यिकीय परीक्षण करना। ये परीक्षण Data टैब में उपलब्ध Data Analysis विकल्प के प्रयोग से किए जा सकते हैं। यद्यपि, MS- Excel के कई संस्करणों में यह विकल्प तत्काल उपलब्ध नहीं होते तथा इन्हें जोड़ना पड़ता है। इस भाग में हम Excel के Analysis Tool पैक को प्रयोग करने की विधि की व्याख्या करेंगे जिससे पाठक आँकड़ों पर विभिन्न सांख्यिकीय परीक्षण कर सकें।

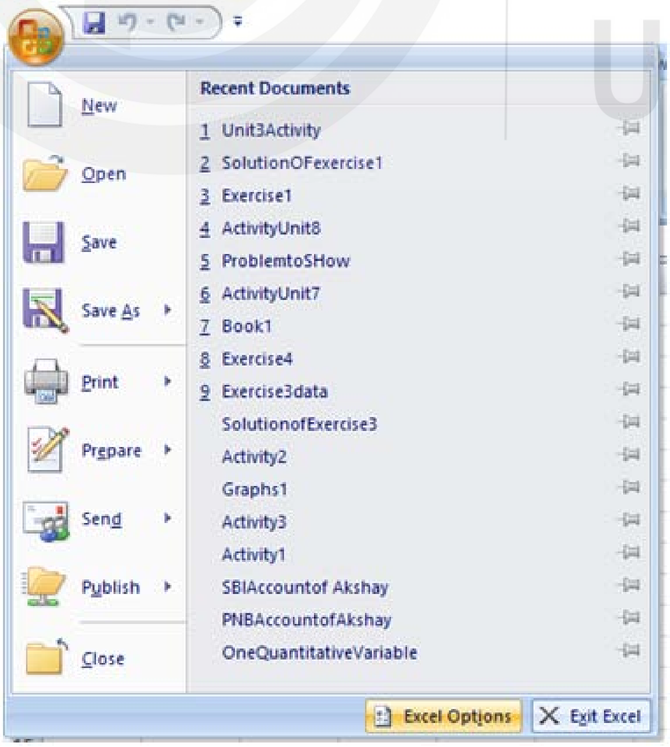
परंतु पहले डेटा टैब में जाकर जाँच कर लें कि Excel का जो संस्करण आपके पास उपलब्ध है उसमें डेटा विश्लेषण विकल्प उपलब्ध है अथवा नहीं। डेटा रिबन के सबसे दाईं ओर के हिस्से की जाँच करें। यदि इसमें डेटा विश्लेषण आइकॉन न दिखे (चित्र 3.18 देखें), तो इस भाग में दिए कदम उठाएँ।



चित्र 3.18 : डेटा विश्लेषण विकल्प के बिना डेटा टैब

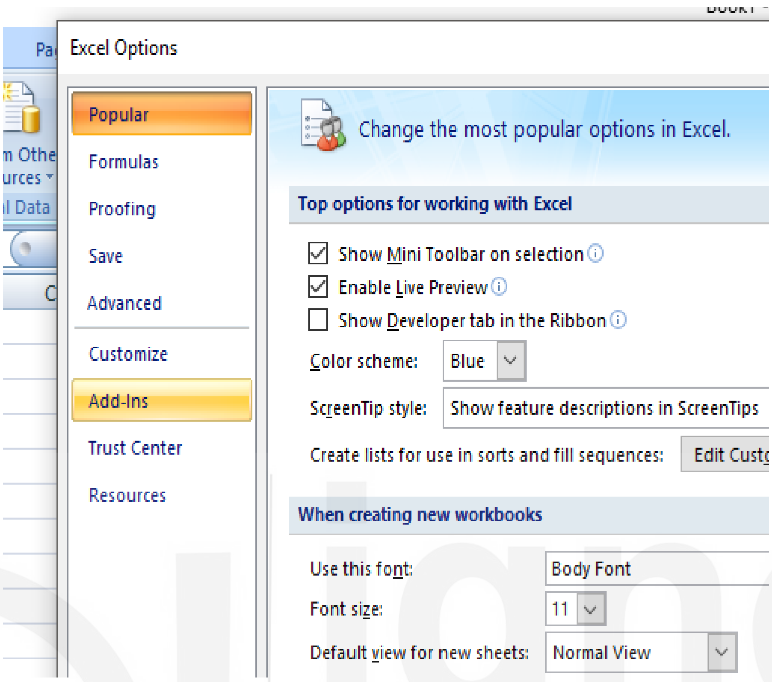
डेटा विश्लेषण पैक को जोड़ना :

MS-Excel Office बटन का चयन करें। इससे निम्नलिखित विकल्प दिखेंगे :



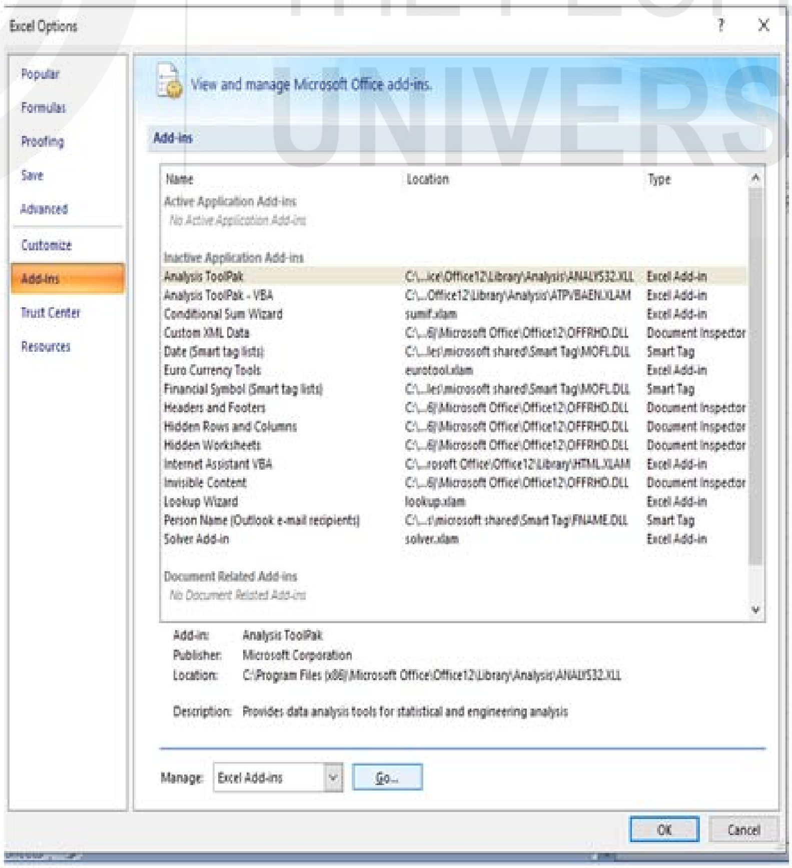
चित्र 3.19 : ऊपरी बाएं भाग में स्थित Office बटन को दबाने पर सामने आने वाले विकल्पों की सूची

कृपया नीचे की ओर दिए Options बटन को दबाएं। ऐसा करने से Excel Options का एक डायलॉग बॉक्स खुल जाएगा। इसके बाएँ फलक (pane) में ***Add-Ins*** को चुनें (चित्र 3.20 देखें) तथा OK बटन दबाएँ जो कि नीचे की ओर दाएं कोने पर स्थित बटनों में से हैं (परंतु इस चित्र में नहीं दिख रहा)।

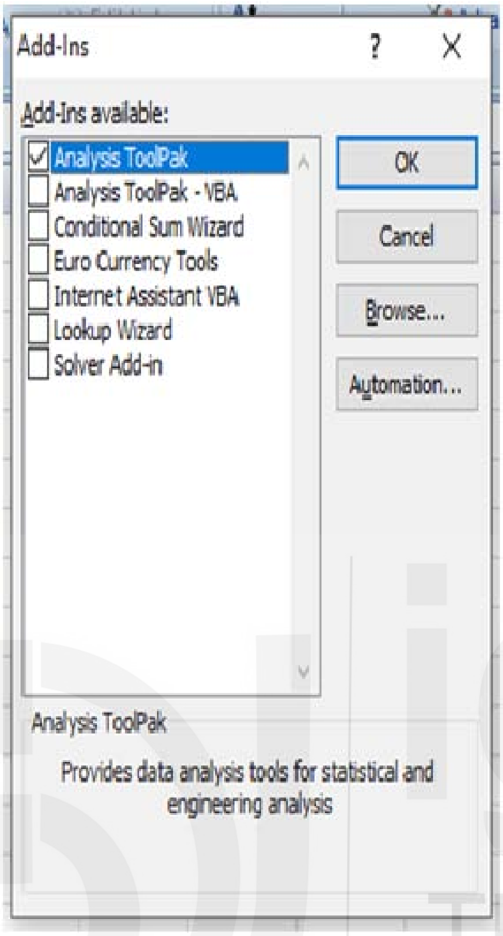


चित्र 3.20 : Add-Ins विकल्प

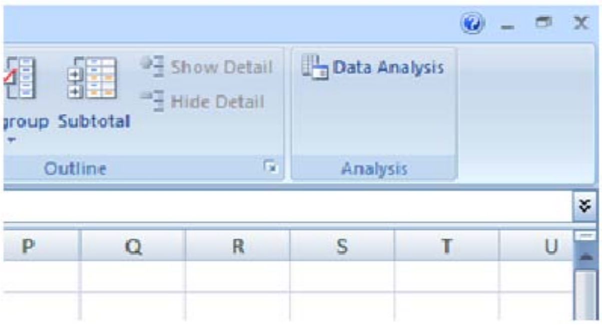
डायलॉग बॉक्स निम्न रूप में परिवर्तित हो जाएगा :



कृपया ध्यान दें कि Analysis Tool **पैक** हाईलाइट हो गया है तथा ***Manage :*** लिस्ट बॉक्स में ***Excel Add-Ins*** चुना गया है। इस लिस्ट बॉक्स के साथ दिए बटन ***Go…*** को दबाएँ। ऐसा करने पर निम्नलिखित डायलॉग बॉक्स सामने आ जाएगा :



Analysis Tool Pak से पहले दिए चैक बॉक्स को चैक करें (चुनें) तथा OK बटन को दबाएँ आप देखेंगे कि डेटा टैब पर डेटा विश्लेषण (Data Analysis) विकल्प खुल जाते हैं जैसा कि ऊपर दिखाया गया है। इन विकल्पों को हम इस पाठ्यक्रम में आगे आने वाली इकाइयों में प्रयोग करते हैं।



चित्र 3.21: डेटा विश्लेषण विकल्प के साथ डेटा टैब

3.7 सार-संक्षेप

इस इकाई में, हमने पाठकों का परिचय सांख्यिकीय सॉफ्टवेयर पैकेजिस् की कुछ महत्त्वपूर्ण विशेषताओं से करवाया। विभिन्न सांख्यिकीय पैकेजिस् में विभिन्न विकल्पों का उपयोग करने की प्रक्रिया भिन्न हो सकती है। इस इकाई के लिए हमने डेटा

हैंडलिंग के कुछ महत्त्वपूर्ण पहलुओं को दर्शाने के लिए हमने MS Excel का प्रयोग किया है। हमने इस इकाई में सूत्रों तथा फलनों की व्याख्या की है जो कि प्रत्येक सांख्यिकीय पैकेज का हिस्सा होते हैं। इस इकाई में, Excel के कुछ सरल फलनों के बारे में चर्चा की।

चार्ट सांख्यिकीय विश्लेषण में अत्यंत उपयोगी सिद्ध होते हैं। इस इकाई में, सरल चार्ट बनाने की विधि का परिचय दिया गया है। आगे आने वाली इकाइयों में हम सांख्यिकीय पैकेजिस् के पहलुओं पर विस्तार से चर्चा करेंगे।

3.8 अभ्यास प्रश्न

1) 10 चर्चित सांख्यिकीय सॉफ्टवेयरस् के नाम ज्ञात कीजिए। इनके महत्त्वपूर्ण फलनों के बारे में जानकारी प्राप्त कीजिए। प्रत्येक सॉफ्टवेयर की कीमत भी ज्ञात कीजिए।

2) पाँच विभिन्न प्रकार के फलनों की सूची बनाइये। चार फलनों की व्याख्या उनके प्राचलों तथा उदाहरण सहित कीजिए।

3) अपने कंप्यूटर में Excel में Analysis Tool Pack जोड़े।

4) अधिकतर सांख्यिकीय सॉफ्टवेयर में पाए जाने वाले विभिन्न प्रकार के चार्ट्स/आलेखों की सूची बनाइये।

5) एक परीक्षा में 25 विद्यार्थियों द्वारा प्राप्तांक नीचे दिए हैं। परीक्षा में अधिकतम अंक 75 हैं।

39	37	59	61	56	63	38	30	56	50
70	65	59	50	32	58	51	42	25	43
38	30	54	61	44					

   i) इन विद्यार्थियों के प्राप्तांकों वाली एक वर्कशीट बनाइये तथा प्रत्येक विद्यार्थी के प्रतिशत अंक ज्ञात कीजिए।

   ii) ऐसे विद्यार्थियों की संख्या ज्ञात कीजिए जिन्होंने 50% से अधिक अंक प्राप्त किए हैं।

   iii) नीचे दिए नियम के आधार पर प्रत्येक विद्यार्थी का ग्रेड ज्ञात कीजिए : 80% और उससे अधिक : A+, 60% तथा उससे अधिक परंतु 80% से कम : B+, 60% से कम : C+.

   iv) ग्रेड आबंटन का Pie चार्ट बनाइये।

3.9 आगे अध्ययन के लिए सुझाव

सांख्यिकीय सॉफ्टवेयर निरंतर विकसित हो रहे हैं। सॉफ्टवेयर निर्माता (Developer) उनमें नित नई विशेषताएँ उपलब्ध करवा रहे हैं। कृत्रिम बुद्धि (Artificial Intelligence) और मशीन लर्निंग उपस्कर के मुख्यधारा में क्रमिक प्रयोग से सांख्यिकीय विश्लेषण के और अधिक बेहतर उपस्कर भविष्य में उपलब्ध होंगे। पाठकों

के लिए किसी भी सॉफ्टवेयर के बारे में अपनी जानकारी को बढ़ाने का सबसे अच्छा तरीका यह है कि सॉफ्टवेयर में उपलब्ध Help का उपयोग करेंगे। अधिकतर सॉफ्टवेयरों से Help सामान्यतः YouTube पर वीडियो के माध्यम से भी उपलब्ध होती हैं। सांख्यिकीय विश्लेषण के लिए बड़े पैमाने पर निःशुल्क ऑनलाइन पाठ्यक्रम (MOOC Massive Online Courses) उपलब्ध हैं।

3.10 बोध प्रश्नों के उत्तर अथवा संकेत

बोध प्रश्न 1

1) बहुत से अच्छे सांख्यिकीय सॉफ्टवेयर मौजूद हैं। इनमें शामिली हैं : आई.बी.एम. का एस.पी.एस.एस. (SPSS), R प्रोग्रामिंग, MATLAB, Octave, SAS, Minitab आदि आदि। इन सभी के आँकड़ों के विश्लेषण के लिए अच्छे अंतराफलक हैं, ये चार्ट/चित्रांकन की सुविधा भी प्रदान करते हैं जिनके आधार पर विवरणात्मक, प्राचल आधारित और गैर-प्राचल परीक्षण एवं प्रतिमान रचना संभव होती है। आप इनके अभिलक्षणों की जानकारी www पर पा सकते हैं।

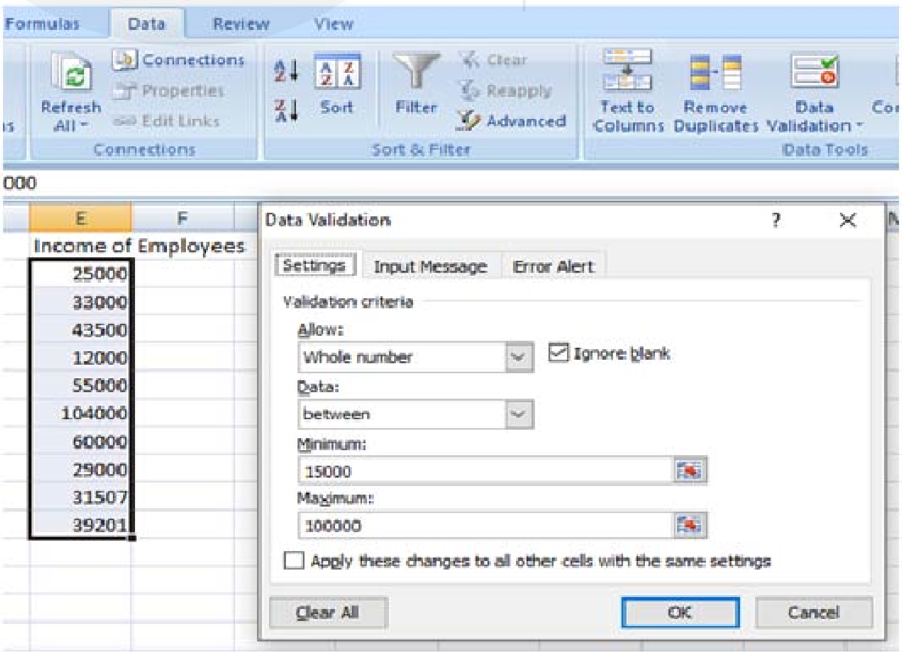
2) विश्लेषण का सबसे निर्णायक अंग आँकड़ों की सटीकता है। यहाँ सटीकता का अभाव गलत परिणामों और निष्कर्षों का कारण बन जाता है। इससे बहुत-सा समय भी बर्बाद होता है।

3) इन कदमों का अनुसरण करें :

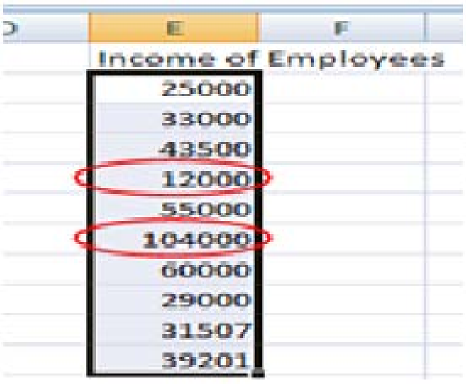
- दिए गए आय के आँकड़े एक स्तंभ (कॉलम) में दर्ज करें;
- इस प्रकार दर्ज आय के आँकड़ों को Select कर एक आँकड़ा सत्यापन सूत्र की रचना करें :

  Data → Data Tools → Data Validation → Data Validation...

- आँकड़ा वैधता (Data Validation) डायलॉग बॉक्स में निम्न चित्र के अनुसार मान चुनकर OK बटन दबाएँ :



- अब चुनें : Data → Data Tools → Data Validation → Circle Invalid Data
- यह उन आँकड़ों को वृत्तांकित कर देगा जो पूर्णांकित नियम की अवमानना कर रहे हैं : देखें निम्न चित्र :



बोध प्रश्न 2

1) आप जिस पैकेज का प्रयोग कर रहे हैं उसके help को इस्तेमाल करते हुए इस तालिका को भरें :

फलन	विन्यास (Syntax)	उद्देश्य
IF	IF (< शर्त या तार्किक वाक्य >, [मान यदि शर्त सही हो]; [मान यदि शर्त सही नहीं हो]	IF फलन का प्रयोग शर्त के अनुसार मान परीक्षण एवं नियत करने के लिए किया जाता है।
VLOOKUP	--	--
MAX		
MIN		
COUNTIF		
SUM		

2) एक कार्यपटल (Worksheet) में सूत्र होने चाहिए – ये सूत्र ही सुनिश्चित करते हैं कि किसी कोष्ठक की प्रविष्टि बदलने पर मानों का पुनः मूल्यांकन भी स्वतः ही हो जाएगा। सूत्रों से कार्यपटल का प्रयोग करना आसान हो जाता है। परम रूप में स्थान निर्धारण उस समय विशेष उपयोगी होता है जब मिलते-जुलते सूत्र विभिन्न कोष्ठकों में कॉपी किए जाते हैं और कुछ कोष्ठक सभी सूत्रों में प्रयोग होते हों। फलन आपको जटिल कार्य सरलता से करने में सहायक रहते हैं। उदाहरण के लिए, VLOOKUP, COUNTIF आदि।

3) आप अपना समाधान/हल स्वयं निकाल सकते हैं। एक संभव हल इस प्रकार है:

Q21 fx

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1		Marks of Student in 5 Subjects						Percentage in Every Subject						TOTAL	TOTAL %	Successful
2	Subject No	1	2	3	4	5		1	2	3	4	5				NOT Successful
3	Maximum Marks	50	50	50	50	50	Min Pass %	40%	40%	40%	40%	40%		250		
4	Student Number															
5	1	30	22	33	23	45		60%	44%	66%	46%	90%		153	61.20%	Successful
6	2	40	43	22	33	43		80%	86%	44%	66%	86%		181	72.40%	Successful
7	3	34	32	47	32	34		68%	64%	94%	64%	68%		179	71.60%	Successful
8	4	45	30	41	44	39		90%	60%	82%	88%	78%		199	79.60%	Successful
9	5	49	21	23	32	38		98%	42%	46%	64%	76%		163	65.20%	Successful
10	6	21	15	50	39	35		42%	30%	100%	78%	70%		160	64.00%	NOT Successful
11	7	24	29	43	41	19		48%	58%	86%	82%	38%		156	62.40%	NOT Successful
12	8	27	45	32	21	24		54%	90%	64%	42%	48%		149	59.60%	Successful
13	9	34	34	25	41	32		68%	68%	50%	82%	64%		166	66.40%	Successful
14	10	23	36	27	27	22		46%	72%	54%	54%	44%		135	54.00%	Successful

उपर्युक्त में प्रयोग किए गए सूत्र इस प्रकार हैं :

B3 fx 50

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1		Mark						Percentag						TOTAL	TOTAL %	Successful
2	Subject	1	2	3	4	5		1	2	3	4	5				NOT Successful
3	Maxim	50	50	50	50	50	M	0.4	0.4	0.4	0.4	0.4		=SUM(B3:F3)		
4	Studen															
5	1	30	22	33	23	45		=B5/B$3	=C5/C$3	=D5/D$3	=E5/E$3	=F5/F$3		=SUM(B5:F5)	=N5/N$3	=IF(AND(H5>=$H$3, I5>=$I$3, J5>=$J$3, K5>=$K$3, L5>=$L$3), $P$1, $P$2)
6	2	40	43	22	33	43		=B6/B$3	=C6/C$3	=D6/D$3	=E6/E$3	=F6/F$3		=SUM(B6:F6)	=N6/N$3	=IF(AND(H6>=$H$3, I6>=$I$3, J6>=$J$3, K6>=$K$3, L6>=$L$3), $P$1, $P$2)
7	3	34	32	47	32	34		=B7/B$3	=C7/C$3	=D7/D$3	=E7/E$3	=F7/F$3		=SUM(B7:F7)	=N7/N$3	=IF(AND(H7>=$H$3, I7>=$I$3, J7>=$J$3, K7>=$K$3, L7>=$L$3), $P$1, $P$2)
8	4	45	30	41	44	39		=B8/B$3	=C8/C$3	=D8/D$3	=E8/E$3	=F8/F$3		=SUM(B8:F8)	=N8/N$3	=IF(AND(H8>=$H$3, I8>=$I$3, J8>=$J$3, K8>=$K$3, L8>=$L$3), $P$1, $P$2)
9	5	49	21	23	32	38		=B9/B$3	=C9/C$3	=D9/D$3	=E9/E$3	=F9/F$3		=SUM(B9:F9)	=N9/N$3	=IF(AND(H9>=$H$3, I9>=$I$3, J9>=$J$3, K9>=$K$3, L9>=$L$3), $P$1, $P$2)
10	6	21	15	50	39	35		=B10/B$3	=C10/C$3	=D10/D$3	=E10/E$3	=F10/F$3		=SUM(B10:F10)	=N10/N$3	=IF(AND(H10>=$H$3, I10>=$I$3, J10>=$J$3, K10>=$K$3, L10>=$L$3), $P$1, $P$2)
11	7	24	29	43	41	19		=B11/B$3	=C11/C$3	=D11/D$3	=E11/E$3	=F11/F$3		=SUM(B11:F11)	=N11/N$3	=IF(AND(H11>=$H$3, I11>=$I$3, J11>=$J$3, K11>=$K$3, L11>=$L$3), $P$1, $P$2)
12	8	27	45	32	21	24		=B12/B$3	=C12/C$3	=D12/D$3	=E12/E$3	=F12/F$3		=SUM(B12:F12)	=N12/N$3	=IF(AND(H12>=$H$3, I12>=$I$3, J12>=$J$3, K12>=$K$3, L12>=$L$3), $P$1, $P$2)
13	9	34	34	25	41	32		=B13/B$3	=C13/C$3	=D13/D$3	=E13/E$3	=F13/F$3		=SUM(B13:F13)	=N13/N$3	=IF(AND(H13>=$H$3, I13>=$I$3, J13>=$J$3, K13>=$K$3, L13>=$L$3), $P$1, $P$2)
14	10	23	36	27	27	22		=B14/B$3	=C14/C$3	=D14/D$3	=E14/E$3	=F14/F$3		=SUM(B14:F14)	=N14/N$3	=IF(AND(H14>=$H$3, I14>=$I$3, J14>=$J$3, K14>=$K$3, L14>=$L$3), $P$1, $P$2)
15																

अभ्यास प्रश्न

1) आपको सर्च इंजन का प्रयोग कर WWW से इस जानकारी को जाँचना होगा। यह सूची बदलती रहती है।

2) आप जिस सॉफ्टवेयर का प्रयोग कर रहे हैं उसका 'help' देखें। अनेक प्रकार के फलन होते है, जैसे कि math, logical आदि।

3) इकाई के अंतिम भाग में संपूर्ण प्रक्रिया बताई गई है। आपको बताए गए कदमों का अनुसरण करना है।

4) आप जिस सांख्यिकी सॉफ्टवेयर का प्रयोग कर रहे हैं उसका विवरण देखें।

5) यह प्रश्न आप स्वयं हल करें।