



Notes

BASIC STATISTICS

Numbers play an important role in our lives. We use numbers everyday to know the temperature, weather conditions, our height, weight, blood pressure etc. These numbers also help us in comparing our selves with others. For example, one person may be taller than the other, but may have lesser weight. Two individuals will not exhibit the same attributes in the same strength when measured on a scale. This makes understanding of basic statistics crucial for us.

In the previous lesson you have learnt about the various methods of psychological research. Some of these are observation, interview, case study etc. The skills involved in carrying out research were also explained. You learnt and reflected on the various ethical concerns that must be kept in mind while doing research. Initially the data is raw and unorganized. We then apply statistics to enhance the understanding of the results. Statistics is a branch of Mathematics which deals with selection, classification, tabulation, organisation and analysis of numerical facts. The data is used for description, comparison and prediction. For statistics to give meaningful results the objective of the research should be clear. Statistics has a wide application. It is indispensable in Psychology as it consists of various techniques to organize, summarize, and interpret the raw data. It is used by researchers to conduct various studies in order to examine and understand the relationship between various variables. Psychologists rely on statistical applications to establish the reliability, validity and norms of a test. It is also used to make the collected information and data more meaningful and comprehensive.

Foundations of Psychology



Notes



LEARNING OUTCOMES

After studying this lesson, learner :

- defines the concept of central tendency;
- explains the purpose and function of three types of central tendency (i.e., mean, median and mode);
- enlists the characteristics of mean, median and mode;
- calculates the mean, median and mode from the given data;
- explains the organization of raw data and presenting it in the form of a frequency distribution;
- describes that data can be presented in a graphical manner, through bar diagram, histogram, frequency polygon and pie chart;
- enumerates the key features of bar diagram, histogram, frequency polygon and pie chart;
- enlists the limitations of graphical representation of data;
- explains the difference between raw data and grouped data;
- describes the various types of grouping of the data;
- enumerates the relevant points to be kept in mind while graphically representing a data;
- draws a bar diagram, histogram, frequency polygon and a pie chart for a frequency distribution; and
- explains the importance and limitations of graphically representing the data.

5.1 MEASURES OF CENTRAL TENDENCY

A Central tendency reflects the significance of the entire data by accurate calculation of the average, or the central score, for a set of scores. It is a representative value of a statistical series of data. The three measures of central tendency are explained as follows.

5.1.1 Mean

It is the simple average of all items in a series. Therefore it is calculated as the sum of all observations, divided by the number of observations. For example, Raghav scored 83, 90, 82, 87, 88, 92 on six tests. We will add the scores $83 + 90 + 82 + 87 + 88 + 92$. This is equal to 522. Now, divide the sum, ie 522, by the number of observations, ie 6.

$$83+90+82+87+88+92=522$$

$$522/6=87$$

Therefore the mean is 87. You will note that the mean is very sensitive to extreme scores. If any score is extremely low or high the mean will also fall or rise respectively. The formula for mean is stated as follows:

$$X=\Sigma F/N$$

Where X =mean, ΣF = sum of all frequencies, and N is total number of entries

Let us take another example.

Neha was able to recall the following number of words, from the list of words shown in various languages.

Table 5.1.1

SUBJECT	NO. OF WORDS LEARNT
Hindi	20
Sanskrit	20
Bengali	18
English	18
Assamese	24

Thus, the average number of words recalled is calculated as follows:

Formula is $X=\Sigma F/N$

$$20+20+18+18+24=100$$



Notes

Foundations of Psychology



Notes

$$100/5=20$$

The average number of words recalled is 20.

The mean has the following characteristics

- It is simple and easy to understand.
- It is also easy to calculate.
- It takes all observations into consideration.
- It is affected very much by extreme values.
- The average of scores may not reflect whether the performance of a person has improved or declined. To understand this, the entire data needs to be viewed. For example, two students obtain the marks as follows in various tests.

	TEST 1	TEST 2	TEST3	TEST4	MEAN
Student A	30	49	55	70	51
Student B	68	54	52	30	51

It is observed that the average score of both students A and B is the same. But, we can see that student A's performance is improving whereas student B's performance is declining.

- The mean value may not figure in the series at all. In the above example the average score of 51 does not exist in any of the observations.

5.1.2 Median

Median is a measure of central tendency which reflects a score in a distribution that divides the distribution into half. It indicates the midpoint of an array of scores in a distribution. In a research study, the number of hours of sleep of each student of the dormitory was noted. Refer to Table 5.1.2 given below.

Table 5.1.2a

Student's name	Number of hours of sleep
A	7
B	10

C	6
D	8
E	8
F	9
G	7
H	7
I	6
J	8



Notes

Arranging the scores in descending order we will get the following distribution.

Table 5.1.2b

Name of student	Hours of sleep in descending order
B	10
F	9
E	8
D	8
J	> 8
G	> 7
A	7
H	7
C	6
I	6

The median is calculated by adding the centre most values and dividing it by 2 (Hence we calculate median as $8+7=15$, which is divided by 2). Thus the median is the score of 7.5 that divides the distribution into half. Here the number of entries is 10. Thus we

Foundations of Psychology



Notes

have the median lying between the 5th and the 6th entry. If we had data of only 9 students, the median value would be the 5th entry. Since 9 is an odd number, it becomes easy to calculate the median.

Activity 5.1.1

Draw and cut paper in the shape of squares of seven different sizes. Now arrange them in descending sizes, start by putting the largest on top, then smaller, and the smallest at the bottom. Now try to locate the median.

Interesting fact about the median

You can quickly find the location of the median by using the expression $n+1$ divided by 2.

Median = $n+1/2$

The letter n is the total number of data values in the distribution. If n is an odd number, the median is the middle value of the ordered data (ordered smallest to largest). In the previous example if $n=9$ (which is an odd number), then the median value is $9+1/2 = 10/2 = 5$ th entry. If n is an even number, the median is equal to the two middle values added together and divided by two after the data has been ordered.

Let us take another example, the number of items recalled from the given list of words is as follows:

Table 5.1.3

Name of the person	Number of words recalled
M	18
N	24
O	22
P	30
Q	30
R	10

S	17
T	5
U	28

Arranging the scores from highest to lowest we get 22 as the median as shown in the table below:

Table 5.1.4

Name	Words recalled
P	30
Q	30
U	28
N	24
O	> 22
M	18
S	17
R	10
T	5

The following are the characteristics of median.

- It is a simple measure of central tendency.
- It is not affected by extreme scores.
- It is easy to understand and calculate.
- It is possible to derive the median value even when the data is incomplete.
- In case of even numbers of observations, the median cannot be determined directly.



Notes

Foundations of Psychology



Notes

5.1.3 MODE

Mode is an important measure of central tendency. The mode has the highest frequency in the distribution. The literal meaning of the word 'mode' refers to something that is in vogue or fashion. Hence, it is determined by the maximum number of entries/scores. It is also a measure of reference and popularity.



ACTIVITY

Fill in the blanks with an approximate value.

- Most women in your village/town have a height of _____ cms .
- _____ is the fruit that most of the people in your village like to eat.
- The brand of mobile phones that most people bought is _____.
- The most watched TV serial is _____.
- The most worn colour of shoes is _____.

Once you have filled these, you will have an understanding of the concept of mode. For example, as shown in the table below, the number of marks that 10 children get in Mathematics is 60, 48, 50, 60, 39, 50, 60, 27, 60, 42. The mode will be 60, as the maximum number of children scored 60.

Table 5.1.5

Serial number	Name of child	Marks scored
1	Debjit	60
2	Arav	48
3	Raghu	50
4	Rekha	60
5	Namita	39
6	Taran	50
7	Nafeeza	60

8	Indra	27
9	Ramesh	60
10	Tanya	48

Now, you will note that the number of children scoring a particular marks is as follows:

Table 5.1.6

Marks	Number of children scoring the particular marks
27	1
39	1
48	2
50	2
60 (is the mode)	4

Characteristics of Mode

- It is easy to calculate the mode. Just by inspection we can know the mode.
- It is easy to understand.
- It is applied in our daily lives.
- It is not affected by extreme values.



ACTIVITY 5.1.3

Make ten paper cards. Choose any nine numbers. Write any of these numbers in bold on each of these 10 paper cards. You can repeat any one number, as you have only 9 numbers to be written on ten cards. Invite ten of your friends, siblings and/or cousins for this activity. Distribute the number card each to your friends and cousins. Now, let each person hold the number, so that it is visible to everyone. Play music and ask them to stand in descending order. Now find the mean, median and mode of the numbers that they are holding.



Notes

**Foundations of
Psychology**



Notes



INTEXT QUESTIONS 5.1

1. State whether the following statements are True or False.
 - a. The mean ,median and mode can never be the same.
 - b. When we calculate a student's average marks,we are calculating the mean.
 - c. Any change in scores will change its mean value but may or may not change the mode.
 - d. The mode can be known by mere inspection.
 - e. The mean can be known by mere inspection.

2. 10 people from Areema's town walk to the old age home everyday to help the senior citizens of the town. The distance travelled by each person is shown in the following table . Calculate the value of the mean, median and mode of the distance travelled.

Person's name	Distance walked in kms (To help the senior citizens in old age home)
Ahmad	2
Bela	4
Celia	2
Deepak	3
Emily	3
Farida	2
Gauri	1
Hemant	3
Isha	3
Jasbir	3

5.2 REPRESENTATION OF DATA

In the previous module, you have learnt about the mean, median and mode. They are the measures of central tendencies. Understanding of measures of central tendencies also helps us in our day to day life situations. By knowing the average number of hours to be devoted to study and sleep one can not only make a plan, but also effectively follow it. This data can also be shown graphically. This module will explain how data can be shown in the form of a frequency distribution, and further in a graphical or pictorial form.

5.2.1 Raw data and data grouping

Researchers collect data adopting various methods of research. Data refers to information in the form of numerical figures or a set of given facts. Data is collected by the researcher which is in an un-organised form. This is called raw data. This data needs to be organised and interpreted. The raw data needs to be divided into various classes/groups to enhance our understanding of the same. It can then be used for further algebraic and graphical representation. Items/scores are arranged into various groups/classes in a tabular manner, on the basis of similarities. This condensed data helps to know the salient features of the data. You will now study various types of data grouping.

5.2.1.1 Individual Series

Individual series is in the form of collected data which is raw. For example, the number of trees planted by 5 students is 6, 2, 1, 3, 4. Another example of individual series is as follows.

The marks of 10 students in Sanskrit are 92, 80, 70, 88, 75, 60, 55, 90, 85 and 92. Thus, in an individual series items are listed singularly. Hence, the name 'individual series'. This can then be arranged in either ascending or descending order. When the number of items is large, this organisation of data is tedious and lengthy. Computers are a big aid in arranging large number of items in ascending or descending order for better understanding of data.

5.2.1.2 Discrete Series

When the frequency of observations is taken into consideration in order to organise



Notes

Foundations of Psychology



Notes

the data, a discrete series can be prepared. A frequency table that is prepared indicates the number of scores within each group. The number of occurrences of observations for each item is noted in the form of tally, and a frequency table is prepared. Discrete series is also known as a frequency table. In this there are no class intervals, and particular item in the series is numbered. Let us understand this through the following example:

20 students of class 12 secured the following marks

10, 15, 10, 12, 16, 10, 16, 17, 15, 17, 18, 20, 15, 17, 20, 20, 15, 17, 20, 17. The table below presents this data in the form of a frequency table. Items are arranged in a descending order. The tally bar represents every single entry. The frequency is noted in front of each item to signify the number of times that item occurs in the series.

Marks	Tally bars	Frequency
10	///	3
12	/	1
15	////	4
16	//	2
17	###	5
18	/	1
20	////	4

5.2.2 Frequency Distribution For Grouped Data

In this frequency distribution, data is classified into different classes with a range. This range determines the class interval. For example 0-5, 6-10 etc. These are the class intervals in which all observations falling within the range are shown. Each item in the series is written against a particular class interval. The number of times an observation occurs is shown as frequency against the class interval. The following table offers an illustration of a frequency distribution with class intervals. For example:

Exposure to various social media sites influences a person's attitude towards physical activities like swimming, walking etc. To know this a researcher noted the number of people aged 15 - 40 years who had facebook accounts. This is shown in Table 5.2.1 which is a frequency distribution.

Table 5.2.1 Frequency distribution for
Number of people on facebook in various age groups.

Class Interval	Frequency
Age in years	Number of people on facebook
15-20	12
20-25	38
25-30	34
30-35	10
35-40	6
Total	100

From the above table you can note that in the class interval 15-20, number of people on Facebook is 12, whereas the same is 38 for range 20-25 years. This shows that the maximum number of people using facebook 38, which is in age group 20 to 25 years old. The frequency (ie number of people having facebook account) can thus be clearly seen from this frequency distribution. This helps us in comparing the various age groups with respect to facebook accounts.



INTEXT QUESTIONS 5.2

1. State whether the following statements are True or False.
 - a. Raw data is very meaningful.
 - b. The data collected by the researcher is in a raw format.
 - c. Raw data can be arranged in both ascending as well as descending order.
 - d. Discrete series is also known as individual series.
2. Fill in the blanks.
 1. The number of times a particular observation occurs is known as its _____.



Notes

Foundations of Psychology



Notes

2. _____ is the group into which raw data is arranged .It consists of the lower limit and the upper limit of the class .

5.2.3 GRAPHICAL REPRESENTATION OF DATA

A frequency distribution can be represented by means of lines, curves or bars on a graph. This is a widely used technique to represent memory, learning, intelligence and many more psychological concepts and relationships. The two axis, i.e. X and Y axis are perpendicular to each other. Showing the frequency against the scores on the graph makes the data meaningful and simple to understand. Just by looking at the graph, the meaning of the observations is clear. Graphical representation acts as a visual aid and the image can be recalled with comparative ease .It also facilitates the comparison of two or more frequency distributions.

Some relevant points to be kept in mind while constructing a graph are as follows:

- The heading should be precise and suitable and should explain the nature of information presented in the graph.
- The points should be plotted from left to right on the X-axis ,and upwards from bottom to top on the Y-axis.

Some of the important types of graphical representation are considered below.

5.2.3.1 Bar Diagram

These are columns or rectangles which represent the data, or the value of the variable, in the form of bars. The following should be kept in mind while presenting the bar diagram

- A bar diagram is drawn on an X and Y axis.
- The length or height of the bars is drawn in accordance to the value of the variable. However, the breadth of each bar remains the same.
- Bars can be both vertical or horizontal. However, they are usually used in the vertical form. Bars should be equidistant from each other.
- The bars may be shaded to make them more attractive.
- Two or more sets of data may be shown in a bar diagram simultaneously. This is known as a multiple bar diagram.

For example

The following frequency distribution can be represented through a bar graph in the following manner.

Table 5.2.2 Frequency distribution for
Number of people on facebook in various age groups.

Class Interval	Frequency
Age in years	Number of people on facebook
15-20	12
20-25	38
25-30	34
30-35	10
35-40	6
Total	100

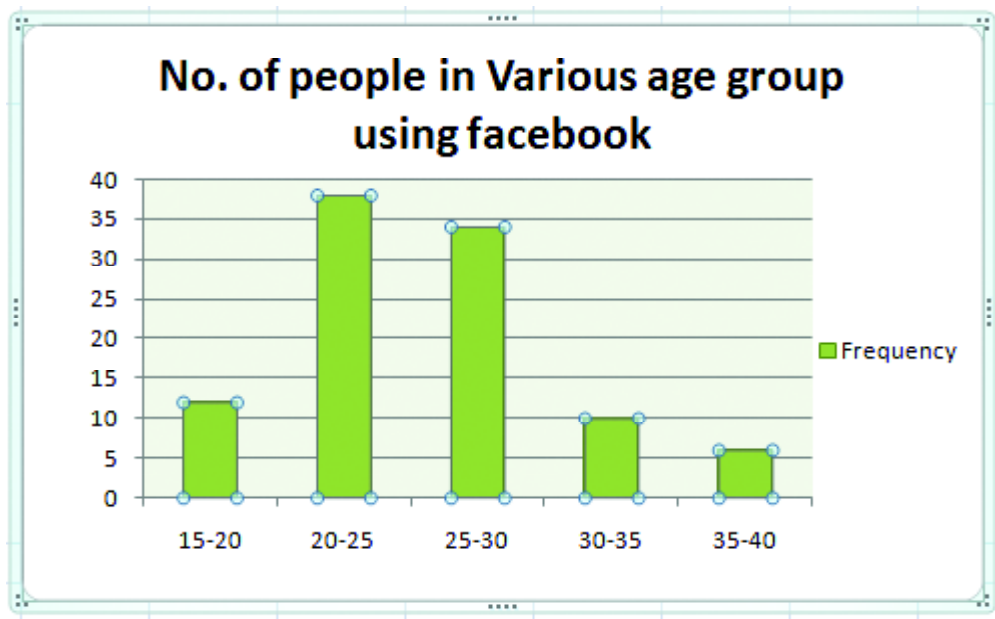


Fig 5.1 : Bar diagram showing no. of people in various age groups using facebook

Notes



Foundations of Psychology



Notes



INTEXT QUESTIONS 5.4

1. State whether the following statements are True or False.
 - a. The height of bars in a bar diagram need not be equal .
 - b. When more than two sets of data is to be presented simultaneously we use the multiple bar diagram.
2. Choose the correct option from the options given below and fill in the blanks.
 - i. A bar diagram is _____.
 - a. a one-dimensional diagram.
 - b. a two dimensional diagram .
 - c. a diagram with no dimension.
 - d. none of the above
 - ii. Bars in a bar diagram can be _____.
 - a. vertical or horizontal
 - b. only drawn as lines
 - c. only horizontal
 - d. only vertical

5.2.3.2 Histogram

A histogram represents a grouped or classified data. The X- axis represents the class limits, or the class boundaries. The Y-axis shows the frequencies. The columns or rectangles are drawn by taking the X-axis as the base and the Y-axis to represent the height ,ie frequencies.

A linkmark (~~~~) may be drawn on the X axis ,to show the continuity. (when there is a gap between 0 and the first class interval) The rectangles may be shaded to give a clear understanding to the histogram. The breadth in a histogram is not arbitrary, it is determined by the class size. If the class intervals are discontinuous, they have to be changed to continuous form for a histogram to be drawn. The following frequency distribution of IQ of 30 students is followed by aHistogram representing this data.



Notes

Class interval	Frequency
20-40	1
40-60	3
60-80	4
80-100	10
100-120	10
120-140	2

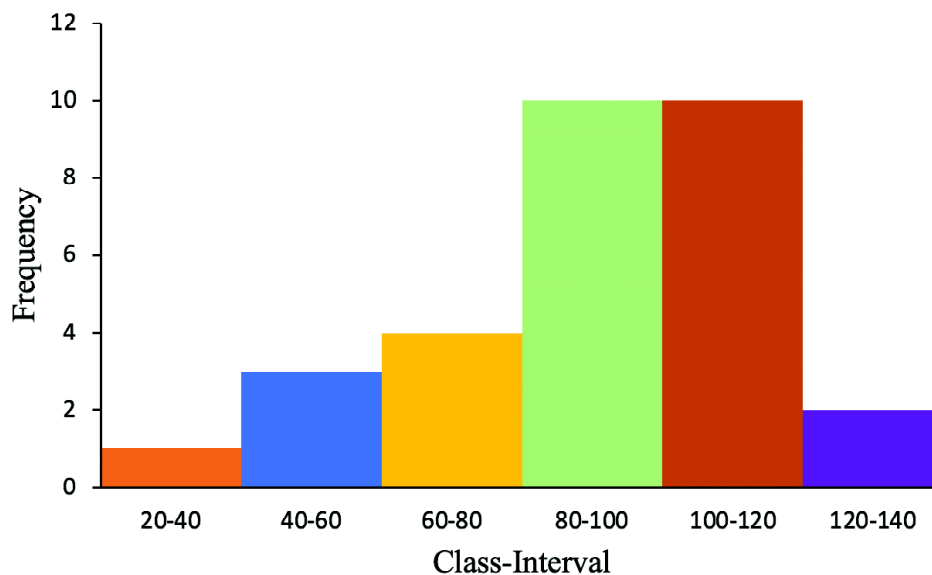


Fig 5.2 : Histogram showing marks distribution of students

5.2.3.3 Frequency Polygon

In a frequency polygon the midpoints of the class intervals are calculated and shown on the X axis. The frequencies are shown on the Y-axis. After plotting the points on the graph, they are joined by straight line segments. Both sides of the frequency polygon are extended to meet the X-axis.

Consider the following example

The following frequency distribution of marks of various students is further represented through a Frequency Polygon by joining the mid points.

Foundations of Psychology



Marks	Number of students
10-20	2
20-30	5
30-40	9
40-50	8
50-60	9
60-70	10
70-80	7

We will first mark the mid points of the class interval.

(Class Interval)

Marks	Mid point	Number of students
10-20	15	2
20-30	25	5
30-40	35	9
40-50	45	8
50-60	55	9
60-70	65	10
70-80	75	7

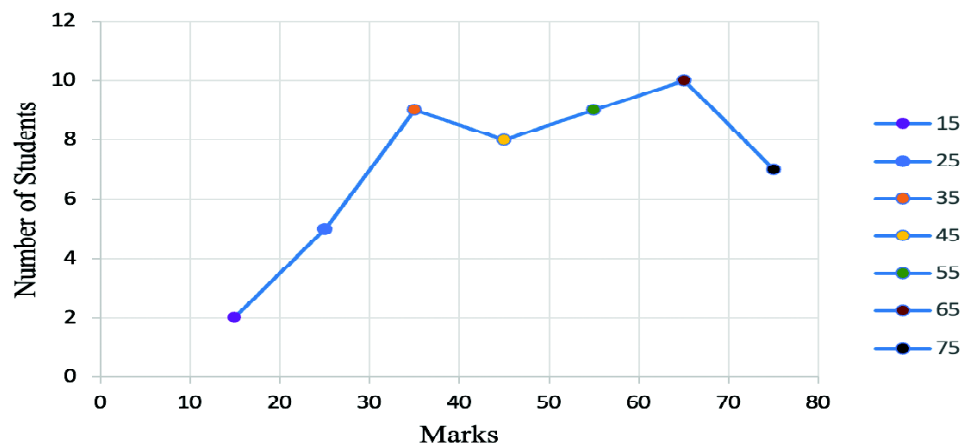


Fig 5.3 : Frequency polygon showing marks distribution of students

**INTEXT QUESTIONS 5.5**

Choose the correct option and fill in the blanks.

- (i) The breadth of columns in a histogram is determined by _____.
- frequency
 - frequency distribution
 - class size
 - all of the above
- (ii) In order to draw a histogram the class interval should be.....
- Large
 - Continuous
 - Discontinuous
 - None of the above
- (iii) In a histogram the frequency is shown as
- a circle
 - a line
 - a curve
 - None of the above
- (iv) In a _____ the mid points of the class intervals are considered.
- Bar diagram
 - Histogram
 - Frequency polygon
 - None of the above
- (v) In order to draw a frequency polygon, _____
- a histogram should be drawn.

**Notes**

Foundations of Psychology



Notes

- b. a bar diagram should be drawn.
- c. Both, histogram and bar diagram, should be drawn.
- d. No other diagram is required.

5.2.3.4 Pie Chart

Pie chart reflects the percentage of frequencies in a particular class interval or a group. A circle is made up of four right angles and is measured as 360°. This 360 degree is also a representation of aggregate of all frequencies. In accordance with the percentage of total frequency each segment of the circle is marked. The pie chart does not show absolute values, it shows percentages.

The steps can be understood from the example below.

There are 40 children in a class . Their interest in various activities is as follows

Area of interest	Frequency (Absolute value)
Music	4
Painting	6
Theatre	8
Sports	16
Poetry	6

Now, the absolute values will be converted to percentage, which is further converted to degrees, to mark the required segment.

Interest	Frequency (Absolute Value)	Value (percentage)	Value (degree)
Music	4	10	36
Painting	6	15	54
Theatre	8	20	72
Sports	16	40	144
Poetry	6	15	54

Interest of students in various activities

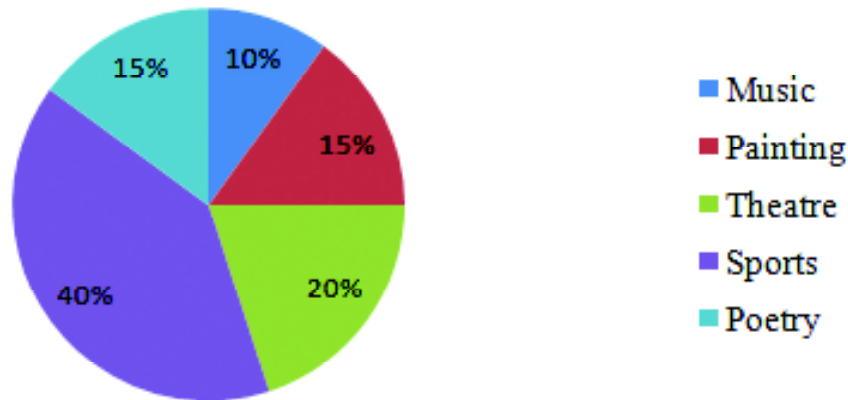


Fig 5.4 : Pie chart showing interest areas of students

Step 1

The absolute values are converted to percentage. For eg, from the table it is observed that four children, from a class of 40 children, like music. This means that 4 upon 40 into 100 is equal to 10% of children like music. Similar calculation is done for all areas of interest.

Step 2

Draw a circle and from the centre point draw perpendicular to find the value of percent score from 360 degree (ie total frequency). We will divide the persons score from hundred and multiply with 360. Thus, 10 children are interested in music which makes it 10 upon hundred into 360 is equal to 36 degrees.

A similar calculation will be done for all domains.

Step 3

Each value in terms of degree is shown on the circle. Hence from the pie chart each component is clearly visible. It can be observed that sports is of interest to 40% percent children, poetry to 15% and music is liked by only 10% of the children. The area in the circle represents each area of interest, in the right proportion.



Notes

Foundations of Psychology



Notes

Limitations of graphical representation

- A small change in the scale of a graph can lead to a lot of difference in structure of the graph. Therefore, a wrong meaning is attributed to the graph.
- Actual values are mostly not clear from the graph. They just give a general tendency of the scores.
- Sometimes the graphical representation is difficult to comprehend. A layman is most likely to misinterpret it.

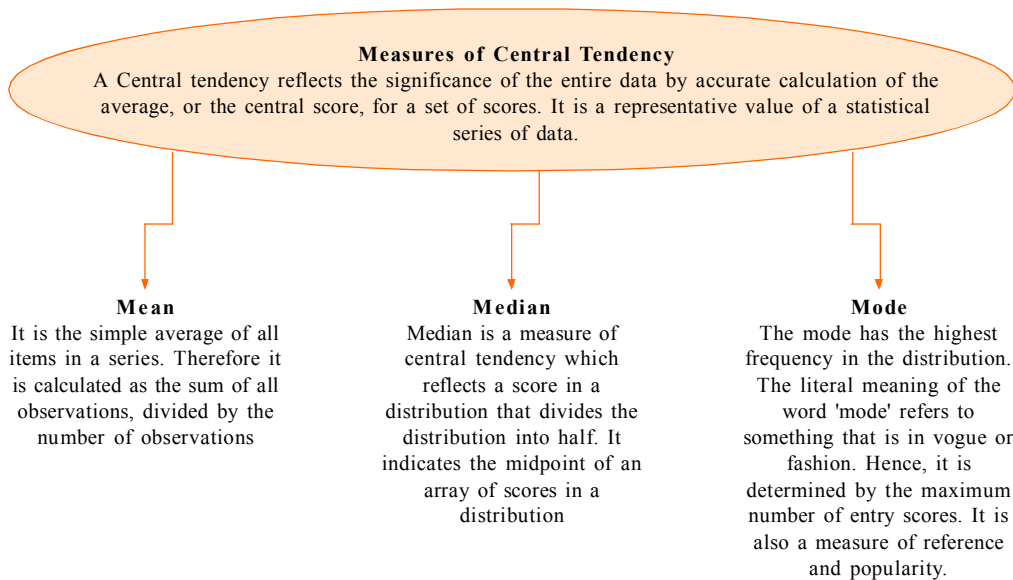


INTEXT QUESTIONS 5.6

- (i) A pie chart depicts the data as a proportional part of a (Fill in the blank).
- (ii) The largest part on the pie chart reflects the minimum frequency.(State whether True or False).



WHAT YOU HAVE LEARNT



**TERMINAL QUESTIONS**

1. Data indicating the number of months a person spends on tours, while holding a job is as follows:

3;15; 16; 4;11; 12; 13

15;16;2;6;4;13;8;11;15,16;6;7

Calculate the mean, median and mode from the above data.

2. An aptitude test was conducted for 10 students. The scores on mathematical comprehension were as follows :

10,8,6,12,8,x,2x, 2,5,7. The mean of the scores is 7. Find the value of x and 2x. Also find the median and mode of the data.

3. By the end of 2021, Facebook had many users in India. The following table shows the number of users in different age groups in a town. Construct a bar graph and pie chart using the following data.

Age groups	Number of Facebook users
0-10	0
10-20	100
20-30	95
30-40	80
40+50	75
50-60	42

4. Twenty-five students who were randomly selected were asked the approximate number of movies they watched every month. The results were as follows.

**Notes**

Foundations of Psychology



Notes

Number of movies	Frequency(number of people)
0-1	5
1-2	8
2-3	6
3-4	3
4-5	1
5-6	2

Draw a frequency polygon and histogram of the above data.

5. The people who visit the neighbouring park are mostly infants, children, college students, working adults, and retirees. The distribution shows the five age groups and the number of people in the park from each age group. Construct a pie chart showing the proportions.

Age groups	Number of people
Infants	22
Children	34
College students	10
Working-age adults	8
Retirees	26

6. What are the important things to keep in mind while presenting a bar diagram?
7. What is a frequency distribution for grouped data and how is it classified into different classes?
8. What is the formula for calculating the mean of a statistical series of data?
9. What are the characteristics of the median as a measure of central tendency? How the median is affected by extreme scores and is it possible to determine the median when the data is incomplete?
10. What are the characteristics of mode as a measure of central tendency and why is it important?

**ANSWERS TO INTEXT QUESTIONS****Foundations of
Psychology****5.1**

- a. False
- b. True
- c. True
- d. True
- e. False

5.2

Mean = 2.6 Median = 3 Mode = 3

5.2.1

- a. False
- b. True
- c. True
- d. False

5.2.2

- a. frequency
- b. Class interval

5.2.3

- a. True
- b. True

5.2.4

- i. a. one dimensional diagram
- ii. a. vertical or horizontal

5.2.5

- i. c. class size
- ii. b. continuous
- iii. d. none of the above
- iv. c. frequency polygon
- v. d. No other diagram is required.

5.2.6.

- i. Circle
- ii. False

MODULE -2

Basic Psychological Processes

To make the learners aware about how information gathered through different sense modalities influences attention and perception, and the laws that govern these processes. The learners will also become aware of how learning and memorisation take place and the influences on these underlying processes. The learners will understand how and why forgetting takes place, along with techniques of enhancing memory. The students will become aware of motivation as extrinsic and intrinsic and how different needs and aspirations the process of achievement and effort. They will also understand how emotions facilitate our well being and will learn to manage emotions better. This module will introduce them to thinking processes and the role of language.

6. Attention and Perceptual Processes
 7. Learning
 8. Human Memory
 9. Motivation
 10. Emotions
 11. Thinking and Problem
-