**Introduction to Statistics**

1. What are the two sources of data?
2. Methods of collecting Primary Data
3. Two meanings of the term Statistics
4. The three forms of graphical representation

**Histogram**

This is a form of representation like the bar graph, but it is used for continuous class intervals.

For instance, consider the frequency distribution representing the weights of 36 students of a class:

|  |  |
| --- | --- |
| **Weights (in kgs)** | **Nos** |
| 30.5-35.5 | 9 |
| 35.5-40.5 | 6 |
| 40.5-45.5 | 15 |
| 45.5-50.5 | 3 |
| 50.5-55.5 | 1 |
| 55.5-60.5 | 2 |
| Total | 36 |

*Draw the histogram*

No gaps in between consecutive rectangles, the resultant graph appears like a solid figure.

This is called a *histogram*, which is a graphical representation of a grouped frequency distribution with continuous classes.

Unlike a bar graph, **the width of the bar plays a significant role in its construction**.

Areas of the rectangles erected are proportional to the corresponding frequencies.

**Data with un-equal class intervals**

A teacher wanted to analyse the performance of students in a mathematics test of 100 marks.

|  |  |
| --- | --- |
| **Marks** | No. of Students |
| 0-20 | 7 |
| 20-30 | 10 |
| 30-40 | 10 |
| 40-50 | 20 |
| 50-60 | 20 |
| 60-70 | 15 |
| 70 – above | 8 |
| Total | 90 |

*Draw a histogram (wrongly)*

**Is it right? What do you see wrong here?**

Adjust the frequency heights such that total area is proportional to the height

Two class intervals require adjustment

1. 20 and 70 – above
2. 20: 7
3. 20: 3.5

70 – above: 9

70- above: 2.67

*Draw a histogram (correctly)*

**Frequency polygon**

There is yet another visual way of representing quantitative data and its frequencies.

This is a polygon.

To see what we mean, consider the histogram represented earlier.

Let us join the mid-points of the upper sides of the adjacent rectangles of this histogram by means of line segments.

|  |  |
| --- | --- |
| Classes | Frequency |
| 140-150 | 5 |
| 150-160 | 10 |
| 160-170 | 20 |
| 170-180 | 9 |
| 180-190 | 6 |
| 190-200 | 2 |

* Start with a space for 130-140
* End with space for 200-210
* Draw the histograms
* Join the mid points

Remove the histogram and present only with polygon.

# Measure of Central Tendency

**Theory**

*What is central tendency?*

Earlier we represented the data in various forms through frequency distribution tables, bar graphs, histograms and frequency polygons.

Now, the question arises if we always need to study all the data to ‘make sense’ of it, or if we can make out some important features of it by considering only certain ***representatives*** of the data.

This is possible, by using measures of central tendency or averages.

*Give examples of Central Tendency in real life.*

**Illustration**

Example: Two students Ramesh and Suresh received their test copies.

The test had five questions, each carrying ten marks. Their scores were as follows

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Question Nos** | **1** | **2** | **3** | **4** | **5** |
| Ramesh | 10 | 8 | 9 | 8 | 7 |
| Suresh | 4 | 7 | 10 | 10 | 10 |

*Who scored higher marks? What basis? Can there be other basis?*

What is the average score?

It is called Mean or Arithmetic mean.

**Median**

Arrange the scores in ascending order now.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Question Nos** | **1** | **2** | **3** | **4** | **5** |
| Ramesh | 7 | 8 | 8 | 9 | 10 |
| Suresh | 4 | 7 | 10 | 10 | 10 |

The middle score is called Median.

**Mode**

Ramesh 8 appears twice

Suresh 10 appears thrice

Which is greater?

The most often appearing score is called Mode.

**Summary**

|  |  |  |
| --- | --- | --- |
|  | Ramesh | Suresh |
| Mean | 42/5 = 8.4 | 41/5 = 8.2 |
| Median | 8 | 10 |
| Mode | 8 | 10 |

**Recap**

The **mean** (or **average**) of a number of observations is the sum of the values of

all the observations divided by the total number of observations.

Median is the middle observation

Mode is the most frequently occurring observation

**Arithmetic Mean**

**Raw Data**

**Example**

5 people were asked about the time in a week they spend in doing social work in their community. They said 10, 7, 13, 20 and 15 hours, respectively. Find the central tendencies of the time in a week devoted by them for social work.