


Descriptive Statistics

Descriptive measures

- The objective is to develop measures that can be used to summarize a data set.
- These descriptive measures are quantities whose values are determined by the data.
- Most commonly used descriptive measures can be categorized as
 - a. **Measures of central tendency:** These are measures that indicate the most typical value or center of a data set.
 - b. **Measures of dispersion:** These measures indicate the variability or spread of a dataset.



Measures of Central Tendency

- The measure of central tendency is a single value that attempts to describe a set of data by identifying the central position within that set of data.
 - As such, measures of central tendency are sometimes called measures of central location.
 - They are also classify as summary statistics.
 - There are three main measures of central tendency:
 1. Mean
 2. Median
 3. Mode
- 

Mean

- The most commonly used measure of central tendency is the mean.
- **Definition** : The mean of a data set is the sum of the observations divided by the number of observations.
- The mean is usually referred to as average.
- Arithmetic average; divide the sum of the values by the number of values (another typical value)
- For discrete observations:
- **Sample mean:** $\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$
- **Population mean:** $\mu = \frac{x_1 + x_2 + \dots + x_N}{N}$

- **Ungrouped or Raw Data :**

If the variable x assumes n values $x_1, x_2 \dots x_n$ then the mean, \bar{x} , is given by,

$$\begin{aligned}\bar{x} &= \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} \\ &= \frac{1}{n} \sum_{i=1}^n x_i\end{aligned}$$

Example : Calculate the mean for 2, 4, 6, 8, 10

Solution :

$$\bar{x} = \frac{2+4+6+8+10}{5} = 6$$

1. 2, 12, 5, 7, 6, 7, 3;
2. 2, 105, 5, 7, 6, 7, 3
3. The marks obtained by ten students in an exam is 68, 79, 38, 68, 35, 70, 61, 47, 58, 66

- **Grouped Data :**

- a) Discrete Series:**

The mean for grouped data is obtained from the following formula:

$$\bar{x} = \frac{\sum fx}{N}$$

where x = corresponding class

f = the frequency of individual class

N = the sum of the frequencies or total frequencies.



1. Example : The following data is the response from 15 individuals.

2, 1, 3, 4, 5, 2, 3, 3, 3, 4, 4, 1, 2, 3, 4

Solution : $\bar{x} = \frac{f_1x_1 + f_2x_2 + \dots + f_nx_n}{n}$

Value(x_i)	Tally mark	Frequency(f_i)	f_ix_i
1		2	2
2		3	6
3		5	15
4		4	16
5		1	5
Total		15	44

Mean=44 /15=2.93

2. Calculate the mean for the data

X:	5	8	12	15	20	24
f:	3	4	6	5	3	2

3. The monthly income of ten families(in rupees) in a certain locality are given below. Calculate the arithmetic average

Family	A	B	C	D	E	F	G
Income(in rupees)	30	70	60	100	200	150	300

b) Continuous Series:

The mean for grouped data is obtained from the following formula:

$$\bar{x} = \frac{\sum fx}{N}$$

Where x = the mid-point of individual class

f = the frequency of individual class

N = the sum of the frequencies or total frequencies.



$$\bar{x} = \frac{f_1 m_1 + f_2 m_2 + \dots + f_n m_n}{n}$$

Class interval	Tally mark	Frequency(f_i)	Mid point(m_i)	$f_i m_i$
30-40		3	35	105
40-50		6	45	270
50-60		18	55	990
60-70		17	65	1105
70-80		4	75	300
80-90		2	85	170
Total		50		2940

$$\text{Average} = \frac{2940}{50} = 58.8.$$

1. Following is the distribution of persons according to different income groups. Calculate arithmetic mean.

Income Rs(100)	0-10	10-20	20-30	30-40	40-50	50-60	60-70
Number of persons	6	8	10	12	7	4	3

2. The following table gives the distribution of the number of workers according to the weekly wage in a company. Obtain the mean weekly wage.

Weekly wage (in Rs.100' s)	0-10	10-20	20-30	30-40
Numbers of workers	5	10	15	18

40-50	50-60	60-70	70-80
7	8	5	3