

DESCRIPTIVE AND PREDICTIVE STATISTICS

Statistics is a mathematical tool to analyse interpret predict the data given.

There are two types

1. Descriptive Statistics
2. Inferential Statistics

DESCRIPTIVE STATISTICS :

Analyse the data within the given

range. TWO TYPES

1. Central Tendency
2. Dispersion

Central Tendency used to find the

middle value of the data.

(i.e) Mean, Median and Mode

Dispersion how long the data is spread out.

Range, Std deviation, Variance and

Quartile deviation.

MEAN : Average Value

Mean is the sum of all the measures divided by the No. of measurements in the set

Also called Arithmetic Mean.

i. Discrete : $\bar{x} = \frac{\sum x}{n}$

ii. Continuous : $\bar{x} = \frac{\sum f(x)}{\sum f}$

PROBLEM :

1. Find the arithmetic mean for the following data 39, 35, 75, 73, 63, 50, 51, 57, 60, 40

$$\bar{x} = \frac{\sum x}{n} = \frac{543}{10} = 54.3$$

$$\bar{x} = 54.3$$

2. The following table shows the frequency for the No. of students per table in 600 colleges. Find the mean of the distribution.

Student x	10	15	20	25	30	35	40	45	50
Frequency f	20	100	150	180	50	40	30	20	10
f(x)	200	1500	3000	4500	1500	1400	1200	900	500

$$\bar{x} = \frac{\sum f(x)}{\sum f}$$

$$\bar{x} = \frac{14700}{600} = 24.5$$

$$\bar{x} = 24.5$$

$$\boxed{\bar{x} = 25} \text{ (approx)}$$

On average $\boxed{25}$ students per teacher.

3. Find the Mean marks of 40 students in Mathematics given in the following table.

Marks	0 - 5	5 - 10	10 - 15	15 - 20	20 - 25	25 - 30
No. of Students f	4	6	10	12	5	3
x	2.5	7.5	12.5	17.5	22.5	27.5
f(x)	10	45	25	210	112.5	82.5

$$\bar{x} = \frac{\sum f(x)}{\sum f} = \frac{585}{40}$$

$$\boxed{\bar{x} = 14.625}$$

4. The mean of 100 items is 46. Later on it was found that an item 16 was misread 61 and another item 43 was 34. It was also found that No. of items was 90 and not 100. Find the correct mean.

$$\text{Old mean value} = 46$$

$$\text{Old } \bar{x} = 46$$

$$\frac{\sum x}{n} = 46$$

$$\text{Substitute or } \sum x = 46 \times 100$$

$$\text{or } \sum x = 4600$$

New

$$\sum x = 4600 - 61 + 16 + 43 - 34 = 4564$$

90

$$\therefore \text{New/ correct mean} = \frac{4564}{90} = 50.71$$

5. Calculate the mean for the data given below

Cls interval	0 - 9	10 - 19	20 - 29	30 - 39	40 - 49
Frequency	10	32	88	70	10
x	4.5	14.5	24.5	34.5	44.5
f(x)	45	464	2156	2415	445

$$\bar{x} = \frac{5525}{210} = 26.30$$

Change Of Scale method

By considering a assumed mean

calculate $d = \frac{x - A}{c}$ where c is the

class interval. Then mean $\bar{x} = A + \frac{cf \sum d}{\sum f}$

Calculate the mean from the following continuous

Weight	100 - 110	110 - 120	120 - 130	130 - 140	140 - 150
No. of persons	4	6	20	32	33

150 - 160	160 - 170	170 - 180
17	8	2

Weight	f	x	f(x)
100 - 110	105	4	420
110 - 120	115	6	690
120 - 130	125	20	2500
130 - 140	135	32	4320
140 - 150	145	33	4785
150 - 160	155	17	2635
160 - 170	165	8	1320
170 - 180	175	2	350
	122		17020

$$\bar{x} = \frac{\sum f(x)}{\sum f}$$

$$= \frac{17020}{122}$$

$$\boxed{\bar{x} = 139.50}$$

Find the mean of the following
discrete data.

x	1 - 5	6 - 10	11 - 15	16 - 20	21 - 25
f	7	10	16	30	24
		26 - 30	31 - 35	36 - 40	41 - 45
		17	10	5	1

Range	f	x	$f(x)$
1 - 5	3	7	21
6 - 10	8	10	80
11 - 15	13	16	208
16 - 20	18	30	540
21 - 25	23	24	552
26 - 30	28	17	476
31 - 35	32	10	330
36 - 40	38	5	190
41 - 45	43	1	43
		120	2440

$$\bar{x} = \frac{\sum f y}{\sum f}$$

$$= \frac{2440}{120}$$

$$\boxed{\bar{x} = 20.33}$$

MEDIUM:

Medium is the middle value in a data which is arranged either in ascending or descending order.

DISCRETE DATA

- i] If No. of observation is odd then the middle value is the medium.
- ii] If No. of observation is even then average of middle two values is the median.

CONTINUOUS:

$$\text{Median} = L + \frac{i}{f} \left[\frac{N}{2} - c \right]$$

Where L \rightarrow the lower limit

$N \rightarrow$ Total frequency

$i \rightarrow$ length of the median class

$f \rightarrow$ Frequency of the median class

$c \rightarrow$ Value previous to the frequency value.

1. Find the median of 25, 16, 45, 60, 54, 12, 18

Sort : 12, 16, 18, 25, 45, 54, 60

4. Calculate

No. of observation is odd.

∴ The median is 25.

2. Find the median of 24, 17, 49, 61, 54, 13, 22, 20

Sort : 13, 17, 20, 22, 24, 49, 54, 61

$$\text{Average of middle two value} = \frac{22+24}{2}$$

$$= 23$$

∴ The median is 23

3. Find the value of the median for the following data.

Marks	10	23	18	38	65	92	40	58
No. of students	8	12	16	12	10	18	4	1
Cumulative Frequency	8	20	36	48	58	76	80	81

$$\text{Here, } \frac{N+1}{2} = \frac{81+1}{2} = \frac{82}{2} = 41$$

The cumulative frequency greater than
41 is 48

∴ The corresponding value is 38,

Marks	No. of students	Cumulative freq
10	8	8
23	12	20
18	16	36
38	12	48
65	10	58
92	18	76
40	4	80
58	1	81

∴ The median is $13\frac{1}{11}$

4. Calculate the median of the marks of 40 students in a subjects given below.

Marks	0 - 5	5 - 10	10 - 15	15 - 20	20 - 25	25 - 30
No. of students	5	7	9	10	6	3
Cumulative frequency	5	12	21	31	37	40

$$\text{Here } \frac{N}{2} = \frac{40}{2} = 20$$

The immediate cumulative frequency greater than 20 is 21

∴ The corresponding value is 10 - 15

$$\text{Median} = l + \frac{i}{f} \left(\frac{N}{2} - C \right)$$

$$= 10 + \frac{5}{9} (20 - 12)$$

$$= 10 + \frac{5}{9} (8)$$

$$= 10 + \frac{40}{9}$$

$$= 14.44$$

Mode:

Mode is the value that occurs most frequently in a set of numbers.

DISCRETE DATA:

the mode is the value of * corresponding to the max frequency.

CONTINUOUS DATA:

$$\text{Mode} = l + \left[\frac{f_1 - f_0}{2f_1 - f_0 - f_2} \right] i$$

$l \rightarrow$ Lower limit

$f_1 \rightarrow$ frequency of modal class

$f_0 \rightarrow$ frequency of class just preceding the modal class

$f_2 \rightarrow$ frequency of class just next to the modal class

$i \rightarrow$ length of the modal class.

1. Find the mode:

i) 45, 36, 28, 42, 45, 40, 44

Descending:

\therefore The mode is 44

ii) 36, 30, 26, 20, 32, 31

∴ No mode

iii) 7, 10, 18, 15, 17, 10, 16, 18, 12

∴ The mode is 10 & 18

2. Find the mode of the frequency distribution

x	1	2	3	4	5	6	7	8
y	4	9	16	25	22	15	7	3

4 occurred 25 times.

∴ The max frequency is 25

∴ Mode is 4

Class	0 - 10	10 - 20	20 - 30	30 - 40	40 - 50
Frequency	10	14	19	17	13

The max frequency is 19

∴ Modal class is 20 - 30

$$l = 20 \quad f_1 = 19 \quad f_0 = 14 \quad f_2 = 17 \quad i = 10$$

$$\text{Mode} = 20 + \left[\frac{19 - 14}{38 - 14 - 17} \right] 10$$

$$= 20 + \left[\frac{5}{7} \right] 10 = 20 + \frac{50}{7} = 27.142$$

RELATIONSHIP BETWEEN MEAN, MEDIAN AND MODE

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

In case of symmetrical distribution

- For a symmetrical distribution if the mean and median are 68 & 62 respectively

Find mode.

$$\text{Mean} - \text{Mode} = 3\text{Mean} - 3\text{Median}$$

$$-\text{Mode} = 3\text{Mean} - 3\text{Median} - \text{Mean}$$

$$= 3 \times 68 - 3 \times 62 - 68$$

$$-M = -50$$

$$\text{Mode} = 50 //$$

DISPERSION:

Refers to how the data value are scattered.

- Range: It is the difference b/w the max and min value in the given data set.

$$\text{Range} = L - S$$

$L \rightarrow$ Max value

$S \rightarrow$ Min value

$$\text{Coefficient of Range} = \frac{L - S}{L + S} \quad | \text{Relative Range}$$

In case of continuous data:

$L \rightarrow$ Highest class interval (middle value)

$S \rightarrow$ Middle value of the lowest class.

10. Calculate the mean, median and mode.

Weight	100-110	110-120	120-130	130-140	140-150	150-160
No. of persons	4	6	20	32	33	17
Cumulative freq.	4	10	30	<u>62</u>	95	102

160-170	170-180
8	2
120	122

Above bar is even.
 $\therefore 122$ is even.

$$\frac{N}{2} = \frac{122}{2} = 61$$

The immediate cumulative frequency greater than 61 is 62.

Corresponding value is 32

$$\text{Median} = l + \frac{i}{j} \left[\frac{N}{2} - c \right]$$

$$= 130 + \frac{10}{32} (61 - 30) = 130 + \frac{310}{32} = 139.69$$

Mode :

$$\text{Mode} = l + \left[\frac{f_1 - f_0}{2f_1 - f_0 - f_2} \right] i$$

Max frequency is 33

∴ Modal class is 140 - 150

$$l = 140, f_1 = 33, f_0 = 32, f_2 = 17, i = 10$$

$$\text{Mode} = 140 + \left[\frac{33 - 32}{2 \times 33 - 32 - 17} \right] 10$$

$$= 140 + \left[\frac{10}{17} \right]$$

$$\text{Mode} = 140.59$$

2. Find median and mode.

X	1-5	6-10	11-15	16-20	21-25	26-30	31-35
F	7	10	16	30	24	17	10
W.M. Freq	7	17	33	63	87	104	114

36-40	41-45
5	1
119	120

$$\frac{N}{2} = \frac{120}{2} = 60$$

Conc. value = 30

$$\text{Median} = 16 + \frac{4}{30} [60 - 43] = 16 + \frac{17}{30}$$

$$= 16 + \frac{108}{30} = 16 + \frac{135}{30}$$
$$= 19.60 \quad : 20.5$$

MODE :

$$\text{Mode} = l + \left[\frac{f_0 - f_o}{2f_1 - f_0 - f_2} \right] h$$

Max freq $\rightarrow 30$

Modal class $\rightarrow 16 - 20$

$l \rightarrow 16$, $f_i = 30$, $f_d = 16$, $2f_1 = 60$, $2f_2 = 43$ $\therefore 16 + \frac{14}{20} \times 4 = 18.80$

$$= 16 + \left[\frac{14}{60 - 16 - 24} \right] h$$

$$= 16 + \left[\frac{14 \times 4}{20} \right] = 16 + \left(\frac{14 \times 5}{20} \right)$$

$$= 16 + \frac{56}{120} = 16 + \frac{70}{20}$$

$$= 18.80 \quad : 19.50$$

1. find the values of range and its coefficient for the following data

6, 8, 5, 10, 11, 12

$$\begin{aligned}\text{Range} &= L - S \\ &= 12 - 5 \\ &= 7\end{aligned}$$

$$\begin{aligned}\text{Coefficient of Range} &= \frac{L - S}{L + S} \\ &= \frac{7}{17} \\ &= 0.41\end{aligned}$$

2. Calculate the range:

Size	20-22	23-25	26-28	29-31	32-34
No	7	9	19	42	34

$$L = 33 \quad S = 21$$

$$\begin{aligned}\text{Range} &= L - S = 33 - 21 \\ &= 12\end{aligned}$$

$$\begin{aligned}\text{Coefficient of Range} &= \frac{12}{54} = 0.22 \\ &= 0.22\end{aligned}$$