STATISTICS: Basics

What is statistics:

- It is a science of Collection, organization, Presentation, Analysis, and Interpretation of data

Collection of Data:

1. Primary data
2. Secondary data

Organization of data

Presentation of Data:

1. Tabular Form
2. Graphical

Analysis of data:

1. Measurement of Central Tendency
2. Dispersion
3. Regression

Interpretation:

1. Conclusion

Descriptive statistics-

- It is a method of collection, organizing, summarizing and presenting data in an informative way
- It explains already known data
- Final result represent by charts, graphs and table
- Mean, mode, median standard deviation and frequency distribution
- Students report card

Inferential statistics-

- To explain the chances of occurrence of an event
- It attempts to reach the conclusion to learn about the population
- Final result based on probability
- It predicting the relationship between death & smoking habit
- It predicting effectiveness of a drug

Data or variable

Qualitative

Which cannot be expressed numerically

Example- fishes in market, intelligence of student, degree in university ,colors of flowers

Quantitative

Which can be expressed numerically

Example-height of students, weight of students ,age of students

Series

Individual
(Only values)

Marks
2
4
6
7
9

Discrete
(values + frequency)

Marks	No. of students
5	4
10	3
15	5
20	8
25	3

Continuous (values + frequency)

Marks	No. of students
5 - 10	4
10 - 15	3
15 – 20	5
20 - 25	8
25 – 30	3

Types of Class Intervals

1. Exclusive
2. Inclusive
3. Less than
4. More than
5. Mid value

Importance of statistics

Government--

- To formulate policy
- Functions of the ministry
- To analyze basic economics problems like poverty, unemployment

Economics--

- Study of consumption, cost
- Economic planning of the best use of natural resources

Business--

- Import, export, transport, banking, advertize, insurance and packing etc

Limitations of statistics

1. Only use in homogeneous data
2. Only use in numerical facts
3. Does not study individuals
4. It can use by experts
5. Statistics law are not exact(predictable)
6. Statistics are not universally applicable
7. It can be misused

Central Tendency

Meaning :

The numerical value of an observation is called central value.
This central values of other observations in the group or cluster is called central tendency

Example-the production manager of a manufacturing firms finds that the daily production of items produced by some workers is

11,9,17,15,12,16,20,15,19 & 16

then the average units produced by each worker is ----

11+9+17+15+12+16+20+15+19+16/10=15, so

15 Is central tendency

Characteristic of good average

1. Easy to understand
2. Simple to compute
3. Based all observations
4. It should be uniquely defined
5. It should be mathematically expressed
6. It should not be affected by extremes values

Measures of central tendency

- Arithmetic mean
- Geometric mean
- Harmonic Mean
- Median
- Mode

AIRTHMATIC MEAN

(Average)

Average is the sum of a list of numbers divided by the no. of numbers in the list.

What is the average of 3,4,5

$$\text{Average} = 3+4+5/3 = 12/3 = 4$$

$$\text{Average} = \frac{\text{sum of observation}}{\text{total no. of observation}}$$

UNGROUPED DATA:-there are two methods for calculating arithmetic mean of ungrouped data

1. **Direct method**

2. **Indirect or shortcut method**

GROUPED DATA:-In grouped data we are using short cut method

OR

Assume mean method

OR

Step deviation method

MEAN

FORMULA :-direct method

1. $\bar{X} = \frac{\sum x}{N}$

2. $\bar{X} = \frac{\sum f x}{N}$ OR $\bar{X} = \frac{\sum f x}{\sum f}$

DIRECT METHOD

Example;- The number of new orders received by a company over the last 25 working days Were recorded as follows;

3,0,1,4,4,4,2,5,3,6,4,5,1,4,2,3,0,2,0,5,4,2,3,3,1

Calculate Arithmetic mean

Solution;---applying formula,

$$\textbf{A.M.} = \frac{\sum x}{N}$$

$$= \frac{3+0+1+4+4+4+2+5+3+6+4+5+1+4+2+3+0+2+0+5+4+2+3+3+1}{25}$$

$$= \frac{71}{25} = 2.84$$

❖**EXAMPLE:-** **Find A.M. of the following frequency distribution**:

X:	1	2	3	4	5	6	7	8	9
f:	7	11	16	17	26	31	11	1	1

x	f	f x
1	7	7
2	11	22
3	16	48
4	17	68
5	26	130
6	31	186
7	11	77
8	1	8
9	1	9
	$\sum f=121$	$\sum f x=555$

$$A.M .= \bar{X} = \frac{\sum f x}{\sum f}$$

$$= \frac{555}{121}$$

$$= 4.59$$

Example:- Determine **mean**

daily wages(in Rs)	: 0-10	10-20	20-30	30-40	40-50	50-60	60-70
No. of workers	: 6	5	8	15	7	6	3

Solution:

wages	Mid value(x)	No. of workers(f)	fx
0-10	5	6	30
10-20	15	5	75
20-30	25	8	200
30-40	35	15	525
40-50	45	7	315
50-60	55	6	330
60-70	65	3	195
		∑f=50	∑f x=1670

$$\text{Mid value } x = \frac{l_1 + l_2}{2}$$

$$\bar{X} = \text{Mean} = \frac{\sum f x}{\sum f}$$

$$= \frac{1670}{50}$$

$$= 33.40$$

SHORT CUT METHOD

QU.- Calculate the average sales per firm from the following data of sales of 100 firms using short cut method

SALES :	10	13	14	19	25	28
NO. OF FIRMS	13	19	30	16	14	8

Solution:-

x	f	d = x-A	f d
10	13	-4	-52
13	19	-1	-19
14	30	0	0
19	16	+5	+80
25	14	+11	+154
28	8	+14	+112
	∑f = 100		**∑f d = 275**

80+154+112 = +346

-52-19= -71

346-71=275

Assume mean= A= 14

$$\bar{X} = A + \frac{\sum f d}{\sum f}$$

$$= 14 + \frac{275}{100}$$

$$= 14+2.75$$

$$= 16.75$$

Qu.-calculate mean--

GROUPED DATA (SHORT CUT METHOD)

Scores	No. of candidates
0 – 10	8
10 – 20	15
20 – 30	17
30 – 40	28
40 – 50	25
50 – 60	24
60 – 70	18
70 – 80	9
80 – 90	6

Solution:

Score	f	Mid-Point=x	d= $\frac{x-A}{i}$	f d
0 – 10	8	5	–4	–32
10 – 20	15	15	–3	–45
20 – 30	17	25	–2	–34
30 – 40	28	35	–1	–28
40 – 50	25	45	0	0
50 – 60	24	55	1	24
60 – 70	18	65	2	36
70 – 80	9	75	3	27
80 – 90	6	85	4	24
	n= 150			∑f d= -28

$$\bar{x} = A + \frac{\Sigma fd'}{n} \times i$$

5-45= - 40
10
= -4

15-45=-30/10=-3

55-45=10/10=+1

$$\bar{x} = A + \frac{\sum f\ d}{\sum f} * i$$

$$= 45 + \frac{-28}{150} * 10$$

$$= 45 - 1.866$$

$$= 43.13$$

Assume Mean = A = 45

∑ f d = -28

∑f = 150 = n

i= 10

Qu. **-find A.M. applying step deviation method of the following distribution**:-

Marks:	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No. of students :	6	5	8	15	7	6	3

Solution:

marks	mid value(x)	f	d=x-A/i	f d
0-10	5	6	-3	-18
10-20	15	5	-2	-10
20-30	25	8	-1	-8
30-40	**35**	15	0	0
40-50	45	7	+1	+7
50-60	55	6	+2	+12
60-70	65	3	+3	+9
		∑f=50		∑ f d= -8

Mid value-

0+10/2=5

10+20=30/2=15

20+30=50/2=25

30+40=70/2=35.....

Assume mean=35

5-35=-30/10=-3

15-35=-20/10=-2

25-35=-10/10=-1

35-35=0

i=10

45-35=10/10=+1

55-35=20/10=+2

65-35=30/10=+3

$$\text{Now A.M.} = \bar{X} = A + \frac{\sum f d}{\sum f} * i$$

$$= 35 + \frac{-8}{50} * 10$$

$$= 35 - 1.6$$

$$= 34.4$$

Qu:- 1. find mean

marks :	10-15	15-20	20-25	25-30	30-35	35-40
no. of students :	11	20	35	20	8	6

ANS = 23.1

Qu.:-2. find A.M.—

Height:	130-134	135-139	140-144	145-149	150-154	155-159	160-164
frequency:	5	15	28	24	17	10	1

ANS = 145.35

Qu:- 3. determine arithmetic mean---

wages	No. of workers
12.5-17.5	2
17.5-22.5	22
22.5-27.5	19
27.5-32.5	14
32.5-37.5	3
37.5-42.5	4
42.5-47.5	6
47.5-52.5	1

ANS = 27.46

Cumulative frequency distribution

Qu:- find the average mean using arithmetic mean method

marks	No. of students
Less than 10	5
Less than 20	17
Less than 30	31
Less than 40	41
less than 50	49

Solution:- first we convert of cumulative frequency distribution to simple(exclusive) frequency distribution-

marks	no. of students
0-10	5
10-20	17-5=12
20-30	31-17=14
30-40	41-31=10
40-50	49-41=8

Marks	Mid value(x)	f	d=x-A/i	f*d	
0-10	5	5	-2	-10	-22
10-20	15	12	-1	-12	
20-30	25 A	14	0	0	
30-40	35	10	+1	+10	26
40-50	45	8	+2	+16	
		∑f =49		∑fd=4	

Applying formula

$$\text{Now A.M.} = \bar{X} = A + \frac{\sum f d}{\sum f} * i$$

$$= 25 + \frac{4}{49} * 10 = 25 + 0.81 = 25.81$$

More than (Ogive) Cumulative frequency distribution

Find mean -

Production yield (in kg)	Number of Firms
More than 50	100
More than 55	98
More than 60	90
More than 65	78
More than 70	54
More than 75	16

Solution:- first we convert of cumulative frequency distribution to simple(exclusive) frequency distribution-

Production Yield	f
50-55	100-98=2
55-60	98-90=8
60-65	90-78=12
65-70	78-54=24
70-75	54-16=38
75-80	16

Solution -

Production Yield	Mid Value x	f	d=x-A/I	f*d	
50-55	52.5	2	-3	-6	
55-60	57.5	18	-2	-36	-54
60-65	62.5	12	-1	-12	
65-70	**67.5 A**	24	0	0	
70-75	72.5	38	+1	+38	
75-80	77.5	16	+2	+32	+70
		∑f=100		∑fd=16	

Now A.M . = X = A + $\frac{\sum f d}{\sum f}$ * i

=67.5+16/100 *5 = 67.5 + 0..81

=68.31

Combined or composite mean

- If there are n_1 values in a distribution whose A.M. is $\bar{x}_1$ and similarly the A.M.of n_2 values of another distribution is x_2 then **combined** mean is given by

$$\bar{X} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$$

Example:- the average daily wage of 30 female workers in a factory is 65 Rs and the average daily wages of 70 male workers is 75 Rs ,find the **combined** average daily wages of total workers.

Solution- let x_1 = the wages of male workers =75

x_2= the wages of female workers =65

n_1= no. of male workers =70

n_2= no. of female workers =30

$$\bar{x} = \frac{70x75 + 30\,x65}{70+30}$$

$$= \frac{5250+1950}{100}$$

$$= \frac{7200}{100}$$

= **72 IS COMBINED MEAN**

Qu **The average marks obtained by two groups of students in an examination are 75 & 85. if the average marks of all the students is 80,find the ratio of student in the two group.**

Solution :-

let first group's average marks is $\bar{x}_1 = 75$

& second group's average marks is $\bar{x}_2 = 85$

total average is $\bar{x} = 80$

number of first group is $= n_1$

number of second group is $= n_2$

$n_1/n_2 = ?$

$$\bar{x} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{n_1 + n_2}$$

now, $$80 = \frac{n_1 (75) + n_2 (85)}{n_1 + n_2}$$

$$80n_1 + 80n_2 = 75n_1 + 85n_2$$

$$80n_1 - 75n_1 = 85n_2 - 80n_2$$

$$5n_1 = 5n_2$$

$$n_1/n_2 = 1/1 = 1:1$$

Qu:- 1 the mean salary for a group of 40 female workers is 5200 per month and that for a group of 60 male workers is 6800 per month ,What is the COMBINED MEAN salary **(ANS = 6160)**

Qu:-2 the average marks of 100 students of one group are 60 and that of another group of 50 students are 80,find COMBINED AVERAGE of all 150 students**.** **(ANS = 66.66)**

Qu:- 3 the mean of 35 observation of a group is 28.4 and the mean of 15 observation of another group is 32, if the two groups of observation are combined ,find the mean of the entire group **(ANS = 23.8)**

Qu:- 4 the mean weight of 150 students in a class is 60 kg. the mean weight of boy students is 70 kg and that of girl students is 55 kg ,find the number of boys and girls **(ANS = 50 &100)**

Qu:- 5 the average salary of employees of a firm is 6800,the average salary of gents and ladies are are 6400 & 7500 respectively .find ratio of gents & ladies . **(ANS = 4: 7)**

Merits of arithmetic mean:-

- it is very simple to calculate.
- it is very easy to understand.
- it is rigidly defined.
- it is based on all items.
- Algebraic treatment is possible.
- it is least affected by sampling fluctuations.
- no arrangement required.
- Accuracy test can be conducted.

Demerits of arithmetic mean:-

- it is inconvenient to calculate.
- it is effected of extreme values.
- it cannot be calculated for qualitative data.
- It cannot be calculated open ended continuous series.
- Laughable conclusion .
- Misleading conclusion.
- It cannot be determined by graphical method.
- more stress on items of higher value.