Dr.V.MAHESWARI

Assistant Professor in Commerce

PS.PT.MGR.GOVT.ARTS & SCIENCE COLLEGE

PUTHUR,SIRKALI-609108

UNIT-I

Introduction – Meaning, Characteristics, Stages and Uses of Statistics – Classification and Tabulation – Diagrams and graphs – Bar and Pie diagrams – Graphs of one and two variables – Graphs of frequency distribution - Measure of central tendency – Arithmetic mean, Median, Mode, Geometric Mean and Harmonic mean.

BUSINESS TOOLS FOR DECISION MAKING

Dr.V.MAHESWARI

Definition of Statistics

Statistics may be defined as the science of collection, presentation, analysis and Interpretation of numerical data".

Various Stages in Statistics

Stages	*Statistical Study*	*Statistical Tools*
Stage-I	Collection of data	Census or sample techniques
Stage-II	Organisation of data	Array of data and tally bars
Stage-III	Presentation of data	Tables, graphs and diagrams.
Stage-IV	Analysis of data	Percentages, averages.
Stage-V	Interpretation of data	Magnitude of percentages, averages, degree of relationship between different variables.

Characteristics of Statistics

*Statistics is a quantitative science

*It never consider a single item

*The values should be different

*Statistical result true on average

Classification

Classification is the process of arranging data into groups or classes according to the common characteristics possessed by the individual items.

Types of Classification

- *Geographical classification
- *Historical classification
- *Qualitative classification
- *Quantitative classification

Tabulation

Tabulation is the process of arranging data systematically in rows and columns of table.

Diagram

Diagram or various geometrical shapes such as bar and circles etc. They are more attractive and easier to understand then graphs and are widely used in advertisements and publicity.

TYPES OF BAR DIAGRAM

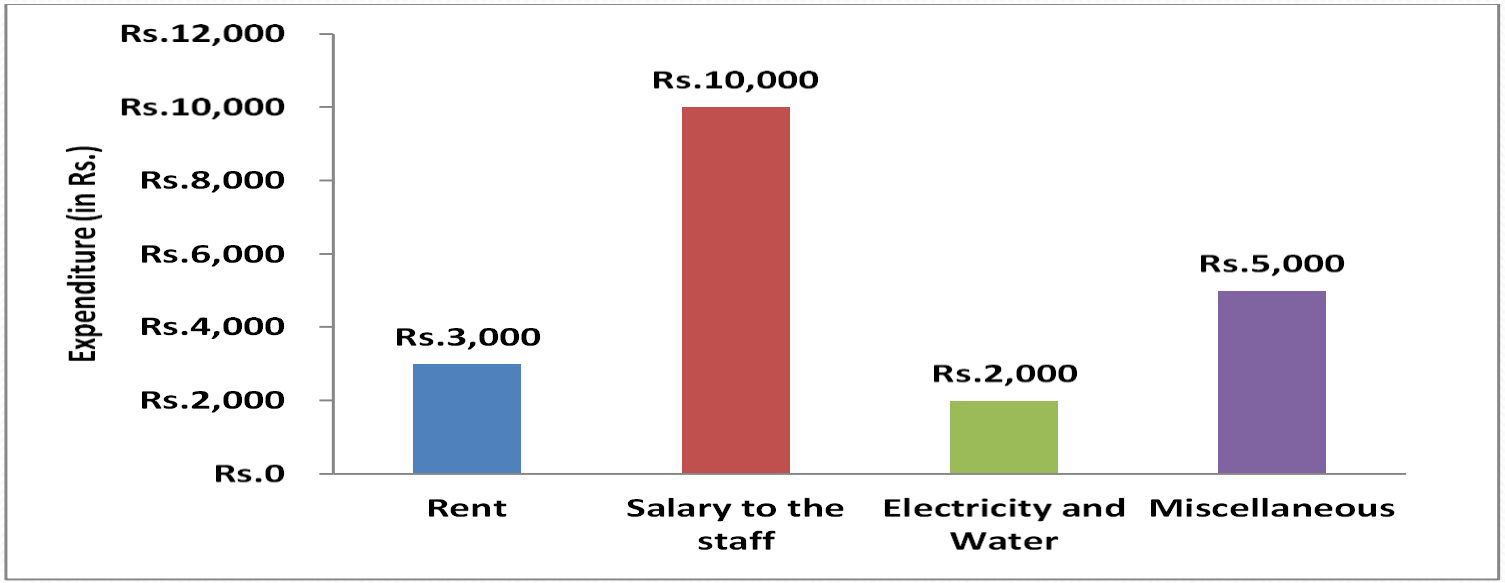
There are 4 types,

- ***Simple bar diagram**
- ***Multi bar diagram**
- ***Subdivided bar diagram**
- ***Percentage bar diagram**

Following table shows the monthly expenditure of a firm:

Item	Rent	Salary to the staff	Electricity and Water	Miscellaneous
Expenditure	Rs.3000	Rs. 10,000	Rs.2000	Rs.5000

Solution

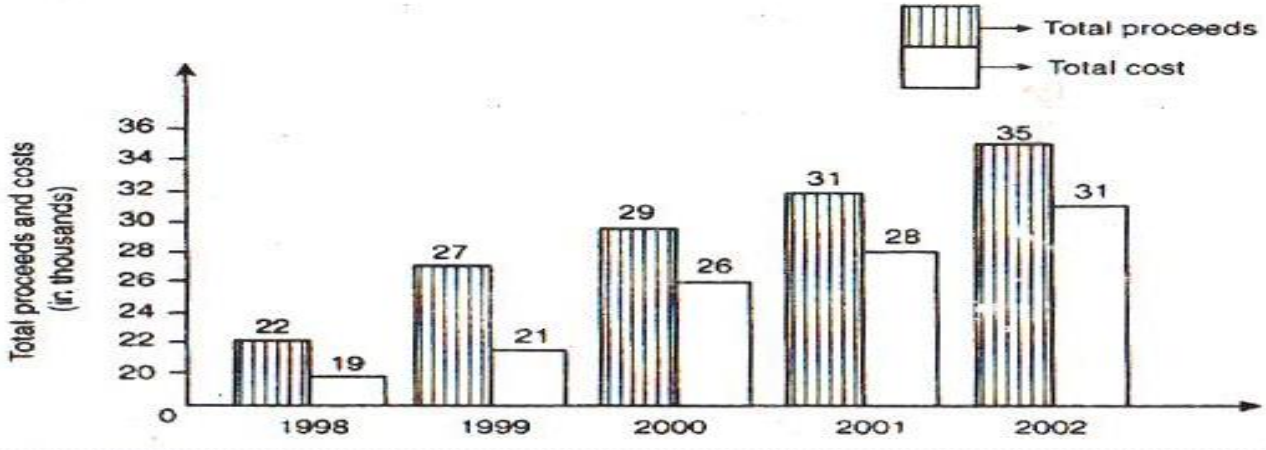


Simple bar diagram showing the monthly expenditure of a firm

Represent following data by a suitable diagram showing the difference between proceeds and costs. (Multiple bar diagram)

Year	Total proceeds (Rs. in thousands)	Total cost (Rs. in thousands)
1998	22	19
1999	27	21
2000	29	26
2001	31	28
2002	35	31

Solution



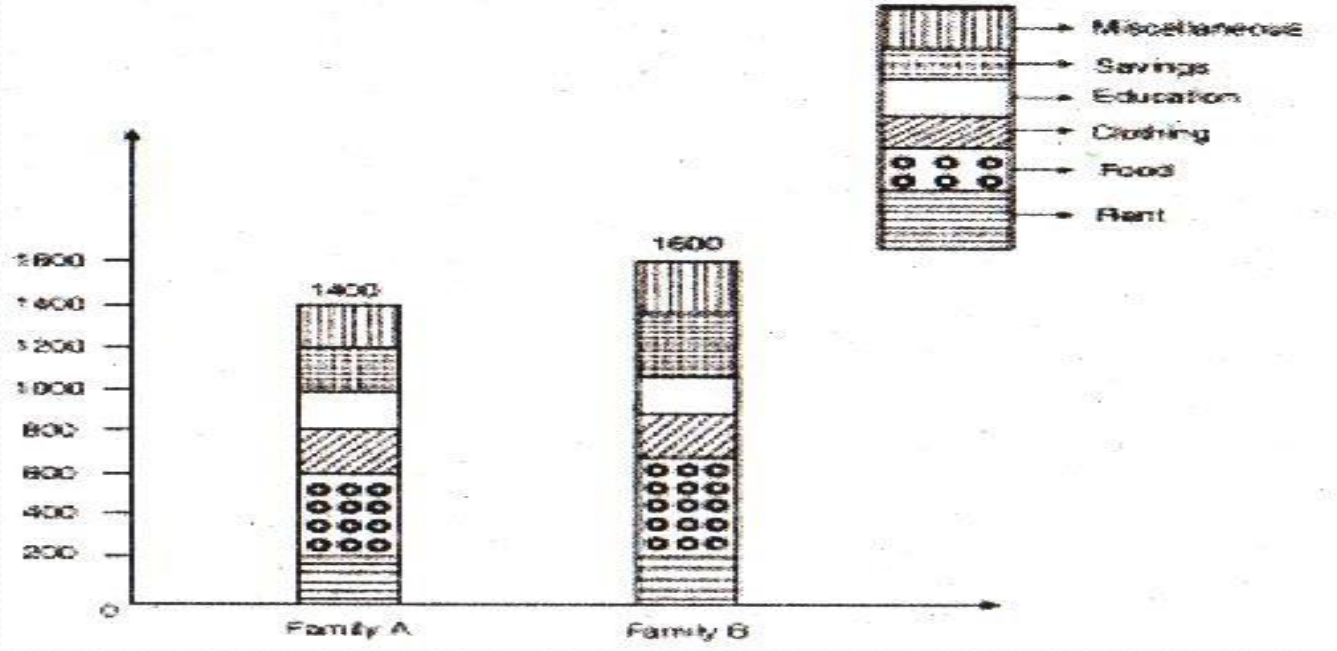
Multiple bar diagram showing the total proceeds and total cost for five years

Represent the following data by a subdivided bar diagram.

Distribution of monthly income of 2 families *A* and *B*

Income	Family A	Family B
Rent	250	300
Food	400	500
Clothing	200	250
Education	100	250
Saving	150	100
Miscellaneous	300	200

Solution



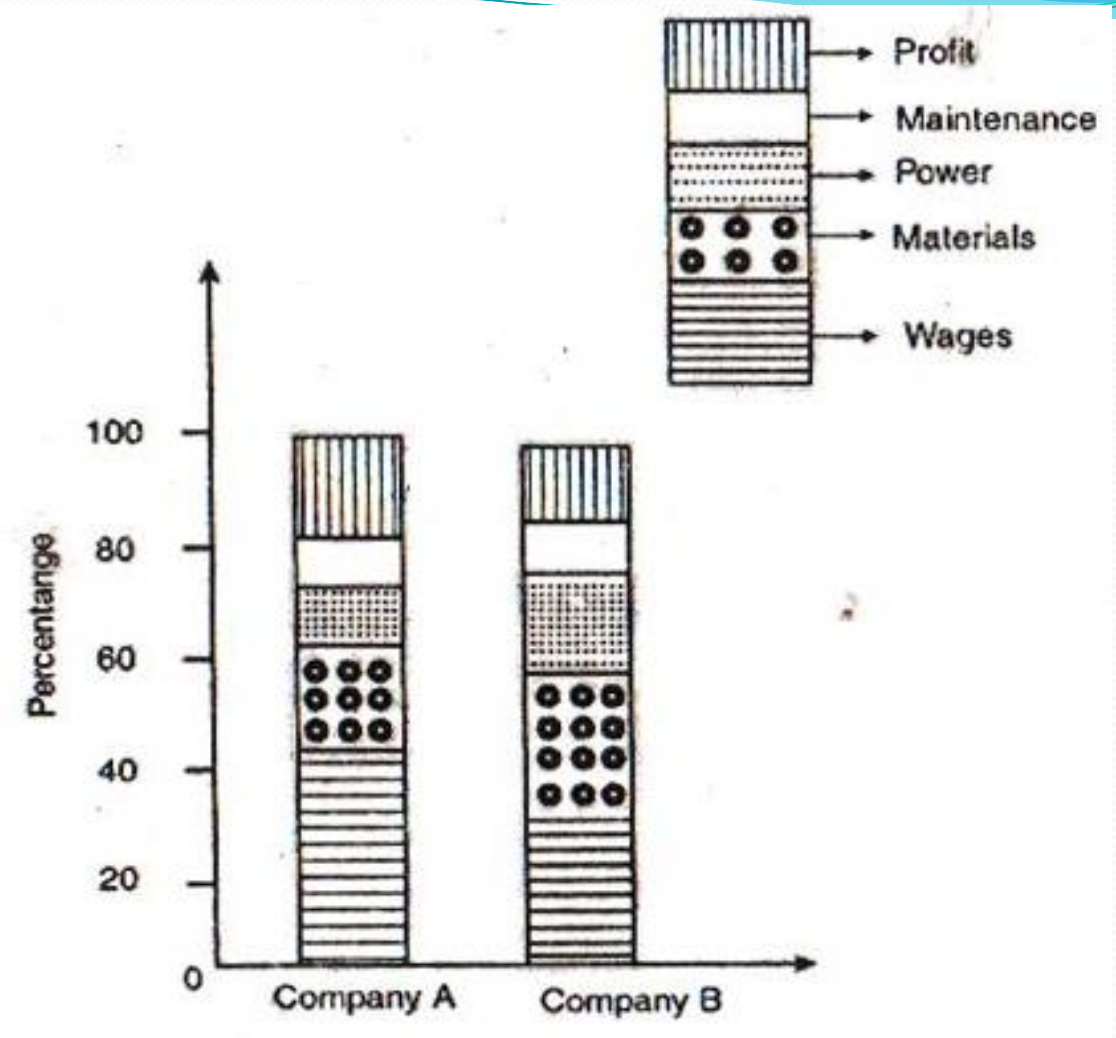
Subdivided bar diagram showing the distribution of monthly income of 2 families

Draw a percentage bar diagram for the following data:

Expenditure	Company A	Company B
Wages	450	700
Materials	200	500
Power	75	350
Maintenance	80	175
Profit	195	275
Total	1000	2000

Solution

Expenditure	Company A in%	Company B in%
Wages	$\frac{450}{1000}\times 100 = 45$	$\frac{700}{2000}\times 100 = 35$
Materials	$\frac{200}{1000}\times 100 = 20$	25
Power	$\frac{75}{1000}\times 100 = 7.5$	17.5
Maintenance	$\frac{80}{1000}\times 100 = 8$	8.75
Profit	$\frac{195}{1000}\times 100 = 19.5$	13.75



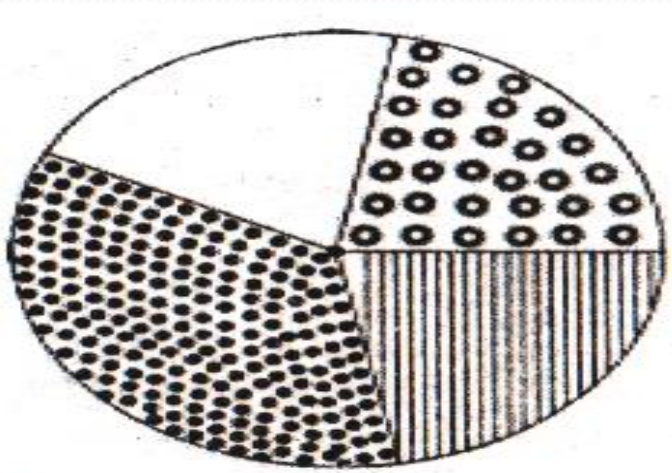
Percentage bar diagram showing the expenditure of two companies

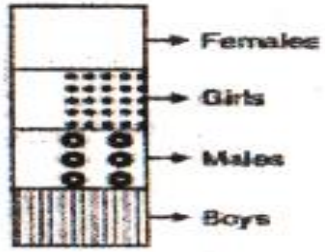
Draw a pie diagram to represent the following population in a town:

Males	Females	Girls	Boys	Total
2000	1800	4200	2000	10000

Solution

	Total Number	In degrees
Males	2000	$\frac{2000}{10000} \times 360 = 72°$
Females	1800	$\frac{1800}{10000} \times 360 = 64.8°$
Girls	4200	151.2°
Boys	2000	72°





Historical Classification

Some data can be classified on the basis of time and arranged chronologically or historically.

TYPES OF GRAPHS

1. HISTOGRAM
2. FREQUENCY POLYGON
3. FREQUENCY CURVE
4. OGIVE
5. LORENZ CURVE

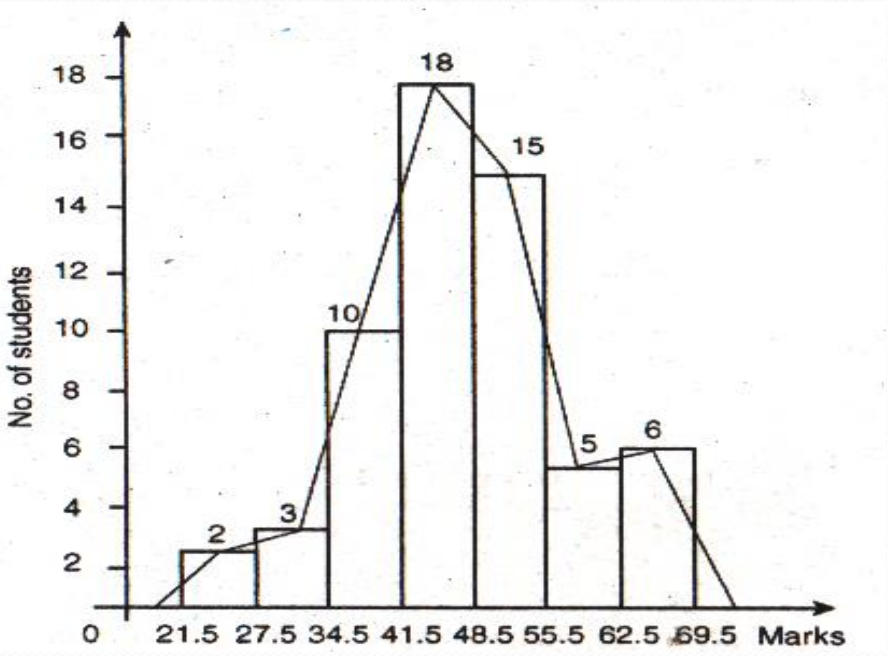
Construct a histogram and frequency polygon for the following distribution:

Marks	21-27	28-34	35-41	42-48	49-55
No. of students	2	3	10	18	15
Marks	56-62	63-69			
No. of students	5	6			

Solution

Since the class limits are not true class limits, first we have to convert into continuous true class limits.

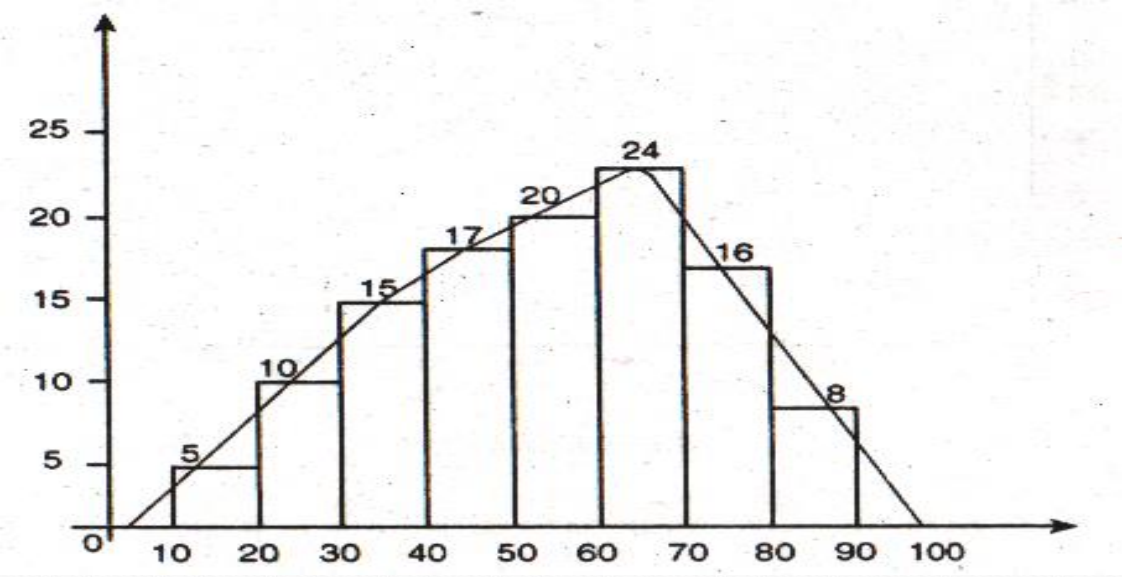
True class limits	21.5-27.5	27.5-34.5	34.5-41.5	41.5-48.5
Frequency	2	3	10	18
True class limits	48.5-55.5	55.5-62.5	62.5-69.5	
Frequency	15	5	6	



For the following table, draw a histogram and frequency curve:

Class interval	10-20	20-30	30-40	40-50	50-60
Frequency	5	10	15	17	20
Class interval	60-70	70-80	80-90		
Frequency	24	16	8		

Solution

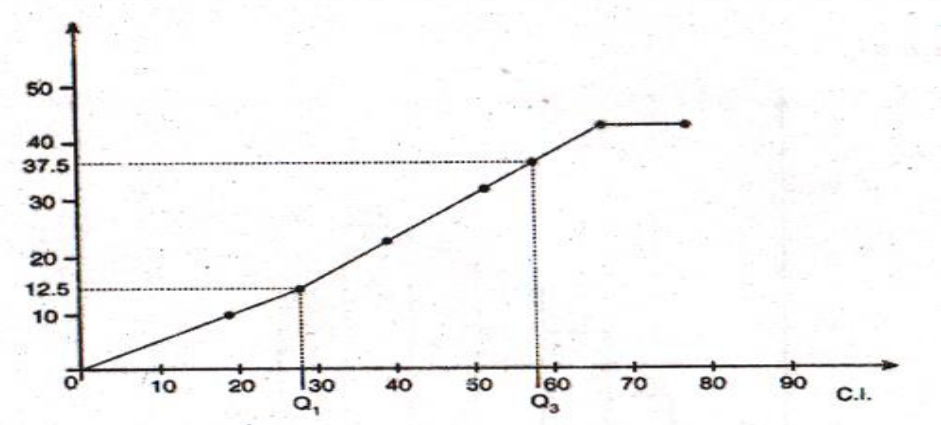


Estimate Q_3 and Q_1 for the following frequency table, using an ogive

Class interval	10-20	20-30	30-40	40-50
Frequency	6	7	9	10
Class interval	50-60	60-70	70-80	
Frequency	8	7	3	

Solution

C.I	Less than	*f*	*c.f*
10-20	20	6	6
20-30	30	7	13
30-40	40	9	22
40-50	50	10	32
50-60	60	8	40
60-70	70	7	47
70-80	80	3	50



Less than ogive

$$Q_1 = \text{Size of } \left(\frac{N}{4}\right)^{th} \text{ item} = \text{Size of } \left(\frac{50}{4}\right)^{th} \text{ item} = 12.5^{th} \text{ item}$$

From the diagram, we see that $Q_1 = 27$

$$Q_3 = \text{Size of } 3\left(\frac{N}{4}\right)^{th} \text{ item} = \text{Size of } 3 \times 12.5^{th} \text{ item} = 37.5^{th} \text{ item}$$

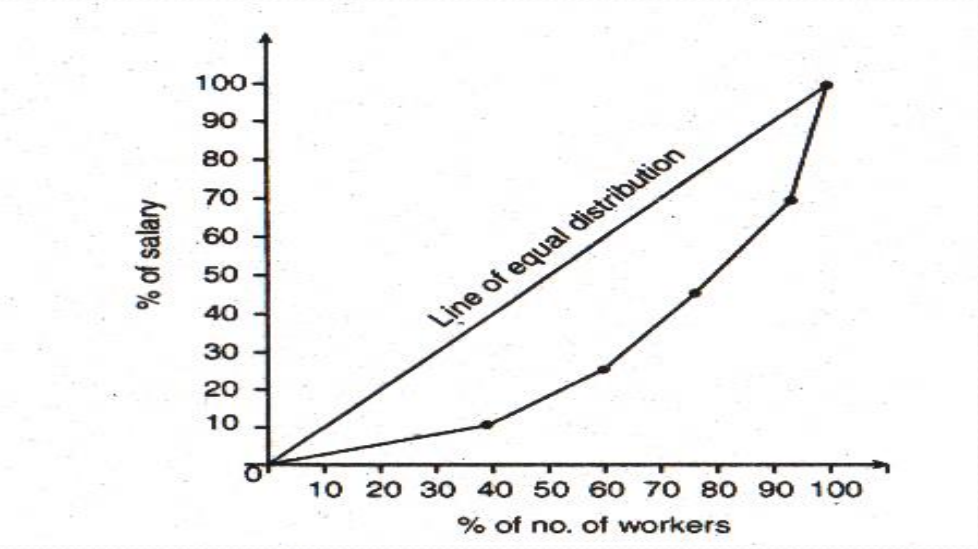
$$Q_3 = 57$$

Draw a Lorenz curve from the following data to study the rent of dispersion graphically:

Salary (in Rs.)	100	150	200	250	300
No. of workers	20	10	8	10	2

Solution

Salary (in Rs.)	Cumulative salary	Cumulative % salary	No. of workers	Cumulative No. of workers	Cumulative %
100	100	10	20	20	40
150	250	25	10	30	60
200	450	45	8	38	76
250	700	70	10	48	96
300	1000	100	2	50	100



Individual Series

Individual Series is a statistical series in which the all the observations are listed out and all the observations have a frequency of 1

- **Example: set of observations 25,32,38,24,31,36**

10,20,25,5,15,10,20,25,25,15,10,5,25,15,10,25,5,10,10,10,15,25,5,15,10,15,20,20,20,10,15

- X F
- 5 4
- 10 9
- 15 7
- 20 5
- 25 6
  - 31

Discrete Series

- **A discrete distribution is one in which the data can only take on certain values, for example integers. Frequency of each value multiplied with the respective size All data have a frequency in this method.**
- **AM= Σfx ÷Σf or N**
- **Age (x) : 8 10 12 15**
- **No.of Workers(F): 2 3 6 10**

Continuous Series

- In continuous frequency distribution each individual frequency distribution is unknown an assumption in frequency distribution concentrate only in the midpoint or class interval
- A continuous distribution is one in which data can take on any value within a specified range (which may be infinite) Ex.10-20,20-30,30-40,40-50 .....It has a class intervals
- **Marks : 0-10 10-20 20-30 30-40**
- **No.of.Students : 5 4 3 6**

Measures Of Central Tendency

In statistics, the three most common **measures of central tendency** are the **mean**, **median**, and **mode.**

Mean

Mean is Adding the numbers in a data set and dividing by how many numbers there are.

Arithmetic Mean

- **It is also called Arithmetic average .AM is the most commonly used method of measure the central tendency. Arithmetic Mean is Adding the numbers in a data set and dividing by how many numbers there are.**
- **AM=Sum of variables /No, of observation**
- **Example: 10,8,5,4**
- **AM=10+8+6+4/4=28/4=7**

MEAN

Individual series = $\bar{X} = \frac{\Sigma x}{n}$

Discrete series = $\bar{X} = \frac{\Sigma fx}{\Sigma f}$

Continuous series = $\bar{X} = A + \frac{\Sigma fd}{\Sigma f} \times c$

Median

Median is the middle number in a data set when the numbers are listed in either ascending or descending order.

Example: 10 15 12 18 20 25 5

Arrange the order small to big numbers

5 10 12 15 18 20 25

Median

Individual series = $\left(\frac{N+1}{2}\right)$

Discrete series = $\left(\frac{N+1}{2}\right)$

Continuous series = $M = L + \frac{\frac{N}{2} - cf}{f} \times c$

Mode

Mode is the value which has higher number of frequencies of value.

Example: 10 16 12 18 16 20 16

Mode: 16

Example: 20 30 50 80 40 10 25

- **Mode: 80**

Mode

Mode:

Individual series = **value that occurs maximum times**

Discrete series = **grouping table method is used**

Continuous series = $$L + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times c$$

Find the Arithmetic mean of the following set of observation:

25,32,28,34,24,31,36,27,29,30 (Individual series)

- Arithmetic Mean = $\bar{X} = \frac{\Sigma X}{N}$

- $\bar{X} = \frac{25+32+28+34+24+31+36+27+29+30}{10}$

$\bar{X} = \frac{296}{10}$

$\bar{X} = 29.6$

(… Sigma Σ = Sum or Add)

Find the median of the set of observation 10,18,25,8,7,5,30

- **Ascending order**
- 5,7,8,10,18,25,30
- Median = Size of $\left\{\frac{N+1}{2}\right\}$ of item
  - Median$=\left\{\frac{7+1}{2}\right\}=\frac{8}{2}=4$
  - Median 4^{th} item $=10$

Mode: set of observations 10,15,16,14,15,10,20,15,30

- **MODE = 15 (Frequency Repeated)**

- **2.set of observation 10,15,20,25,45,55,80,40,30**

- **Mode = 80 (higher value)**

Find the Arithmetic mean of the following set of observation:

25,32,28,34,24,31,36,27,29,30 (Individual series)

- Arithmetic Mean = $\bar{X} = \frac{\Sigma X}{N}$

- $\bar{X} = \frac{25+32+28+34+24+31+36+27+29+30}{10}$

$\bar{X} = \frac{296}{10}$

$\bar{X} = 29.6$

(… Sigma Σ = Sum or Add)

From the following table, find the mean height. (discrete series)

Height (inches)	60	61	62	63	64
No. of children	2	3	5	8	7

Solution:

x	f	fx
60	2	120
61	3	183
62	5	310
63	8	504
64	7	448
	25	1565

$$\bar{X} = \frac{\sum fx}{\sum f} = \frac{1565}{25} = 62.6$$

(… Sigma Σ = Sum or Add)

The following is the age distribution of 100 persons in a street. Calculate the arithmetic mean. (continuous series)

Age group	0-10	10-20	20-30	30-40	40-50	50-60
No. of persons	5	10	25	30	20	10

Solution:

$x \rightarrow$ Age group A = Average mid = middle

$f \rightarrow$ No. of persons mid x = 0+10 =10 / 2 =5 c = class interval(0 – 10) =10

x	f	mid x	$d = \frac{x - A}{c}$ $A = 35, c = 10$	fd
0-10	5	5	– 3	–15
10-20	10	15	– 2	–20
20-30	25	25	– 1	–25
30-40	30	35	0	0
40-50	20	45	1	20
50-60	10	55	2	20
	100			–20

$\sum fd = -20, \ \sum f = 100, \ A = 35, \ c = 10$

$$\text{Mean, } \bar{X} = A + \frac{\sum fd}{\sum f} \times c$$

(… Sigma Σ = Sum or Add)

$$= 35 + \frac{(-20)}{100} \times 10$$

$$= 33$$

Find the median of marks of 9 students:
70, 60, 75, 90, 65, 80, 42, 65, 72

Solution:

Let us first arrange the data in ascending order:

42, 60, 65, 65, 70, 72, 75, 80, 90

$$\text{Median} = \text{Size of } \left(\frac{N+1}{2}\right)^{th} \text{ item}$$

Where *N* is the No. of observations.

$$\therefore \text{Median} = \text{Size of } \left(\frac{9+1}{2}\right)^{th} \text{ item}$$

$$= 5^{th} \text{ item}$$

$$= 70$$

Find the median of the following data:
84, 91, 72, 68, 87, 78 (individual series)

Solution:

Let us arrange the data in ascending order:

68, 72, 78, 84, 87, 91

$N = 6$

$$\text{Median} = \text{Size of } \left(\frac{N+1}{2}\right)^{th} \text{ item}$$

$$= \text{Size of } \left(\frac{7}{2}\right)^{th} \text{ item}$$

$$= \frac{1}{3}(3\text{rd and } 4\text{th item})$$

$$= \frac{1}{2}(78+84)$$

$$= \frac{162}{2} = 81$$

Calculate the median of the following distribution:

Heights (cms)	120	122	124	126	128	130
No. of students	5	7	9	6	4	10

Solution:

Heights x	No. of students	
	f	c,f
120	5	5
122	7	12
124	9	21
126	6	27
128	4	31
130	10	41
	$N = 41$	

CF - Cumulative Frequency (discrete series)

$$\text{Median} = \text{Size of } \left(\frac{N+1}{2}\right)^{th} \text{ item}$$

$$= \text{Size of } \left(\frac{41+1}{2}\right)^{th} \text{ item}$$

= 21st item

= 124

Calculate the median of the following data: (continuous series)

Savings (in Rs) less than	10	20	30	40	50	60	70	80
Cumulative frequency	15	35	64	84	96	120	192	256

Solution:

Since the less than values are given, we have to find the true class limits and their corresponding frequencies.

0 – 10 = 15

10 – 20 = 35 – 15 = 20

True class intervals

x	*f*	*c,f*
0-10	**15**	**15**
10-20	**20**	**35**
20-30	**29**	**64**
30-40	**20**	**84**
40-50	**12**	**96**
50-60	**24**	**120**
60-70	**72**	**192**
70-80	**64**	**256**

$$\text{Median} = \text{Size of } \left(\frac{N}{2}\right)^{th} \text{ item}$$

$$= \text{Size of } \left(\frac{256}{2}\right)^{th} \text{ item}$$

$$= \mathbf{128^{th}\ item}$$

$\therefore$ **Median class** $= \mathbf{60 - 70}$

$$\text{Median} = L + \frac{\frac{N}{2} - c.f}{f} \times i$$

$$\boldsymbol{L = 60,\quad c.f = 120,\quad f = 72,\quad i = 10}$$

$$\text{Median} = 60 + \frac{128 - 120}{72} \times 10$$

$$= \mathbf{60 + 1.111}$$

$$= \mathbf{61.111}$$

1. Find the mode for the set of numbers:

2, 2, 3, 5, 6, 8, 5, 9, 5

Solution:

Since 5 appear maximum number of times

Mode = 5

2. Calculate the mode for the following data: (discrete series)

x	3	5	7	9	11	13	15	17
f	2	5	7	8	15	7	5	1

Solution:

By inspection method, the value corresponding to the maximum frequency i.e. 15 is 11

∴ Mode = 11

Calculate the mode from the following distribution: (continuous)

x	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45
f	3	6	10	20	15	5	4	2

Solution:

x	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45
f	3	6	10	20	15	5	4	2
			F_0	F_1	F_2			

By inspection method, the model class is the class corresponding to the highest frequency.

i.e. Model class = 20 – 25

Exact value of is given by the formula

$$\text{Mode} = \text{L} + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \text{X c}$$

$$L = 20,\ f_1 = 20,\ f_0 = 10,\ f_2 = 15,\ c = 5$$

$$\therefore \text{Mode} = 20 + \frac{20 - 10}{2(20) - 10 - 15} \times 5$$

$$= 20 + \frac{10}{15} \times 5 = 20 + 3.33$$

= 23.33

Calculate the mode from the following data: (discrete series)

x	25	30	35	40	45	50	55
f	7	11	17	15	14	10	11

Solution:

In this problem, inspection method fails because the difference between the maximum frequency and the next is very small. So we form grouping and analysis table.

GROUPING TABLE

x	*f*(Col I)	Col II	Col III	Col IV	Col V	Col VI
25	7					
		18				
30	11			35		
			28			
35	17				43	
		32				
40	15					46
			29			
45	14			39		
		24				
50	10				35	
			22			
55	11					

ANALYSIS TABLE

x	I	II	III	IV	V	VI	Total
25							0
30					1		1
35	1	1			1	1	4
40		1	1	1	1	1	5
45			1	1		1	3
50				1			1
55							0

In the total column, maximum is 5 and the corresponding value of the variable is 40

$\therefore$ Mode = 40

GM-GEOMETRIC MEAN-

Geometric mean of N values is the Nth root of the product of the N

Find the G.M of the following data:

82, 93, 50, 54 72

Solution:

x	$\log x$
82	1.9138
93	1.9685
50	1.6990
54	1.7324
72	1.8573
	9.1710

$\sum \log x = 9.1710 \qquad n = 5$

$$\text{G.M} = \text{Antilog}\left[\frac{\sum \log x}{n}\right]$$

$$= \text{Antilog}\left[\frac{9.1710}{5}\right]$$

$$= \text{Antilog}\,(1.8342)$$

$$= 68.26$$

Compute the geometric mean from the data given below:

Category	I	II	III	IV	V	VI	VII	VIII
Monthly income	500	3750	3000	750	600	400	300	200
No. of employees	2	4	6	8	6	100	10	50

Solution:

Category	x Monthly income	f No. of Employees	log x	f log x
I	5000	2	3.6990	7.398
II	3750	4	3.5740	14.296
III	3000	6	3.4771	20.863
IV	750	8	2.8751	23.001
V	600	6	2.7782	16.669
VI	400	100	2.6021	260.210
VII	300	10	2.4771	24.771
VIII	200	50	2.3010	115.050

$\sum \log x = 482.258$ $\sum f = 186$

$$G.M = \text{Antilog}\left[\frac{\sum f \log x}{\sum f}\right]$$

$$= \text{Antilog}\left[\frac{482.258}{186}\right]$$

$$= \text{Antilog}\,[2.5938]$$

$$= 392.40$$

Compute the G.M for the following data:

Class	0-10	10-20	20-30	30-40	40-50
Frequency	5	7	15	25	8

Solution

Class x	Frequency f	Mid m	log m	f log m
0-10	5	5	0.6990	3.4950
10-20	7	15	1.1761	8.2327
20-30	15	25	1.3979	20.9685
30-40	25	35	1.5441	38.6025
40-50	8	45	1.6532	13.2256
				84.5243

$\sum f \log m = 84.5243$ $\sum f = 60$

$$\text{G.M} = \text{Antilog}\left[\frac{\sum f \log m}{\sum f}\right]$$

$$= \text{Antilog}\left[\frac{84.5243}{60}\right]$$

$$= \text{Antilog}\ (1.4087)$$

$$= 25.63$$

HM-HARMONIC MEAN- is the reciprocal of the mean of the reciprocals of the values.

Find the H.M for the following individual data:

6,15,35,40,900,520,300,400,400,1800,2000

Solution

Harmonic mean is calculated by the formula,

$$H.M = \frac{n}{\sum\left(\frac{1}{x}\right)}$$

x	$\frac{1}{x}$
6	0.1667
15	0.0667
35	0.0286
40	0.0250
900	0.0011
520	0.0019
300	0.0033
400	0.0025
1800	0.0006
2000	0.0005
	0.2969

$$H.M = \frac{n}{\sum \frac{1}{x}}$$

$$\sum \frac{1}{x} = 0.2969 \qquad n = 10$$

$$= \frac{10}{0.2969}$$

$$= 33.6814$$

Calculate the harmonic mean for the following data:

x	10	12	14	16	18	20
f	5	18	20	10	6	1

Solution

x	f	$\frac{1}{x}$	$f\left(\frac{1}{x}\right)$
10	5	0.1	0.5
12	18	0.0833	1.5
14	20	0.0714	1.428
16	10	0.0625	0.625
18	6	0.0556	0.333
20	1	0.05	0.05
	60		4.436

$$\sum f = N = 60, \qquad \sum f\left(\frac{1}{x}\right) = 4.436$$

$$\text{Harmonic mean} = \frac{\sum f}{\sum f\left(\frac{1}{x}\right)}$$

$$= \frac{60}{4.436}$$

$$= 13.5257$$

Find the H.M from the data given below:

Marks	15-25	25-35	35-45	45-55	55-65	65-75
No. of students	4	11	19	14	0	2

Solution

$x \rightarrow$ Marks, $f \rightarrow$ No. of students

x	f	Mid x m	$\frac{1}{m}$	$f\left(\frac{1}{x}\right)$
15-25	4	20	0.0500	0.2000
25-35	11	30	0.0333	0.3663
35-45	19	40	0.0250	0.4750
45-55	14	50	0.0200	0.2800
55-65	0	60	0.0166	0.0000
65-75	2	70	0.0148	0.0296
	50			1.3509

$$\sum f = 50, \qquad \sum f\left(\frac{1}{m}\right) = 1.3509$$

$$\text{Harmonic mean} = \frac{\sum f}{\sum f\left(\frac{1}{m}\right)}$$

$$= \frac{50}{1.3509}$$

$$= 37.01$$

Thank you