Unit 3
Tables and Charts

For
BCA Sixth Semester
Pokhara University

Syllabus

- Unit 3:- Tables and Charts 4 hrs.
- Frequency distribution
- Cumulative frequency distribution
- Relative frequency distribution
- Polygons
- Time plots
- Steam-and-leaf display

Data Presentation (Methods of describing data)

- After arranging the data in a proper order they have to classify and present according to their characteristics.
- The systematically collected data can be presented by the following ways:
- Frequency Distribution: Individual, Discrete and Continuous (inclusive and exclusive)
- Cumulative Frequency Distribution
- Relative Frequency Distribution
- Tables (simple and Complex)
- Charts and Diagrams

Individual Frequency Distribution

Days	Sun	Mon	Tue	Wed	Thu	Fri	Sat
Temperatures(°C)	32	33	31	34	30	29	35

- Discrete frequency distribution

Marks	Tally Bars	Frequency
15		7
16		9
17		5
18		3
19		1
Total		25

Continuous (grouped) frequency distribution

- (i) Exclusive type

Marks	Tally Bars	Frequency
30 – 35		5
35 – 40		7
40 – 45		10
45 – 50		12
50 – 55		3
55 – 60		3
Total		40

(ii) inclusive type

Marks	Tally Bars	Frequency
10 – 19		5
20 – 29		7
30 – 39		10
40 – 49		12
50 – 59		3
60 – 69		3
Total		40

Open – ended classes

Marks	Below 50	50 – 60	60 – 70	70 – 80	80 – 90	90 – 100
No. of Students	10	20	30	35	30	5

Marks	40 – 50	50 – 60	60 – 70	70 – 80	80 – 90	Above 90
No. of Students	10	20	30	35	30	5

Marks	Below 50	50 – 60	60 – 70	70 – 80	80 – 90	Above 90
No. of Students	10	20	30	35	30	5

Construction of frequency distribution

Step 1: The number of classes should be decided by Struge's rule:

no of classes(k) = 1 + 3.322 Log N(i. e. generally 5 to 15). Where N = total no. of observations.

- Step 2: Size of class interval(w) = R/K,

  Where, Range(R)= Largest item – smallest item.

- Step 3: The frequencies have to be counted by tally bars.
- Step 4: Classes should be non overlapping, open ended classes should be avoided.
- Step 5: The frequency table can be made by two methods; (i) exclusive (ii) inclusive

Example

- Construct a group frequency distribution using Struge's rule from the following data.

23	24	16	12	6	10	32	11
18	26	30	17	17	14	17	25
20	8	33	19	22	7	16	20
29	12	18	27	13	7	**5**	16
15	19	20	21	22	15	10	37
9	25	**39**	36	23	24	24	22
14	28	23	26	21	21	22	11
28	23	13	34	21	18	9	31

Solution: No. of observations(n) = 64

No. of classes(k) = 1 + 3.322 Log N = 1 + 3.322 log 64 = 7

Class interval(w) = $\frac{R}{k} = \frac{L-S}{7} = \frac{39-5}{7} = 4.85 \approx 5$

Construction of frequency distribution

Construction of frequency of inclusive class			
Class interval	Tally Bars	Frequency	Cumulative frequencies
5 – 9		7	7
10 – 14		10	10 + 7 = 17
15 – 19		13	17 + 13 = 30
20 – 24		18	18 + 30 = 48
25 – 29		8	8 + 48 = 56
30 – 34		5	5 + 56 = 61
34 – 39		3	3 + 61 = 64
Total		N = 64	

Construction of frequency of exclusive class			
Class interval	Tally Bars	Frequency	Cumulative frequencies
5 – 10		7	7
10 – 15		10	10 + 7 = 17
15 – 20		13	17 + 13 = 30
20 – 25		18	18 + 30 = 48
25 – 30		8	8 + 48 = 56
30 – 35		5	5 + 56 = 61
35 – 40		3	3 + 61 = 64
Total		N = 64	

Cumulative frequency distribution

Marks obtained	Number of students
0 – 10	4
10 – 20	21
20 – 30	28
30 – 40	10
40 – 50	7
Total	70

Marks less than	No. of students	Cumulative frequency
Less than 10	4	4
Less than 20	4 + 21 = 25	25
Less than 30	25 + 28 = 53	53
Less than 40	53 + 10 = 63	63
Less than 50	63 + 7 = 70	70

Marks more than	No. of students	Cumulative frequency
More than 0	66 + 4 = 70	70
More than 10	45 + 21 = 66	66
More than 20	17 + 28 = 45	45
More than 30	7 + 10 = 17	17
More than 40	7	7

Relative Frequency Distribution

Marks obtained	Number of students	Cumulative frequency	Relative cumulative frequency	Percent frequency
0 – 10	4	4	4/70 = 0.06	6%
10 – 20	21	4 + 21 = 25	21/70 =0.30	30%
20 – 30	28	28 + 25 = 53	28/70 =0.40	40%
30 – 40	10	10 + 53 = 63	10/70 =0.14	14%
40 – 50	7	7 + 63 = 70	7/70 =0.10	10%
Total	70		1	100%

Class work

- The following are the weekly production in units of 60 workers in a manufacturing company. Construct a group frequency distribution using Struge's rule from the following data.

23	35	41	82	75	46	48	88	72	44
85	51	77	62	75	68	50	64	25	49
56	55	16	82	39	32	51	52	62	19
52	54	63	45	56	33	48	67	59	40
54	50	64	46	69	57	40	39	49	55
53	20	55	72	57	52	57	42	75	51

The manager has decided to give bonus of Rs. 5, 10, 15, 20, 25, 30, 35 and 40 to each worker in the respective output group. Calculate mean bonus
Find the average bonus received the workers. (ii) Find S. D.& C. V. bonus.
Solution: k = 1+3.322logN = 1+3.322log60 = 6.9 $\approx$ 7

$$w = \frac{72}{7} = 10.28 \approx 10$$

Calculation Table					
Class	Frequency (f)	(m) bonus	(f x m)	(m^2)	(f x m^2)
15 – 25	4	5	20	25	100
25 – 35	3	10	30	100	300
35 – 45	8	15	120	225	1800
45 – 55	18	20	360	400	7200
55 – 65	14	25	350	625	8750
65 – 75	5	30	150	900	4500
75 – 85	5	35	175	1225	6125
85 – 95	2	40	80	1600	3200
	N = 60		$\sum fm = 1285$		$\sum fm^2 = 31975$

- Mean$(\overline{X}) = \frac{\sum fm}{n} = \frac{1285}{60} = 21.41$
- Standard Deviation$(\sigma) = \sqrt{\frac{\sum fm^2}{n} - \left(\frac{\sum fm}{n}\right)^2} = \sqrt{\frac{31975}{60} - \left(\frac{1285}{60}\right)^2} = 8.61$
- Coefficient of Variation (CV) = $\frac{\sigma}{\bar{X}} x100 = \frac{8.61}{21.41} x100 = 40.21\%$

Class work

- The marks obtained by 50 students in Statistics test are given below.

40	45	48	49	54	58	36	76	76	85
48	48	48	47	56	34	39	72	28	67
50	48	47	54	32	38	71	26	67	67
52	52	55	55	38	58	22	84	27	15
57	57	57	58	59	62	64	64	10	66

- With reference with this table:
- Find the highest and the lowest marks and the range of marks.
- Construct a group frequency distribution having equal class width of 10 (exclusive class intervals).
- How many students received 60 marks and above. Ans: 13
- How many students received marks below 40. Ans: 12
- What percentage of students passed the test taking 40 marks as the minimum pass mark? Ans: 76%
- What percentage of student's secured first division taking 60 and above marks for getting the first division. Ans: 26%
- What percentage of student's secured marks 70 and above? Ans: 12%

Charts and Diagrams

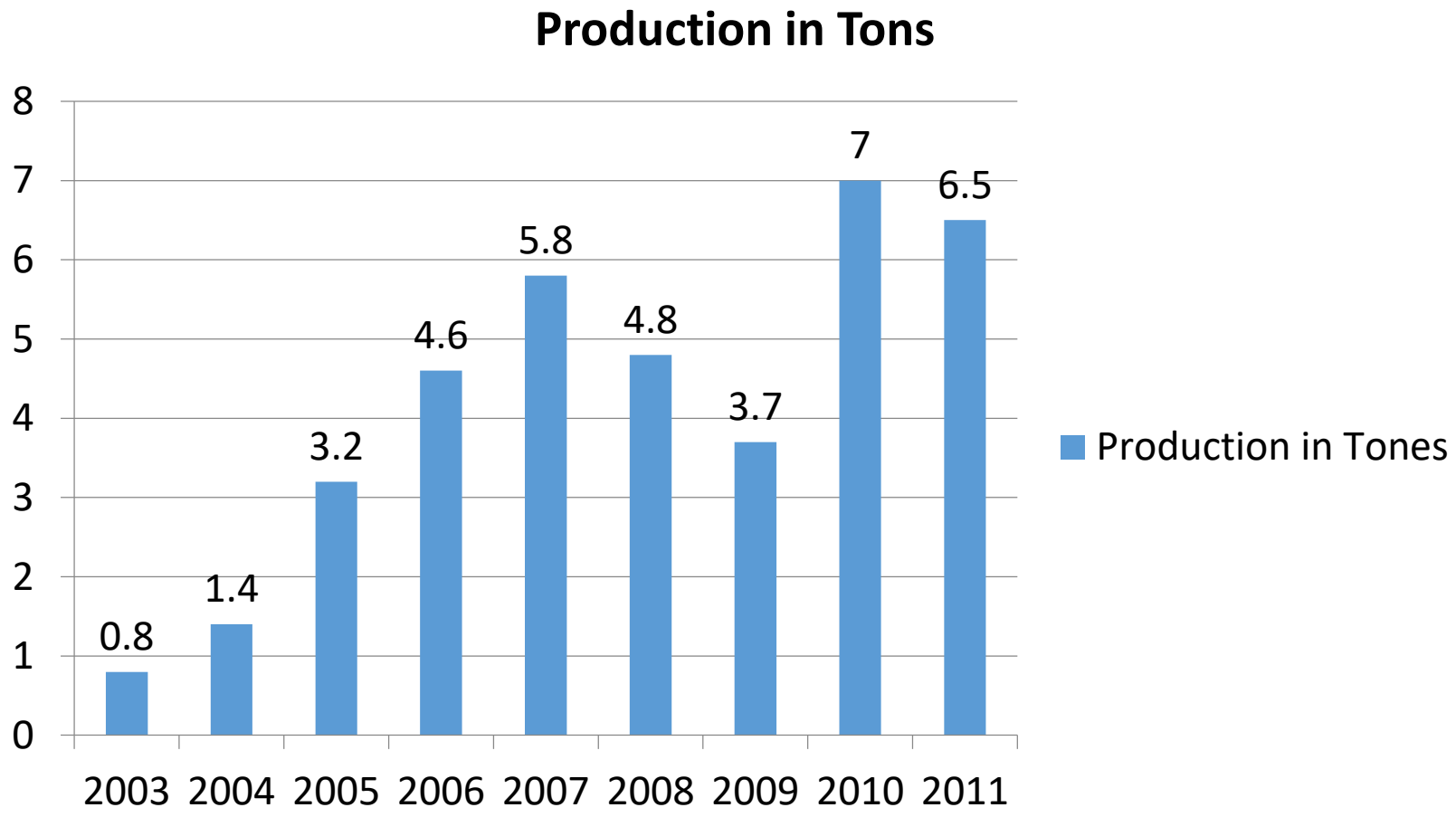
- Diagrams presents the data in a simple and interesting way and it is easy to understand them.
- Diagrams look attractive and arouse the interest of the readers.
- Diagram have visual appeal and hence are quite impressive. They are not only remember but impression left by them on the mind lasts much longer than that left by figures in a table.
- Diagrams help us in making quick comparison of data relating to different time and places.
- Diagram helps to study the relation between two or more sets of data easily and quickly.
- Diagram saves a lot of trouble and time. Figures are not easily understood. One must make an effort to grasp their meaning and draw proper conclusion from them. They give a clear picture of the data at a single glance and no time to effort is lost.

Types of Diagrams

i. Simple bar diagram
ii. Multiple bar diagram
iii. Sub-divided bar diagram
iv. Percentage bar diagram
v. Pie chart (circular bar diagram)

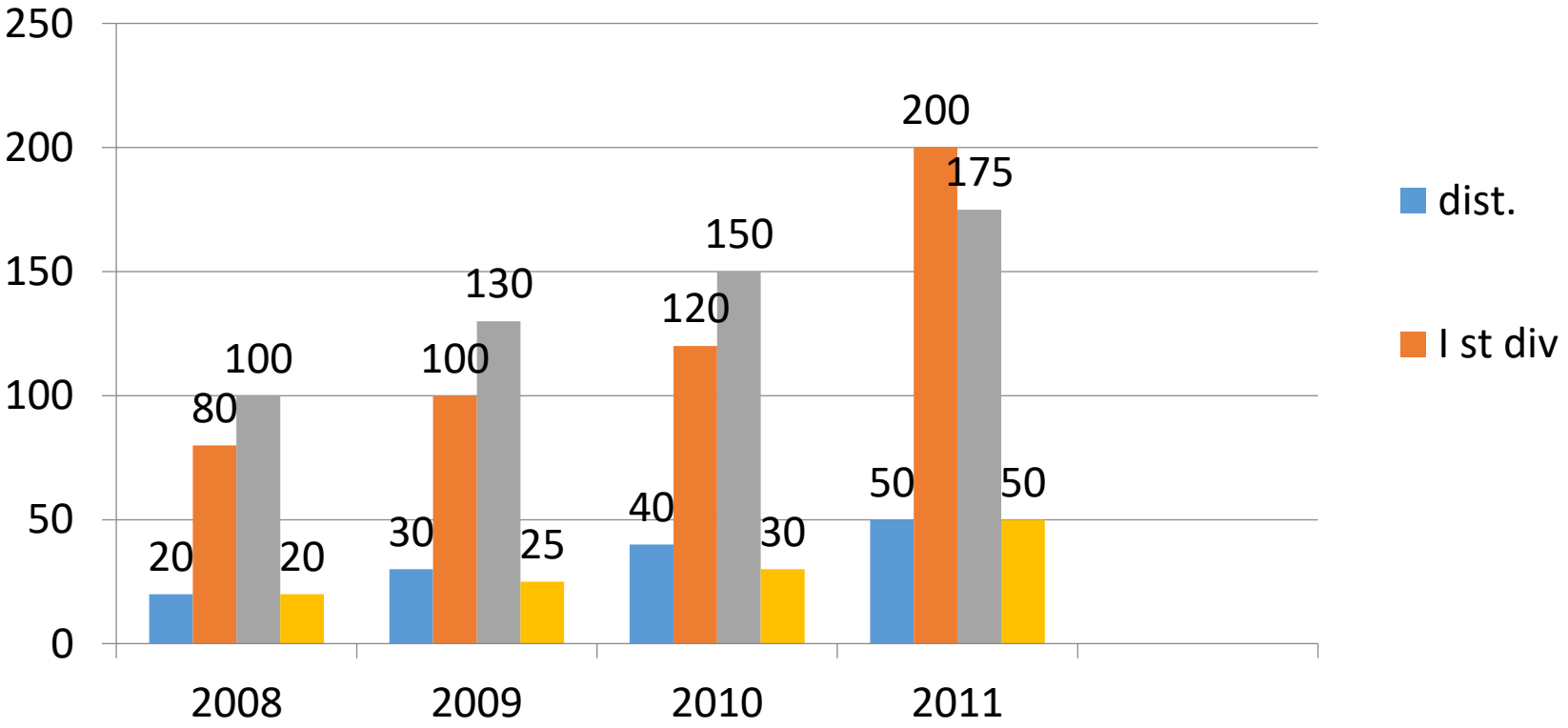
Simple bar diagram

Years	2003	2004	2005	2006	2007	2008	2009	2010	2011
Production (in 100 tons)	0.8	1.4	3.2	4.6	5.8	4.8	3.7	7	6.5



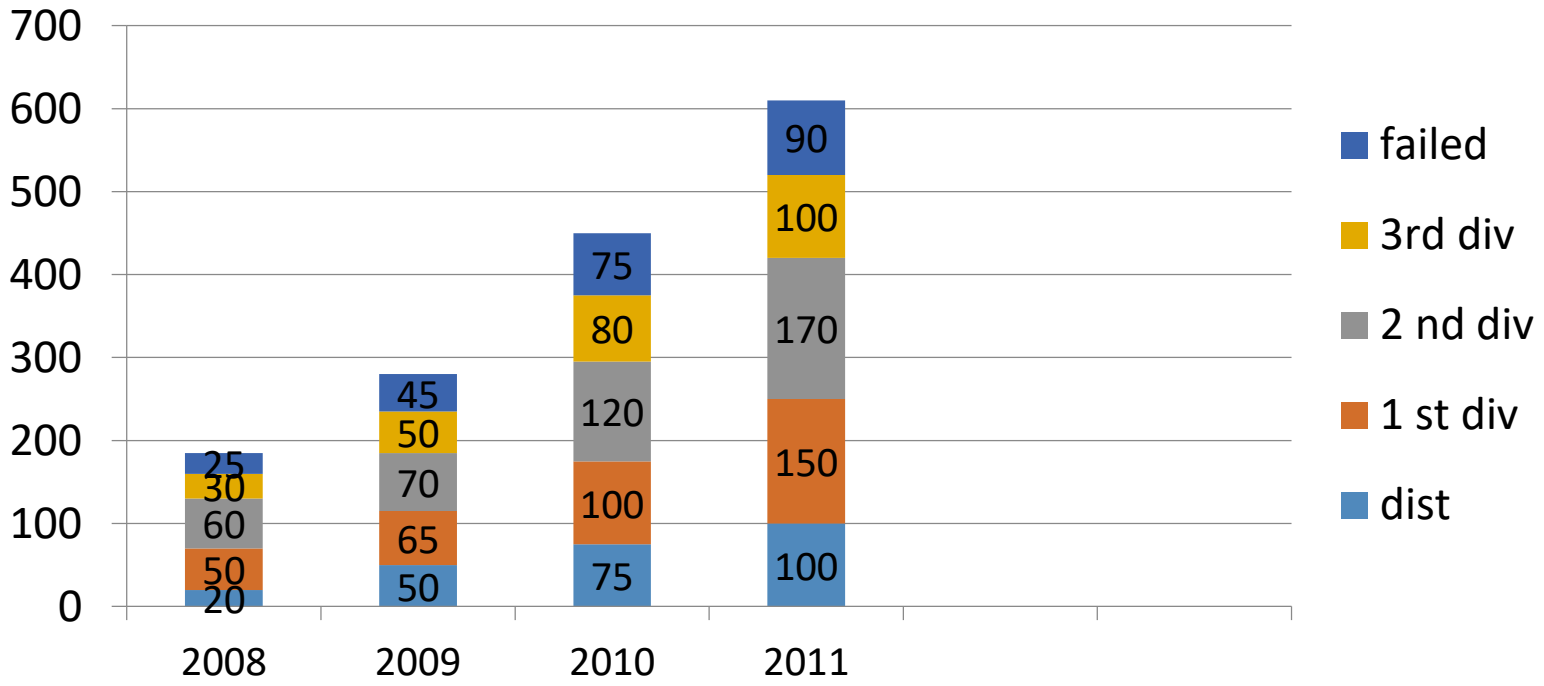
Multiple Bar Diagram

Year	Dist	1 st div	2 nd div	3 rd div
2008	20	80	100	20
2009	30	100	130	25
2010	40	120	150	30
2011	50	200	175	50

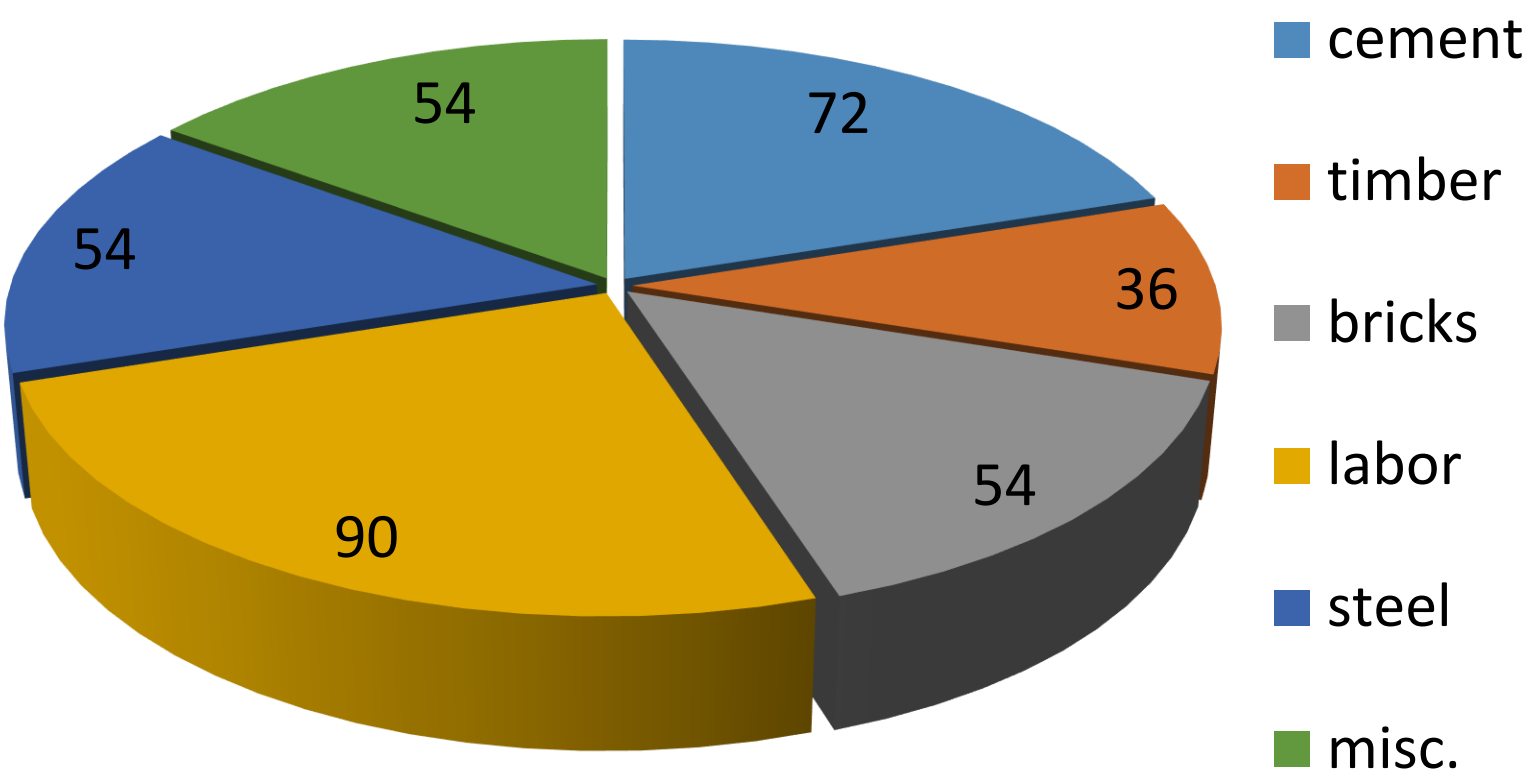


Sub- Divided Bar Diagram

Year	Dist	1 st div	2 nd div	3 rd div	failed
2008	20	50	60	30	25
2009	50	65	70	50	45
2010	75	100	120	80	75
2011	100	150	170	100	90



Circular bar diagram

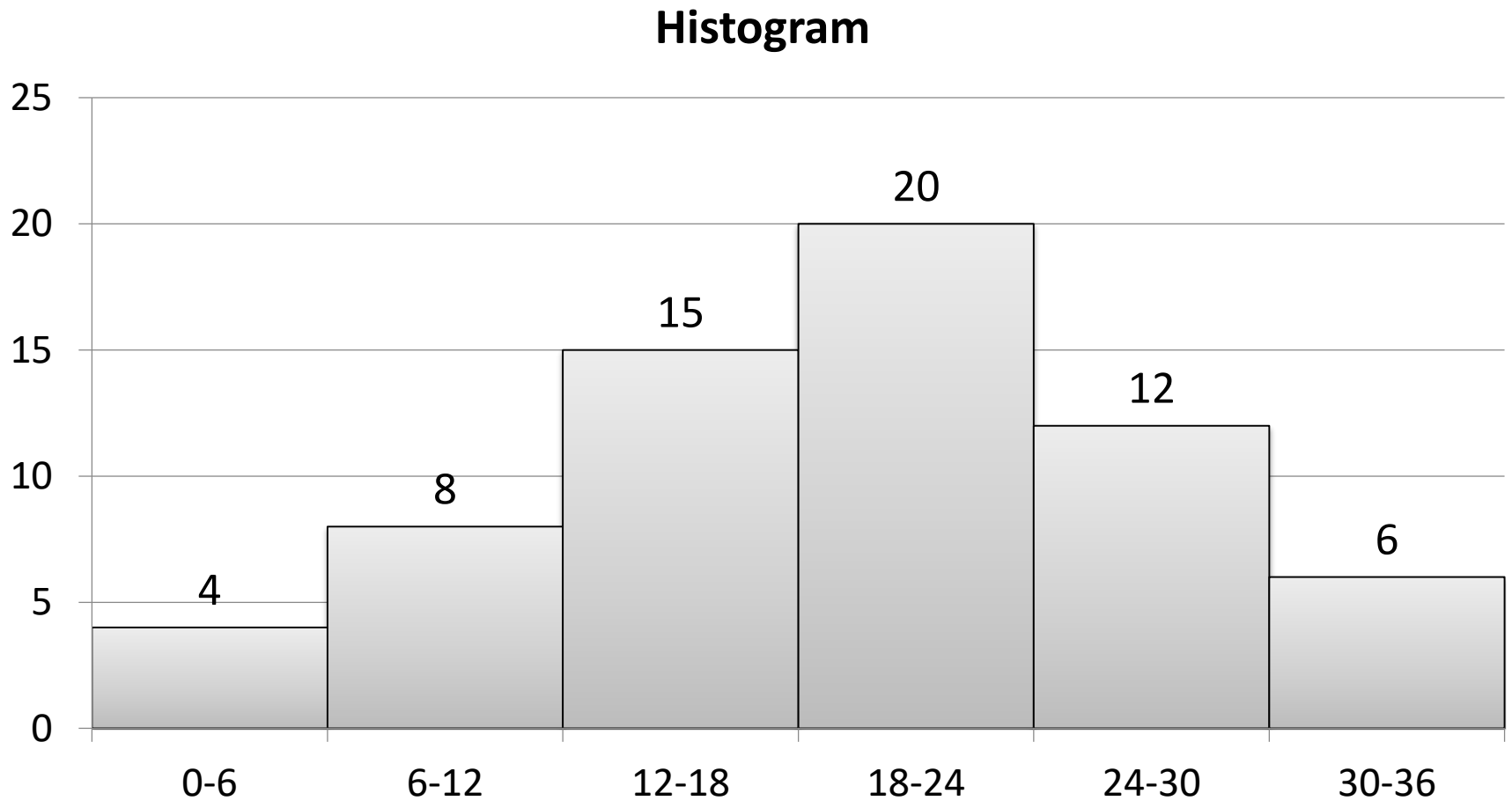


Comparison between histogram and frequency polygon

Histogram	Frequency Polygon
1. It is essentially the bar graph of the given frequency distribution. 2. It does not provide better conception of the contour of the distribution. 3. It is poorly useful. 4. The histogram gives a very clear as well as accurate picture of the relative proportions of the frequency from interval to interval.	1. The frequency polygon is a line graph of the frequency distribution. 2. It gives much better conception of the contour of the distribution. 3. It is more useful and practicable. 4. In the frequency polygon, it is assumed that the frequencies are concentrated at the mid-points of the each classes. It merely points out the graphical relationship between the mid-points and frequencies.

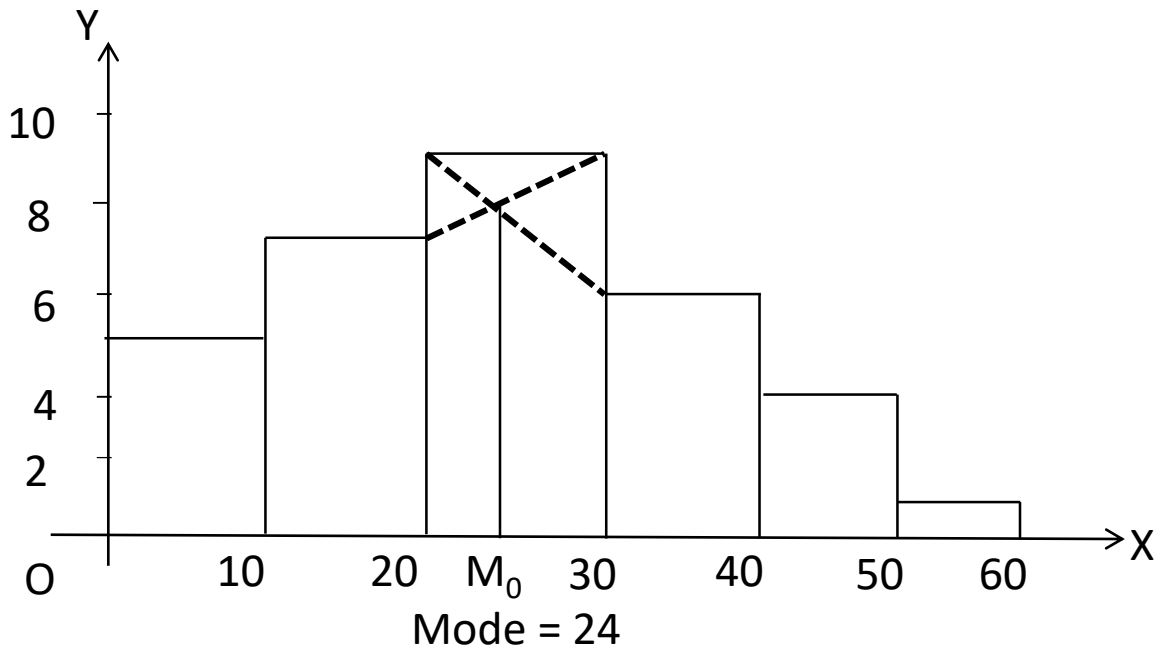
Histogram

Class	0 - 6	6 - 12	12 - 18	18 - 24	24 – 30	30 - 36
Frequency	4	8	15	20	12	6



Histogram and Mode

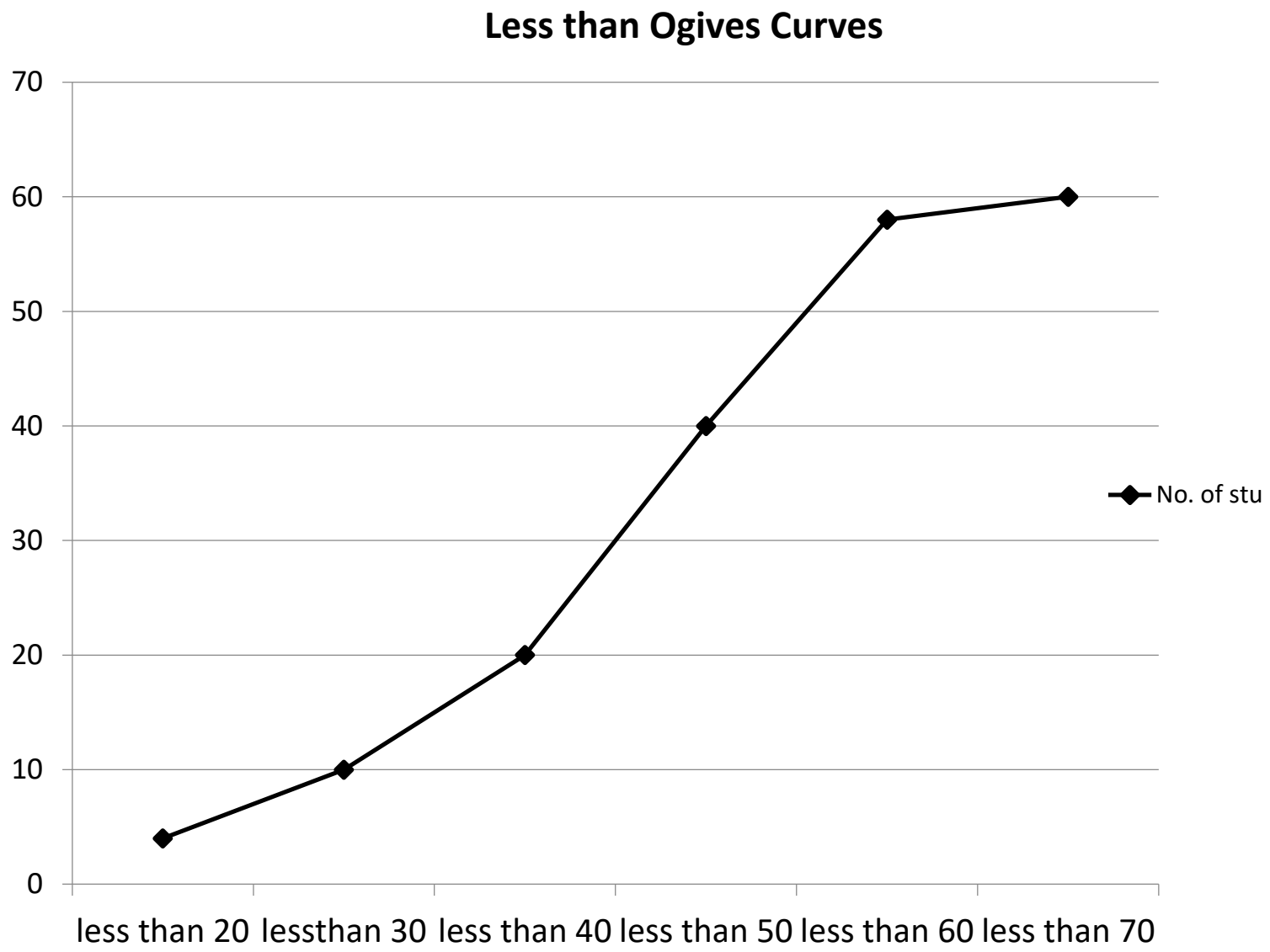
Class	0 – 10	10 – 20	20 – 30	30 – 40	40 – 50	50 – 60
Frequency	5	7	9	6	4	1



Ogives Curves

Marks	No of students
10 – 20	4
20 – 30	6
30 – 40	10
40 – 50	20
50 – 60	18
60 – 70	2

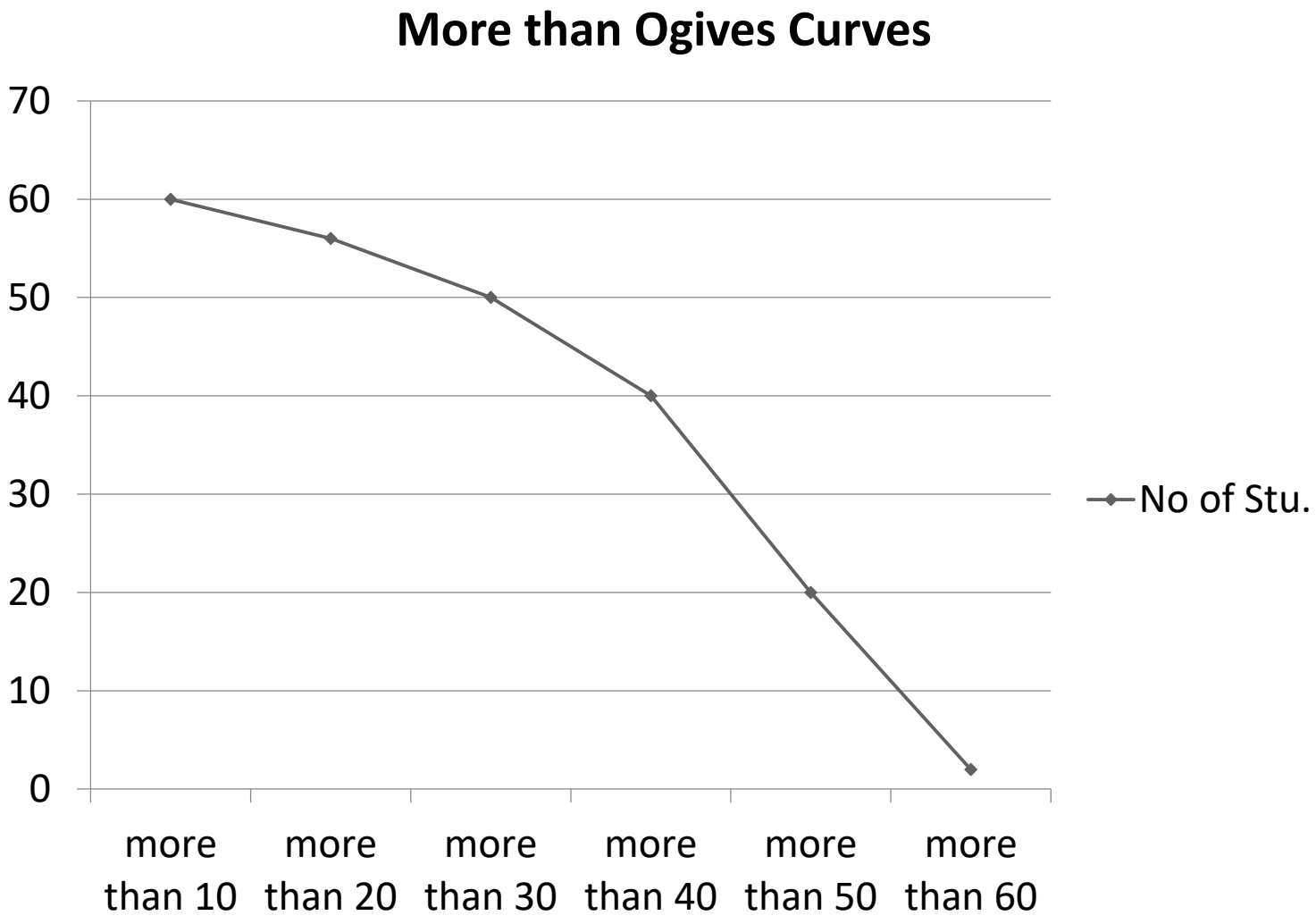
Marks	No. of Students
Less than 20	4
Less than 30	10
Less than 40	20
Less than 50	40
Less than 60	58
Less than 70	60


Less than Ogives Curves
70
60
50
40
30
20
10
0
No. of stu
less than 20
lessthan 30
less than 40
less than 50
less than 60
less than 70

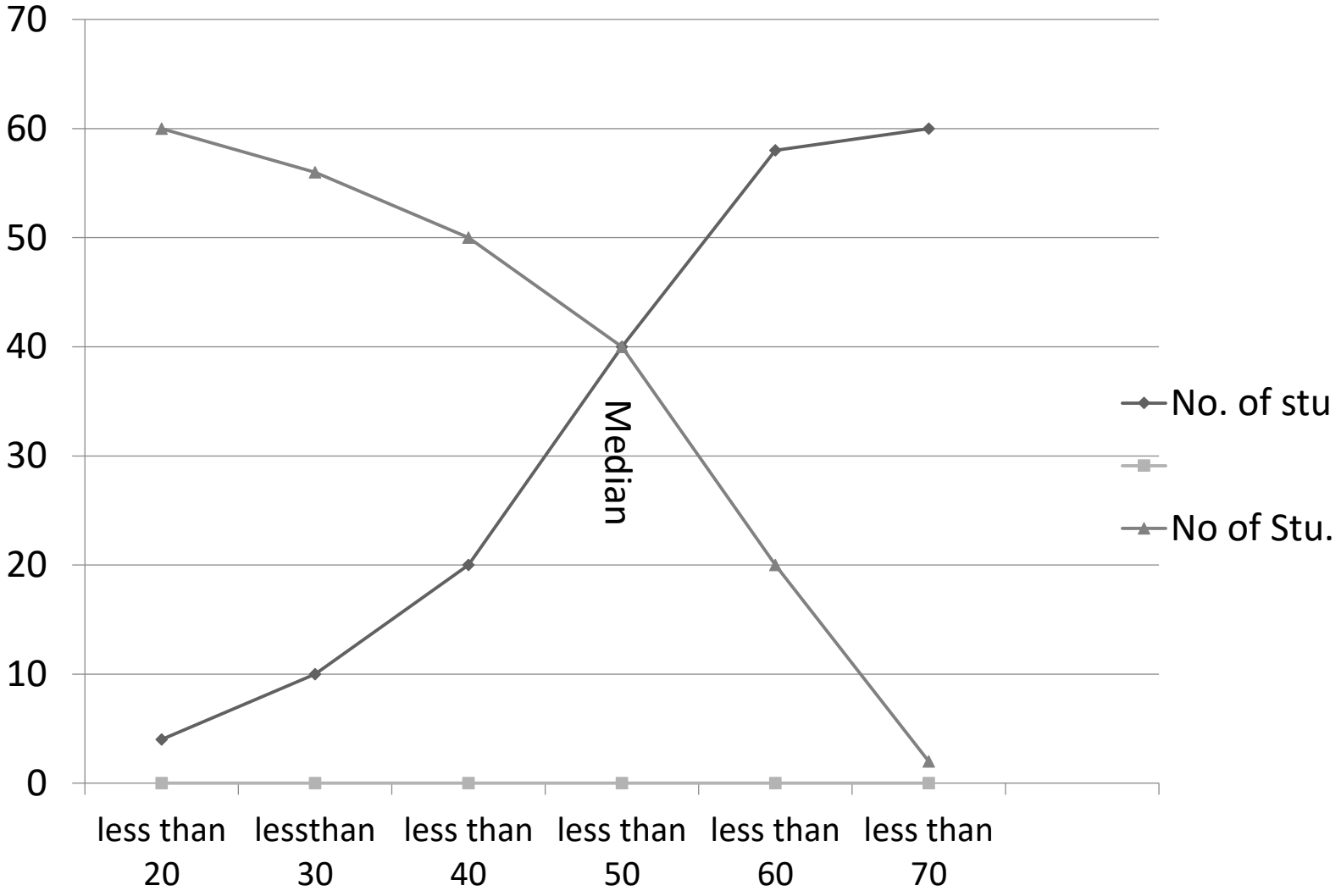
More than Ogives curve

Marks	No of students
10 – 20	4
20 – 30	6
30 – 40	10
40 – 50	20
50 – 60	18
60 – 70	2

Marks	No. of Students
More than 10	60
More than 20	56
More than 30	50
More than 40	40
More than 50	20
More than 60	2


More than Ogives Curves
70
60
50
40
30
20
10
0
more than 10
more than 20
more than 30
more than 40
more than 50
more than 60
No of Stu.

More than and less than Ogives curves



Classwork

- Following are the marks obtained by 50 students in Statistics.

(i) Construct the frequency distribution using the Struge's rule.

(ii) Draw histogram, frequency polygon, frequency curve and calculate mode using histogram.

(iii) Construct Ogives and find median using Ogives.

70	55	51	42	57	45	60	47	63	53
33	65	39	82	55	64	58	61	65	42
50	52	53	45	45	25	36	59	63	39
65	45	49	54	64	75	42	41	52	35
30	35	48	15	26	20	40	55	46	18

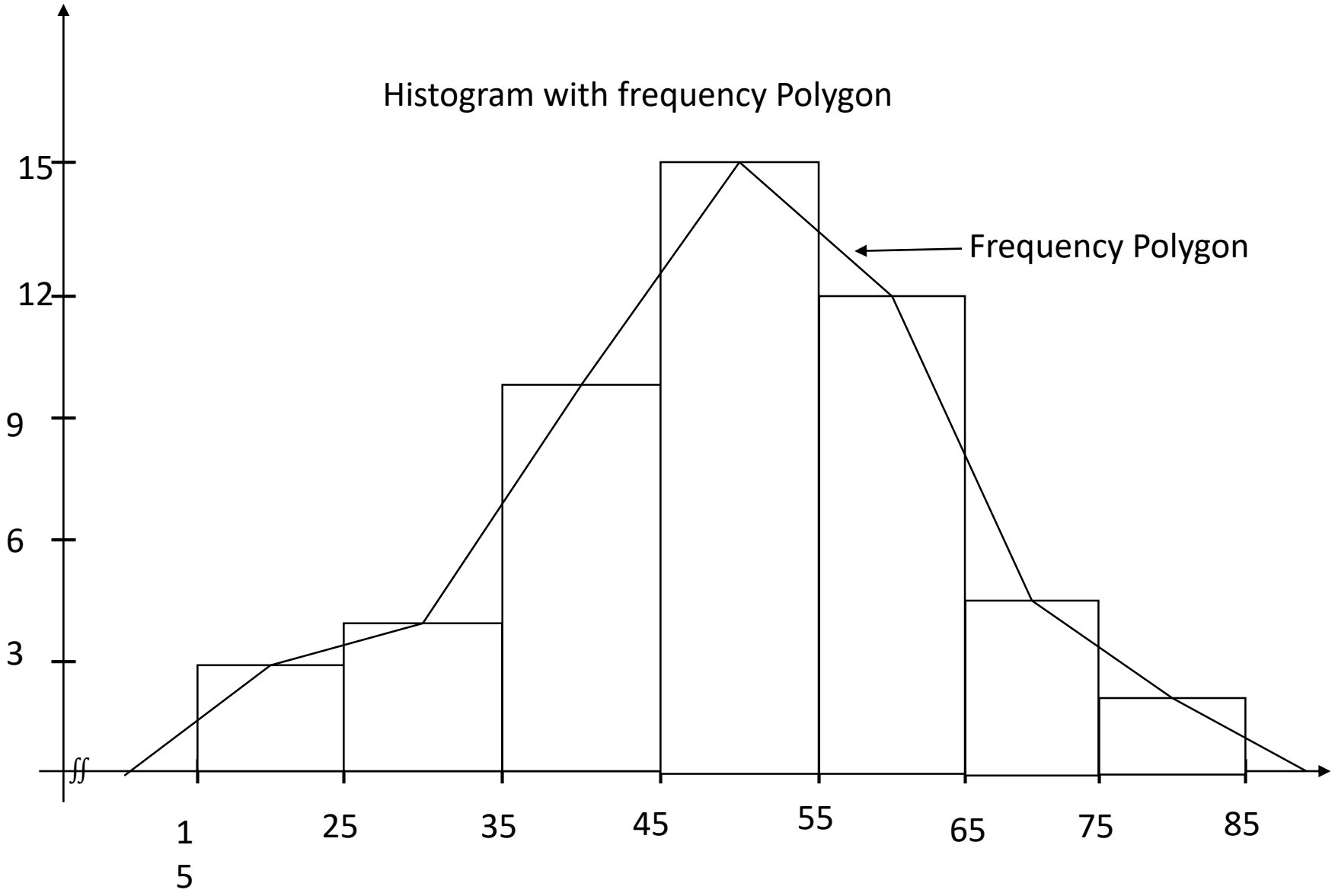
Solution: No. of observations(n) = 50

No. of classes(k) = 1 + 3.322 Log N = 1 + 3.322 log 50= 6.64 ≈ 7

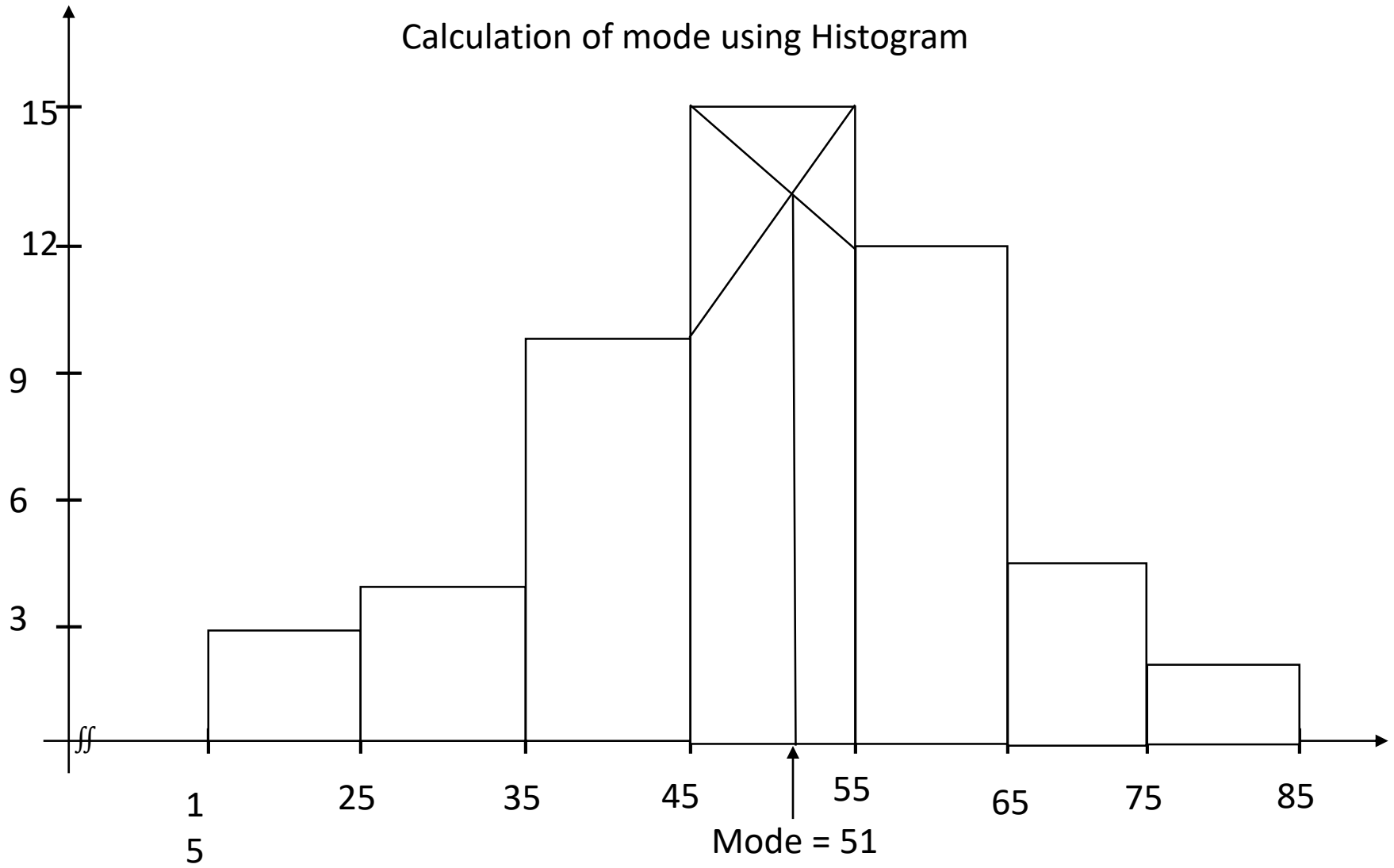
Class interval(w) = $\frac{R}{k} = \frac{L-S}{7} = \frac{82-15}{7} = 9.57 \approx 10$

Construction of frequency distribution

Class	Tally Bars	Frequency	C. F.	R. F.	% of R.F.
15 – 25		3	3	3/50 = 0.06	6%
25 – 35		4	4+3 = 7	4/50 = 0.08	8%
35 – 45		10	10 + 7 = 17	10/50 = 0.2	20%
45 – 55		15	15 + 17 = 32	15/50 = 0.3	30%
55 – 65		12	12 + 32 = 44	12/50 = 0.24	24%
65 – 75		4	4 + 44 = 48	4/50 = 0.08	8%
75 – 85		2	2 + 48 = 50	2/50 = 0.04	4%
Total		N = 50		1	

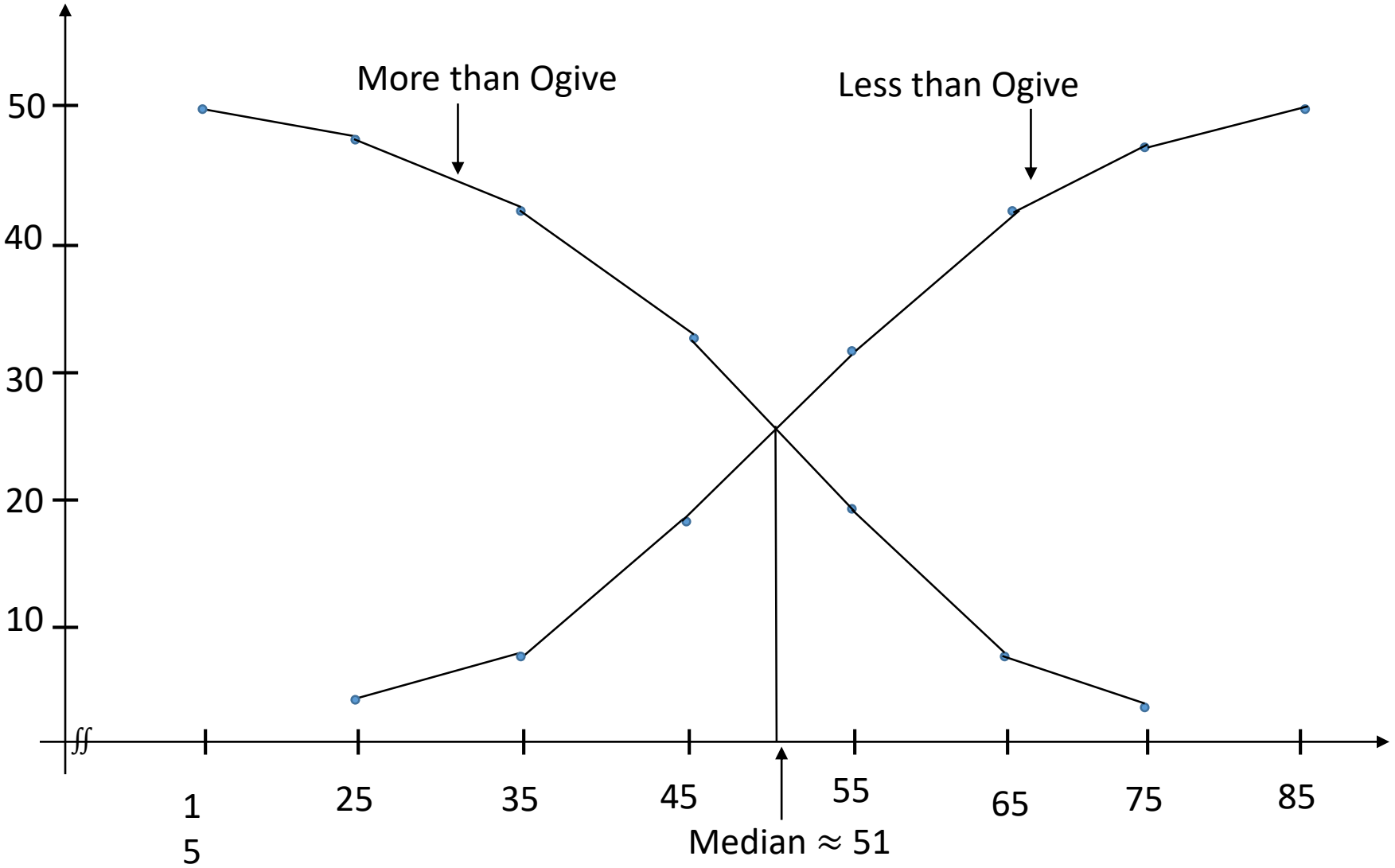

Histogram with frequency Polygon
15
12
9
6
3
Frequency Polygon
1
5
25
35
45
55
65
75
85

Calculation of mode using Histogram



Class	Frequency	Less than	Frequency	More than	Frequency
15 – 25	3	Less than 25	3	More than 15	50
25 – 35	4	Less than 35	4+3 = 7	More than 25	50 – 3 = 47
35 – 45	10	Less than 45	10 + 7 = 17	More than 35	47 – 4 = 43
45 – 55	15	Less than 55	15 + 17 = 32	More than 45	43 – 10 = 33
55 – 65	12	Less than 65	12 + 32 = 44	More than 55	33 – 15 = 18
65 – 75	4	Less than 75	4 + 44 = 48	More than 65	18 – 12 = 6
75 – 85	2	Less than 85	2 + 48 = 50	More than 75	6 – 4 =2
Total	N = 50				

Calculation of median using Ogives



Classwork

- 1. The marks in Statistics of 300 students are recorded in the following table:

Marks	30 – 40	40 – 50	50 – 60	60 – 70	70 – 80	80 – 90	90 – 100
No. of Students	15	30	55	90	60	33	17

- Draw histogram and estimate mode from the graph.
- 2. The records of the sales in a small grocery store for the 360 days that it opened during the year.

Sales ($100)	< 2	2 – 3	3 – 4	4 – 5	5 – 6	6 – 7	7 – 8	> 8
No. of days	15	27	64	72	86	70	16	10

- Days for which sales fall below $325 are classified as ‘poor’ and those for which the sales exceed $775 are classified as ‘good’ in 2002.
- (i) Draw less than cumulative frequency graph.
- (ii) Use graph to estimate the number of poor days. (Ans: 58 days)
- (iii) Use graph to estimate the number of good days. (Ans: 360 – 346 = 14 days)

Classwork

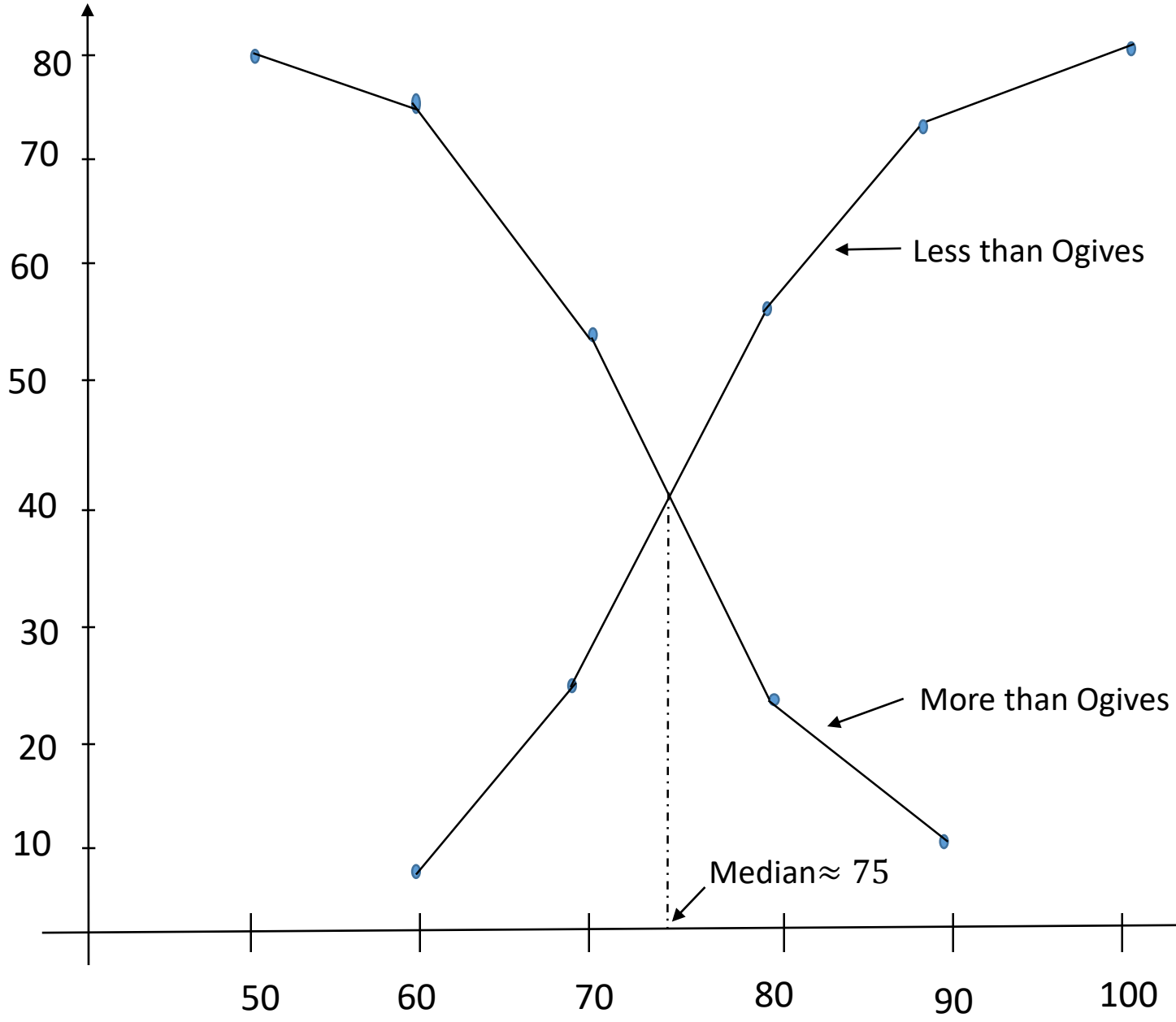
The final grades in Statistics of 80 students is recorded in the comparing table as:

68	84	75	82	68	90	62	88	76	93
73	79	88	73	60	93	71	59	85	75
61	65	75	87	75	62	95	78	63	72
66	78	82	75	94	77	69	74	68	60
96	78	89	61	75	95	60	79	83	71
79	62	67	97	78	85	76	65	71	75
65	80	73	57	88	78	62	76	53	74
86	67	73	81	72	63	76	75	85	77

a. Construct a group frequency distribution having class width of 10.
b. Draw histogram, frequency polygon and frequency curve in the same plane.
c. Does it possible to locate the modal mark with the help of histogram? If yes, locate the value.
d. Draw less than and more than Ogive in a same plane. Locate the median graphically.
e. Draw less than relative cumulative frequency curve.

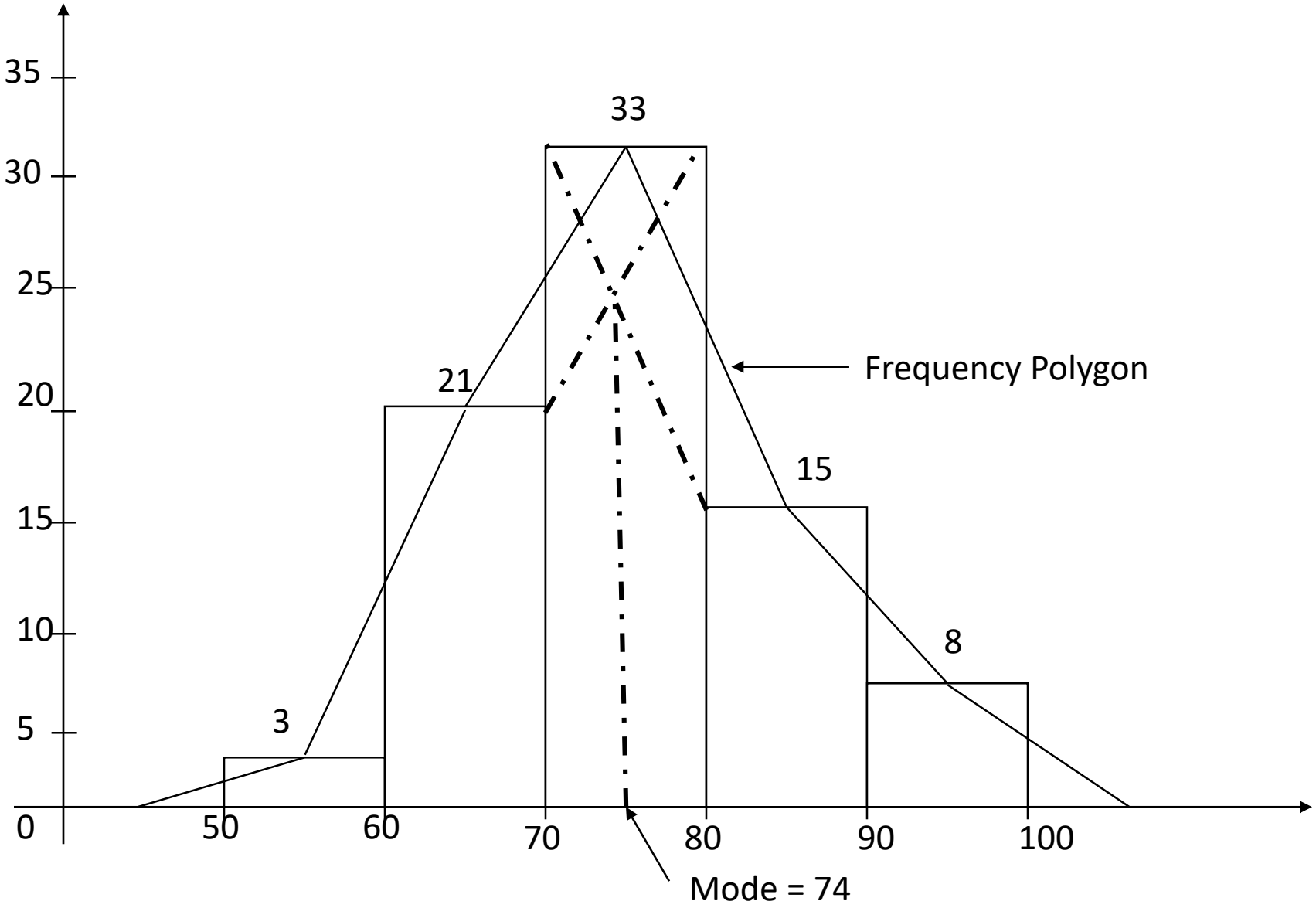
Solution

Class	Frequency	Less than	Frequency	More than	Frequency
50 – 60	3	Less than 60	3	More than 50	80
60 – 70	21	Less than 70	3 + 21 = 24	More than 60	80 – 3 = 77
70 – 80	33	Less than 80	33 + 24 = 57	More than 70	77 – 21 =56
80 – 90	15	Less than 90	15 + 57 = 72	More than 80	56 – 33 = 23
90 – 100	8	Less than 100	8 + 72 = 80	More than 90	23 – 15 = 8



Less than	Frequency
Less than 60	3
Less than 70	3 + 21 = 24
Less than 80	33 + 24 = 57
Less than 90	15 + 57 = 72
Less than 100	8 + 72 = 80

More than	Frequency
More than 50	80
More than 60	80 – 3 = 77
More than 70	77 – 21 =56
More than 80	56 – 33 = 23
More than 90	23 – 15 = 8


35
30
25
20
15
10
5
0
50
60
70
80
90
100
33
21
15
8
3
Frequency Polygon
Mode = 74

Class work

- A National Association of the Real State sellers has collected these data on a sample of 130 salespeople representing their total commission earnings annually.

Commission ($'000)	< 5	5 – 10	10 – 15	15 – 20	20 – 30	30 – 40	40 – 50	> 50
Frequency	5	9	11	33	37	19	9	7

- Construct an Ogive that will help your answer these questions:

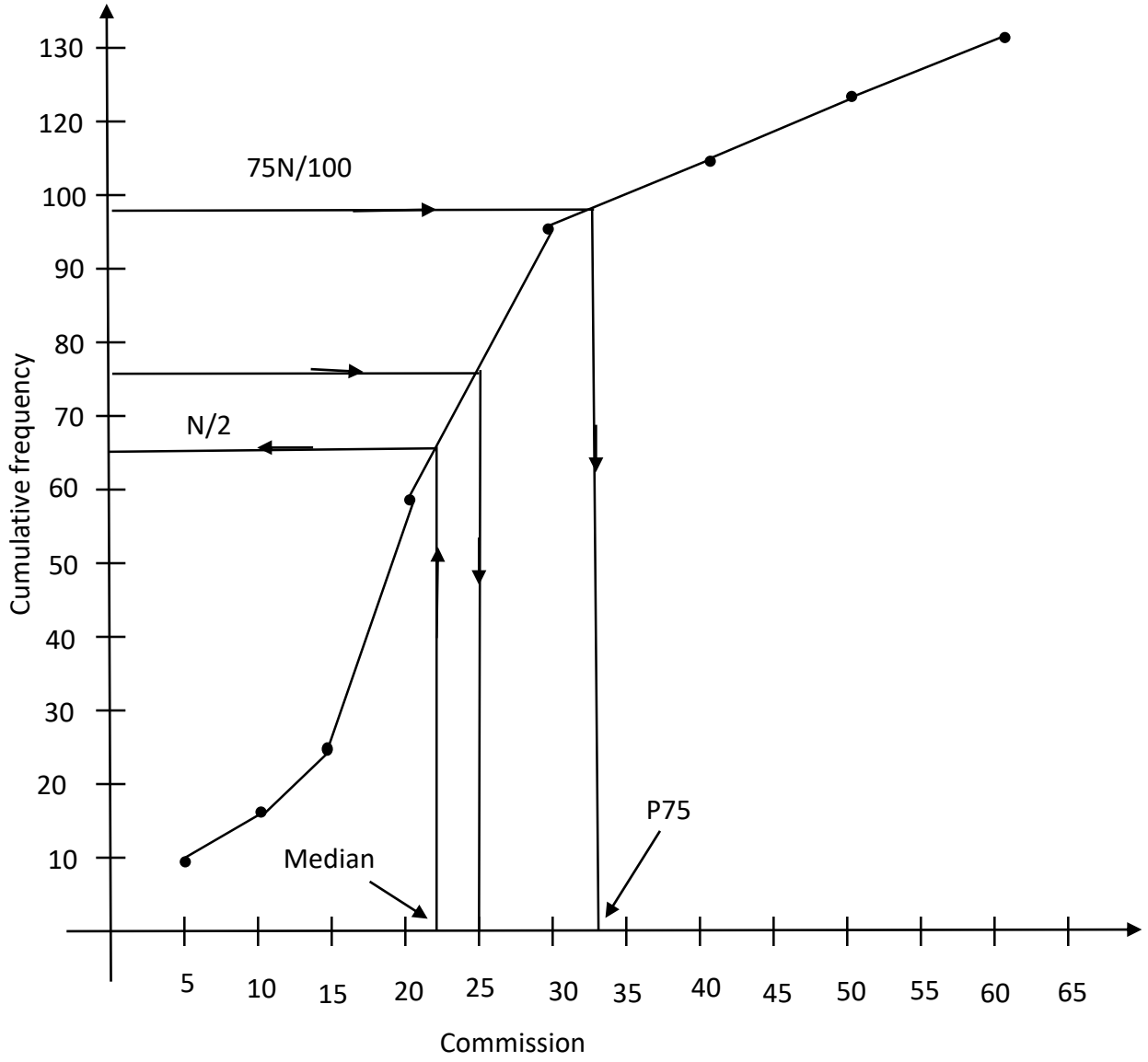
i. About what portion of the salesperson earns more than $ 25,000?

ii. About what does the middle salesperson in the same earn?

iii. Approximately how much could a real estate salesperson whose performance was about 25% from the top expect to earn annually?

Construction of Ogive curve	
Commission ($1000)	Frequency (c. f.)
Below 5	5
Below 10	14
Below 15	25
Below 20	58
Below 30	95
Below 40	114
Below 50	123
Below 60	130

• Less than Ogive curve

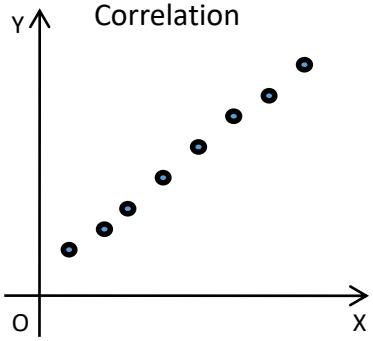


Contd…

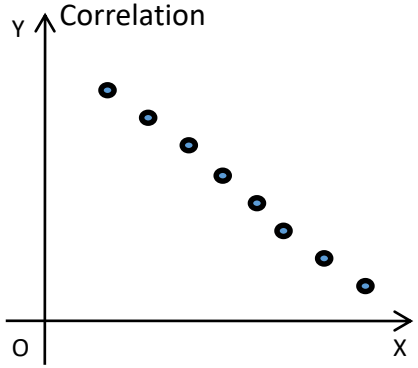
- (i). From Ogive: The number of sales people earns less than $ 25,000 = 76 (approximately). Therefore number of sales persons earns more than $25,000 = 130 – 76 = 54. Required portion of sales people earns more than $ 25,000 = $\frac{54}{130}.100 = 42.54\% \approx 42\%$
- (ii). Commission of the middle sales person in the sample = Median = (N/2) th item = 130/2 = 65. From the Ogive curve Median = 22 (approximately) = $ 22,000.
- (iii). The real estate sales person whose performance was about 25% from the top expect to earn annually = P75 th.
- Now P75 th =$\frac{75\ x\ 130}{100} = 97.5$. from the Ogive P75 = 32(approximately) = $ 32,000.

Scatter plot

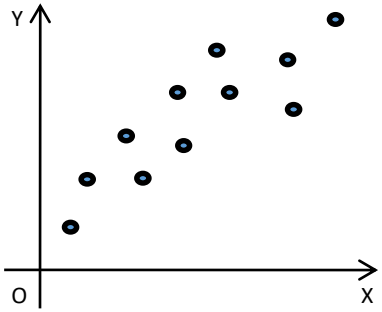
Perfect Positive Correlation



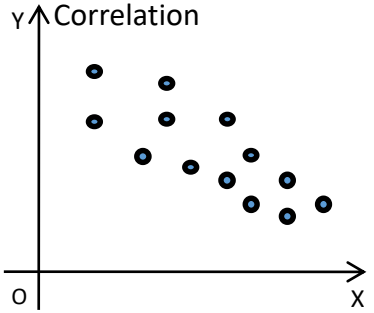
Perfect Negative Correlation



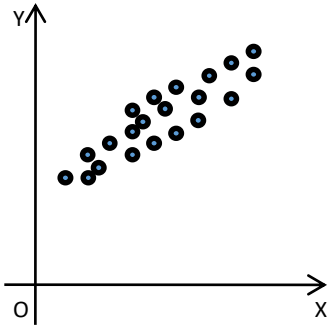
Low Degree of Positive Correlation



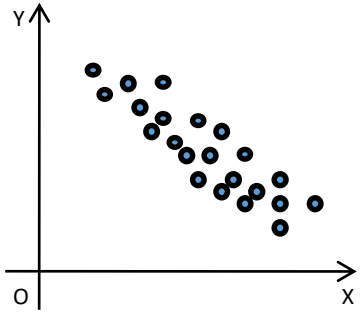
Low Degree of Negative Correlation



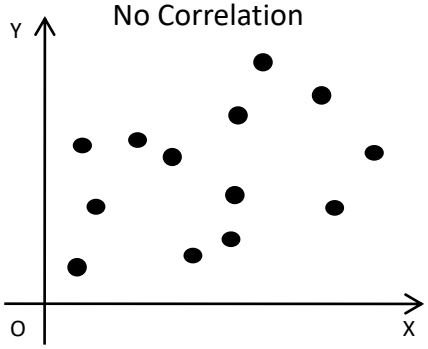
High Degree of Positive Correlation



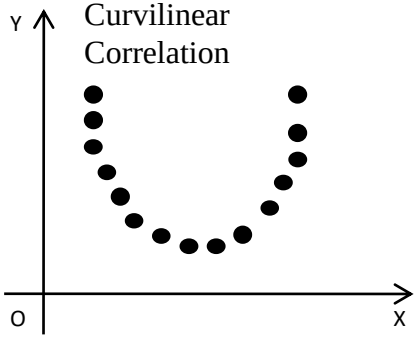
High Degree of Negative Correlation



No Correlation



Curvilinear Correlation



Data Recording

- *Tabulation of data*:
- Tabulation is the process of arranging the data in an orderly manner into rows and columns.
- The purpose of the tabulation is to simplify the presentation of data and to facilitate comparison between related information so collected.
- It is the final steps of collection of data and is a stepping stone to the analysis and interpretation of facts and figures.
- Advantage:

i. helps easy to understand the data.

ii. facilitates comparison between two or more homogeneous information.

iii. huge mass of data can be reduced to rows and columns, tabulation saves time and money.

iv. It helps to calculate statistical measure and interpretation

Different parts of table

1. *Table number*: to identify tables, every table should be numbered normally kept at the top of the table.
2. *Title of the table*: title should be short, clear, simple and non-ambiguous.
3. *Caption(column heading*): title of the column is known as caption. The word caption should also be simple and brief. Caption are usually written at the middle of the column.
4. *Stubs (Row heading*): Stubs means row heading. They are at the extreme left.
5. *Body of the table*: It contains data. The data are entered from the top to bottom in columns and from left to right in rows.
6. *Head note*: a head note is kept just below the title of the table indicating the unit of measurement applicable to the data displayed. It is normally kept within brackets.
7. *Foot note*: are written directly below the body of the table whenever necessary. Captions and stubs should be explained in foot note.
8. *Sources*: The sources from which the data have been taken should be mentioned. The sources should be given at the bottom of the table, just below the foot note.

Format of table

A format of blank table is given below:

Table no.:

Title (in capital letter)

(Head note if any)

Stub heading	Caption (column heading)
Stub entries	Body of table

Foot note:

Sources:

Simple table

1. *Simple table*: A simple table is one in which the data are presented according to only one character.

Table No.

Title: AREAS OF SAARC COUNTRIES

Head note:(in square kilometers)

Countries	Area
Bangladesh	1,43,998
Bhutan	46,500
India	32,87,590
Maldives	9,000
Nepal	1,47,181
Pakistan	7,69,099
Srilanka	65,610

Source: Statistical pocket book 1988

Two way table

Table No.

Title: STUDENTS OF SECONDARY SCHOOL, ACCORDING TO GENDER FOR 2011 – 2012

	Example 3			
Patients	Treatment			Total
		DOTS	No DOTS	
	Cured	150	100	250
	Not cured	15	70	85
	Total	165	170	335

Sources:

Age of wife (years)	Age of Husband (years)				
	20 – 30	30 – 40	40 – 50	50 – 60	Total
15 – 25	5	7	4	6	22
25 – 35	3	4	6	7	20
35 – 45	9	12	14	8	43
45 – 55	2	5	6	9	22
Total	19	28	30	30	N= 107

Three ways table:

Table No.

Title: DISTRIBUTION OF STUDENTS OF T. U. ACCORDING TO GENDER, AGE & FACULTY

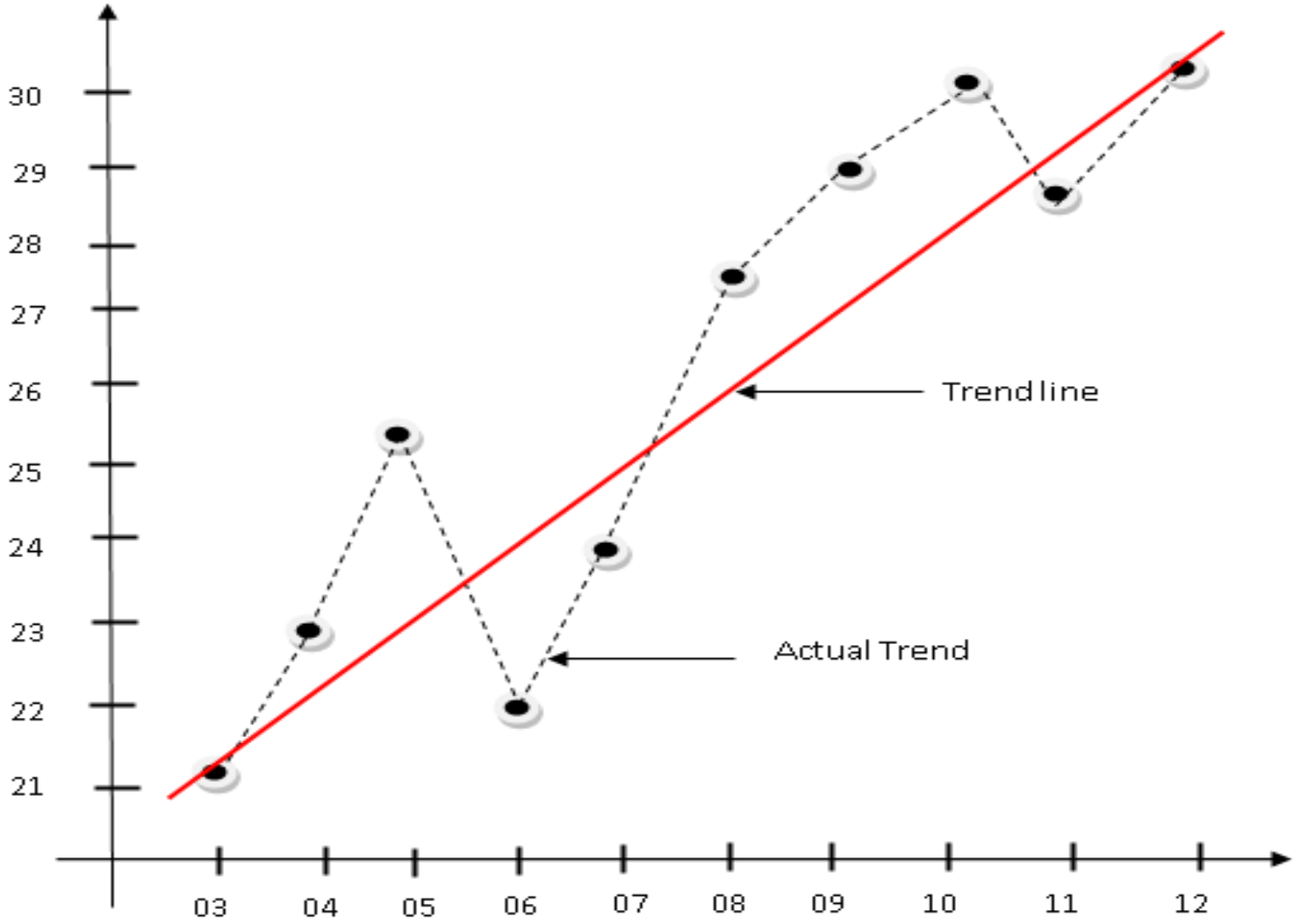
Age	Faculty																				
	Humanity				Management				Science				Education				Others				Total
	B	G	T	To	B	G	T	To	B	G	T	To	B	G	T	To	B	G	T	To	
19-22																					
22-25																					
27-above																					
Total																					

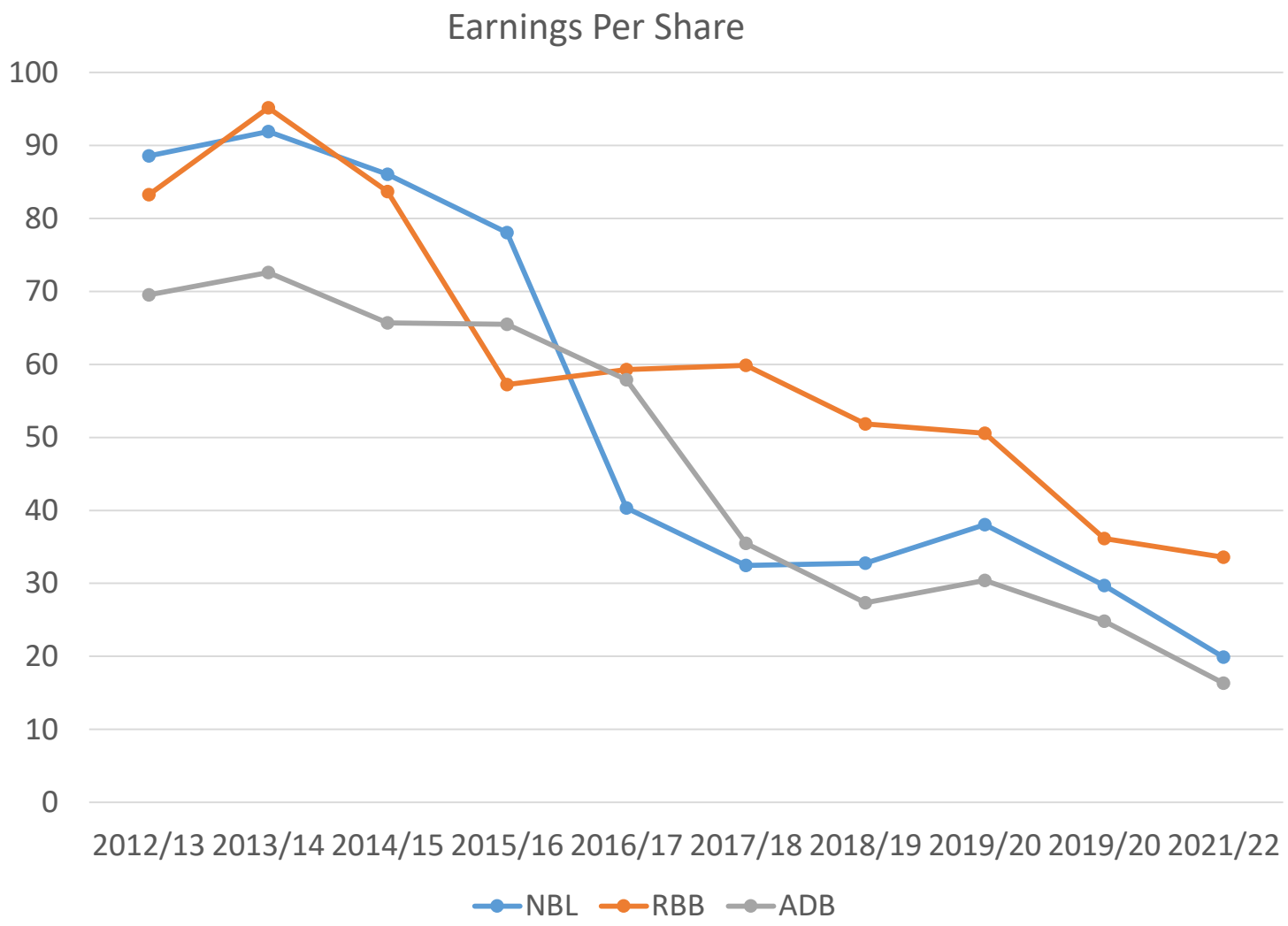
Foot note: B = Boys, G = girls, T = (transgender) , To = Total

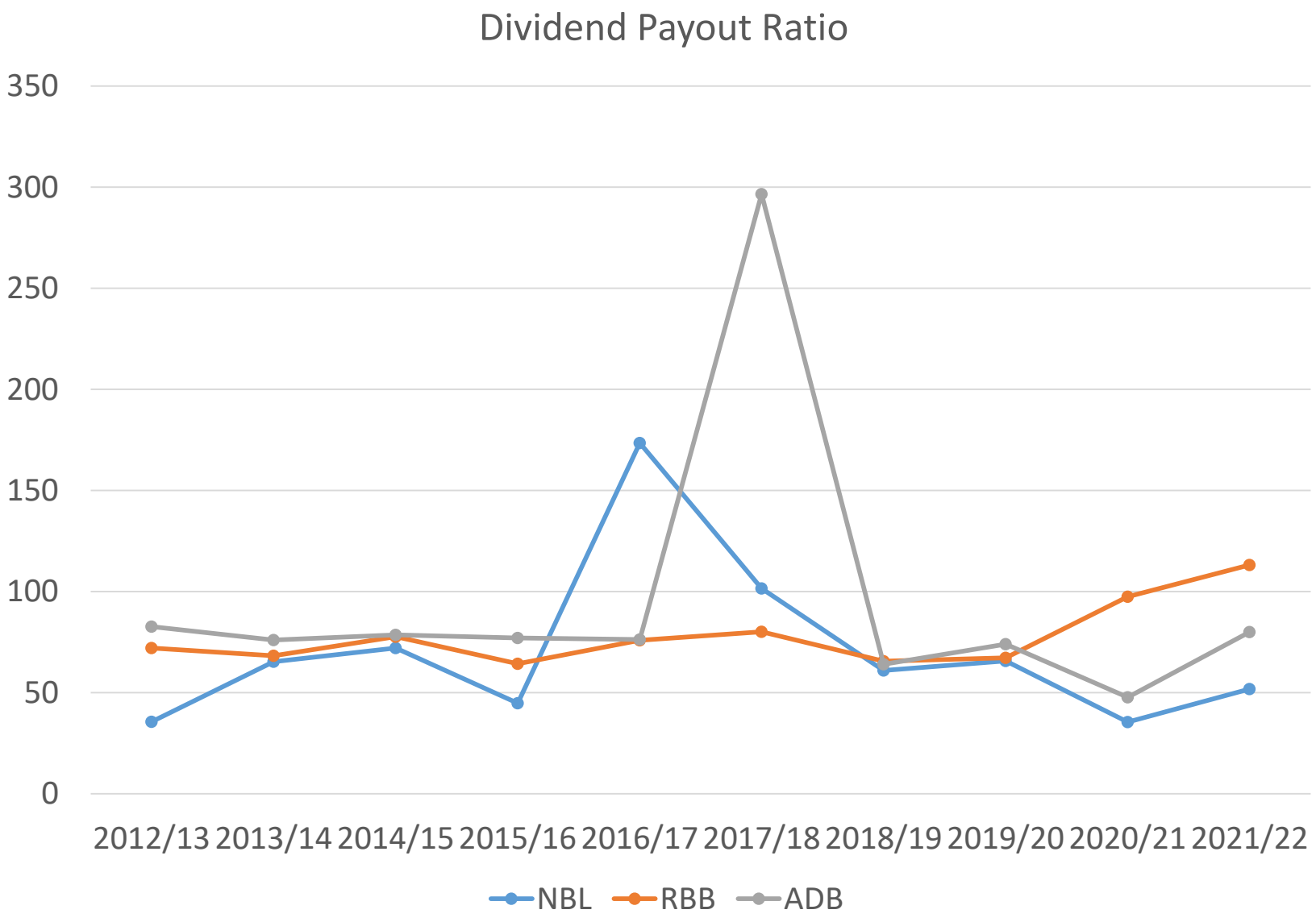
Source:

Time Plots

Year	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Sales	21.6	22.9	25.5	21.9	23.9	27.5	29	29.7	28.6	29.9

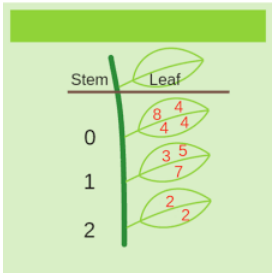



Earnings Per Share
100
90
80
70
60
50
40
30
20
10
0
2012/13
2013/14
2014/15
2015/16
2016/17
2017/18
2018/19
2019/20
2019/20
2021/22
NBL
RBB
ADB

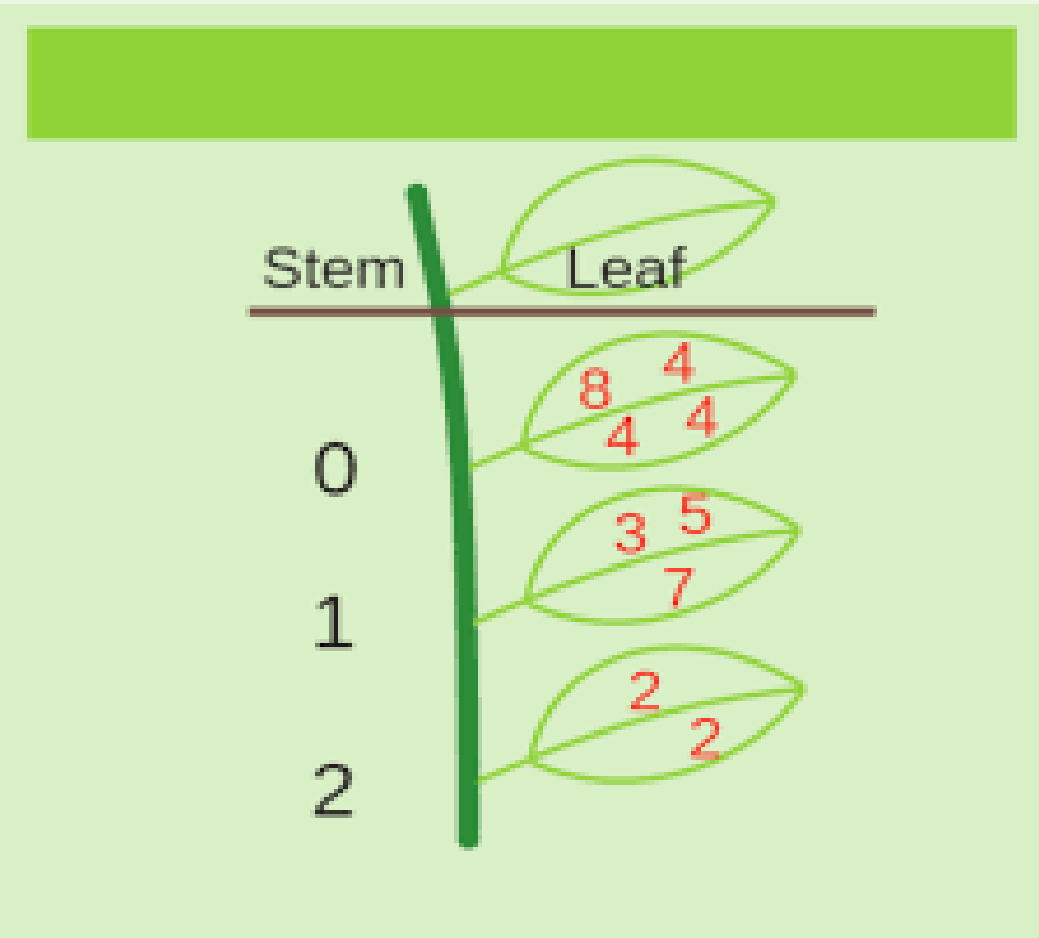

Dividend Payout Ratio
350
300
250
200
150
100
50
0
2012/13
2013/14
2014/15
2015/16
2016/17
2017/18
2018/19
2019/20
2020/21
2021/22
NBL
RBB
ADB

Stem and leaf display

- It is an order array of a sequence of the data in a rank order from smallest value to the largest of the given data set.
- The greatest common place value of the data set is used to form the stem.
- The next greatest common place value is used to form the leaves.
- It is useful to see how the data are distributed and where the concentrations of the data occur.
- steam and leaf display to have an idea about the shape of the distribution of the given data



Contd…



Example

- Construct the stem and leaf display for the following data: 70, 72, 75, 64, 58, 83, 80, 76, 75, 68, 65, 57, 78, and 85
- Solution: Arranging the given data set in ascending order of magnitude: 57, 58, 64, 65, 68, 70, 72, 75, 75, 76, 78, 80, 83 and 85.
- Now stem and leaf display of the data as:

Stem (unit 10)	Leaf (unit one)
5	7, 8,
6	4, 5, 8,
7	0, 2, 5, 5, 6, 8,
8	0, 3, 5

Example

- The stem and leaf display represent the bounced check fee for a sample of 25 banks for direct deposit customers.

Stem	Leaves
1	5, 5, 8, 8
2	0, 0, 0, 0, 1, 2, 2, 5, 5, 5, 5, 8, 8, 9
3	4, 5, 6, 7, 7
4	1, 2

What are the two information's obtained from this stem and leaf display?

Solution:

i. Minimum and Maximum bounced check fee are 15 and 42 respectively and the maximum concentration on the bounced check group is "20 – 29".

ii. The most of the bounced check fee (modal bounced checks) are 20 and 25 since each of them are maximum number of times i. e. 4 times.

Thank You