1. The value of π till 50 decimal places is given below:

3.14159265358979323846264338327950288419716939937510

a. Make a frequency table of the digits from 0 to 9 after the decimal point.

b. What are the most and the least frequently occurring digits?

Solution:

a.

Digit الخانة	Frequency التكرار
0	2
1	5
2	5
3	8
4	4
5	5
6	4
7	4
8	5
9	8
Total	50

ليكن لديك قيمة الثابت

π لغاية ٥٠

نزلة بعد الفاصلة العشرية

المطلوب

جدول تكراري يحوي المنازل
الى ٩ بعد الفاصلة العشرية

ما هي المنزلة الاكثر تكرارا

و المنزلة الاقل تكرارا

b. The least frequency digit is 0

The most frequency digits are 3 and 9

2. Using the data shown in the following table:

No. of subjects in which student failed	Frequency
0	8
1	1S
2	12
3	2
Total	40

a. Represent them graphically using pie chart and Bar chart.

b. Compute the range of this data.

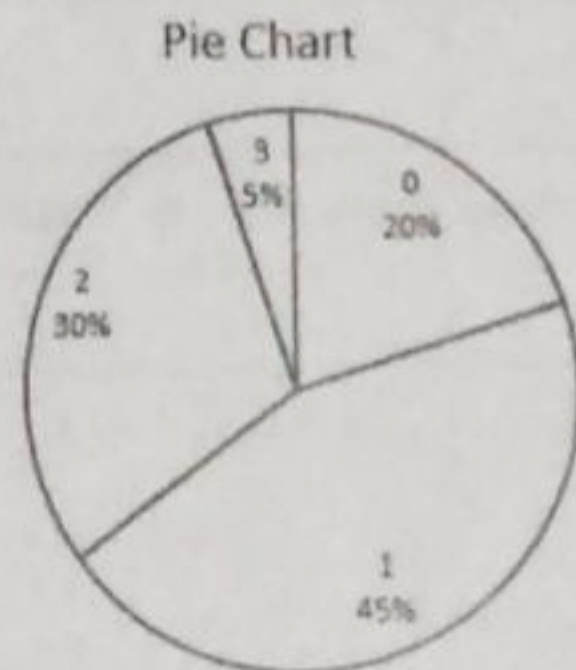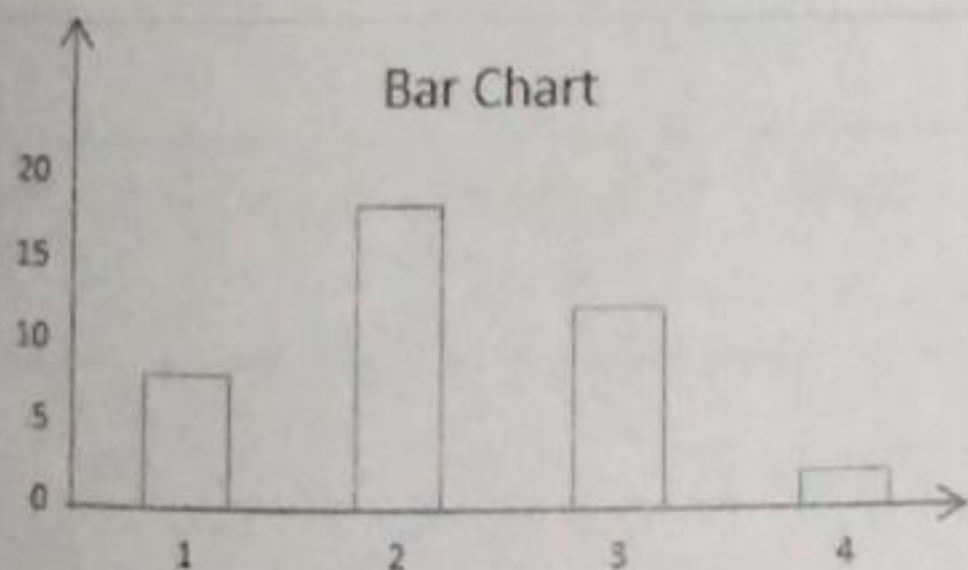**c.** Compute the mean of tis data.

Solution:

a. The angle of category $0 = \dfrac{8}{40} \times 360 = 0.2 \times 360 = 72$ degree

The angle of category $1 = \dfrac{18}{40} \times 360 = 0.45 \times 360 = 162$ degree

The angle of category $2 = \dfrac{12}{40} \times 360 = 0.3 \times 360 = 108$ degree

The angle of category $3 = \dfrac{2}{40} \times 360 = 0.05 \times 360 = 18$ degree

Bar Chart

Pie Chart

b. Range $R = x_l - x_s = 3 - 0 = 3$

c. Mean

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \frac{0 \times 8 + 1 \times 18 + 2 \times 12 + 3 \times 2}{40} = \frac{0 + 18 + 24 + 6}{40} = \frac{48}{40} = 1.2$$

3. The following data give the results of a sample survey. The letters A, B and C represent the three categories:

A	C	B	A	C	B	C	C	C	B
C	B	C	B	C	C	B	C	C	C
A	B	C	C	B	C	B	A	C	C

a. Prepare a frequency table of tis data.

b. Calculate the relative frequencies and percentages for all symbols.

c. What percentage of the elements belongs to category B?

d. Draw a bar chart and pie chart for the frequency table.

Solution:

a. and b.

Relative frequency = $\dfrac{frequncey}{total\ frequency}$

التكرار النسبي $= \dfrac{\text{تكرار المشاهدة}}{\text{مجموع التكرارات}}$

$Percentage\ frequency = Relative\ frequency \times 100\%$

التكرار المنوي $=$ التكرار النسبي $\times 100\%$

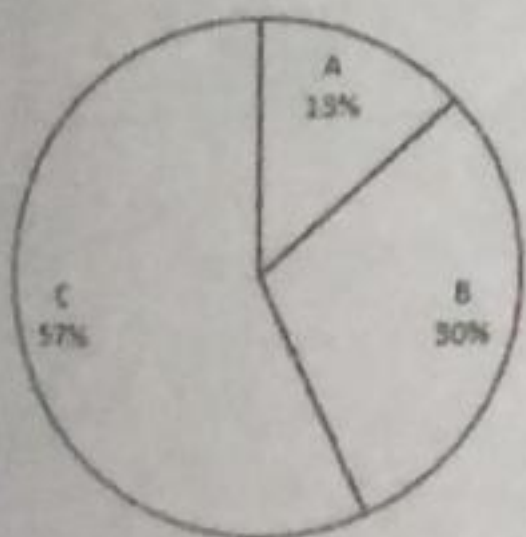| Category المصنف	Frequency التكرار	Relative frequency التكرار النسبي	Percentages frequency التكرار المئوي
A	4	$\dfrac{4}{30} \approx 0.13$	13%
B	9	$\dfrac{9}{30} = 0.3$	30%
C	17	$\dfrac{17}{30} \approx 0.57$	57%
Total	30	1	100%

c. The percentage of **B** is 30%

زاوية القطاع = التكرار النسبي مضروب بـــ ٣٦٠

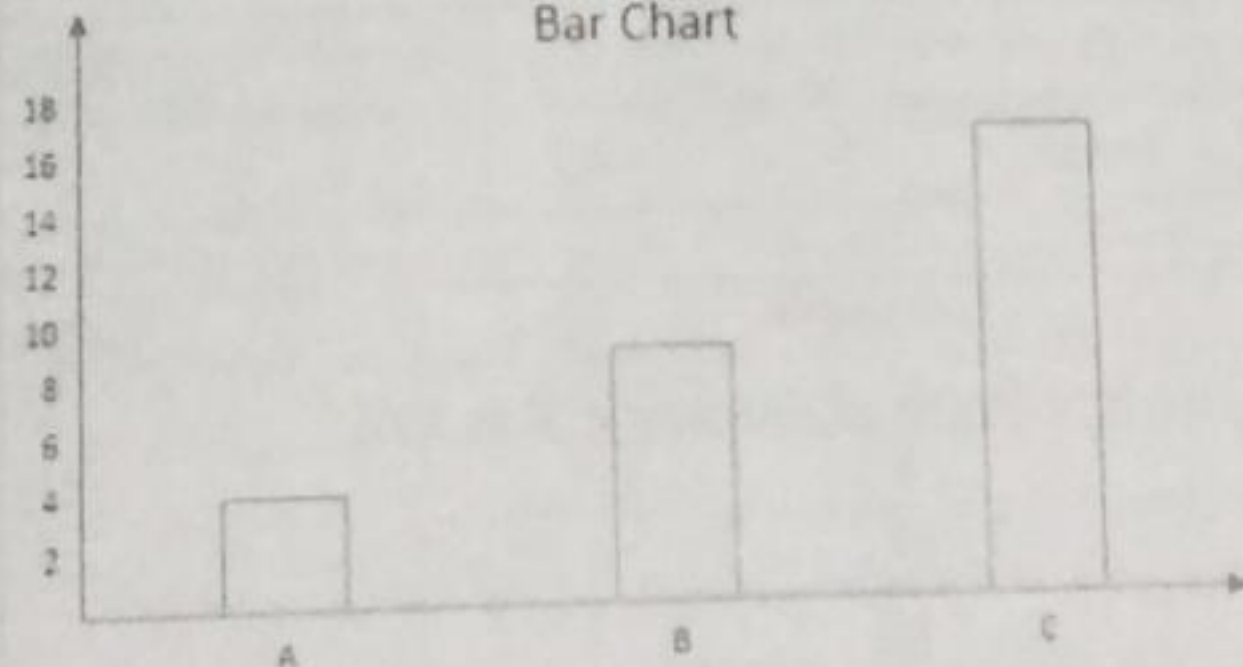d. The angle of category **A** = $\dfrac{4}{30} \times 360 = 48$ degree

The angle of category **B** = $\dfrac{9}{30} \times 360 = 108$ degree

The angle of category **C** = $\dfrac{17}{30} \times 360 = 204$ degree

Pie Chart

Bar Chart

4. The following data give the results of a sample survey. The letters Y, N and D represent the three categories:

N	N	N	Y	Y	Y	N	Y	D	N
Y	Y	Y	Y	N	Y	Y	N	N	D
D	Y	Y	D	D	N	N	N	Y	N
Y	Y	.N	N	Y	Y	N	N	D	Y

a. Prepare a frequency distribution table.

b. Calculate the relative frequencies and percentages for all categories.

c. What percentage of the elements belongs to category Y?

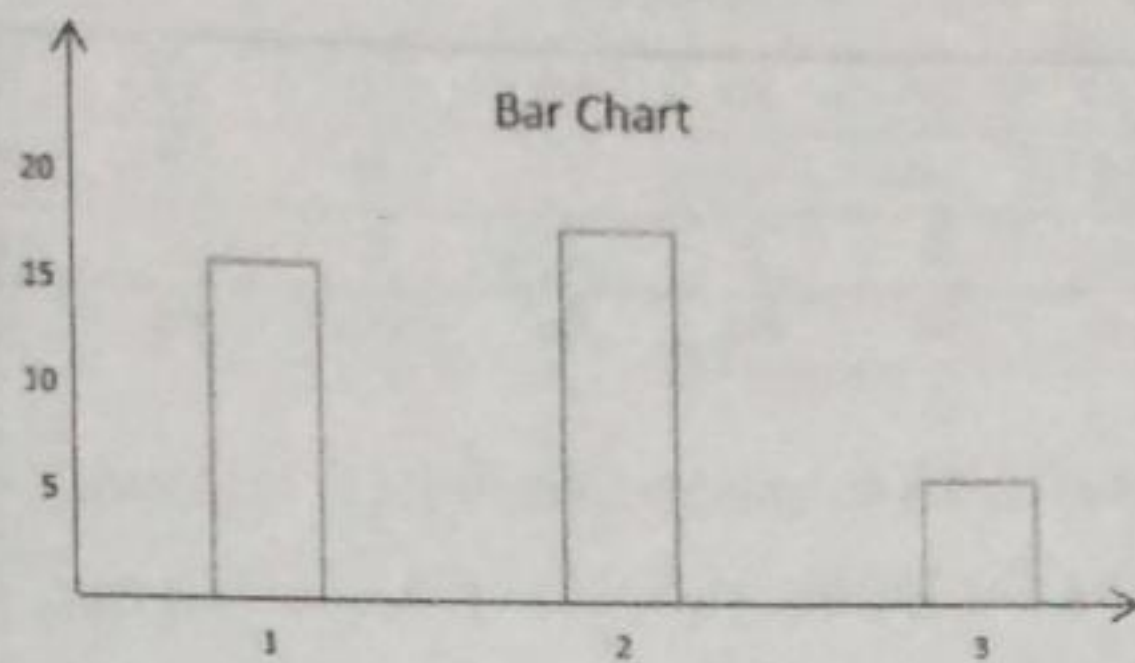**d.** Draw a bar chart and pie chart for the given data.

Solution:
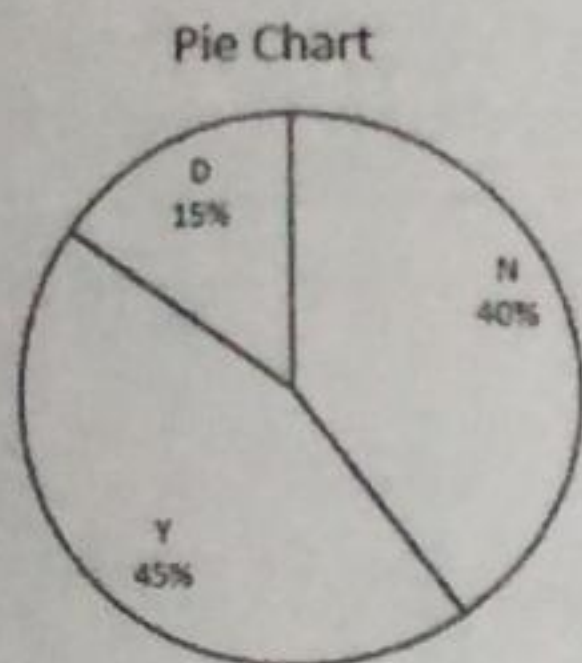
    a. and b.

Category	Frequency	Relative frequency	Percentage frequency
N	16	0.4	40%
Y	18	0.45	45%
D	6	0.15	15%
Total	40	1	100%

    c. The percentage of category Y is 45%

d.

Pie Chart

Bar Chart

5. A company manufactures car batteries of a particular type. The lives (in years) of 40 such batteries were recorded as follows:

2.6	3.0	3.7	3.2	2.2	4.1	3.5	4.5	4.6	3.8
3.5	2.3	3.2	3.4	3.8	3.2	4.6	3.7	2.9	3.6
2.5	4.4	3.4	3.3	2.9	3.0	4.3	2.8	3.5	4.2
3.5	3.2	3.9	3.2	3.2	3.1	3.7	3.4	3.2	2.6

Construct a frequency distribution table for this data, using class intervals of size 0.5 starting from the interval 2 − 2.5 .

Solution:

ملاحظة: العدد 2.5 يوضع مع الفئة 2.5 → 3.0   وليس مع الفئة 2.5 → 2

Class boundaries	Frequency
2.0 → 2.5	2
2.5 → 3.0	6
3.0 → 3.5	14
3.5 → 4.0	11
4.0 → 4.5	4
4.5 → 5.0	3
Total	40

6. The distance (in km) of 40 engineers from their residence to their place of work were found   as follows:

5	3	10	20	25	11	13	7	12	31
19	10	12	17	18	11	32	17	16	2
7	9	7	8	3	5	12	15	18	3
12	14	0.5	9	6	15	15	7	6	12

a. Construct a frequency distribution table with class size 5 for the data given.

b. Draw the histogram for the data of frequency distribution table.

c. Draw the polygon for the data of frequency distribution table.

d. Draw the ogive for the data of frequency distribution table.

e. How many engineers have residence at distance less than 20 km from their workplace?

f. How many engineers have residence at distance more than 15 km from their workplace?
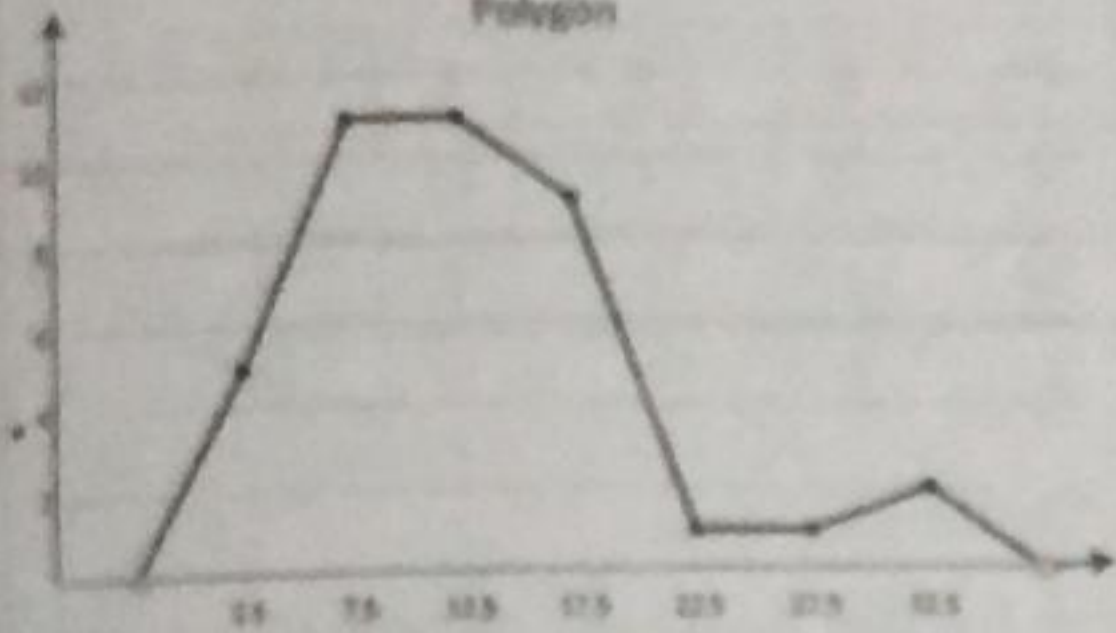
Solution:

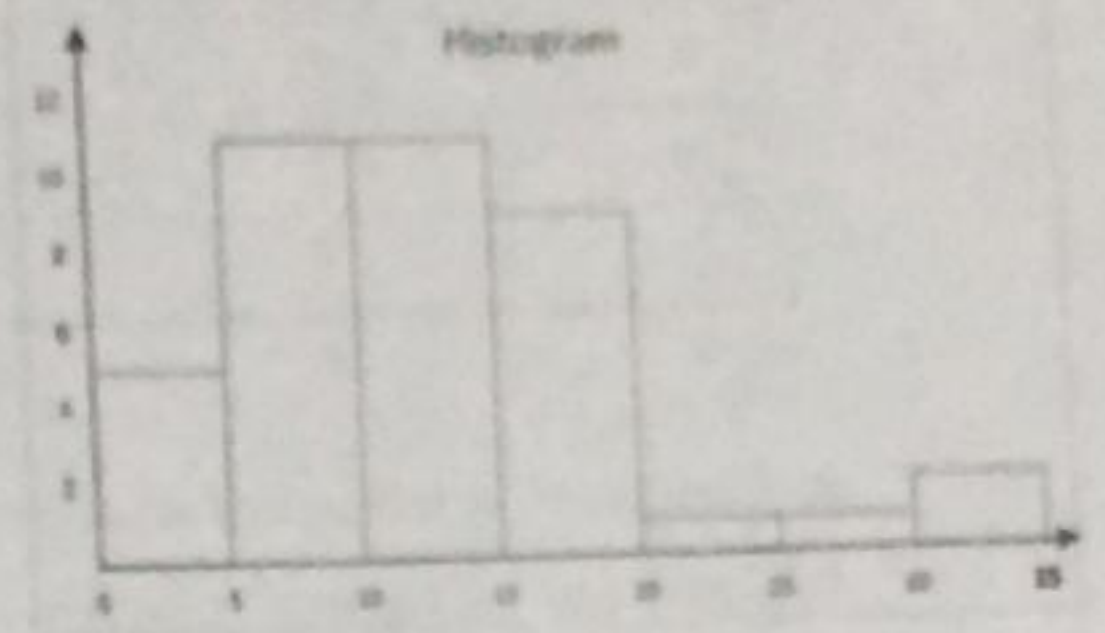$$C = \frac{R}{K} \Rightarrow K = \frac{R}{C}$$

$$K = \frac{31.5}{5} = 6.3 \approx 7 \text{ Classes}$$ اذا نحتاج الى ٧ فئات

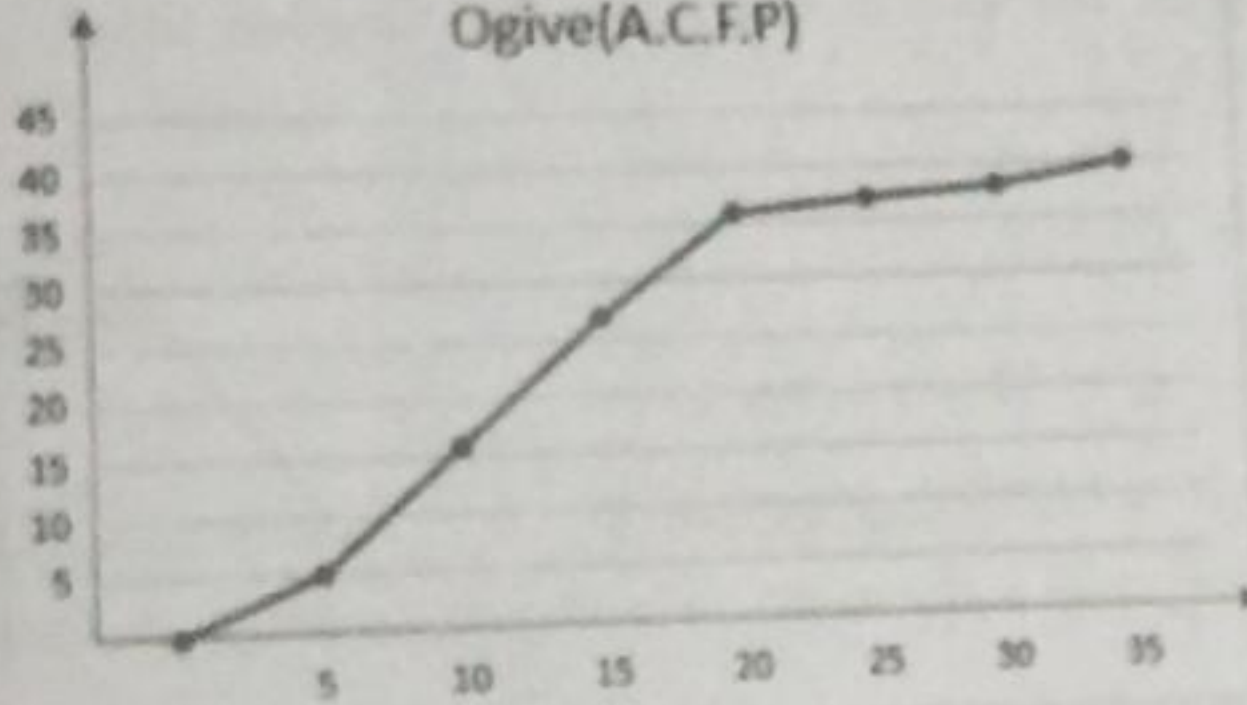| Class Limit	Class Boundaries	Class Midpoint	Frequency	Relative Frequency	Less than	Ascending Cumulative Frequency(AC
0.5 – 4.5	0 → 5	2.5	5	0.125	5	5
5.5 – 9.5	5 → 10	7.5	11	0.275	10	16
10.5 – 14.5	10 → 15	12.5	11	0.275	15	27
15.5 – 19.5	15 → 20	17.5	9	0.225	20	36
20.5 – 24.5	20 → 25	22.5	1	0.025	25	37
25.5 – 29.5	25 → 30	27.5	1	0.025	30	38
30.5 – 34.5	30 → 35	32.5	2	0.05	35	40
Total			40	1		

e.

f.

Polygon

Histogram

Ogive(A.C.F.P)

e. 36

f. $40 - 27 = 13$ OR $9 + 1 + 1 + 2 = 13$

7. Thirty children were asked about the number of hours they watched TV programs in the previous week. The results were found as follows:

8	10	12	14	12	10	8	6	4	2
10	3	4	12	2	8	15	1	17	6
1	6	2	3	5	12	5	8	4	8
3	2	8	5	9	6	8	7	14	12

a. Construct a frequency distribution table for this data.

b. Draw the histogram, polygon and ogive for the frequency distribution table.

c. Draw the polygon for the data of frequency distribution table.

d. Draw the ogive for the data of frequency distribution table.

Solution:

لتكوين جدول التوزيع التكراري

$$Range\ (R) = x_l - x_s = 17 - 1 = 16$$

١- نجد المدى

$$Numer\ of\ class\ K = \lfloor 3.322 Log(40) \rfloor = 5$$

٢- نجد عدد الفئات المطلوب حسب القاعده

$$Class\ length\ C = \frac{R}{K} = \frac{16}{5} = 3.2 \approx 4$$

٣- نجد طول الفئة (مع الانتباه الى ان التقريب يكون للعدد الصحيح الاكبر)
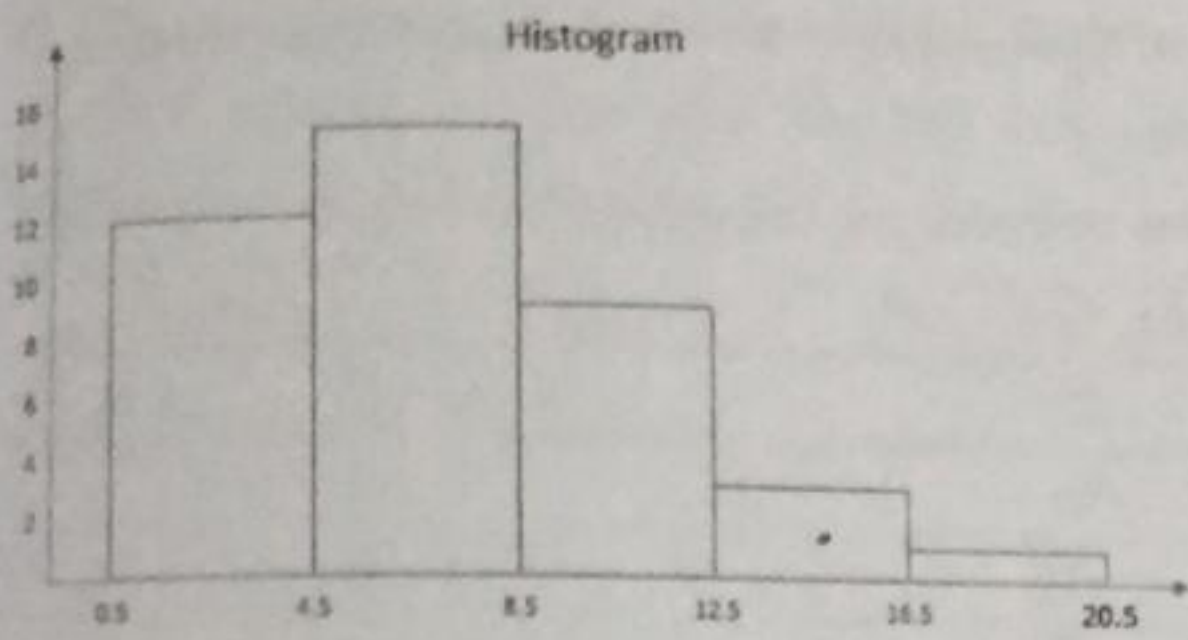
$$first\ class\ boundary\ is\ (\ 0.5 - 4.5\ )$$

٤- نبدأ بتكوين الفئات
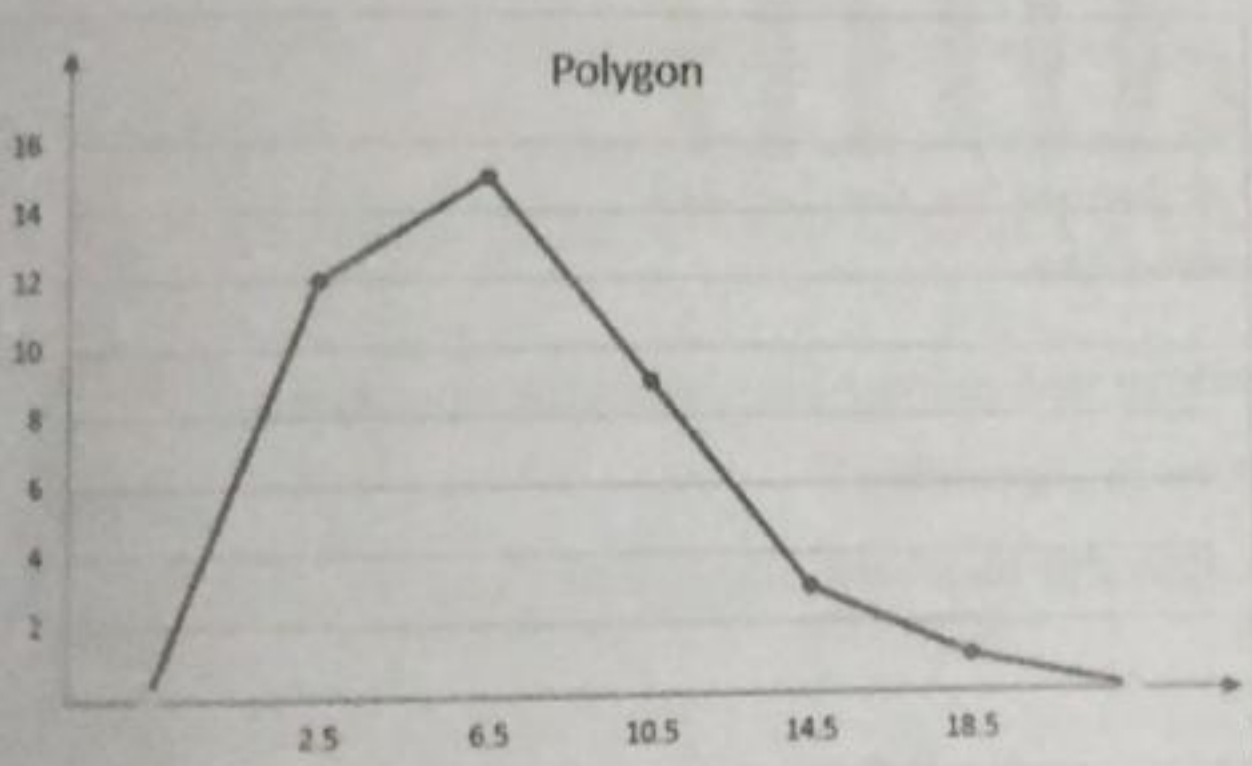
عند تفريغ البيانات في الجدول نفرغها حسب ال Class boundaries

Class limit	Class boundaries	Class midpoint	frequency	Relative frequency	Ascending cumulative frequency (A.C.F)
1 – 4	0.5 – 4.5	2.5	12	0.3	12
5 – 8	4.5 – 8.5	6.5	15	0.375	12+15=27
9 – 12	8.5 – 12.5	10.5	9	0.225	9+15+12=36
13 – 16	12.5 – 16.5	14.5	3	0.075	3+9+15+12=39
17 – 20	16.5 – 20.5	18.5	1	0.025	1+3+9+15+12=40
TOTAL	-------	----------	40	1	----------

b.

من ان مراكز الفئات صحيحة يجب ان يكون الفرق بين كل مركز والذي يليه يساوي طول الفئة
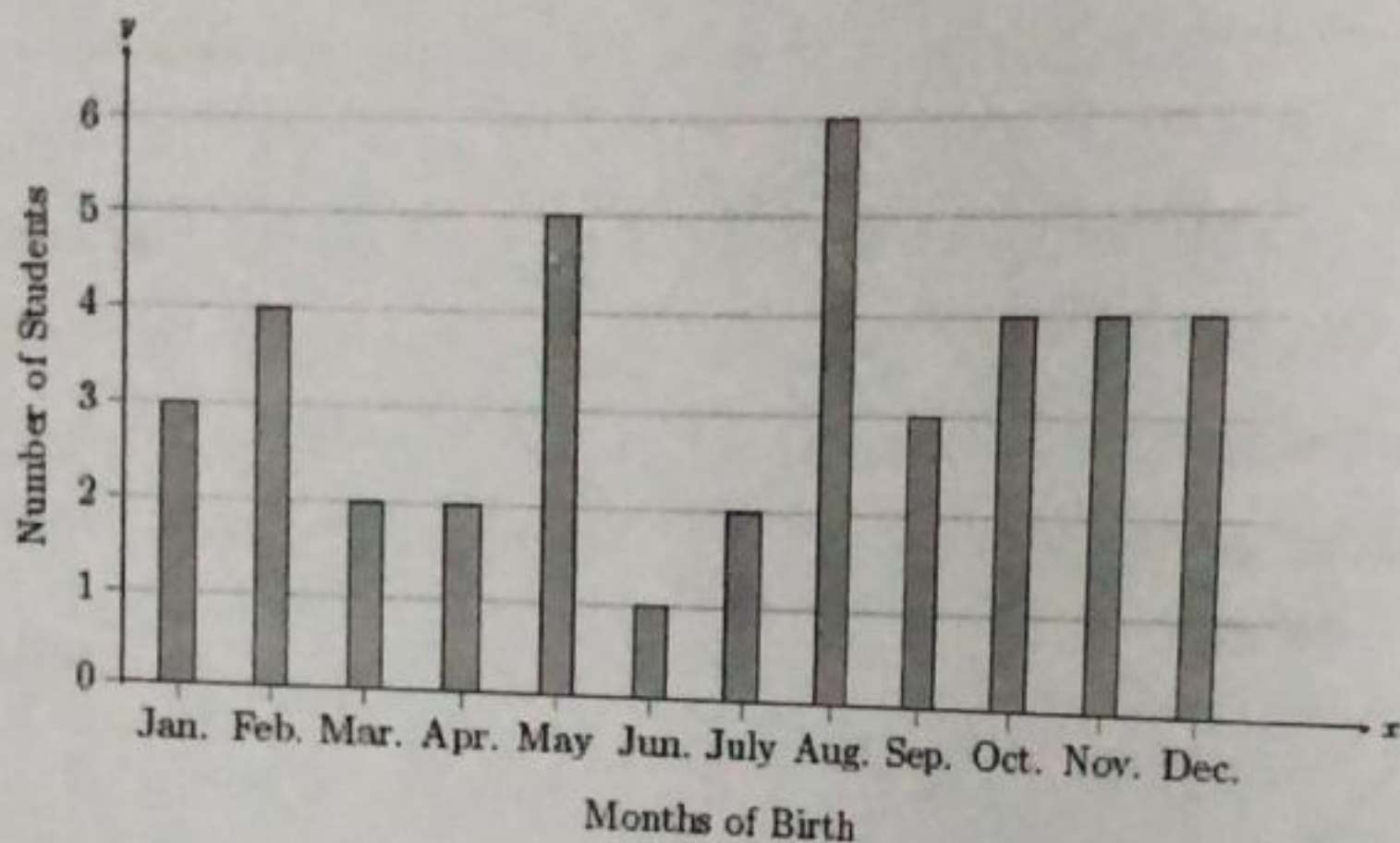
Histogram

c.

Polygon

d.

Ogive

8. In a particular section of Class X, 40 students were asked about the months of their birth and the following graph was prepared for the data so obtained:

Months of Birth

Observe the bar graph given above and answer the following questions:
a. How many students were born in November?
b. In which month were the maximum numbers of students born?

Solution:

   a. 4 students
   b. August

9. A sample of 100 children was asked how many times they play computer games for a period of one week. The following table gives the frequency distribution of their responses

Number of times of playing	Number of children
0 – 3	23
4 – 7	40
8 – 11	28
12 – 15	6
16 – 19	3

a. Find the class midpoints?

b. Do all classes have the same width? If so, what is this width?

c. Prepare the relative frequency and percentage distribution columns?

d. What percentage of these children plays 8 or more times a week?

Solution:

a. and c.

Classes	frequency	Midpoint	Relative frequency	Percentage frequency
0-3	23	1.5	$\frac{23}{100} = 0.23$	23%
4-7	40	5.5	$\frac{40}{100} = 0.40$	40%
8-11	28	9.5	$\frac{28}{100} = 0.28$	28%
11-15	6	13.5	$\frac{6}{100} = 0.06$	6%
16-19	3	17.5	$\frac{3}{100} = 0.03$	3%
Total	100		1	100%

b) Yes, and the length is 4

The length of the first class is 3-0+1=4

The length of the second class is 7-4+1=4

The length of the third class is 11-8+1=4

The length of the fourth class is 15-12+1=4

The length of the fifth class is 19-16+1=4

> ملاحظة ١) طول الفئة الفعلية (Class boundaries) = الحد الاعلى للفئة – الحد الادنى للفئة
>
> ١) طول الفئة العملية (Class limit) = الحد الاعلى للفئة – الحد الادنى للفئة + ١

d) The percentage of children plays 8 or more is $28\% + 6\% + 3\% = 37\%$

10. Consider the following frequency distribution, representing the weights of 36 students of
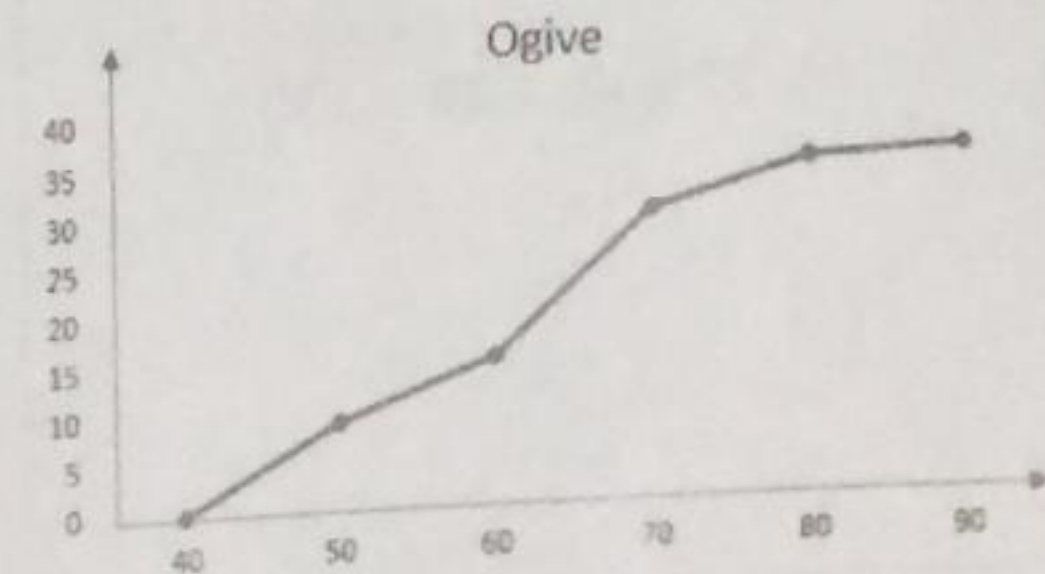
a class:

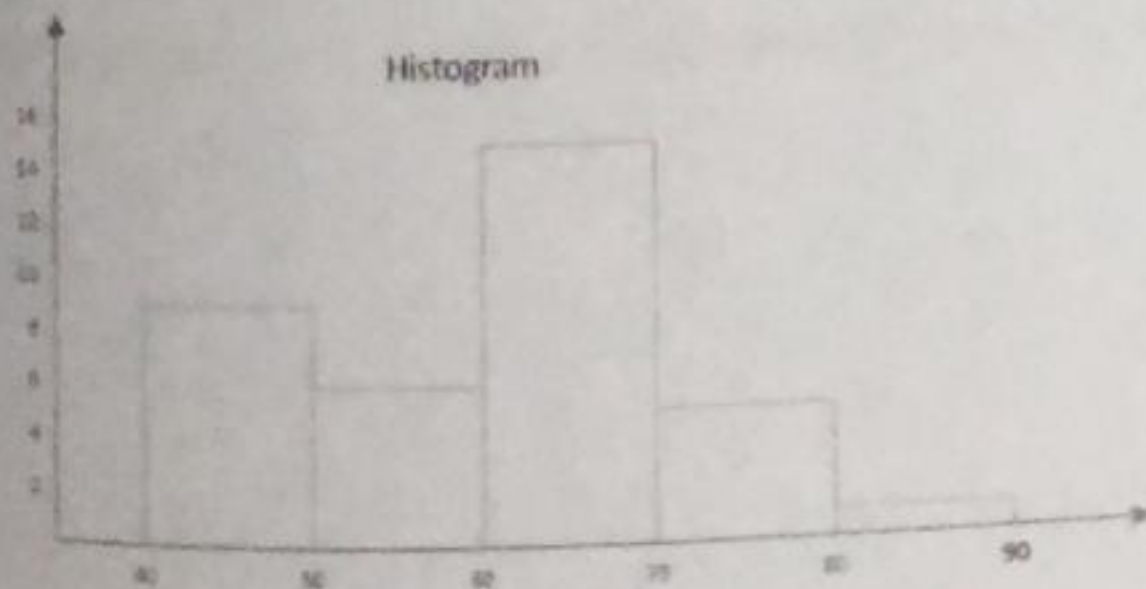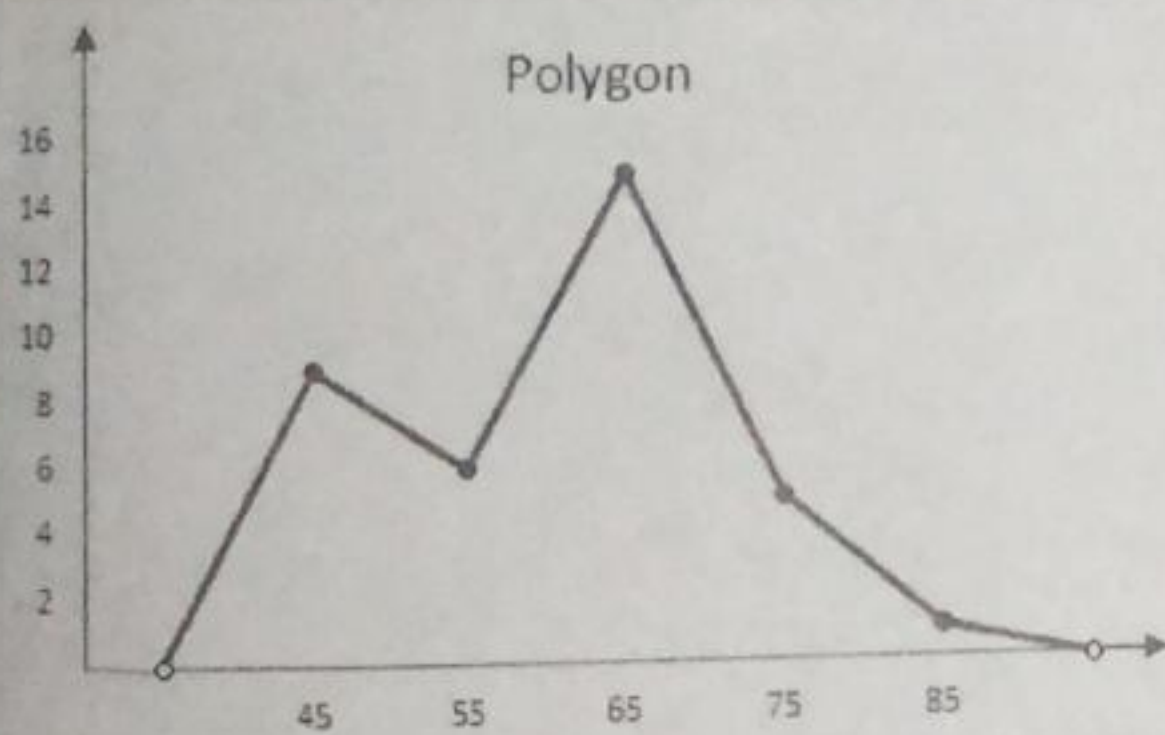| Weights (in kg)	Number of students
40 → 50	9
50 → 60	6
60 → 70	15
70 → 80	5
80 → 90	1
Total	36

a. Draw the histogram, polygon and ogive for the above table.

b. How many students have weights less than 70 Kg?

Solution:

Class Boundaries	Frequency	Mid Point	Less than	Ascending Cumulative Frequency (ACF)
40 → 50	9	45	50	9
50 → 60	6	55	60	15
60 → 70	15	65	70	30
70 → 80	5	75	80	35
80 → 90	1	85	90	36
Total	36			

Polygon

Histogram

Ogive
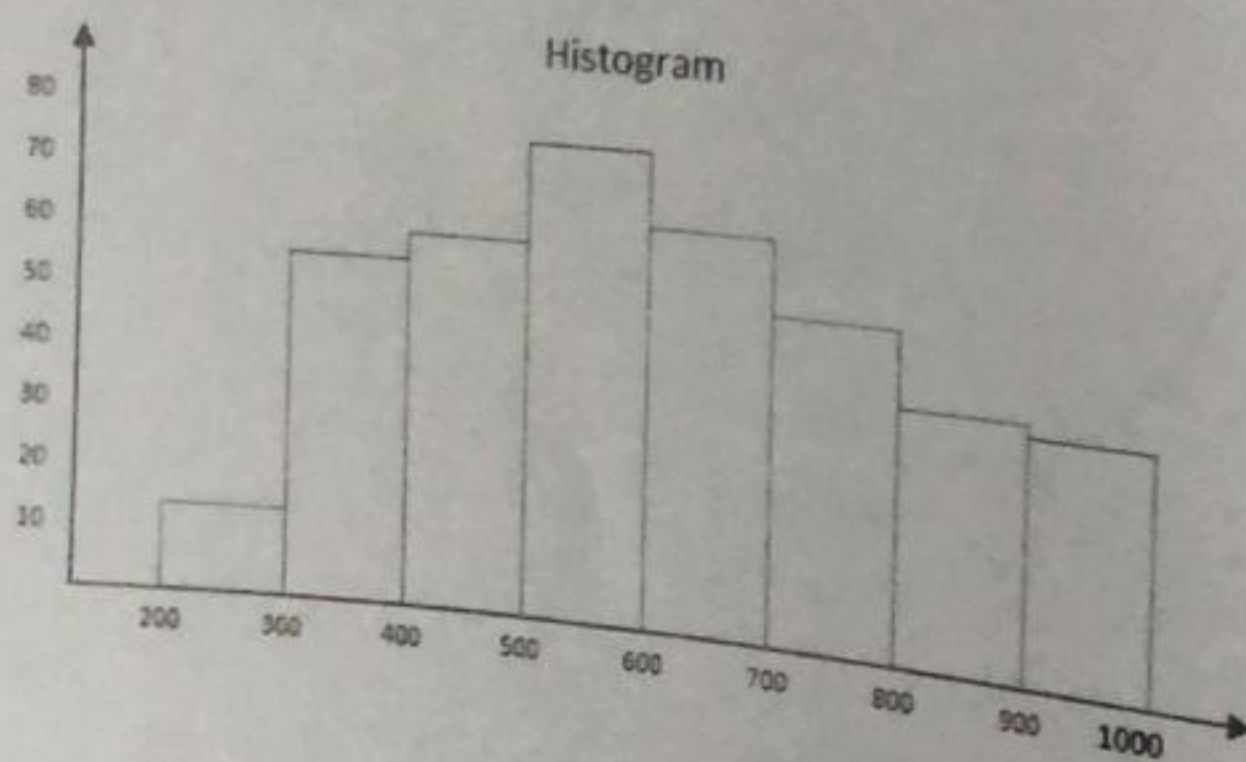
b) 30 students

11. The following table gives the life times of 400 neon lamps:

N.o.C.	Lifetime (in hours)	Number of lamps
1	$200 \rightarrow 300$	14
2	$300 \rightarrow 400$	56
3	$400 \rightarrow 500$	60
4	$500 \rightarrow 600$	76
5	$600 \rightarrow 700$	64
6	$700 \rightarrow 800$	52
7	$800 \rightarrow 900$	40
8	$900 \rightarrow 1000$	38

a. Represent the given information with the help of a histogram.

b. How many lamps have a life time of more than 700 hours?
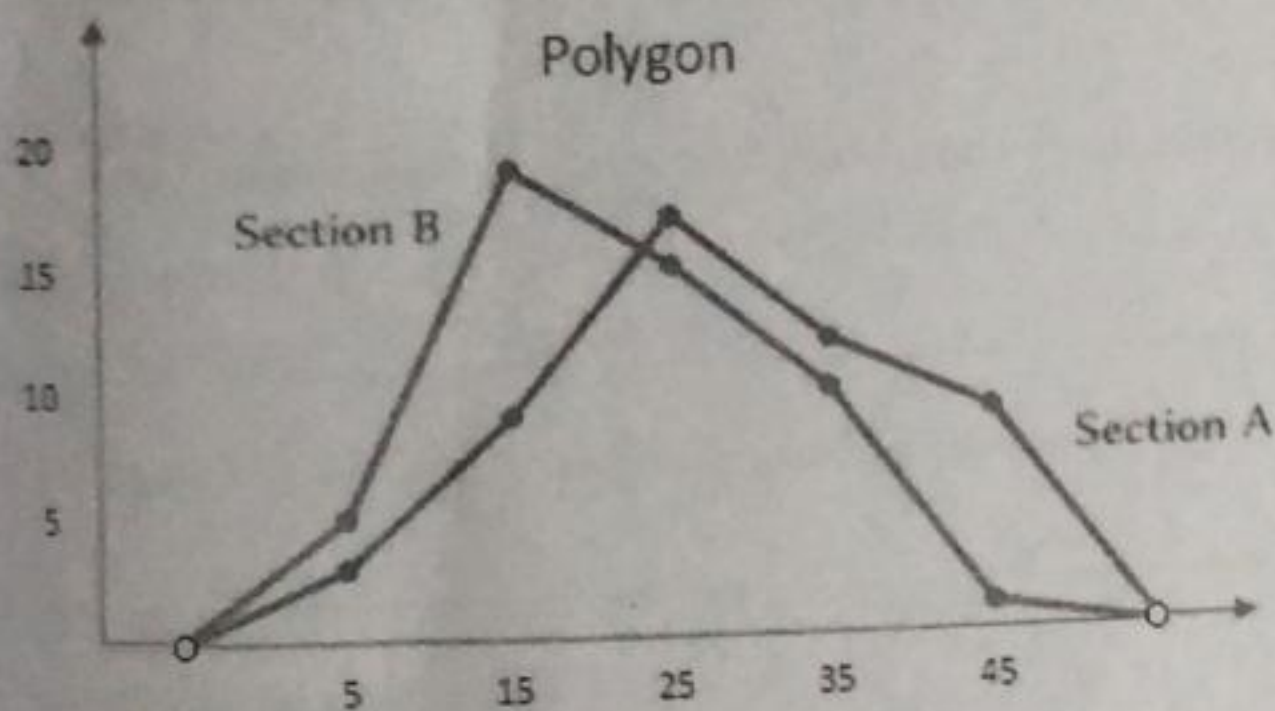
Solution:

Histogram

b) $52 + 40 + 38 = 130$

12. The following table gives the distribution of students of two sections according to the marks obtained by them:

Marks	Section A Frequency	Section B Frequency
0 → 10	3	5
10 → 20	9	19
20 → 30	17	15
30 → 40	12	10
40 → 50	9	1

Represent the marks of the students of both the sections on the same graph by two frequency polygons. From the two polygons compare the performance of the two sections.

Solution:

The performance of section A is **better** than the performance of section B.

13. Consider the following frequency distribution, representing the degree of examination of 50 students of a class:

Class Limit	Class Boundaries	Class Midpoint	Frequency	Relative Frequency	Ascending Cumulative Frequency (ACF)
2 - 6			6		
7 - 11				0.24	
12 - 16					36
17 - 21				0.12	
22 - 26			8		
Total			50		

Then:

a. Complete the above frequency distribution table.

b. Draw the histogram, polygon and ogive for this frequency distribution table.

c. Calculate the mean, median and mod for the above frequency distribution table.

d. Calculate the standard deviation of the above frequency distribution table.

Solution:

Relative frequency = $\dfrac{frequency\ of\ class}{Total\ of\ frequencies}$ $\Rightarrow\Rightarrow$ frequency of class = Relative frequency $\times$ Total of fre

Class limit	Class boundaries	Class midpoint	Frequency	Relative frequency	Ascending cumulative frequency (A.C.F)
2 - 6	1.5 - 6.5	4	6	0.12	6
7 - 11	6.5 - 11.5	9	(0.24)(50)=12	0.24	18
12 - 16	11.5 - 16.5	14	36 - 18 = 18	0.36	36
17 - 21	16.5 - 21.5	19	(0.12)(50)=6	0.12	42
22 - 26	21.5 - 26.5	24	8	0.16	50
Total	------	------	50	1	------

c)

Th

M

Th

m

d.

c) $mean = \bar{x} = \frac{\sum x_i \cdot f_i}{\sum f_i} = \frac{(4)(6)+(9)(12)+(14)(18)+(19)(6)+(24)(8)}{50} = 13.8$

<div dir="rtl">

الفئة الوسيطية هي الفئة التي يكون تكرارها التراكمي اكبر من او يساوي

</div>

$$\frac{n}{2} = \frac{50}{2} = 25$$

The median class is 11.5-16.5

$Median = \check{x} = L + \dfrac{\frac{\sum f_i}{2}-(F-f)}{f} * c$

$\qquad = 11.5 + \dfrac{25-(36-18)}{18} * 5 \approx 13.44$

<div dir="rtl">

الفئة المنوالية هي الفئة الاكثر تكرارا

</div>

The mode class is 11.5-16.5

$mode = \hat{x} = L + \dfrac{d_1}{d_1 + d_2} * c$

$\qquad = 11.5 + \dfrac{6}{6+12} * 5 \approx 13.17$

d. for standard deviation (S)

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$f_i * (x_i - \bar{x})^2$
4	−9.8	96.04	576.24
9	−4.8	23.04	276.48
14	0.2	0.04	0.72
19	5.2	27.04	162.24
24	10.2	104.04	832.32
Total	0		1847.55

$$s = \sqrt{\frac{\sum f_i * (x_i - \bar{x})^2}{\sum f - 1}} = \sqrt{\frac{1847.55}{50-1}} = \sqrt{37.71} \approx 6.14$$

14. The points scored by a team in a series of matches are as follows:

$$17, 2, 7, 27, 15, 5, 14, 8, 10, 24, 48, 10, 8, 7, 18, 28$$

Then:

a. Calculate the mean and standard deviation of the given data.

b. Calculate the standard score of the value (7) in the given data.

c. Calculate the coefficient of variation for the given data.

d. Calculate $Q_1, Q_2 and Q_3$

solution:

a.

Mean:

$$\bar{x} = \frac{\sum x_i}{n}$$

$$= \frac{17 + 2 + 7 + 27 + 15 + 5 + 14 + 8 + 10 + 24 + 48 + 10 + 8 + 7 + 18 + 28}{16} = \frac{248}{16} = 15.5$$

Variance:

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

$$= \frac{(17-15.5)^2 + (2-15.5)^2 + (7-15.5)^2 + (27-15.5)^2 + (15-15.5)^2 + ... + (28-15.5)^2}{15}$$

$$= \frac{2038}{15}$$

Standard deviation $S = \sqrt{\frac{2038}{15}} = 11.66$

ويمكن عمل جدول لتسهيل الحسابات اذا كان حجم البيانات كبير كالتالي

x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
17	$17 - 15.5 = 1.5$	2.25
2	$2 - 15.5 = -13.5$	182.25
7	$7 - 15.5 = -8.5$	72.25
27	$27 - 15.5 = 11.5$	132.25
15	$15 - 15.5 = -0.5$	0.25
5	$5 - 15.5 = -10.5$	110.25
14	$14 - 15.5 = -1.5$	2.25
8	$8 - 15.5 = -7.5$	56.25
10	$10 - 15.5 = -5.5$	30.25
24	$24 - 15.5 = 8.5$	72.25
48	$48 - 15.5 = 32.5$	1056.25
10	$10 - 15.5 = -5.5$	30.25
8	$8 - 15.5 = -7.5$	56.25
7	$7 - 15.5 = -8.5$	72.25
18	$18 - 15.5 = 2.5$	6.25
28	$28 - 15.5 = 12.5$	156.25
$\sum_{i=1}^{16} x_i = 248$	$\sum_{i=1}^{16}(x_i - \bar{x}) = 0$	$\sum_{i=1}^{16}(x_i - \bar{x})^2 = 2038$

Mean :

$$\bar{x} = \frac{\sum x_i}{n} = \frac{248}{16} = 15.5$$

Standard deviation

$$S = \sqrt{\frac{\sum_{i=1}^{16}(x_i - \bar{x})^2}{15}} = \sqrt{\frac{2038}{15}} \approx 11.66$$

b.

$$z = \frac{x - \bar{x}}{S} = \frac{7 - 15.5}{11.66} \approx -0.73$$

c.

$$CV = \frac{S}{\bar{x}} \times 100\% = \frac{11.66}{15.5} \times 100\% \approx 75.23\%$$

d.

$$2, 5, 7, 7, 8, 8, 10, 10, 14, 15, 17, 18, 24, 27, 28, 48$$

$q_1 = \dfrac{16+1}{4} = \dfrac{17}{4} = 4.25$

$q_2 = \dfrac{2(16+1)}{4} = \dfrac{34}{4} = 8.5$

$Q_1 = x_4 + 0.25(x_5 - x_4)$

$\quad = 7 + 0.25(8 - 7)$

$\quad = 7 + 0.25 = 7.25$

$Q_2 = x_8 + 0.5(x_9 - x_8)$

$\quad = 10 + 0.5(14 - 10)$

$\quad = 10 + 2 = 12$

$q_3 = \dfrac{3(16+1)}{4} = \dfrac{51}{4} = 12.75$

$Q_3 = x_{12} + 0.75(x_{13} - x_{12})$

$\quad = 18 + 0.75(24 - 18)$

$\quad = 18 + 4.5 = 22.5$

15. Consider the marks obtained (out of 100 marks) by 30 students of Class X of a school:

10	20	36	92	95	40	50	56	60	70
92	88	80	70	72	70	36	40	36	40
92	40	50	50	56	60	70	60	60	88
92	88	80	70	72	70	36	40	36	40
92	40	50	50	56	60	70	60	60	88

Then:

a. Calculate the mean and standard deviation of the given data.
b. Calculate P_{10}, P_{50} and P_{93}.
c. Calculate D_3, D_5 and D_8.
d. Calculate Q_1, Q_2 and Q_3.

Solution:

10	20	36	36	36
36	36	40	40	40
40	40	40	40	50
50	50	50	50	56
56	56	60	60	60
60	60	60	60	70
70	70	70	70	70
70	72	72	80	80
88	88	88	88	92
92	92	92	92	95

a) mean

$$\bar{x} = \frac{\Sigma x}{n} =$$
$$\frac{10+20+(5)(36)+(7)(40)+(5)(50)+(3)(56)+(7)(60)+(7)(70)+(2)(72)+(2)(80)+(4)(88)+(5)(92)+95}{50}$$

$$\frac{3029}{50} = 60.58$$

b) $P_{10} = \frac{10(50+1)}{100} = 5.1$

$$P_{10} = X_5 + (0.1)(X_6 - X_5) = 36 + 0.1(36 - 36) = 36$$

$$P_{50} = \frac{50(50+1)}{100} = 25.5$$

$$P_{50} = X_{25} + (0.5)(X_{26} - X_{25}) = 60 + 0.1(60 - 60) = 60$$

OR

$$P_{93} = \frac{93(50+1)}{100} = 47.43$$

$$P_{93} = X_{47} + (0.43)(X_{48} - X_{47}) = 92 + 0.43(92 - 92) = 92$$

c)

$$d_3 = \frac{3(50+1)}{10} = 15.3$$

$$D_3 = X_{15} + (0.3)(X_{16} - X_{15}) = 50 + 0.3(50 - 50) = 50$$

$$d_5 = \frac{5(50+1)}{10} = 25.5$$

$$D_5 = X_{25} + (0.5)(X_{26} - X_{25}) = 60 + 0.5(60 - 60) = 60$$

OR

$$D_5 = P_{50} = 60$$

$$d_8 = \frac{8(50+1)}{10} = 40.8$$

$$D_8 = X_{40} + (0.8)(X_{41} - X_{40}) = 80 + 0.8(88 - 80) = 86.4$$

$$q_1 = \frac{1(50+1)}{4} = 12.75$$

$$Q_1 = X_{12} + (0.75)(X_{13} - X_{12}) = 40 + 0.8(40 - 40) = 40$$

$$q_2 = \frac{2(50+1)}{4} = 25.5$$

$$Q_2 = X_{25} + (0.5)(X_{26} - X_{25}) = 60 + 0.5(60 - 60) = 60$$

OR

$$D_5 = P_{50} = Q_2 = 60$$

$$q_3 = \frac{3(50+1)}{4} = 38.25$$

$$Q_3 = X_{38} + (0.25)(X_{39} - X_{38}) = 72 + 0.25(80 - 72) = 74$$

16. Consider a small unit of a factory where there are 5 employees: a supervisor and four laborers. The laborers draw a salary of \$ 5,000 per month each while the supervisor gets \$ 15,000 per month. Calculate the mean, median and mode of the salaries of this unit of the factory.

<u>Solution:</u>

The salaries are 5000, 5000, 5000, 5000, 15000

Mean

$$\bar{x} = \frac{5000 + 5000 + 5000 + 5000 + 15000}{5} = \frac{35000}{5} = 7000$$

Median $\tilde{x} = x_{\frac{n+1}{2}} = x_3 = 5000$

Mode $\hat{x} = 5000$

17. The daily sale of sugar (kg) in a certain grocery shop is given below:

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
75	120	12	50	70.5	140.5

a. Calculate the average daily sale.

b. Calculate the variance and the standard deviation of the above data.

c. Determine the coefficient of variation.

Solution:

a. $average = \bar{x} = \frac{\sum x_i}{n} = \frac{75+120+12+50+70.5+140.5}{6} = \frac{468}{6} = 78$

b. $variance = S^2 = \frac{\sum(x_i-\bar{x})^2}{n-1} =$

$\frac{(75-78)^2+(120-78)^2+(12-78)^2+(50-78)^2+(70.5-78)^2+(140.5-78)^2}{6-1} = \frac{10875.5}{5} = 2175.1$

$standard\ deviation = S = \sqrt{S^2} = \sqrt{2175.1} = 46.64$

c. $coefficient\ of\ variation = CV = \frac{S}{\bar{x}} \times 100\% = \frac{46.64}{78} \times 100\% = 59.8\%$

18. Let the following data be marks obtained (out of 100) by 10 students in a test:

$$45, 45, 63, 76, 67, 84, 75, 48, 62, 65$$

Then:

a. Calculate $Q_1, Q_2 and Q_3$.

b. Calculate the IQR.

c. Have the given data extreme values?

d. Construct the box plot for the given data.

b) $p_{10} = \frac{10(50+1)}{100} = 5.1$

$$P_{10} = X_5 + (0.1)(X_6 - X_5) = 36 + 0.1(36 - 36) = 36$$

$$p_{50} = \frac{50(50 + 1)}{100} = 25.5$$

$$P_{50} = X_{25} + (0.5)(X_{26} - X_{25}) = 60 + 0.1(60 - 60) = 60$$

$$p_{93} = \frac{93(50 + 1)}{100} = 47.43$$

$$P_{93} = X_{47} + (0.43)(X_{48} - X_{47}) = 92 + 0.43(92 - 92) = 92$$

C)

$$d_3 = \frac{3(50 + 1)}{10} = 15.3$$

$$D_3 = X_{15} + (0.3)(X_{16} - X_{15}) = 50 + 0.3(50 - 50) = 50$$

$$d_5 = \frac{5(50 + 1)}{10} = 25.5$$

$$D_5 = X_{25} + (0.5)(X_{26} - X_{25}) = 60 + 0.5(60 - 60) = 60$$

OR

$$D_5 = P_{50} = 60$$

$$d_8 = \frac{8(50 + 1)}{10} = 40.8$$

$$D_8 = X_{40} + (0.8)(X_{41} - X_{40}) = 80 + 0.8(88 - 80) = 86.4$$

$$q_1 = \frac{1(50 + 1)}{4} = 12.75$$

$$Q_1 = X_{12} + (0.75)(X_{13} - X_{12}) = 40 + 0.8(40 - 40) = 40$$

$$q_2 = \frac{2(50 + 1)}{4} = 25.5$$

$$Q_2 = X_{25} + (0.5)(X_{26} - X_{25}) = 60 + 0.5(60 - 60) = 60$$

OR

$$D_5 = P_{50} = Q_2 = 60$$

$$q_3 = \frac{3(50 + 1)}{4} = 38.25$$

$$Q_3 = X_{38} + (0.25)(X_{39} - X_{38}) = 72 + 0.25(80 - 72) = 74$$

16. Consider a small unit of a factory where there are 5 employees: a supervisor and four laborers. The laborers draw a salary of $ 5,000 per month each while the supervisor gets $ 15,000 per month. Calculate the mean, median and mode of the salaries of this unit of the factory.

Solution:

The salaries are 5000, 5000, 5000, 5000, 15000

Mean

$$\bar{x} = \frac{5000 + 5000 + 5000 + 5000 + 15000}{5} = \frac{35000}{5} = 7000$$

Median $\tilde{x} = x_{\frac{n+1}{2}} = x_3 = 5000$

Mode $\hat{x} = 5000$

17. The daily sale of sugar (kg) in a certain grocery shop is given below:

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
75	120	12	50	70.5	140.5

a. Calculate the average daily sale.

b. Calculate the variance and the standard deviation of the above data.

c. Determine the coefficient of variation.

Solution:

a. $average = \bar{x} = \frac{\Sigma x_i}{n} = \frac{75+120+12+50+70.5+140.5}{6} = \frac{468}{6} = 78$

b. $variance = S^2 = \frac{\Sigma(x_i-\bar{x})^2}{n-1} =$

$\frac{(75-78)^2+(120-78)^2+(12-78)^2+(50-78)^2+(70.5-78)^2+(140.5-78)^2}{6-1} = \frac{10875.5}{5} = 2175.1$

$standard\ deviation = S = \sqrt{S^2} = \sqrt{2175.1} = 46.64$

c. $coefficient\ of\ variation = CV = \frac{S}{\bar{x}} \times 100\% = \frac{46.64}{78} \times 100\% = 59.8\%$

18. Let the following data be marks obtained (out of 100) by 10 students in a test:

$$45, 45, 63, 76, 67, 84, 75, 48, 62, 65$$

Then:

a. Calculate $Q_1, Q_2 and Q_3$.

b. Calculate the IQR.

c. Have the given data extreme values?

d. Construct the box plot for the given data.

Solution:

a.

$$45, 45, 48, 62, 63, 65, 67, 75, 76, 84$$

$q_1 = \dfrac{10+1}{4} = \dfrac{11}{4} = 2.75$	$q_2 = \dfrac{2(10+1)}{4} = \dfrac{22}{4} = 5.5$
$Q_1 = x_2 + 0.75(x_3 - x_2)$	$Q_2 = x_5 + 0.5(x_6 - x_5)$
$= 45 + 0.75(48 - 45)$	$= 63 + 0.5(65 - 63)$
$= 45 + 2.25 = 47.25$	$= 63 + 1 = 64$

$$q_3 = \frac{3(10+1)}{4} = \frac{33}{4} = 8.25$$

$$Q_3 = x_8 + 0.25(x_9 - x_8)$$

$$= 75 + 0.25(76 - 75)$$

$$= 75 + 0.25 = 75.25$$

b.

$$IQR = Q_3 - Q_1$$
$$= 75.25 - 47.25 = 28$$

c.

$$LF = Q_1 - 1.5(Q_3 - Q_1) = 47.25 - 1.5(28) = 47.25 - 42 = 5.25$$

$$HF = Q_3 + 1.5(Q_3 - Q_1) = 75.25 + 1.5(28) = 75.25 + 42 = 117.25$$

The data have no extrema values

d.

Xs=45   Q1=47.25   Q2=64   Q3=75.25   Xl=84

45   50   55   60   65   70   75   80   85

19. Consider the following data:

$$40, 40, 40, 60, 65, 65, 70, 70, 75, 75, 75, 80, 85, 90\ 90, 150$$

Then:

a. Calculate Q_1, Q_2 and Q_3.

b. Calculate the IQR

c. Have the given data extreme values?

d. Construct the box plot for the given data.

Solution:

a.

نلاحظ ان البيانات مرتبة تصاعديا وبالتالي نبدأ الحل مباشرة

$$q_1 = \frac{1(16 + 1)}{4} = 4.25$$

$$Q_1 = X_4 + (0.25)(X_5 - X_4) = 60 + 0.25(65 - 60) = 61.25$$

$$q_2 = \frac{2(16 + 1)}{4} = 8.5$$

$$Q_2 = X_8 + (0.5)(X_9 - X_8) = 70 + 0.5(75 - 70) = 72.5$$

OR

$$Q_2 = \check{X} = \frac{70 + 75}{2} = 72.5$$

$$q_3 = \frac{3(16 + 1)}{4} = 12.75$$

$$Q_3 = X_{12} + (0.75)(X_{13} - X_{12}) = 80 + 0.75(85 - 80) = 83.75$$

b.

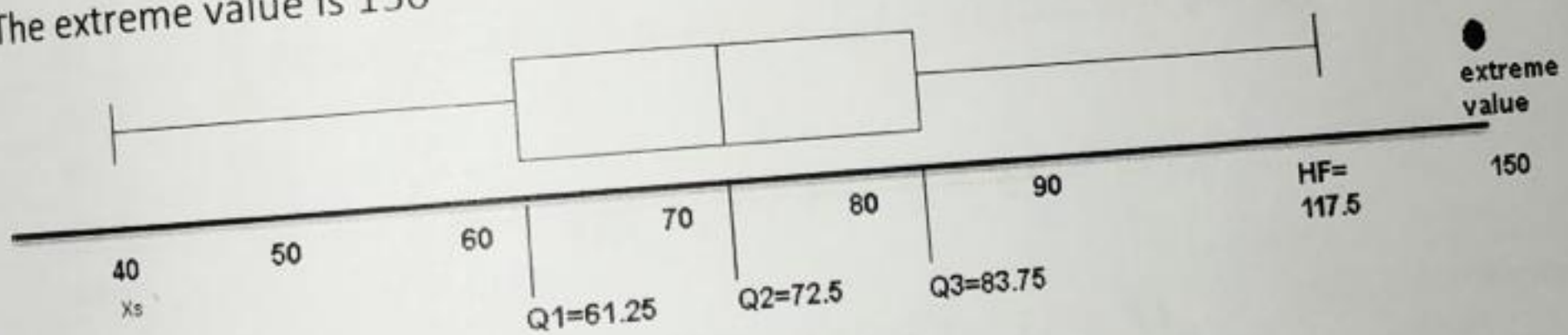$$IQR = Q_3 - Q_1 = 83.75 - 61.25 = 22.5$$

c.

$$LF = Q_1 - 1.5(Q_3 - Q_1)$$

$$LF = 61.25 - 1.5(83.75 - 61.25) = 27.5$$

$$HF = Q_3 + 1.5(Q_3 - Q_1)$$

$$HF = 83.75 + 1.5(83.75 - 61.25) = 117.5$$

The extreme value is 150

20. Consider the following data:

$40, 45, 55, 65, ?, ?, 75, 75, 78, 83$

Then use the suitable measure to calculate the average and dispersion for the given data.

Solution:

For the average we use the mode and it is 75

For the dispersion we use the range and it is $R = 83 - 40 = 43$

21. Consider the following data:

$$-15, 20, 40, 50, 65, 65, 70, 73, 75, 137$$

a. Have the given data extreme values?

b. Use the suitable measure to calculate the average and dispersion for the given data.

<u>Solution:</u>

$$q_1 = \frac{1(10+1)}{4} = 2.75$$

$$Q_1 = X_2 + (0.75)(X_3 - X_2) = 20 + 0.75(40 - 20) = 35$$

$$q_3 = \frac{3(10+1)}{4} = 8.25$$

$$Q_3 = X_8 + (0.25)(X_9 - X_8) = 73 + 0.25(75 - 73) = 73.5$$

$$LF = Q_1 - 1.5(Q_3 - Q_1)$$

$$LF = 35 - 1.5(73.5 - 35) = -22.75$$

$$HF = Q_3 + 1.5(Q_3 - Q_1)$$

$$HF = 73.5 + 1.5(73.5 - 35) = 131.25$$

Extreme value is 137

For dispersion

وجود قيم متطرفة نستخدم IQR كافضل مقياس للتشتت

$$IQR = Q_3 - Q_1 = (73.5 - 35) = 38.5$$

For average

Mean $\bar{x} = \frac{\sum x}{n} = \frac{-15+20+40+50+65+65+70+73+75+137}{10}$

$$=58$$

45
46
51
43
44

Prepare a

<u>Solution:</u>

41
42
43
44
44

$$q_1 = \frac{25+}{4}$$

$$Q_1 = x_6 +$$

$$= 45 +$$

$$= 45 +$$

22. The following data give the number of computer keyboards assembled at a company

for a sample of 25 days

45	52	48	41	56
46	44	42	48	53
51	53	51	48	46
43	52	50	54	47
44	47	50	49	52

Prepare a box-plot and then comment the skewness of these data.

Solution:

41	45	48	50	52
42	46	48	51	53
43	46	48	51	53
44	47	49	52	54
44	47	50	52	56

$$q_1 = \frac{25+1}{4} = \frac{26}{4} = 6.5$$

$$Q_1 = x_6 + 0.5\left(x_7 - x_6\right)$$

$$= 45 + 0.5\left(46 - 45\right)$$

$$= 45 + 0.5 = 45.5$$

$$q_2 = \frac{2(25+1)}{4} = 13$$

$$Q_2 = x_{13} + 0\left(x_{14} - x_{13}\right)$$

$$= 48 + 0 = 48$$
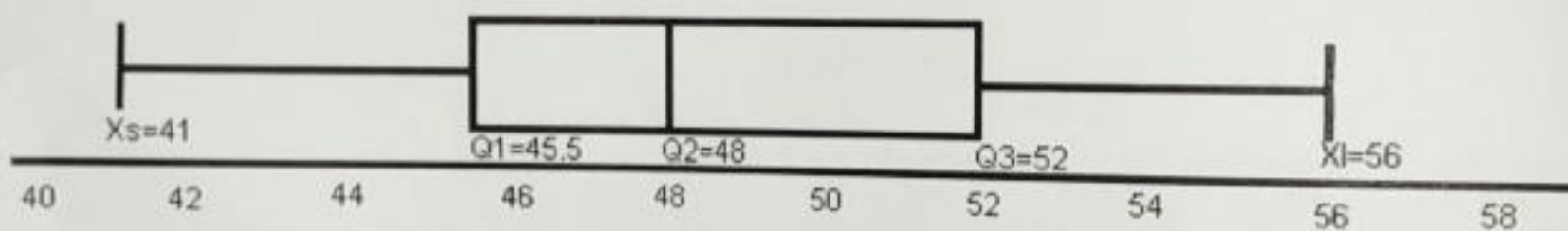
$$q_3 = \frac{3(25+1)}{4} = \frac{78}{4} = 19.5$$

$$Q_3 = x_{19} + 0.5\left(x_{20} - x_{19}\right)$$

$$= 52 + 0.5\left(52 - 52\right)$$

$$= 52 + 0.5(0) = 52$$

$$LF = Q_1 - 1.5\left(Q_3 - Q_1\right) = 45.5 - 1.5\left(52 - 45.5\right) = 45.5 - 9.75 = 35.75$$

$$HF = Q_3 + 1.5\left(Q_3 - Q_1\right) = 52 + 1.5\left(52 - 45.5\right) = 52 + 9.75 = 61.75$$

Left skewed

23. Create a dot plot for the following data set

1	2	0	5	1
1	3	2	0	5
2	1	2	1	2
0	1	3	1	2

Solution:

24. The mean age of six persons is 49 years. The ages of five of these six persons are 55, 39, 44, 51, and 45 years respectively. Find the age of the sixth person.

Solution:

Let a be the age of the sixth person

$$\bar{x} = \frac{\sum_{i=1}^{6} x_i}{6}$$

$$49 = \frac{55 + 39 + 44 + 51 + 45 + a}{6}$$

$$49 = \frac{234 + a}{6} \rightarrow 234 + a = 49(6) \rightarrow a = 294 - 234 = 60$$

OR

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \rightarrow \sum_{i=1}^{n} x_i = n\,\bar{x}$$

$$55 + 39 + 44 + 51 + 45 + a = 6(49)$$

$$234 + a = 294 \rightarrow a = 294 - 234 = 60$$

25. The following data give the masses in grams to the nearest gram, of 10 eggs.

$$46, 51, 48, 62, 54, 56, 58, 60, 71, 75$$

a. Calculate the mean, median, and standard deviation of this data.

b. Calculate the five numbers summary and construct the box plot of this data.

Solution:

a) mean $= \frac{\sum x}{n} = \frac{46 + 48 + 51 + 54 + 56 + 58 + 60 + 62 + 71 + 75}{10} = \frac{581}{10}$

$$= 58.1$$

$46,48,51,54,56,58,60,62,71,75$

$$median = \frac{x_5 + x_6}{2} = \frac{56 + 58}{2} = 57$$

X_i	$X_i - \bar{x}$	$(X_i - \bar{x})^2$
46	-12.1	146.41
48	-10.1	102.01
51	-7.1	50.41
54	-4.1	16.81
56	-2.1	4.41
58	-0.1	0.01
60	1.9	3.61
62	3.9	15.21
71	12.9	166.41
75	16.9	285.61
Sum	0	790.9

$$s = \sqrt{\frac{\sum(X_i - \bar{x})^2}{n-1}} = \sqrt{\frac{790.9}{9}} = 9.37$$

b)

$$q_1 = \frac{10+1}{4} = \frac{11}{4} = 2.75$$

$$Q_1 = x_2 + 0.75\left(x_3 - x_2\right)$$

$$= 48 + 0.75\left(51 - 48\right)$$

$$= 48 + 2.25 = 50.25$$

$$q_2 = \frac{2\left(10+1\right)}{4} = \frac{22}{4} = 5.5$$

$$Q_2 = x_5 + 0.5\left(x_6 - x_5\right)$$

$$= 56 + 0.5\left(58 - 56\right)$$

$$= 56 + 1 = 57$$

$$q_3 = \frac{3\left(10+1\right)}{4} = \frac{33}{4} = 8.25$$

$$Q_3 = x_8 + 0.25\left(x_9 - x_8\right) = 62 + 0.25\left(71 - 62\right)$$
$$= 62 + 2.25 = 64.25$$

$$LF = Q_1 - 1.5\left(Q_3 - Q_1\right) = 50.25 - 1.5\left(64.25 - 50.25\right) = 50.25 - 21 = 29.25$$

$HF = Q_3 + 1$

Five numbe

xs=46

45

26. The fo

29, 32

Now,

a. Calcula

b. Calcula
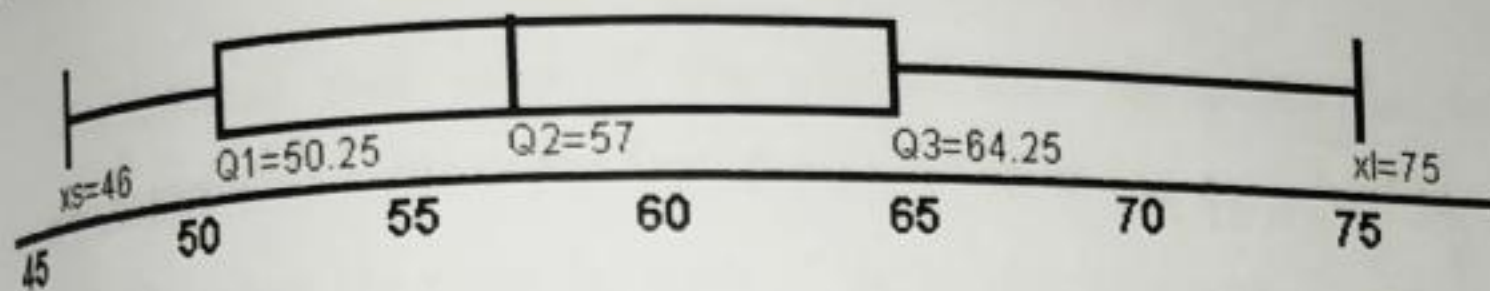
c. Find th

Solution

a.

$median =$

$2x + 2 = 1$

b.

$$HF = Q_3 + 1.5(Q_3 - Q_1) = 64.25 + 1.5(64.25 - 50.25) = 64.25 + 21 = 85.25$$

Five numbers are $\{x_s = 46, \ Q_1 = 50.25, \ Q_2 = 57, \ Q_3 = 64.25, \ x_l = 75\}$

26. The following observations have been arranged in ascending order.

29, 32, 48, 50, x, x + 2, 72, 78, 84, 95

Now, if the median of the data is 63, then:

a. Calculate the value of x.

b. Calculate the mean, and standard deviation of this data.

c. Find the five numbers summary and construct the box plot of this data.

Solution

a.

$$median = \frac{x_5 + x_6}{2} = \frac{x + x + 2}{2} = 63$$

$$2x + 2 = 126 \rightarrow 2x = 124 \rightarrow x = 62$$

b.

x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
29	29 − 61.4 = -32.4	1049.76
32	32 − 61.4 = -29.4	864.36
48	48 − 61.4 = -13.4	179.56
50	50 − 61.4 = -11.4	129.96
62	62 − 61.4 = 0.6	0.36
64	64 − 61.4 = 2.6	6.76
72	72 − 61.4 = 10.6	112.36
78	78 − 61.4 = 16.6	275.56
84	84 − 61.4 = 22.6	510.76
95	95 − 61.4 = 33.6	1128.96
$\sum_{i=1}^{10} x_i = 614$	$\sum_{i=1}^{16}(x_i - \bar{x}) = 0$	$\sum_{i=1}^{16}(x_i - \bar{x})^2 = 4258.4$

Mean: $\bar{x} = \dfrac{\sum\limits_{i=1}^{10} x_i}{10} = \dfrac{614}{10} = 61.4$

Standard deviation

$S = \sqrt{\dfrac{\sum\limits_{i=1}^{10}(x_i - \bar{x})^2}{9}} = \sqrt{\dfrac{4258.4}{9}} \approx 21.75$
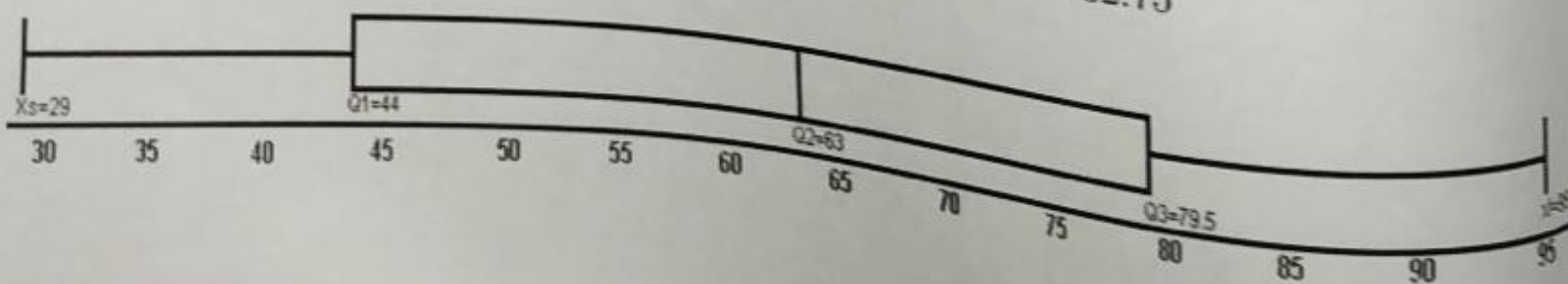
c.

$q_1 = \dfrac{10+1}{4} = \dfrac{11}{4} = 2.75$

$Q_1 = x_2 + 0.75(x_3 - x_2)$

$\quad = 32 + 0.75(48 - 32)$

$\quad = 32 + 12 = 44$

$q_2 = \dfrac{2(10+1)}{4} = 5.5$

$Q_2 = x_5 + 0.5(x_6 - x_5)$

$\quad = 62 + 0.5(64 - 62) = 62 + 1 = 63$

OR $Q_2 = $ median $= 63$

$q_3 = \dfrac{3(10+1)}{4} = \dfrac{33}{4} = 8.25$

$Q_3 = x_8 + 0.25(x_9 - x_8) = 78 + 0.25(84 - 78)$
$\quad = 78 + 0.25(6) = 78 + 1.5 = 79.5$

The five numbers

$x_s = 29, x_l = 95, Q_1 = 44, Q_2 = 63$ and $Q_3 = 79.5$

$LF = Q_1 - 1.5(Q_3 - Q_1) = 44 - 1.5(79.5 - 44) = 44 - 53.25 = -9.25$
$HF = Q_3 + 1.5(Q_3 - Q_1) = 79.5 + 1.5(79.5 - 44) = 79.5 + 53.25 = 132.75$

Xs=29   Q1=44   Q2=63   Q3=79.5

30   35   40   45   50   55   60   65   70   75   80   85   90

27. Consider t

Notice th

correspon

a. Calcul
   betwe

b. Calcul
   relatic

c. Calculat

d. Calculat
   them.

Solution:

a.

b.

standar

$= \sqrt{\dfrac{(12}{}}$

$= \sqrt{214}$

27. Consider the following two data sets.

$$\text{Data set I: } 12, 25, 37, 8, 41$$

$$\text{Data set II: } 19, 32, 44, 15, 48$$

Notice that each value of the second data set is obtained by adding 7 to the corresponding value of the first data set. Then:

a. Calculate the mean for each of these two data sets. Comment on the relationship between the two means.

b. Calculate the standard deviation for each of these two data sets. Comment on the relationship between the two standard deviations.

c. Calculate the standard score of the value (37) in data set I.

d. Calculate the coefficient of variation for each of these two data sets, then compare them.

Solution:

a.

$$mean\ of\ set\ I = \frac{\sum x_i}{n} = \frac{12 + 25 + 37 + 8 + 41}{5} = 24.6$$

$$mean\ of\ set\ II = \frac{\sum x_i}{n} = \frac{19 + 32 + 44 + 15 + 48}{5} = 31.6$$

$$mean\ of\ set\ II = 7 + mean\ of\ set\ I$$

b.

$$standard\ deviation\ of\ set\ I$$

$$= \sqrt{\frac{(12 - 24.6)^2 + (25 - 24.6)^2 + (37 - 24.6)^2 + (8 - 24.6)^2 + (41 - 24.6)^2}{5 - 1}}$$

$$= \sqrt{214.3} \approx 14.64$$

standard deviation of set II

$$= \sqrt{\frac{(19-31.6)^2 + (32-31.6)^2 + (44-31.6) + (15-31.6) + (48-31.6)^2}{5-1}}$$

$$= \sqrt{214.3} = 14.64$$

standard deviation of set I = standard deviation of set II

$$Z = \frac{x - \bar{x}}{s} = \frac{37 - 24.6}{14.64} = 0.84699$$

$$CV(of\ set\ I) = \frac{s}{\bar{x}} \times 100\% = \frac{14.64}{24.6} \times 100\% = 59.5\%$$

$$CV(of\ set\ II) = \frac{s}{\bar{x}} \times 100\% = \frac{14.64}{31.6} \times 100\% = 46.3\%$$

$$CV(of\ set\ II) < CV(of\ set\ I)$$

The variation of Data in set II is more than the variation of data in set I

28. Consider the following two data sets.

Data set I: 4, 8, 15, 9, 11

Data set II: 12, 24, 45, 27, 33

Notice that each value of the second data set is obtained by multiplying the corresponding value of the first data set by 3.

a. Calculate the mean for each of these two data sets. Comment on the relationship between the two means.

b. Calculate the standard deviation for each of these two data sets. Comment on the relationship between the two standard deviations.

c. Calculate the standard score of the value (27) in data set II.

d. Calculate the coefficient of variation for each of these two data sets, then Compare them.

Solution:

a.

The mean of the Data set I

$$\bar{x} = \frac{4+8+15+9+11}{5} = \frac{47}{5} = 9.2$$

The mean of the Data set II

$$\bar{x} = \frac{12+24+45+27+33}{5} = \frac{141}{5} = 28.2$$

The mean of set II is equal the mean of set I multiplying by 3.

b.

The standard deviation of the Data set I

$$S = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

$$= \sqrt{\frac{(4-9.4)^2 + (8-9.4)^2 + (15-9.4)^2 + (9-9.4)^2 + (11-9.4)^2}{4}}$$

$$= \sqrt{\frac{65.2}{4}} = \sqrt{16.3} \approx 4.037$$

The standard deviation of the Data set II

$$S = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

$$= \sqrt{\frac{(12-28.2)^2 + (24-28.2)^2 + (45-28.2)^2 + (27-28.2)^2 + (33-28.2)^2}{4}}$$

$$= \sqrt{\frac{586.8}{4}} = \sqrt{146.7} \approx 12.111$$