Chapter (2)

Descriptive statistics

0798208683 Arwa M. Bader Principles of statistics-JU

Sheet (1)

Q1) If three supermarkets are selling a chocolate bar for 1.25 JD, five are selling it for 1.50 JD, and ten are selling it for 1.15 JD, then the median price of the chocolate bar is:

A) 130 B) 120 C) 150 D) 1.15 E) 1.25

Solution:

X	1.15	1.25	1.5
F	10	3	5
C.F	10	13	18
I	0 – 10	11 – 13	14 - 18

$Q_2 = \frac{n}{2} = \frac{18}{2} = 9 \rightarrow \frac{9th+10th}{2} = \frac{1.15+1.15}{2} = 1.15 \rightarrow D$

Q2) On an exam given to 5 students, the mean grade is 78, the grades of 4 of them are 87, 81 , 76 and 53. Then the grade of the 5^{th} student is :

A) 65 B) 93 C) 71 D) 85 E) 99

Solution: Suppose that the grade of the 5^{th} student is Y

$\sum Xi = \bar{x} * n \rightarrow 87 + 81 + 76 + 53 + y = 5*78 \rightarrow 297 + y = 390$

$\therefore\ y = 93 \rightarrow B$

Q3) In a quiz, 3 students got 1, 5 students got 2 and 2 students got 5. The average score of these students in this quiz is:

A) 3.30 B) 3.00 C) 2.80 D) 3.11 E) 2.30

Solution: We must construct frequency table first:

X	1	2	5
Frequency	3	5	2

Now find F.X :

F.X	1*3 = 3	2*5 = 10	5*2 = 10

$\bar{X} = \frac{\Sigma f * x}{n} = \frac{3+10+10}{10} = 2.3 \rightarrow E$

Q4) The value of the mean times the number of observations equals =

A) The median B) the sum of the data C) the mode D) IQR

Solution: $\sum Xi = \bar{x} * n$ → the sum of the data → B

Q5) If the mean of 9 students is 15, a new student joined the class with mark 20 , Find the new sum , new number of students and new mean.

Solution: When n = 9 we had $\bar{X} = 15$ we can get sum of data values by:

$\sum Xi_{old} = \bar{X} * n_{old}$ → $\Sigma x = 15 * 9 = 135$

→135 is the old sum "sum for 9 students "

→ the new sum is $\sum Xi_{new} = 135 + 20 = 155$

→ the new number of students is: $n_{new} = 1 + 9 = 10$

→ the new mean is $\bar{X} = \frac{\sum Xi_{new}}{n_{new}} = \frac{155}{10} = 15.5$

Q6) If the median of observations 0 , 3 , x , 12 is 5, then the mean of these observations will be :

A) 5.75 B) 5 C) 4.75 D) 5.25 E) 5.5

Solution: $Q_2 = \frac{n}{2} = \frac{4}{2} = 2$ (whole number) → $\frac{2^{nd} + 3^{rd}}{2} = 5$ → $\frac{3 + x}{2} = 5$ → $x = 7$

Now $\bar{X} = \frac{\sum Xi}{n} = \frac{0+3+7+12}{4} = 5.5$ → E

Q7) Consider the following data:

I	3-5	6-8	9-11	12-14
F	5	2	2	1

Then the mode is:

A)13 B)10 C)4 D)7 E)10

Solution: Model class $3 - 5$ → Mode $= \frac{5+3}{2} = 4$ → C

Q8) Consider the following grouped sample data

I	0-4	5-9	10-14	15-19
F	2	3	5	2

Then the mean is:

A)8.1 B)5.7 C)7.2 D)4.9 E)9.92

Solution:

x	2	7	12	17
f. x	4	21	60	34

$$\overline{X} = \frac{\sum f.x}{n} = \frac{4+21+60+34}{12} = 9.92 \rightarrow E$$

Q9) The average salary of 15 male employees is 550 JD, and the average salary of 10 female employees is 420 JD, then the average salary among all these employees is:

A) 484 JD B) 498 JD C) 510 JD D) 515 JD E) 485 JD

Solution: $\sum x = n * \bar{X}$

$$\sum X_m = 15 * 550 = 8250$$

$$\sum X_f = 10 * 420 = 4200$$

$$\overline{X}_c = \frac{\sum X_m + \sum X_m}{n_1 + n_2} = \frac{8250+4200}{10+15} = 498 \rightarrow B$$

Q10) If the mean of 15, X, 2X+3 is 33, then the value of X is:

A) 18 B) 16 C) 25 D) 22 E) 27

Solution: $\bar{X} = \frac{\sum X}{n} \rightarrow 33 = \frac{15+X+2X+3}{3} \rightarrow X = 27 \rightarrow E$

Q11) From January to September, the mean number of car accidents per month was 630. From October to December, the mean was 1350 accidents per month. The mean number of car accidents per month for the whole year was:

A) 720 B) 675 C) 690 D) 810 E) 630

Solution: Jan to Sep → 9 months and Oct to Dec → 3 months

$\sum x = n * \bar{X}$

$\sum X_{J-S} = 9 * 630 = 5670$

$\sum X_{O-D} = 3 * 1350 = 4050$

$\bar{X}_c = \frac{\sum X_{J-S} + \sum X_{O-D}}{n_1 + n_2} = \frac{5670+4050}{12} = 810 \rightarrow D$

Q12) The mean of 50 observations is 85. If an observation was incorrectly recorded 150 instead of 15 , then the correct mean equals ?

A) 85.3 B) 82.3 C) 89.1 D) 9.1 E) 6.7

$\sum X_{old} = n * \bar{X} = 85 * 50 = 4250$

Since it was recorded 150 instead of 15 , then we must subtract 135.

$\sum X_{new} = 4250 - 135 = 4115$

Now $\bar{X} = \frac{\sum X}{n} = \frac{4115}{50} = 82.3 \rightarrow E$

Q13) The grades of 15 students have mean 40 .if the grade of a student is changed from 42 to 48 ,the new mean is =

A) 40.7 B) 40.6 C) 40.4 D) 40.8 E) 44.1

$\sum X = n * \bar{X} = 15*40 = 600$

Since the grade 42 to 48 , then we should add 6 to the old sum .

$\sum X_{new} = 600 + 6 = 606$

$\bar{X} = \frac{\sum X}{n} = \frac{606}{15} = 40.4 \rightarrow C$

Q14) Let 8 be the median of the observations: X , 9 , 7 , 1 , 1 , 11 , 2 , 14.
Then an acceptable value of X is:

A)8 B)7 C)10 D)6 E)5

Since $Q_2 = 8$, then if we order our data, it will be : 1,1,2,$\boxed{7,9}$,11,14

To get $Q_2 = \frac{7+9}{2} = 8$, then X should be more than or equal to 9 → C

Q15) For the following data (9 , -3 , X , X+3 , 11) the mean is 6,
then the median is:

A) 8 B) 6 C) 7 D) 10 E) 9

$\sum X = \overline{X} * n = 6*5 = 30 \rightarrow 30 = 9+(-3) + X + (X+3) + 11 \rightarrow X=5$

Now the data is : (9 , -3 , 5 , 8, 11)

Ordered: (-3 , 5 , 8 , 9 , 11)

$Q_2 = \frac{n}{2} = \frac{5}{2} = 2.5$ (fraction) 3^{rd} value

$\therefore\ Q_2 = 8 \rightarrow$ A

Q16) If the mean of the following observations is 3 (4 , 0 , 8 , X , 1) ,the median is :

A) 3 B) 4 C) 2 D) 5 E) 1

$$\overline{X} = \frac{\Sigma x}{n} \rightarrow 3 = \frac{4+0+8+x+1}{5}$$

$15 = 13 + x \rightarrow x = 2$

$\Rightarrow 4, 0 , 8 , 2 , 1 \rightarrow$ arrange $\rightarrow$ 0,1,2,4,8

$Q_2 = \frac{n}{2} = \frac{5}{2} = 2.5$ (fraction) $\rightarrow 3^{rd}$ value

$Q_2 = 2 \rightarrow C$

Q17) The marks of 8 students are given as follows: 6, 3, 9,10,8,4,8,9. A ninth student also takes the test, then the median will:

A) Increase to 10 B) Increase to 9 C) Decrease to 7

D) Remain 8 E) Decrease to 6

Our data is : 3,4,6,8,8,9,9,10

The median is 8

If the added value is 8 → the median remains 8

If the added value less than 8 → the median remains 8

If the added value more than 8 → the median remains 8 → D

Q18) The table below shows the marks gained in a test by a group of students

Mark	**1**	**2**	**3**	**4**	**5**
Frequency	**5**	**12**	**k**	**6**	**2**

The median is 3 and the mode is 2, the possible values of k are:

A) 11 and 12 **B) 8,9 and 10** **C) 8 and 9**
D) 9 and 10 **E) 10 and 11**

The mode is 2 → K <12

X	1	2	3	4	5
F	5	12	k	6	2
C.F	5	17	17 + k	23+k	25+k
Intervals	0-5	6 – 17	18 – (17+k)	-	-

$Q_2 = \frac{n}{2} = \frac{25+k}{2} \geq 18$ Since $Q_2 = 3$

$\therefore K \geq 11$

If k = 11 → $Q_2 = \frac{25+11}{2} = 18$ (Whole no.) → $\frac{18^{th}+19^{th}}{2} = \frac{3+3}{2} = 3$

If k = 10 $Q_2 = \frac{25+10}{2} = 17.5$ (Fraction) → 18^{th} value → $Q_2 = 3$

$\therefore k = 10 \; and \; 11 \; \rightarrow E$

Sheet (2)

Q1 The following is the age distribution for a random sample of 20 school students

Age in years (X)	10	12	15	18
Frequency	2	6	7	5

The range of the 20 ages in this sample is

A) 28 **B)35** **C)5** **D)8** **E)18**

Solution: Range = Max – Min = 18 – 10 = 8 → D

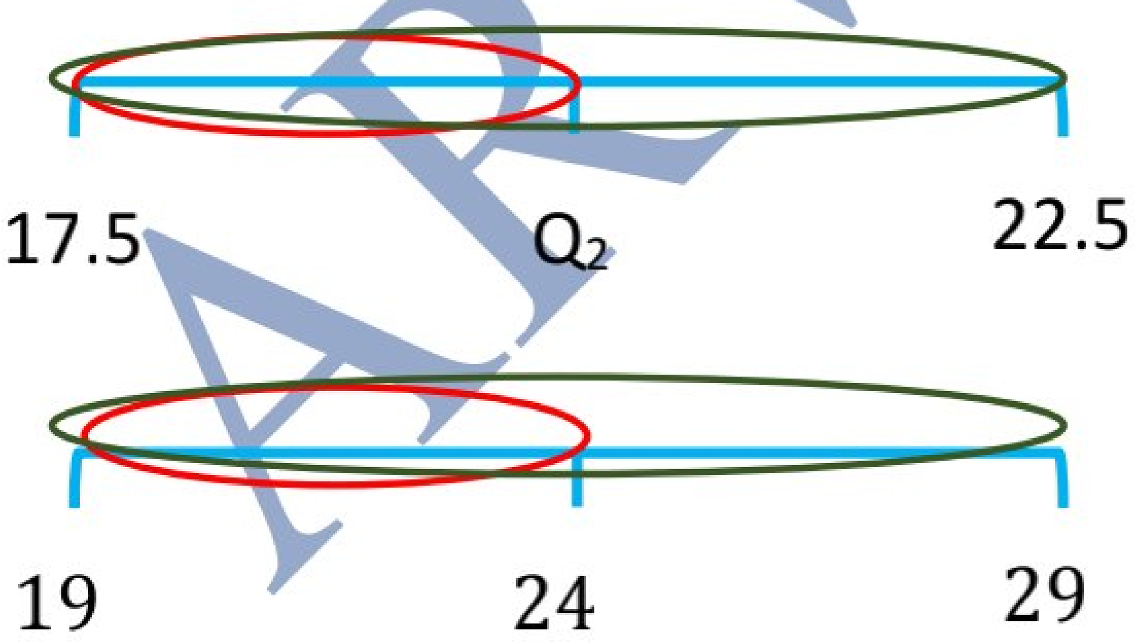
Q2 The 60^{th} percentile of the data presented in the table is:

Class	Frequency
8-12	7
13-17	12
18-22	10
23-27	8
28-32	3

A) 22 **B) 19** **C) 21** **D) 20** **E) 18.5**

Solution:

c. f	7	19	24	29	37	40
U. R. B	12.5	17.5	P_{60}	22.5	27.5	32.5



$$\rightarrow \frac{P60 - 17.5}{22.5 - 17.5} = \frac{24 - 19}{29 - 19} \Rightarrow P60 = 20 \ \rightarrow D$$

Q3 If the range is 30 , then the range of y after coding Y = -2X+3 will be:

A) 60 **B) -60** **C) 57** **D) 63** **E) -57**

Solution: Range (y) $= |-2|$ Range(x) $= 2 \times 30 = 60 \ \rightarrow A$

Q4 Consider the following frequency table of 20 observations:

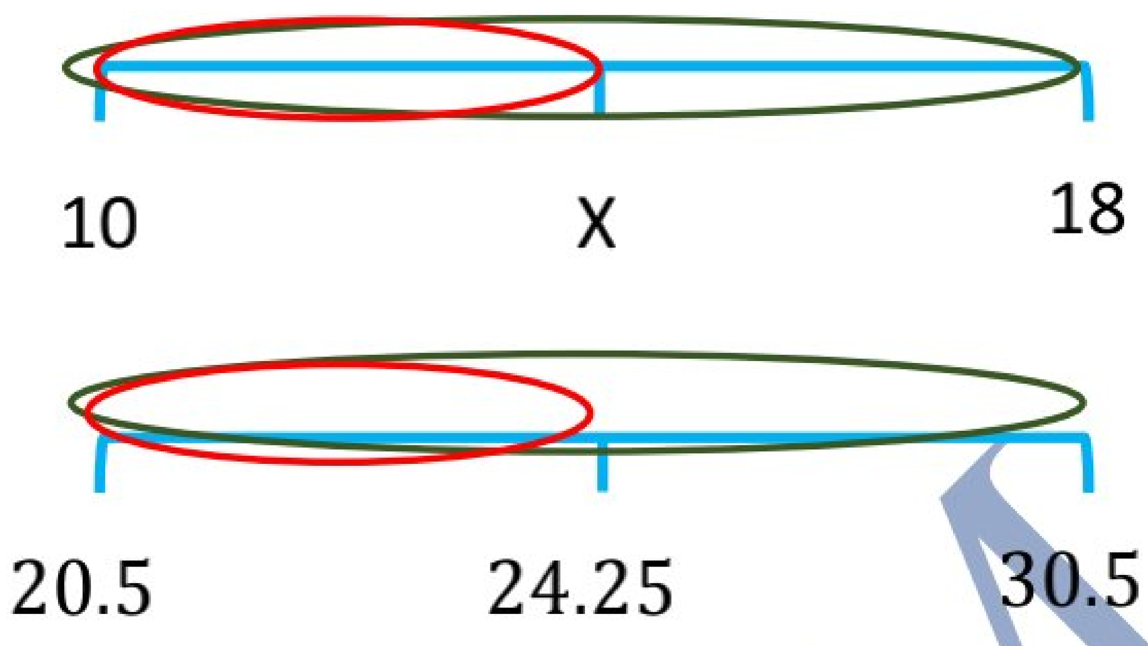
Class	Frequency
1-10	4
11-20	6
21-30	8
31-40	2

The number of observations such that each one of them has value greater than 24.25 is

A) 3 B) 5 C) 7 D) 9 E) 11

Solution:

C.F	4	10	x	18	20
U.R.B	10.5	20.5	24.25	30.5	40.5



$$\frac{X-10}{18-10} = \frac{24.25-20.5}{30.5-20.5} \to X = 13$$

No. of observations less than 24.25 = 13

No. of observations more than 24.25 = 20-13 = 7 → C

Q5 Find the 66th percentile of this sample data (as shown in the following picture)

Grade	3	8	13	18
Frequency	3	6	7	4

A) 13.5 B) 13.9 C) 12.8 D) 13 E) 13.2

C.F	3	9	16	20
Intervals	0 – 3	4 – 9	10 – 16	17 – 20

Solution: $P_{66} = \frac{k}{100} * n = \frac{66}{100} * 20 = 13.2$ (fraction) → 14th value

∴ $P_{66} = 13 \to$ D

Q6 Consider the following data :

Grade	1	2	3	4	5
Freq.	6	5	3	3	3

The 40% to 70% inter – percentile – range is :

A)3.5 B)4.5 C)3 D)1.5 E)2.5

<u>Solution:</u> We should find P_{40} & P_{70}

X	1	2	3	4	5
F	6	5	3	3	3
C.F	6	11	14	17	20
Intervals	0 – 6	7 – 11	12 – 14	15 – 17	18 – 20

$P_{40} = \frac{k}{100} * n = \frac{40}{100} * 20 = 8$ (whole no.) $\rightarrow P_{40} = \frac{8^{th} + 9^{th}}{2} = \frac{2+2}{2} = 2 \rightarrow P_{40} = 2$

$P_{70} = \frac{k}{100} * n = \frac{70}{100} * 20 = 14$ (whole no.) $\rightarrow P_{70} = \frac{14^{th} + 15^{th}}{2} = \frac{3+4}{2} = 3.5$

$P_{70} = 3.5$

$IPR = P_{70} - P_{40} = 3.5 - 2 = 1.5 \rightarrow D$

Q7 For the following grouped frequency distribution

Class	0-5	6-11	12-17	18-23
Frequency	2	8	7	3

The number of observations that lies below 21.5 is:

A) 19 B) 8 C)6 D)14 E)10

C.F	2	10	17	X
U.R.B	5.5	11.5	17.5	21.5

17	X	20

17.5	21.5	23.5

✓ $\frac{X-17}{20-17} = \frac{21.5-17.5}{23.5-17.5} \rightarrow X = 19 \rightarrow A$

Q8 Consider the following data :

Grade	1	2	3	4	5
Freq.	6	5	3	3	3

The 70th percentile (p_{70}) is :

A)3.5 B)4.5 C)3 D)4 E)2.5

Solution:

c.f	6	11	14	17	20
I	0 – 6	7 – 11	12 – 14	15 – 17	18 – 20

$$P_{70} = \frac{70}{100} \times n = \frac{70}{100} \times 20 = 14 \text{ (whole no.)} \rightarrow P_{70} = \frac{14^{th}+15^{th}}{2} = \frac{3+4}{2} = 3.5 \rightarrow A$$

Q9 What is the interquartile range for the following set of numbers?

4, 5, 6, 8, 9, 11, 13, 16, 16, 18, 20, 21, 25, 30, 31, 33, 36, 37, 40, 41.

A) 20 B) 22 C) 24 D) 37

4, 5, 6, 8, 9, 11, 13, 16, 16, 18, 20, 21, 25, 30, 31, 33, 36, 37, 40, 41.

$Q_2 = \frac{n}{2} = \frac{20}{2} = 10 \text{ (whole no.)} \rightarrow \frac{10^{th} + 11^{th}}{2} = \frac{18+20}{2} = 19$

استخدم الطريقة الجديدة

Left: 4, 5, 6, 8, 9, 11, 13, 16, 16,18→ Q1 = 10

Right: 20,21, 25, 30, 31, 33, 36, 37, 40, 41. → Q3 = 32

IQR $= Q_3 - Q_1$ = 32-10 = 22 → B

Q10 Consider the following data set:

(24 , 23 , 29 , 21 , 20 , 21 , 21 , 25 , 29 , 22 , 20 , 21) .The 25th percentile is :

A)25 B)23.5 C) 20 D) 21 E)23

arrange → 20,20,21,21,21,21,22,23,24,25,29,29

$$P_{25} = \frac{25}{100} \times n = \frac{25}{100} \times 12 = 3 \text{(whole no.)} \rightarrow \frac{3^{rd} + 4^{th}}{2} = \frac{21+21}{2} = 21 \rightarrow D$$

Q11 The following is the age distribution for a random sample of 20 school students:

Age in years (X)	**10**	**12**	**15**	**18**
Frequency	**2**	**6**	**7**	**5**

The third quartile Q_3 is:

A) 9.3 B)12 C)4.5 D)16.5 E)6.5

Solution:

C.F	2	8	15	20
Intervals	0 – 2	3 – 8	9 – 15	16 – 20

$Q_3 = \frac{3n}{4} = \frac{3*20}{4}$ =15 (whole number) $\rightarrow \frac{15^{th} + 16^{th}}{2} = \frac{15+18}{2} = 16.5 \rightarrow$ D

Q12 The age distribution of a sample of 30 persons is as follows:

Age class	**10-14**	**15-19**	**20-24**	**25-29**	**30-34**
Freq.	**3**	**7**	**10**	**7**	**3**

The 90^{th} percentile:

A)27 B)29.5 C)28.5 D)14.5 E)0

Solution:

c. f	3	10	20	27	30
U. R. B	14.5	19.5	24.5	29.5	34.5

$P_{90} = \frac{90}{100} \times n = \frac{90}{100} \times 30 = 27 \quad \Rightarrow P_{90} = 29.5 \ \rightarrow B$

Sheet (3)

Q1 The following is the frequency table of a sample data . the first quartile $Q_1 = 7$ & the third quartile $Q_3 = 17$, the number of outliers in this sample is:

X	frequency
3	2
7	6
17	11
18	2
33	3

A) 3 B) 2 C) 0 D) 5 E) 4

Solution: $IQR = Q_3 - Q_1 = 17 - 7 = 10$

Less than $Q_1 - 1.5*IQR \rightarrow 7 - 1.5*10 = -8$
More than $Q_3 + 1.5*IQR \rightarrow 17 + 1.5*10 = 32$

33 is an outlier , and since it was repeated 3 times , then the number of outliers is 3 → A

Q2 Given the sample data: -8,-6,3,4,4,6,8,10,20,26

All the outlier (s) for this sample data is (are)

A)24,26 B)-8,26 C)-8,24,26 D) -8,-6 E) No outliers

Solution: 5

-8,-6,3,4,4↓6,8,10,20,26

$Q_2 = \frac{n}{2} = \frac{10}{2} = 5\ (whole\ no.) \rightarrow \frac{5th + 6th}{2} = \frac{4+6}{2} = 5$

Left: -8, -6, 3, 4, 4 → $Q_1 = 3$

Right: 6,8,10,20,26 → $Q_3 = 10$

$IQR = Q_3 - Q_1 = 10 - 3 = 7$

Less than $Q_1 - 1.5IQR = 3 - 1.5(7) = -7.5$ (-8 is an outlier)

More than $Q_3 + 1.5IQR = 10 + 1.5(7) = 20.5$ (26 is an outlier) → B

Q3 A sample of size 50 has a first quartile Q_1=15 and third quartile Q_3=35 , some of the observations are (1 , 5 , 23 , 36 , 55 , 70) , the possible outliers are:

A) 1& 5 B) 55 & 70 C) 70 only D) 5 only

Solution:

Outliers: more than $Q_3 + 1.5*IQR = 35 + 1.5*(35-15) = 65$

∴ 70 is an outlier.

Less than $Q_1 – 1.5*IQR = 15 – 1.5*(35-15) =-15$

∴ No outlier → C

Q4 Carlos recorded his friends' scores while playing the video game "Golden Eye Commander". Most of his friends' scores were between 9 and 12. One score, however, was 28, and Carlos identified it as an outlier. What should Carlos do with the score of 28 when recording this data?

A) Ignore the outlier since its so far from the average scores.

B) Ignore the outlier because he may have recorded the score incorrectly.

C) Eliminate the outlier and ask that friend to play again to obtain a new score.

D) Keep the outlier as it may help to explain a new strategy for playing the game.

Solution: The answer is D

Q5 The areas of the 46 villages in Jordan, in square meters, are listed in order below. The areas range from 392 to 1228 with Q_1 = 507, M = 660.5, and Q_3 = 795

392	**394**	**395**	**397**	**407**	**411**	**413**	**462**	**463**	**485**	**494**	**507**
511	**512**	**516**	**555**	**557**	**563**	**567**	**576**	**577**	**586**	**647**	**674**
682	**696**	**696**	**700**	**710**	**724**	**740**	**757**	**758**	**772**	**795**	**804**
806	**814**	**819**	**916**	**937**	**1080**	**1128**	**1133**	**1134**	**1228**		

According to the 1.5 X IQR rule, are there any potential outliers in the data set?

A) Yes, 392 and 1228 are potential outliers.

B) Yes, 1228 is a potential outlier.

C) Yes, 1080, 1128, 1133, 1134, and 1228 are potential outliers.

Solution: $IQR = Q_3 – Q_1 = 795 – 507 = 288$

Less than $Q_1 – 1.5 IQR = 507 – 1.5(288) = 75$ (no noutliers)

More than $Q_3 + 1.5 IQR = 795 + 1.5 (288) = 1227$ (1228 is an outlier)→ *B*

Sheet (4)

Q1 A sample of 10 observations has mean = 30 , standard deviation = 5 . the sum of the squares is

A) 9350 B) 9320 C) 9000 D) 9225 E) 9100

Solution: $\sum Xi = n * \bar{X} = 10 * 30 = 300$

$$S^2 = \frac{\sum X^2}{n-1} - \frac{(\sum X)^2}{n(n-1)} \rightarrow 25 = \frac{\sum X^2}{10-1} - \frac{300^2}{10*(10-1)} \rightarrow \sum X^2 = 9225 \rightarrow D$$

Q2 The variance of the numbers 1,1,1,3,5 is:

A)3.2 B)2.1 C)2.7 D)4.1 E)1.9

Solution:

x	1	1	1	3	5
x^2	1	1	1	9	25

$\Sigma x = 11$, $\Sigma x^2 = 37$

$$S^2 = \frac{\Sigma x^2}{n-1} - \frac{(\Sigma x)^2}{n(n-1)} = \frac{37}{4} - \frac{(11)^2}{5\times 4} = 3.2 \rightarrow A$$

Q3 If you are told a population has a mean of 25 and a variance of 0, what must you conclude?

A) Someone has made a mistake.

B) There is only one element in the population.

C) There are no elements in the population.

D) All the elements in the population are 25.

Solution: Since the variance is zero this means all the data are equal , & the mean is 25 then all the data will be equal to 25 so we can get mean 25→ **D**

Q4 For a certain data set. you told that the variance S=0. what else can you say about the data set?

A) Median = 0 B) Mean = 0 C) Range = 0 D) a + b

Since S = 0 , then all the data values are equal →Range = Max – Min = 0 → C

Q5 In a sample of 10 students the grades have mean 60 and variance 36. If two students with grades 50 and 40 left the class, then the new sum of square of grades $\sum(\boldsymbol{xi})^2$ =

A) 30211 B) 31332 C)31799 D)32224 E)32599

Solution: $\sum X_{\text{old}} = n * \bar{X} = 10*60 = 600$

$$S^2_{\text{old}} = \frac{\sum X^2}{n-1} - \frac{(\sum X)^2}{n*(n-1)} \rightarrow 36 = \frac{\sum X^2}{9} - \frac{600^2}{10*9}$$

$\therefore \sum X^2_{\text{old}} = 36324$

✓ $\sum X^2_{\text{new}} = 36324 - (40)^2 - (50)^2 = 32224 \rightarrow D$

Q6 Consider the following sample data:

class	frequency
1-5	2
6-10	4
11-15	4

If the mean of this sample is 9, then the variance is :

A)12.25 B)15.56 C)14.47 D)16.67 E) 13.36

Solution: $\bar{X} = \frac{\sum F.X}{n} \rightarrow \sum F.X = \bar{X}*n = 9*10 = 90 \rightarrow \sum F.X = 90$

X(mid-point)	3	8	13
X^2	9	64	169
$F.X^2$	18	256	676

$\sum F.X^2 = 18+256+676 = 950$

✓ $$S^2 = \frac{\sum F*X^2}{\sum F-1} - \frac{(\sum F*X)^2}{\sum F*(\sum F-1)} = \frac{950}{10-1} - \frac{90^2}{10*(10-1)} = 15.56 \rightarrow B$$

Q7 When the standard deviation is negative, then:

A) most scores were above the mean.
B) most scores were below the mean.
C) the distribution is badly skewed.
D) someone made a mistake.
E) none of these because the standard deviation can never be negative.

Solution: The answer is E

Q8 The mean of a sample data of size 100 is 45 and the variance is 200. When an observation is deleted, $\sum x^2$ of this sample becomes 220700. Then this observation is:

A) 44 B) 40 C) 42 D) 36 E) 38

Solution: $\sum X = n * \bar{X} = 100*45 = 4500$

$$S^2_{old} = \frac{\sum X^2}{n-1} - \frac{(\sum X)^2}{n*(n-1)} \rightarrow 200 = \frac{\sum X^2}{99} - \frac{(4500)^2}{100*99}$$

$\therefore \sum X^2_{old} = 222300$

The difference between $\sum X^2_{old}$ and $\sum X^2_{new}$ is the value squared :

$X^2 = 222300 - 220700 = 1600$

$\therefore X = \sqrt{1600} = 40 \rightarrow$ B

Q9 If the standard deviation of a data set is zero, then all entries in the data must equal zero.

A) False; the standard deviation can never be zero because it measures the distance from the mean and distances are always greater than zero.

B) False; if the standard deviation is zero, then all of the data values are equal.

C) True; since the standard deviation is equal to the mean, all the data values must be zero.

D) True; since the standard deviation is a measure of how spread out the data are, a standard deviation of zero means that all data values must be zero.

Solution: because $S=\sqrt{\frac{\sum(X-\bar{X})^2}{n-1}}=0$, this means there is no difference between X & $\bar{X}$, thus all the data will be equal $\rightarrow$ B

Q10 The following table gives the number of registered mobile lines for a randomly selected sample of 9 students, the variance of this sample is:

Number of mobile lines	Number of students
1	2
2	2
3	2
5	3

A)2 B) 1 C) 1.5 D) 2.5 E) 2.75

Solution:

$f.x$	$(1)(2) = 2$	$(2)(2) = 4$	$3(2) = 6$	$5(3) = 15$
x^2	$(1)^2 = 1$	$(2)^2 = 4$	$(3)^2 = 9$	$(5)^2 = 25$
$f.x^2$	$2(1) = 2$	$2(4) = 8$	$2(9) = 18$	$3(25) = 75$

$$S^2 = \frac{\Sigma f.x^2}{n-1} - \frac{(\Sigma f.x)^2}{n(n-1)} = \frac{103}{8} - \frac{(27)^2}{9 \times 8} = 2.75 \rightarrow E$$

Q11 A sample data of 10 numbers has mean 12 and variance $S^2 = 16$ If one number in this sample was changed from 8 to 10 , then the new standard deviation is:

A) 11.067 B)3.82 C)1584 D)3.33 E)14.62

Solution:

$\overline{X} = 12$, $n = 10 \rightarrow \Sigma x_{Old} = n \times \overline{X} = 12 \times 10 = 120$

$\Sigma x_{New} = 120 + (10 - 8) = 122$

$$S^2_{Old} = \frac{\Sigma x^2}{n-1} - \frac{(\Sigma x)^2}{n(n-1)} \rightarrow 16 = \frac{\Sigma x^2_{Old}}{9} - \frac{(120)^2}{10 \times 9}$$

$\rightarrow \Sigma x^2_{Old} = 1584$

$\Sigma x^2_{New} = 1584 - 8^2 + 10^2 = 1620$

$$S^2_{New} = \frac{1620}{9} - \frac{(122)^2}{10 \times 9} = 14.62$$

$S = \sqrt{14.62} = 3.82 \rightarrow B$

Q12 Of the following Dot plots, which represents the set of data that has the greatest standard deviation?

A) * * * * * * * * * *

B) ***** *****

C) *** **** ***

D) **** ** ****

E) ** ** ** ** **

Solution: The answer is B

Q13 Suppose X in the table beside is the number of emails that eight persons received per hour, then the variance of X is:

x	Frequency
2	2
3	4
4	2

A) 0.571 **B) 0.756** **C) 3.0** **D) 0.50** **E) 0.707**

Solution:

F.X	X^2	$F.X^2$
2*2=4	4	2*4 = 8
3*4 = 12	9	4*9 =36
4*2 =8	16	2*16 = 32

$\sum F.X = 4 + 12 + 8 = 24$

$\sum F.X^2 = 8 + 36 + 32 = 76$

$S^2 = \frac{\sum F*X^2}{\sum F-1} - \frac{(\sum F*X)^2}{\sum F*(\sum F-1)} = \frac{76}{7} - \frac{(24)^2}{8*7} = 0.571 \rightarrow A$

Sheet (5)

Q1 Let y=2-3x if the inter-quartile range (IQR) of y is 30 and Q1 of x is 20, then Q3 of x equals:

A)30 B)-30 C)31 D)40 E)50

Solution:
IQR of Y = |-3|*(IQR of X) → 30 = 3*(IQR of X) → IQR of X =10
IQR = $Q_3 - Q_1$ → $10 = Q_3 - 20$ → $Q_3 = 30$ → A

Q2 In a sample, the standard deviation is 4, if the observation (X) is changed to Y= -4x-1, then the new variance is:

A) -16 B)20 C)256 D)12 E)64

Solution:

$S = 4 \rightarrow S^2 = 16$

$S_y^2 = (-4)^2 S_x^2 = 16*16 = 256 \rightarrow C$

Q3 In a sample , the 1st quartile is 12 , the median is 35 & the 3rd quartile is 42 , if each observation x is changed to Y= -2X-1 ,then the new 3rd quartile will be:

A)-85 B) -84 C) -25 D) -27 E) -2

Solution:

$Q_{3(y)} = -2(Q_{1(x)}) - 1 = -2(12) - 1 = -25 \rightarrow C$

Q4 The variance of the monthly salaries in a company is 400 dinars. if the salary of each employee in this company has been multiplied by 1.2, then the new variance =
A)376 B) 576 C)488 D) 334 E) 48

Solution:

$S_x^2 = 400 \quad , \quad y = 1.2\,x$

$S_y^2 = (1.2)^2 . S_x^2$

$= (1.2)^2 \times 400 = 576 \rightarrow B$

Q5 The variance of the yearly salaries of teachers in a private school is 250000. At the end of year, the school management decides to award each teacher a bonus of 100 dinars and 10% of the yearly salary. The standard deviation of the teacher yearly salary after this bonus becomes:

A) 600 B) 50 C) 100 D) 550 E) 500

Solution:

$y = (1.1)x + 100$

$S_x = \sqrt{250000} = 500$

$S_y = (1.1) \times 500 = 550 \rightarrow D$

Q6 Suppose the mean of a math exam score is 64 with a standard deviation of 10, if each score is increased by 6 and then each result is multiplied by 11, then new mean and standard deviation are:

A) $\mu = 770$ and $\sigma = 110$ B) $\mu = 770$ and $\sigma = 101$ C) $\mu = 777$ and $\sigma = 126$ D) $\mu = 706$ and $\sigma = 100$ E) $\mu = 770$ and $\sigma = 120$

Solution:

$\bar{X} = 64$, S = 10

Y = (X + 6) * 11 = 11 X + 66 → Y = 11X + 66

$\mu_y = 11 * 64 + 66 = 770$

$\sigma_y = |a| * \sigma_x = 11*10 = 110 \rightarrow A$

Q7 In a sample the observations are symmetric with first quartile 12 and median 16. If each observation X in this sample is updated to Y = 6 – 3X , then the new first quartile is:

A)-50 B)-31 C)-36 D)-54 E)-60

Solution:

Symmetric means : $Q_2 - Q_1 = Q_3 - Q_2 \rightarrow 16 - 12 = Q_3 - 16 \rightarrow Q_3 = 20$

For X : $Q_1 = 12$, $Q_2 = 16$ and $Q_3 = 20$

Y = -3 X + 6

For Y : $Q_1 (y) = -3 *(Q_3 (x)) + 6 = -3 * 20 + 6 = -54 \rightarrow D$

Sheet (6)

Q1 Given the five number summary, determine if there are any outliers in the data set. Min =3 Q1=6 Q2 =9 Q3 =12 Max =20:

A) It's not possible to determine if there are any outliers based on the information given

B) There is no outliers

C) There are at least one on the low end & one on the high end

D) There are at least one on the low end or at least one on the high end

E) There are at least on the low end and no outliers on the high end

Solution:

less than $Q_1 - 1.5\text{ IQR} = 6 - 1.5(6) = -3$

No outliers because $\text{Min} = 3$

More than $Q_3 + 1.5\text{ IQR} = 12 + 1.5(6) = 21$

No outliers because $\text{Max} = 20 \rightarrow \text{B}$

Q2 Which of the following measurements cannot be obtained from the box plot

A) mean. B) median. C) maximum. D) first quartile. E) third quartile.

Solution:

The answer is A.

Q3 Given a data set consisting of 32 whole number observations, it is five number summary is (12, 24, 39 , 54, 64) , how many observations are less than or equal to 24 ?

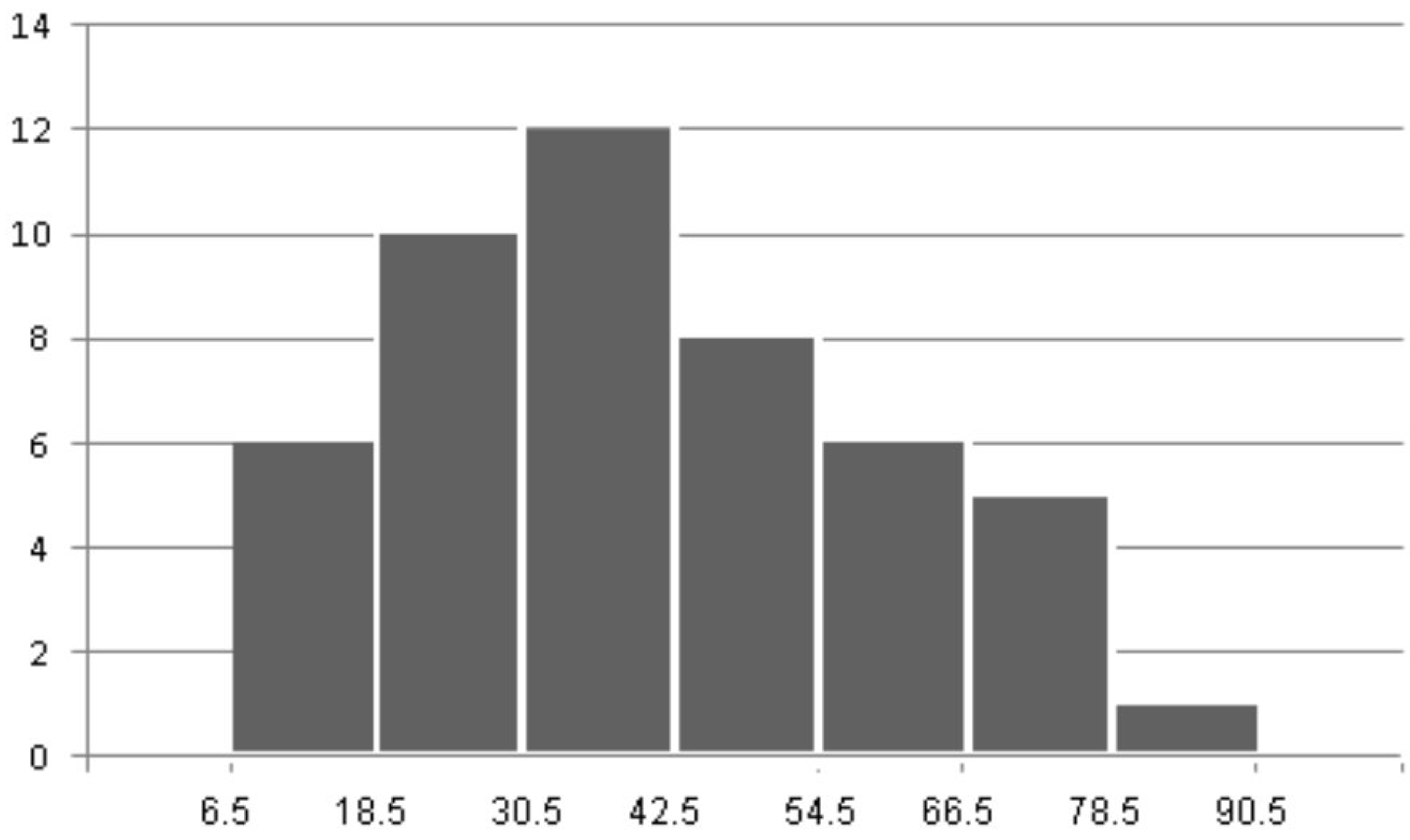
Solution:

24 is the first quartile (Q_1) , so 25% of the data less than 24

Number of observations = $\frac{25}{100}$ *32 = 8 → there is 7 values below it.

Q4 Consider the following histograms graph, How many observation are less than 30.5?

A) 16 B) 14 C) 17 D) 30 E) 9



<u>Solution:</u> $6 + 10 \rightarrow 16$ observations $\rightarrow$ A

Q6 A set of data has the following five number summary

Minimum	First quartile	median	Third quartile	maximum
17	37	40	49	90

Which of the following contains all the outliers in the distribution?

A) 83 , 85 , 90 , 95
B) 17, 81 , 80 , 85 , 90
C) 64, 80, 85
D) 2, 3, 85 , 90
E) 0, 80 , 84 ,89

<u>Solution:</u>

- اي خيار من الخيارات يحتوي على قيمة اقل من الـ Min او أعلى من الـ Max مباشرة بنرفضه،لإنو القيمة مش موجودة ضمن قيم البيانات ، لهيك A ,D , E مرفوضين تماماً.

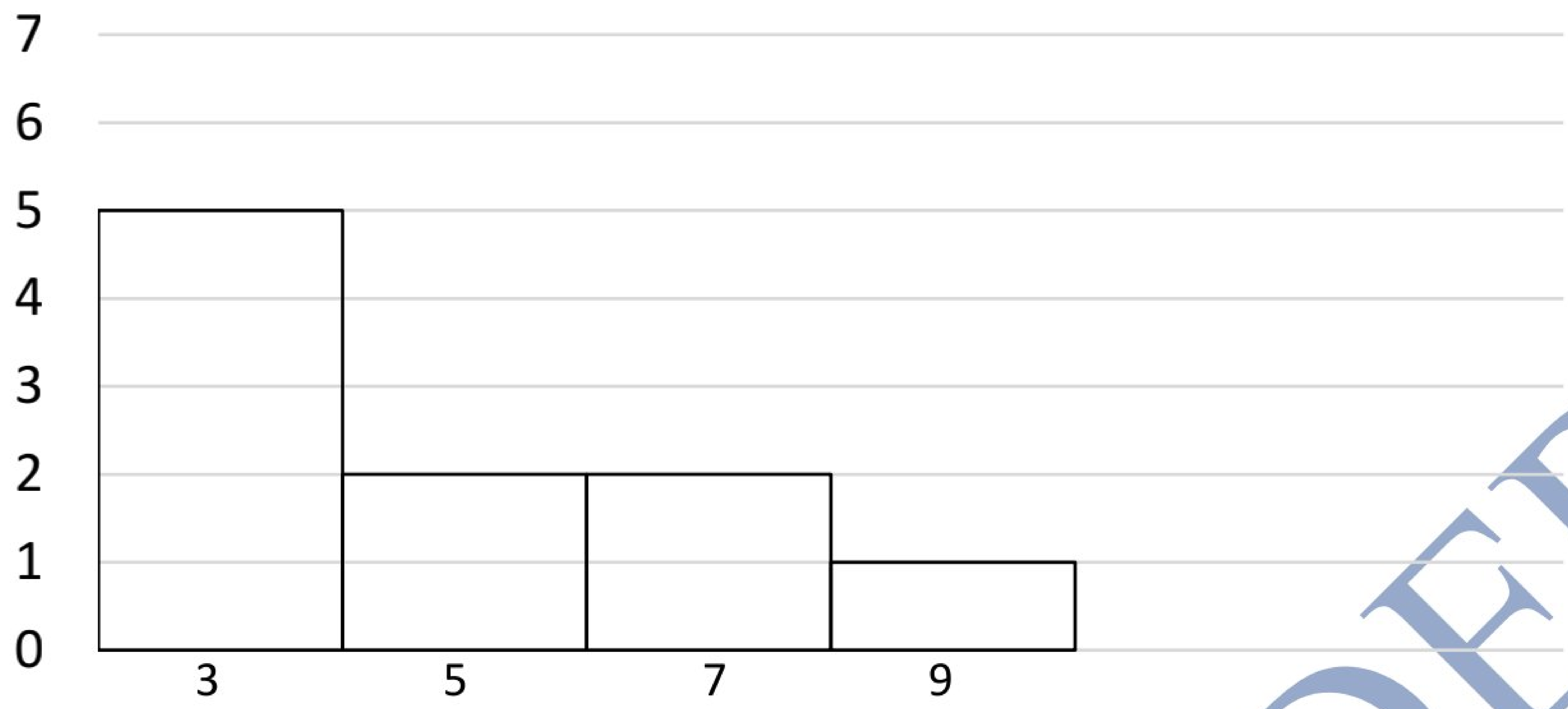
$IQR = Q_3 - Q_1 = 49 - 37 = 12$

Less than $Q_1 - 1.5\ IQR = 37 - 1.5*12 = 19$

More than $Q_3 + 1.5\ IQR = 49 + 1.5*12 = 67$

The answer is B

Q7 The following is the histogram of a grouped sample data. The mean equals:



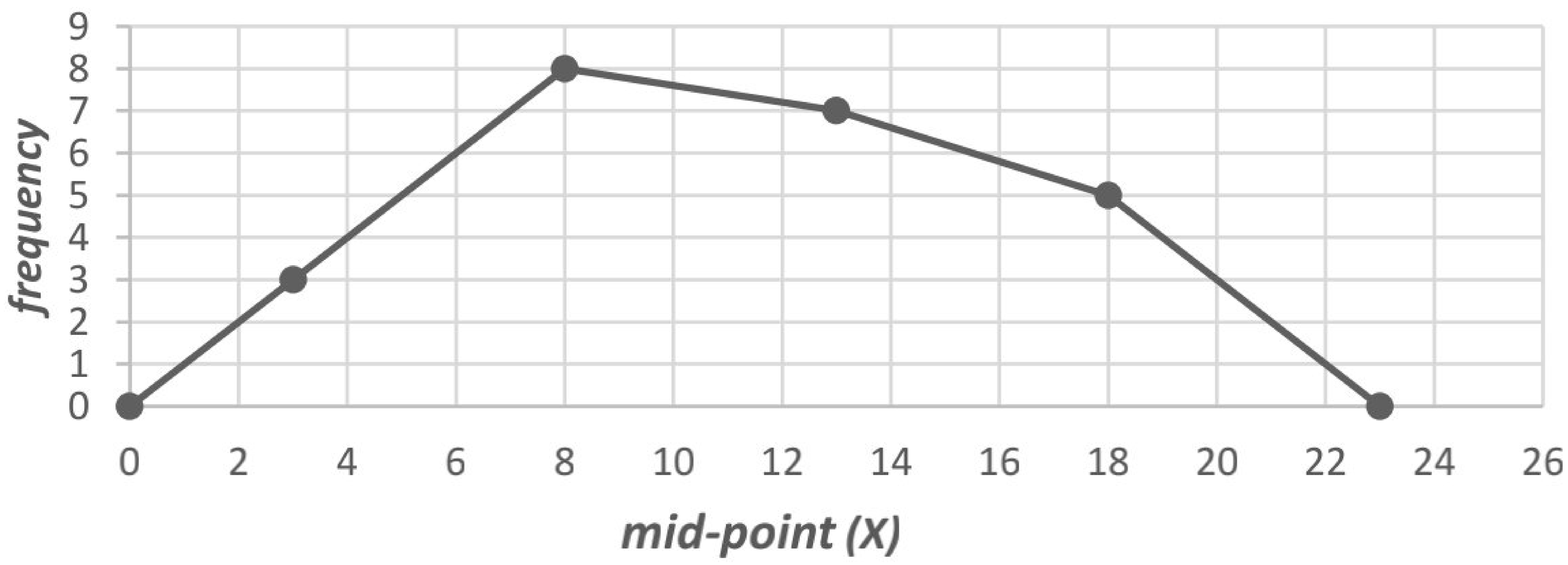
A) 2.2 B)3.4 C) 3.2 D) 3 E) 4.8

Solution:

x	3	5	7	9
f	5	2	2	1
f. x	15	10	14	9

$\overline{X} = \frac{\Sigma f.x}{n} = \frac{15+10+14+9}{10} = 4.8 \rightarrow E$

Q8 For the following polygon, find the mean?

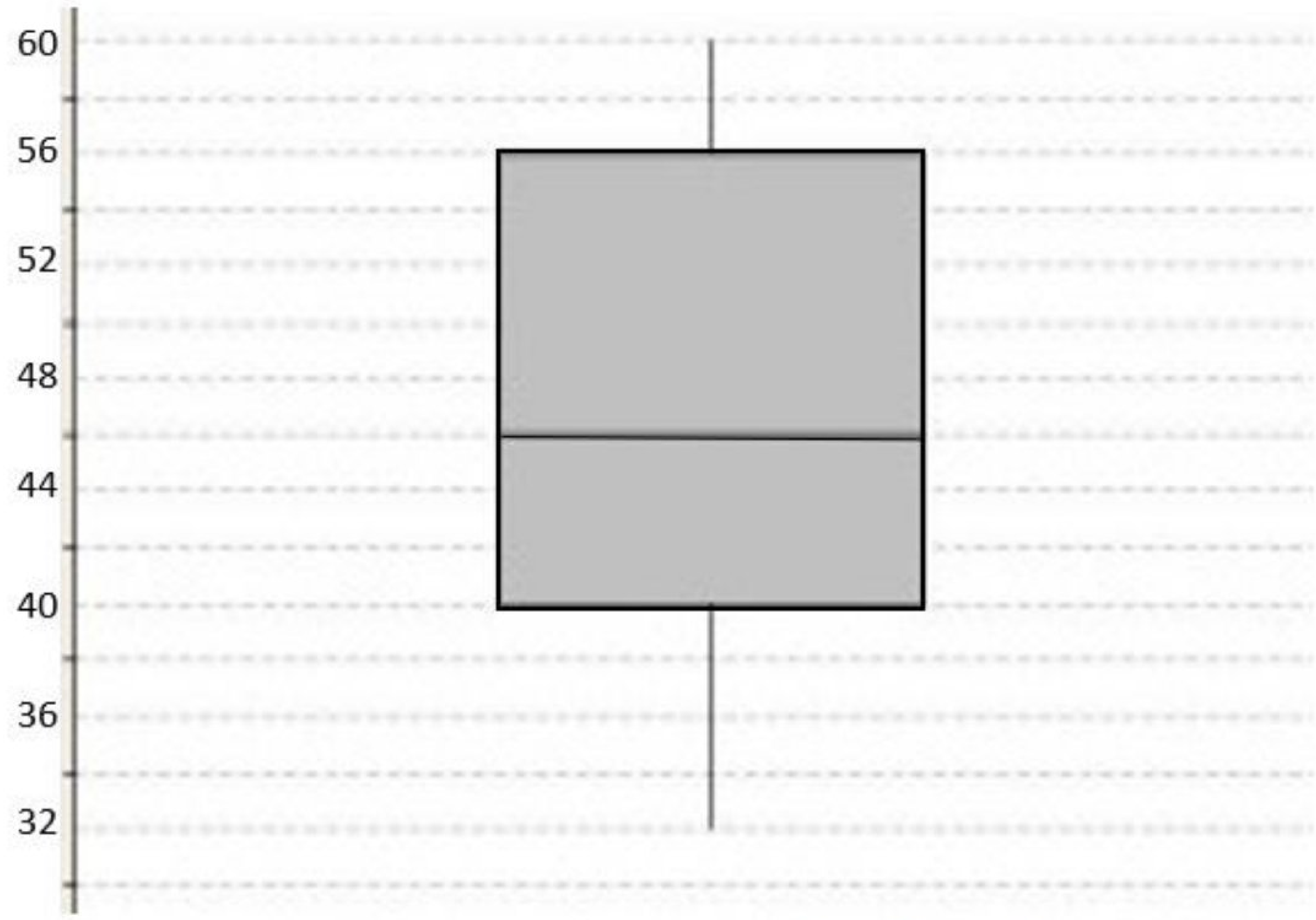


Solution: We must convert our graph to table:

X	3	8	13	18
F	3	8	7	5
F.X	9	64	91	90

$\bar{X} = \frac{\sum F.X}{\sum F} = \frac{9+64+91+90}{23} = 11.04$

Q9 The times that statistics students needed to finish an exam, are shown in the boxplot below.



Find the following:

(1) IQR (2) outliers

<u>Solution:</u>

1) IQR= $Q_3 - Q_1 = 56 - 40 = 16$

2) Outliers : less than $Q_1 - 1.5*IQR = 40 - 1.5*16 = 16$ (No outliers)

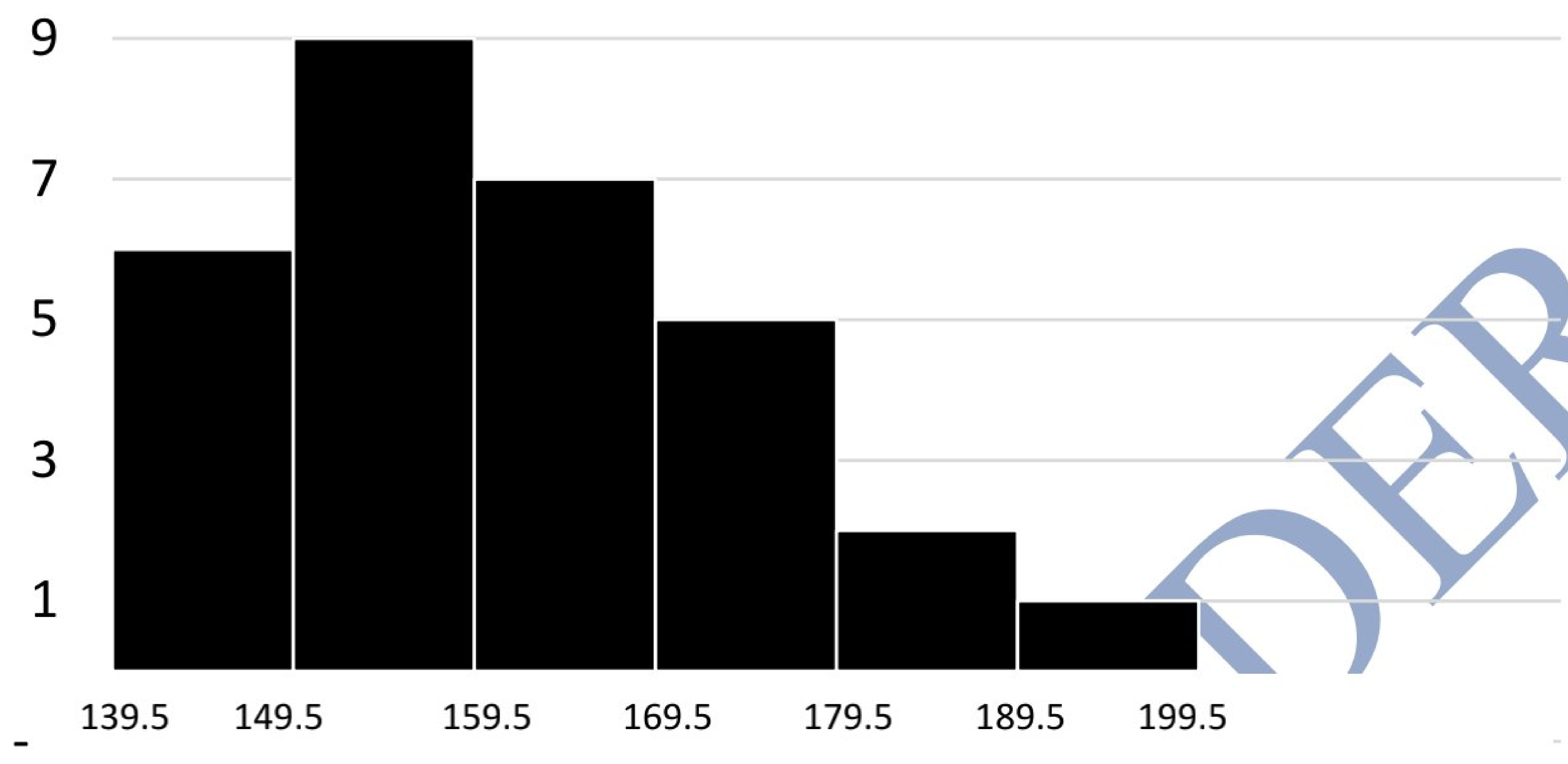
More than $Q_3 + 1.5*IQR = 56 + 1.5*16 = 80$ (No outliers)

Q10 the length of the box in a boxplot represent:
A) Range B) IQR C) variance D) Q3

<u>Solution:</u> The answer is B

Q11 The histogram below shows the heights of 30 people, the percentage of people with heights greater than or equal 159.5 cm is :

a) 50% **b) 60%** **c) 40%** **d) 55%** **e) 45%**



<u>Solution:</u> Percentage $= \frac{F}{\sum F} * 100\% = \frac{7+5+2+1}{6+9+7+5+2+1} * 100\% = 50\% \rightarrow$ A

Sheet (7)

Q1 If the frequency curve of a set data has a bell shape with mean 200 and variance 100, then the percentage of observations that lies in the interval (170, 190) approximately equals to:

A) 81.5% B) 83.5% C) 2% D 13.5% E)15.5%

Solution:

$170 = \bar{X} - K.s \rightarrow 170 = 200 - k * 10 \rightarrow K = 3$

$190 = \bar{X} - K.s \rightarrow 190 = 200 - k * 10 \rightarrow K = 1$

Per = 0.495 – 0.34 = 0.155 = 15.5% → E

Q2 In a sample of 400 observations, the mean is 60 and the variance is 16.

The number of observations outside the interval (40,80) are:

A) at most 100 B) at most 16 C) at most 25

D)at least 384 E) at least 356

Solution:

$40 = \bar{X} - K.S \rightarrow 40 = 60 - k*4 \rightarrow K = 5$

Outside = at most $\frac{1}{K^2} = \frac{1}{5^2} = 0.04$

No. of observations = 0.04 * 400 = 16 → B

Q3 The mean & standard deviation of a sample of size 100 are 12 and 1 ,the smallest possible number of observations that are between 10 & 14 is :

A)61 B) 84 C) 75 D) 80 E) 65

Solution:

$n = 100 \quad , \quad \bar{X} = 12 \quad , \quad S = 1$

$10 = \bar{X} - k.s \rightarrow 10 = 12 - k(1) \rightarrow k = 2$

The percentage between is at least $1 - \frac{1}{k^2} = 1 - \frac{1}{4} = 0.75$

# of observation $= 0.75 \times 100 = 75 \rightarrow C$

Q4 A sample of size 200 observations has means 40 and standard deviation 9 . An interval that contains at least 150 observations is:

A) (22, 58) **B) (26.5, 53.5)** **C) (28.75, 51.25)**

D) (13, 67) **E) (17.5, 62.5)**

Solution:

$$\text{Proportion of observation} = \frac{150}{200} = 0.75$$

$\text{at least } 1 - \frac{1}{k^2} = 0.75 \rightarrow \frac{1}{k^2} = 0.25 \rightarrow k^2 = 4 \rightarrow k = 2$

$\Rightarrow (\bar{X} - k.s, \bar{X} + k.s) = (40 - 2(9), 40 + 2(9)) = (22,58) \rightarrow A$

Q5 A sample data has mean = 30 & Std =14. An interval that contains at least 51% of the data in this sample:

A) (20, 40) **B) (16, 44)** **C) (22, 38)** **D) (10, 50)** **E) (18, 42)**

Solution:

$\text{at least } 1 - \frac{1}{k^2} = 0.51 \rightarrow k = 1.43$

$(\bar{X} - k.s, \bar{X} + k.s) \rightarrow (30 - 1.43\ (14), 30 + 1.43\ (14)) = (10,50) \rightarrow D$

Q6 In a sample of 500 observations. the percentage of observations above which 400 observations is:

A) P_{20} **B)** P_{30} **C)** P_{40} **D)** P_{15} **E)** P_{35}

Solution:

Percentage below = $\frac{400}{500} = 0.80$

Percentage above =100-80 = 20 → A

Q7 The grade of 1000 students are bell shaped with mean = 50 and S = 13. Then What is P_{84}:

A) 56 B)63 C) 37 D) 73 E) 89

Solution:

$P_{84} = \bar{X} + S = 50 + 13 = 63 \rightarrow B$

Q8 The average rain fall in Irbid is 600 m/m, with a standard deviation of 80 m/m over a period of 60 years. Give approximate number of years with rain fall over 680 m/m. (Assume a bell-shaped distribution).

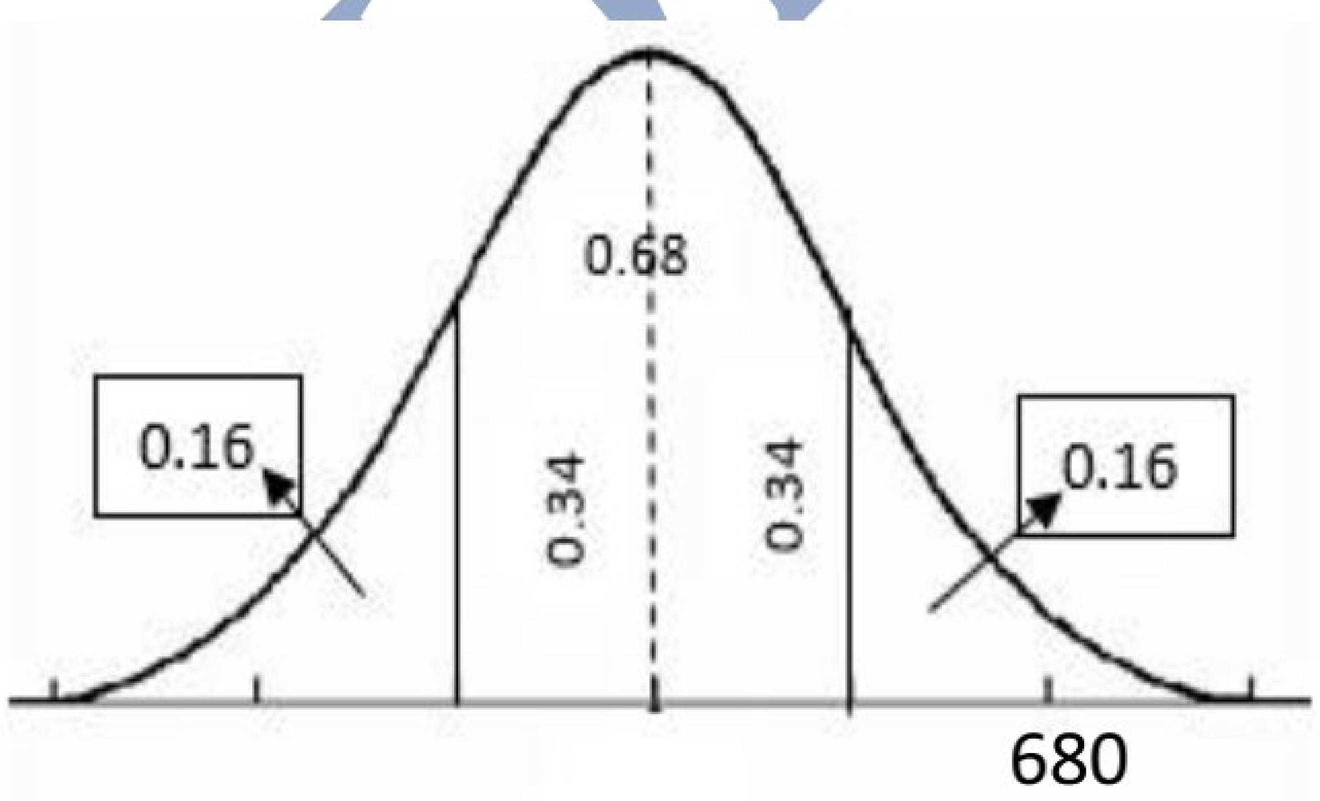
A) 2 B) 4.5 C) 10 D) 7 E) 3

Solution:

$\bar{X} = 600\,, S = 80\,, n = 60$

$680 = \bar{X} + k.S \rightarrow 680 = 600 + k * 80 \rightarrow k = 1$

$\therefore 680 = \bar{X} + S$



Percentage above 680 is 16%

No. of years = 0.16 * 60 = 9.6 ≈ 10 → C

Q9 The distribution of the grades of a Mathematics exam is bell shaped with mean 60 and standard deviation 13. The percentage of students with grades less than 73 is

A) 97.5% B) 68% C) 95% D) 84% E)16%

Solution:

$73 = \bar{X} + k.S \rightarrow 73 = 60 + k * 13 \rightarrow k = 1$

$\therefore 73 = \bar{X} + S$

Percentage below 73 is 84% → D

Q10 According to the Chebyshev's rule, the approximate proportion of observations within 3.5 standard deviations of the mean is:

A) At least 8% B) At least 92% C) At most 92%
D) Exactly 8% E) At most 8%

Solution:

Within = at least $1 - \frac{1}{k^2} = 1 - \frac{1}{3.5^2} = 0.92 = 92\% \rightarrow B$

Q11 If a given distribution is known to be symmetric, what percent of the observations are expected to fall within two standard deviations of the mean?

A) 68 B) 75% C) 89% D) 95% E) almost all.

Solution:

Symmetric = bell-shaped

Percentage within 2S = 0.95 = 95% → D

Q12 In 2013, the prices of homes sold in Amman was skewed to right with mean $ 255,000 & standard deviation $105,000. For 2013, the minimum percentage of homes sold in Amman with selling prices between $45,000 and $465,000 is:

A) 75% B) 84% C) 88.89% D) 95% E) 99.7%

Solution:

$\bar{X}$-K*S = 45,000 → 255,000 – k*105,000 = 45,000 → K=2

Minimum → $1-\frac{1}{k^2} = 1-\frac{1}{4} = 0.75 = 75\%$ → A

Q13 The mean of the sample data of 180 observation is 40 & the standard deviation is 5. An interval that contains at least 160 of these observations is:

A) (50,70) B) (25,55) C) (160,69) D) (68,80)

Solution:

n = 180, $\bar{X}$ = 40 , S =5

$\text{percentage} = \frac{\text{number of obs in interval}}{\text{total number of observations}} = \frac{160}{180} = 0.889$

At least $1-\frac{1}{k^2} = 0.889$ → K=3

($\bar{X}$-K*S, $\bar{X}$+K*S) = (40 - 3*5, 40 + 3*5) = (25, 55) → B

Q14 For a bell-shaped sample data 16% of observation are greater than 42 and 16% of them are less than 30. The mean $\bar{X}$ of this sample equals:

A) 18 B) 16 C) 36 D) 42 E) 30

<u>Solution:</u>

$42 = \bar{X} + k.S$

$30 = \bar{X} - k.S$

$\rightarrow \bar{X} = \frac{42+30}{2} = 36 \rightarrow$ C

Q15 In a certain city, the yearly rainfall has mean 500 mm and standard deviation 70 mm over a period of 200 years.

Assuming that the distribution is bell shaped, the approximate number of years with rainfall between 290 mm and 360 mm are:

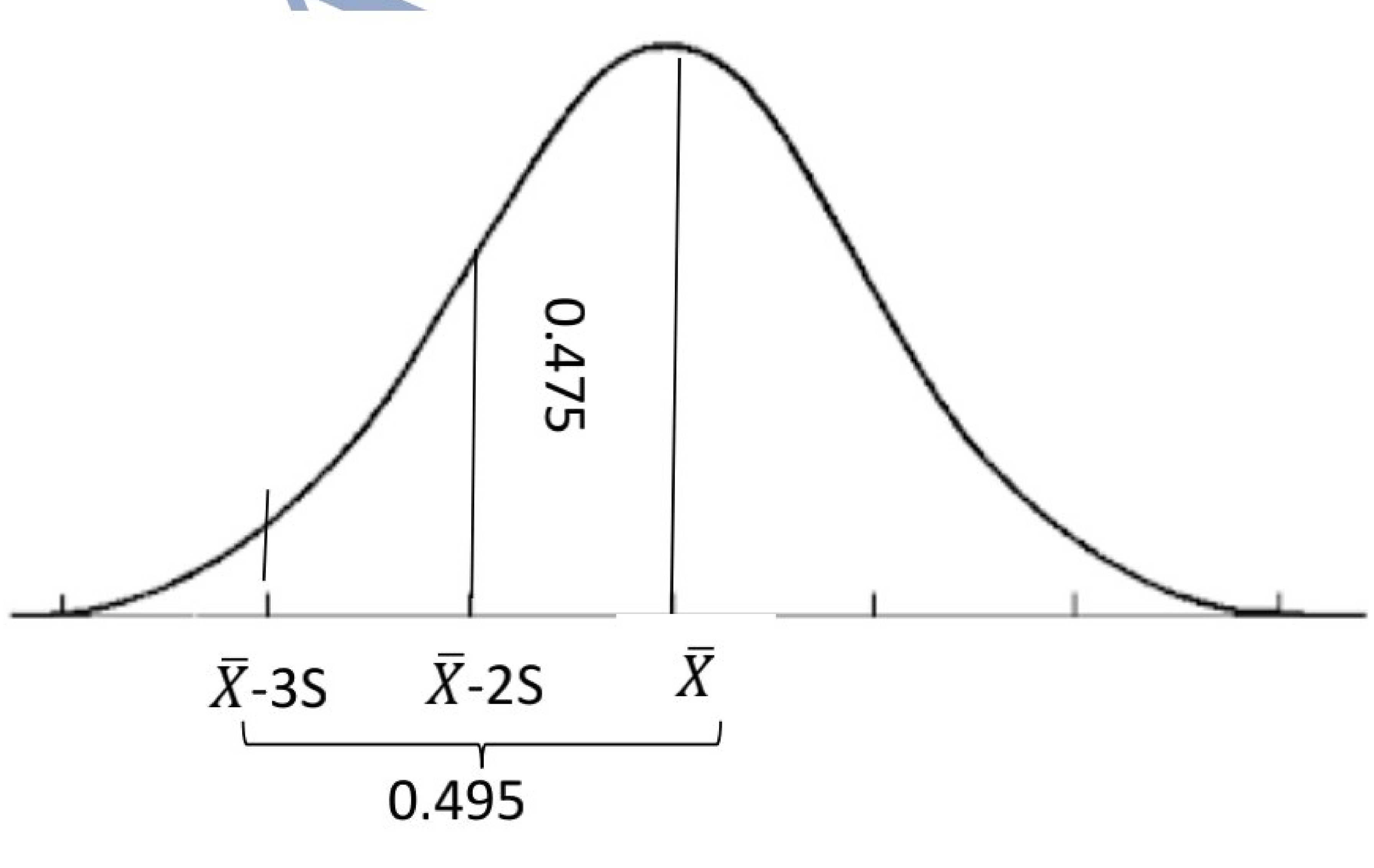
A)8 B)27 C)31 D)4 E)22

<u>Solution:</u>

Since it is bell-shaped, we should find the values of K .

$290 = \bar{X} - K*S \rightarrow 290 = 500 - k(70) \rightarrow K = 3$

$360 = \bar{X} - K*S \rightarrow 360 = 500 - k(70) \rightarrow K = 2$



Percentage = 0.495 – 0.475 = 0.02

∴ No. of years = 0.02 * 200 = 4 → D