

# Introduction to Statistics

## 1.1- An Overview of Statistics

### 1.2- Data Classification

#### True Or False Questions:

1- Nominal level of measurement consists of categories only where data cannot be arranged in an ordering scheme:

تتكون

الترتيب غير محتم

True

False

2- A population is a subset of a sample

العينة جزء المجتمع

True

false

3- The relative frequency of a class is the sample size divided the frequency of the class

على تكرار فئة مقسوم بحجم العينة هو التكرار النسبي

True

False

4- The sum of percentage frequency is 1

التكرار النسبي مجموع

True

False

5- A population is the collection of some outcomes, responses, measurement, or counts that are of interest.

النواتج بعض مجموعة هو المجتمع

True

False

نقطة المنتصف  
 المجموعة كلها  
 الحد الأدنى  
 الحد الأعلى

6- The midpoint of a class is the sum of its lower and upper limits

True

False

choose the correct answer Questions:

1- The branch of statistics that involves using a sample to draw conclusions about a population.

data

statistical inference

descriptive statistics

variable

2- The example of a qualitative data is

number of accidents

blood type

age of patients

weight of students

الذي يشمل الاحصاء فرع للحصول على معالم المجتمع  
 استخدام لعينة

مثال على البيانات النوعية

فحيرة الدم

## Descriptive Statistics

### 2.1- Frequency Distributions & Their Graphs

**Q1-** For the following frequency distribution choose the correct answer:

Class intervals	Frequency
40 - 49	10
50 - 59	5
60 - 69	7
70 - 79	3

1- The class width is:

- a) 10    b) 20    c) 7    d) 5

2- The midpoint of the third class is:

- a) 30    b) 24.5    c) 54.5    d) 64.5

3- The class boundary (true class) of first class is:

- a) 40.5-49.5    b) 39.5-49.5    c) 50.5-59.5    d) 70.5-79.5

4- The relative frequency of the last class is:

- a) 0.5    b) 0.12    c) 0.25    d) 0.15:

5- The cumulative frequency of a second class is:

- a) 25    b) 22    c) 10    d) 15

اختاری

توزیع التکرار لپہاں ہے

Q2- Choose the correct frequency distribution for these data:

ABCAD  
BCFDC  
AFDFB  
DCACC

لہذا لپیٹا ہے

(a)

class	Frequency
A	4
B	3
C	6
D	4
F	3

(b)

class	Frequency
A	4
B	4
C	6
D	4
F	1

(c)

class	Frequency
A	4
B	4
C	5
D	4
F	2

(d)

class	Frequency
A	4
B	4
C	6
D	4
F	2

Q3- From the distribution table below: The midpoint in A is:

Class intervals	True class intervals	Frequency	Midpoints X	The percentage frequency
4 - 7	3.5 - 7.5	4		
8 - 11	7.5 - 11.5	6		B
12 - 15	11.5 - 15.5	2		
16 - 19	15.5 - 19.5	3	A	
Total		15		

$$\frac{16+19}{2} = 17.5$$

From the distribution table below: The percentage frequency in B is:

Class intervals	True class intervals	Frequency	Midpoints X	The percentage frequency
4 - 7	3.5 - 7.5	4		
8 - 11	7.5 - 11.5	6		B
12 - 15	11.5 - 15.5	2		
16 - 19	15.5 - 19.5	3	A	
Total		15		

$$B. \left( \frac{6}{15} \times 100 \right) \% = 40\%$$

المكرر النسبي مجموع

Q4-The sum of relative frequency in a frequency distribution table is:

Zero

0.5

10

1

المكرر النسبي

Q5- For the class 8-18, the lower-class limit is 8:

True

False

كنت لو قال  
\* lower - class Boundary  $\rightarrow 7.5$   
يعني بطرح 0.5

Choose the correct answer.

Using the following frequency table to answer questions, 1, 2, 3, 4, 5 and 6

Interval	Frequency	Midpoint	Relative frequency
0 – 9	5	4.5	0.125
10 – 19	11	14.5	0.275
20 – 29	10	24.5	0.250
30 – 39	9	34.5	.....
40 – 49	2	44.5	0.05
50 – 59	1	.....	.....
60 – 69	2	64.5	0.05
Total	40		.....

1- Percentage frequency of (30 – 39) is:

$$\left(\frac{9}{40} \times 100\right) \% = 22.5\%$$

- a) 5%      b) 22.5%      c) 2.25%      d) 25%

2- Class width is:

- a) 20      b) 11      c) 9      d) 10

3- Midpoint of (50 – 59) is:

$$\frac{50 + 59}{2} = 54.5$$

- a) 54.5      b) 79.5      c) 84      d) 4.5

4- Upper boundary class of (20 – 29) is:

- a) (19 – 30)      b) 20.5      c) 29.5      d) 19.5

5- the sum of relative frequencies is:

- a) 1      b) 100      c) 0      d) 1.5

6- the number of intervals is:

- a) 5      b) 6      c) 8      d) 7

المتكرر النسبي مجموع

الفترات عدد

Lecture 2

## Descriptive Statistics

### 2.1- Frequency Distributions & Their Graphs

### 2.2 More Graphs and Displays

#### True Or False Questions:

البيانات الكمية

1- we use steam and leaf plot to represent quantitative data.

True

False

حجم العينة

أكبر قيمة

أصغر قيمة

2- The sample size is 20, the maximum data is 94 and the minimum data 36

Key = 6|4 = 64

3	6
4	
5	6 6 8
6	0 4 4 6 6 5
7	6 8 8 8
8	0 2 4 6 8
9	4

True

false

choose the correct answer Questions:

لتمثيل  
- يجب رسمه أي

1- Which graph should be used to represent "sales of five types of fruits in Riyadh"?

استخدامها

مبيعات خمس أنواع من إغاكه

A) Histogram

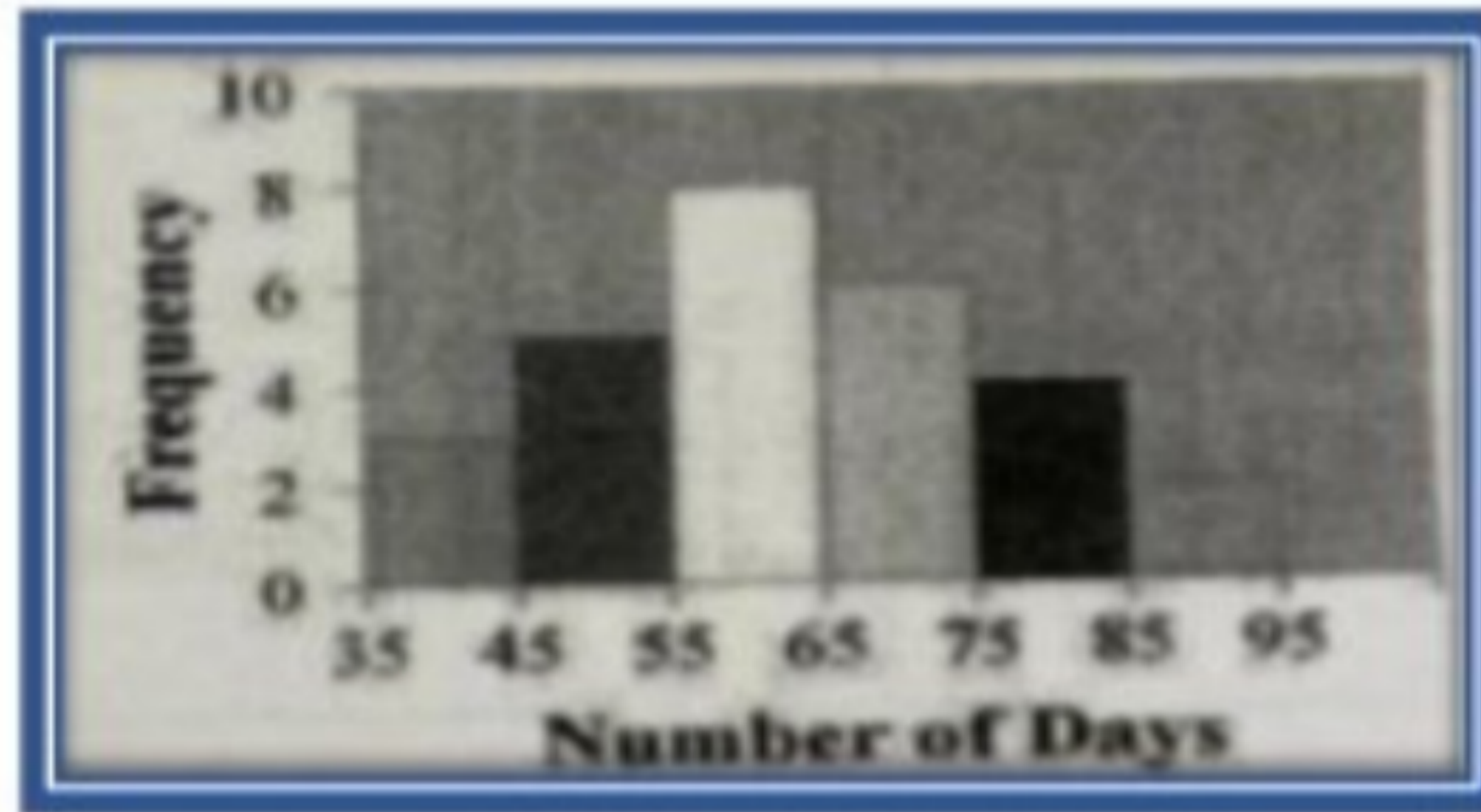
**b) Bar Chart**

c) Ogive

d) Steam and Leaf

أوجدى لفتره التي لها أعلى قيمه

2- Find the class with the highest number of data values:



a) 75-85

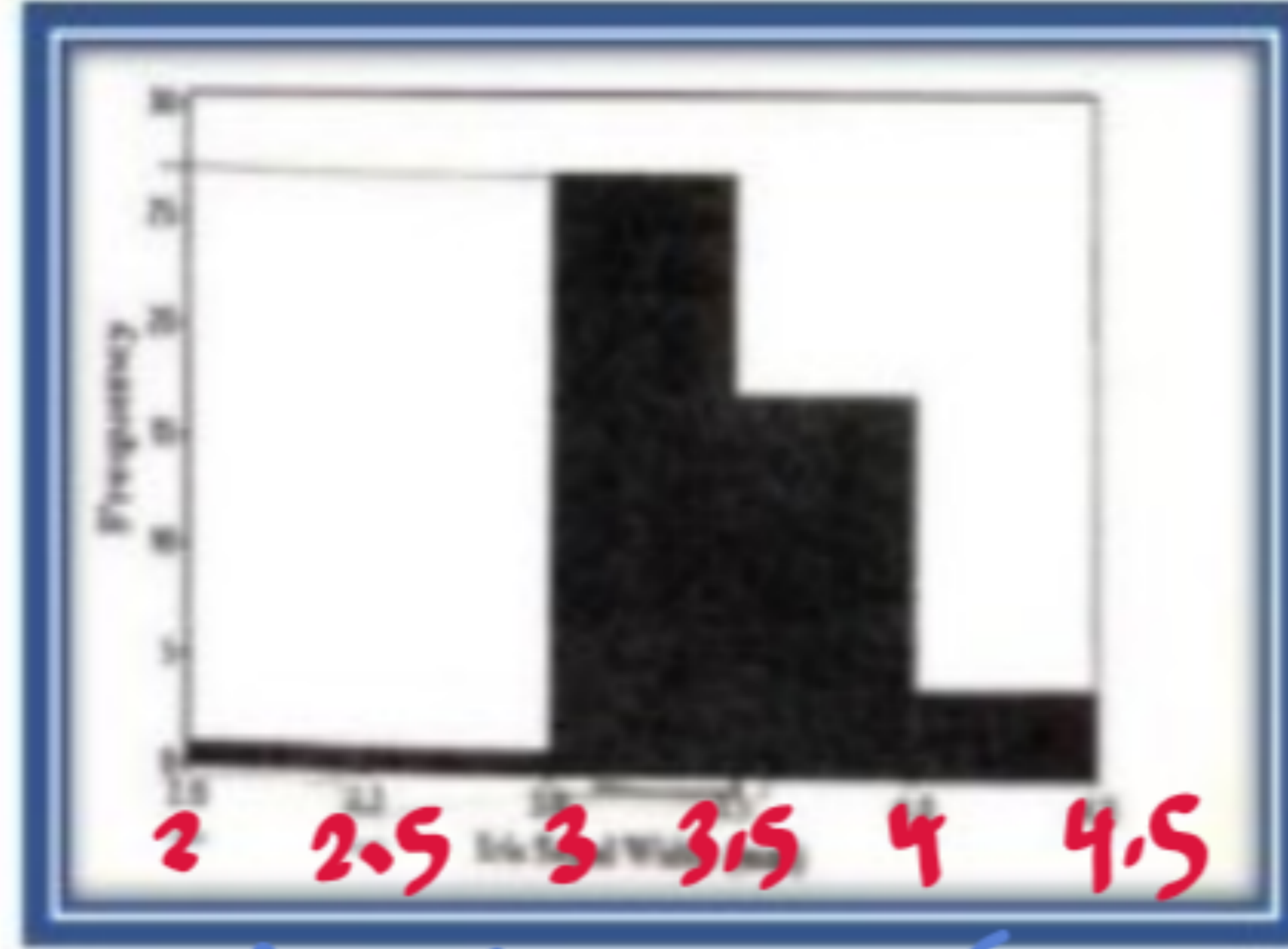
**b) 55-65**

c) 65-75

d) 8

## Questions:

Q1- Based on the following histogram:



آی فتره لها أعلى تكرار

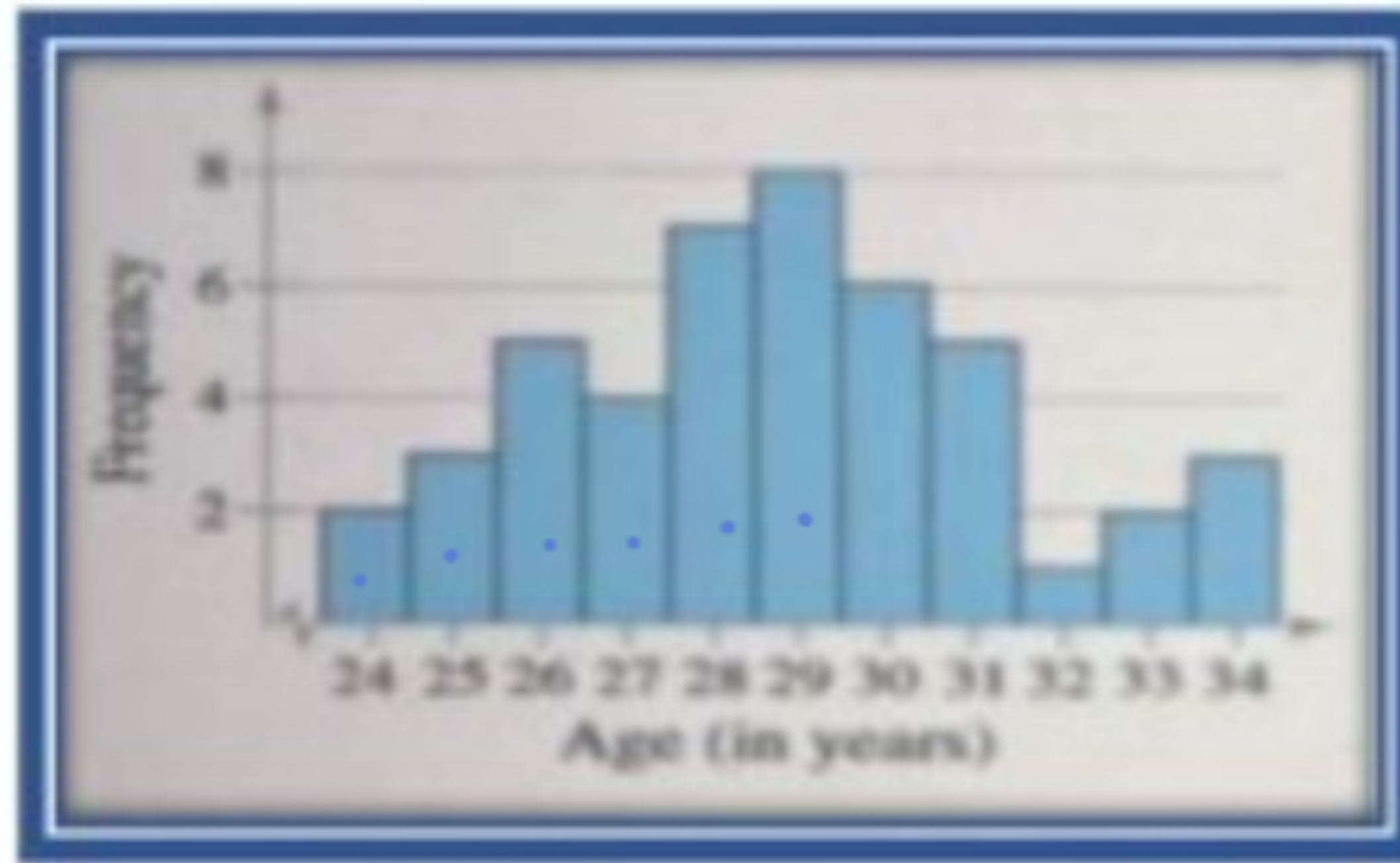
1- Which class corresponds to the highest frequency?

3 - 3.5

2- What are the approximate lower- and upper-class limits of the first class?

2 - 2.5

Q2- Bride's Age at first marriage



تعدد الفترات

1- What is the number of classes?

11

2- What is the number of the fourth class?

4

Q3- If 92 patients were given a blood test to determine their blood type, if the frequency of type O! blood is 23 then the central angle of it in a pie chart is:

فصيلة الدم  
لقدية  
الخبثا دم  
عريف

$$\frac{23}{92} * 360 = 90^\circ$$

## Descriptive Statistics

### 2.3- Measures of Central Tendency

### 2.4- Measures of Variation

#### True Or False Questions:

القيمة

الأكثر تكرار

1- The value most often repeated in a data set is called the mode. ✓

True

False

2- Some data sets have no mode.

بعض البيانات ليس لها mode

True

false

الوسيط

يتأثر

3- The median is affected by the extreme value.

القيمة المتطرفة

True

False

هو المتوسط

4- The mean is one of the measures of variation

أحد مقاييس التشتت

True

False

الوسيط

لا يتأثر

5- The median is not affected by the outliers

بالقيمة المتطرفة

True

False

choose the correct answer Questions:

1- Find the Mean for the following data

1, 3, 6, 4, 5, 6, 3, 4, 6, 3, 6

a) 4.273

b) 47

c) 2.5

d) 5.8

$$\bar{x} = \frac{1+3+6+\dots+6}{11} = 4.273$$

2- The mean of the following data is

1,2,5,3,7,9,12,10

a) A. 10

b) B. 6.12

$$\bar{x} = \frac{1+2+\dots+10}{8} = 6.12$$

3- The median of student's marks 20, 60,80,40,30 is

a) 40

b) 30

c) 60

d) 80

~~20~~, ~~30~~, 40, ~~60~~, ~~80~~

صین الڈیکر

4- In a right-skewed distribution, which is greater?

a) The mean

b) The media

c) The mode

d) The mean= the median= the mode.

ملا حینہ  
left  
mode → الڈیکر

5- The range of the values 10, 15, 40, -5, 0, 99 is:

a) 105

b) 104

c) 99

d) 94

$$\begin{aligned} \text{Range} &= \text{max} - \text{min} \\ &= 99 - (-5) = 104 \end{aligned}$$

ايه اكثر حاجة مناسبة؟

6- What is the most appropriate measure of central for the following data set?

red, blue, blue, red, red, black, yellow

a) The mean

b) The mode

c) The median

d) The midrange

الارخاف الطيارين

7- The standard deviation of the first four pulse rates for males 73, 73, 73, 73 is:

a) 73

b) 0

c) 70

d) 292

المحاضرة الرابعة

[ تساوي هيز لادن كلهم متساويين ]

البيانات كم نسبة

9- What percent of the data lies above the Median?

a) About 50%

b) About 25%

c) About 75%

d) Can't tell without looking

التي تقع فوق

تکرار  
تکرارِ قیَمہ

10- The value most often repeated in a data set is called

- a) mean
- b) median
- c) mode
- d) variance

11- If  $\bar{x} = 5$ ,  $S = 0$  and  $n = 6$ , then the data is

- a) 6, 6, 6, 6, 6
- b) 5, 5, 5, 5, 5, 5
- c) 0, 0, 0, 0, 0, 0
- d) 5, 5, 5

المجاہدۃ، جرابعدہ

12- The mean of the frequency distribution is:

(a)  $\bar{x} = \frac{\sum(wx)}{\sum w}$

(b)  $\bar{x} = \frac{\sum(fx)}{\sum f}$

(c)  $\mu = \frac{\sum(x)}{n}$

(d)  $\bar{x} = \frac{\sum(x)}{N}$

پلٹوسے فی حالہ  
جدول پلٹوزیع پلٹراری لہو:-

## Questions:

**Q1-** Using the following data to find 1, 2 and 3:

3 9 15 10 12 14 5 9

1- The mean is

- a) 8.6   **b) 9.6**   c) 10.6   d) 11.6

$$\bar{x} = \frac{3 + \dots + 9}{8}$$

2- The median is:

- a) 8.5   **b) 9.5**   c) 10.5   d) 11.5

~~3~~ ~~5~~ ~~9~~ **9** **10** ~~12~~ ~~14~~ ~~15~~

3- The mode is

- a) 8   **b) 9**   c) 10   d) 11

**Q2-** For the following sample, answer questions 14-17:

16, 10, 9, 46, 15, 10, 16

14. The sample mean is:

- a) 122   b) 15   c) 7   **d) 17.43**

$$\bar{x} = \frac{16 + 10 + \dots + 16}{7}$$

15. The median is:

- a) 15**   b) 10   c) 16   d) 46

~~9~~ ~~10~~ ~~10~~ **15** ~~16~~ ~~16~~ ~~46~~

أفضل مقياس لقياس لهذه البيانات هو:-

17. The best measure of central tendency to represent this data is:

- a) The mean   b) The mode   **c) The median**   d) None

ایہ آگے وچہرہ قیاس مناجبہ:

**Q3-** What is the most appropriate measure of central tendency for the following data set

male, female, female, male, male, male, female

- The median
- The mode
- The weighted mean
- The mean



**Q4-** The mode of the following data is

1,3,6,1,8,3,9,4,1

- A. 3
- B. one mode 1
- C. two mode 1, 3
- D. 9

**Q5-** For the following data:

Status	Married	Single	No response
Frequency	35	25	5

the mode is:

- a) Married       b) 35       c) No mode       d) Single

باستخدام

هذا الرسم

Q6-Using the following graph, stem and leaf, answer questions 1, 2 and 3.

Key 4 | 4 = 4 4

1	0 4
2	0 3 4 5 5 6 7 9
3	2 3 5 7
4	4 4 6 9
5	3 6 9
6	0 1
7	5 6 9
8	0



1- The range of data is:  $80 - 10 = 70$

a) 70

b) 75

c) 66

d) 61

عدد البيانات

2- The number of data are:

a) 25

b) 22

c) 27

d) 28

3. The Mode of data (is/are):

a) 2

b) 25 & 44

c) 0, 4, 5, 6 & 9

d) no mode

Q8- For the following frequency distribution

Class intervals	Frequency
4 - 7	4
8 - 11	6
12 - 15	2
16 - 19	3
Total	15

mid point	$x \cdot f$
5.5	22
9.5	57
13.5	27
17.5	52.5
	<b>158.5</b>

Find the sample mean

$$\bar{x} = \frac{\sum x \cdot f}{\sum f} = \frac{158.5}{15} = 10.567$$

Q9- Find the mean for the following data

Class interval	Frequency f	Midpoints x	$x \cdot f$
10-14	2	12	24
15-19	5	17	58
20-24	3	22	66
25-29	2	27	54
30-34	7	32	224
35-39	5	37	185
Total	24		638

a) 26.58

b) 106.3

c) 27.47

d) 256

$$\bar{x} = \frac{638}{24} = 26.58$$

الصفات الدراسية من ثلاث مواد درجات الطلاب

**Q10-** The student's marks in three courses are 10, 20, 30, and study hours for these Courses are 2, 3, 4, respectively. Then, the weighted mean  $\bar{x}_w$  is

a) 36

b) 22.22

c) 10

d) 2

$$\bar{x} = \frac{\sum x \cdot w}{\sum w} = \frac{10 \cdot 2 + 20 \cdot 3 + 30 \cdot 4}{2 + 3 + 4} = 22.22$$



تسمى التكرار الأكثر القيمة

**Q15-** the value most often repeated in a data set is called

(a) mean

(b) median

(c) mode

(d) variance