

4

Ch. 2 - Part 2

- Other Types of Graphs

- * Pareto Chart.
- * The time series graph.
- * The Pie graph.

STAT.110

جمال السعدي
رياضيات - إحصاء

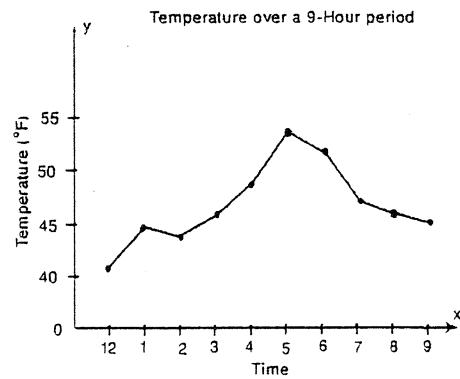
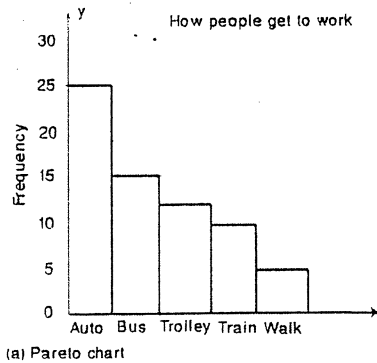


CH.2 Part 2

جمال السعدي

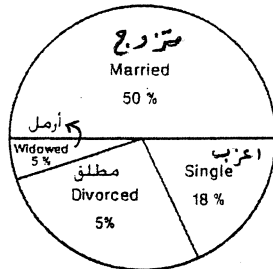
Other types of graphs:

In addition to the histogram, the frequency polygon, and the Ogive, several other types of graphs are often used in statistics. They are the Pareto chart, the time series graph, and the pie graph. Figure shows an example of each type of graph.



الحالة الاجتماعية

Marital status of Employees
at Brown's Department store



Pareto chart

is used to represent a frequency distribution for a categorical variable and the frequencies are displayed by the heights of vertical bars, which are arranged in order from highest to lowest.

On a Pareto chart must be:

1. Make the bars the same width.
2. Arrange the data from largest to smallest according to frequency.
3. Make the units that are used for the frequency equal in size.

● Example:

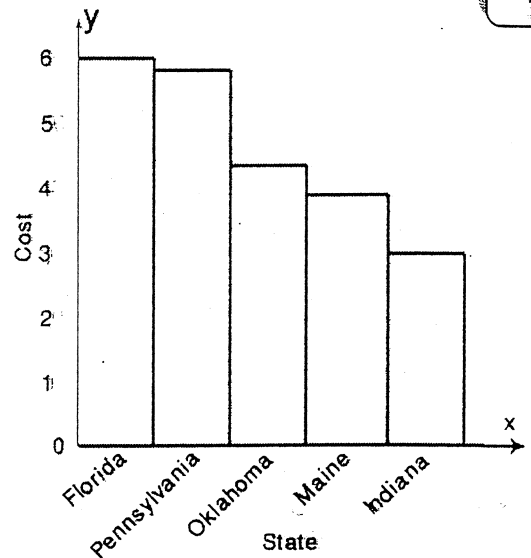
The table shown here is the average cost per mile for passenger vehicles on state turnpikes. Construct and analyze a Pareto chart for the data.

State	Number
Indiana	2.9c
Oklahoma	4.3c
Florida	6.0c
Maine	3.8c
Pennsylvania	5.8c

● Solution:

Arrange the data from the largest to smallest according to frequency.

State	Number
Florida	6.0c
Pennsylvania	5.8c
Oklahoma	4.3c
Maine	3.8c
Indiana	2.9c



The Pareto chart shows that Florida has the highest cost per mile: the cost is more than twice as high as the cost for Indiana.

السلاسل الزمنية

The time series Graph:

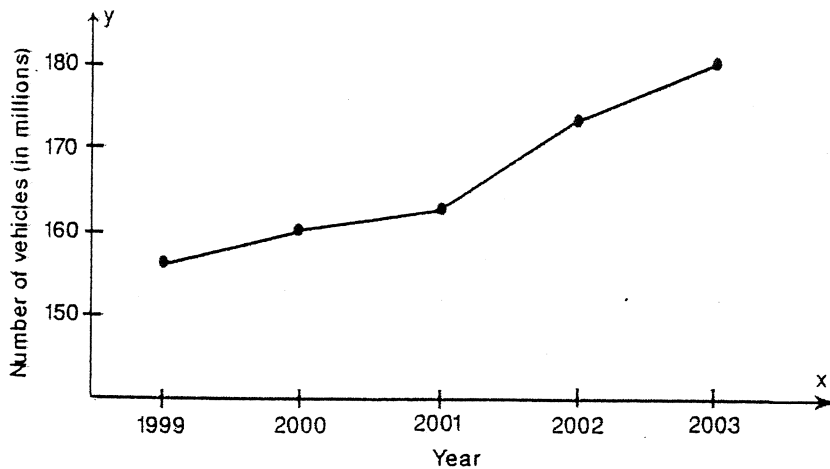
When data are collected over a period of time: they can be represented by a time series graph.

البيانات التي تحدث في فترة زمنية محددة
A time series graph; represents data occur over a specific period of time.

Example:

The number (in millions) of vehicles both ^{المسافر} passenger and ^{تجاري} commercial, that used the Pennsylvania ^{طريق رئيسي} Turnpike for the years 1999 through 2003 is shown. Construct and analyze a time series graph for the data.

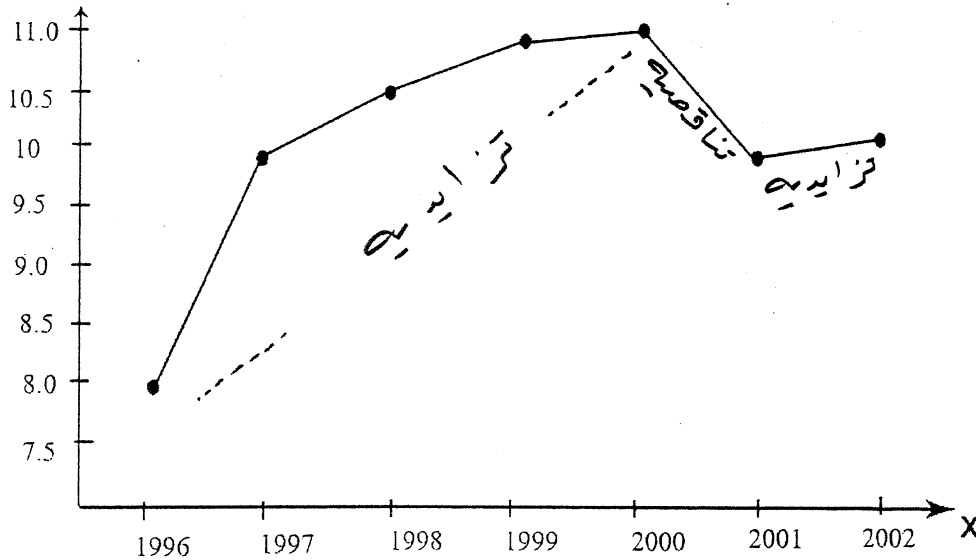
Year	Number
1999	156.2
2000	160.1
2001	162.3
2002	172.8
2003	179.4

A
L
S
A
A
D
I

Example:

Draw a time series graph to represent the data for the number of airline departure (in millions) for the given years. Over the years is the number of departure increasing, decreasing or about the same?

Year	1996	1997	1998	1999	2000	2001	2002
Number of departure	7.9	9.9	10.5	10.9	11.0	9.8	10.1



Departures increased until 2000, decreased from 2000 to 2001, and increased from 2001 to 2002.

The Pie Graph:

- تعريف pie graph
-missing part
- How many degrees كسر - نسبة

is a circle that is divided into sections according to the percentage of frequencies in each category of the distribution.

- * The variable is nominal or categorical.

Example:

Construct a pie graph showing the blood types of the army inductees described in the frequency distribution is repeated here.

Class	Frequency	Percent
A	5	20
B	7	28
O	9	36
AB	4	16
	25	100

Find the number of degrees for each class.

Solution

Using the formula

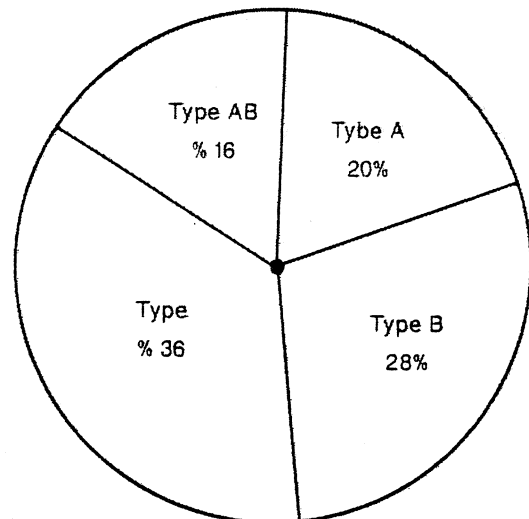
$$\text{Degrees} = \frac{F}{n} \times 360^\circ$$

$$A : \frac{5}{25} \times 360^\circ = 72^\circ$$

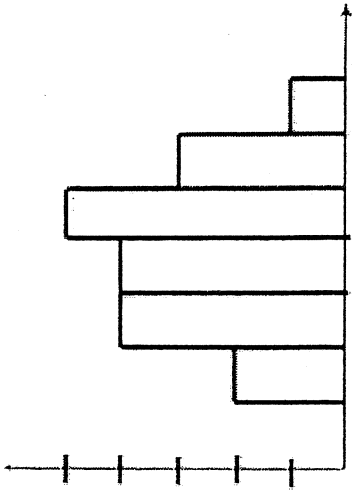
$$B : \frac{7}{25} \times 360^\circ = 100.8^\circ$$

$$O : \frac{9}{25} \times 360^\circ = 129.6^\circ$$

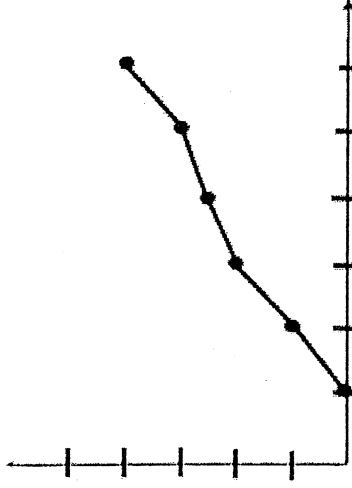
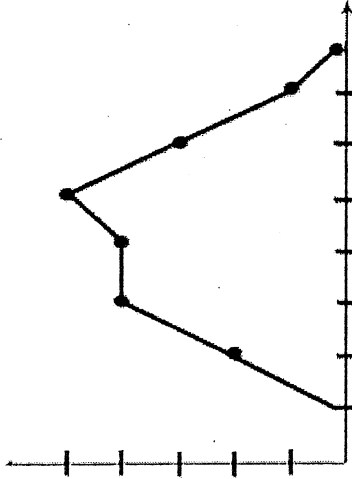
$$AB : \frac{4}{25} \times 360^\circ = 57.6^\circ$$



Summary of Graphs and used of Each

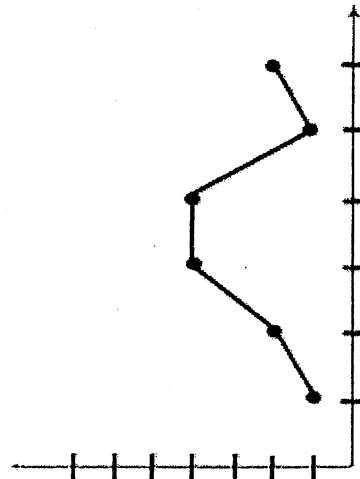


(a) Histogram; frequency polygon: ogive
Used when the data are contained in a grouped frequency distribution



(b) Pareto chart

Used to show frequencies for nominal or qualitative variables



(c) Time series graph

Used to show a pattern or trend that occurs over a period of time

(d) Pie graph

Used to show the relationship

between the parts and the whole
(Most often uses percentage)

مخطط الساق والورقة

Stem and leaf plot:

is a data plot that uses part of the data value as the stem and part of the data value as the leaf.

Example:

Construct a stem and leaf plot for the data.

25	31	20	32	13
14	43	02	57	23
36	32	33	32	44
32	52	44	51	45

Solution

- * Arrange the data from L to H --- رتب البيانات تصاعدياً

02	13	14	20	23	25	31	32	32	32
32	33	36	43	44	44	45	51	52	57

- * Separate the data according to the first digit from the left.

Stem	Leaf
0	2
1	3 4
2	0 3 5
3	1 2 2 2 2 3 6
4	3 4 4 5
5	1 2 7

Example:

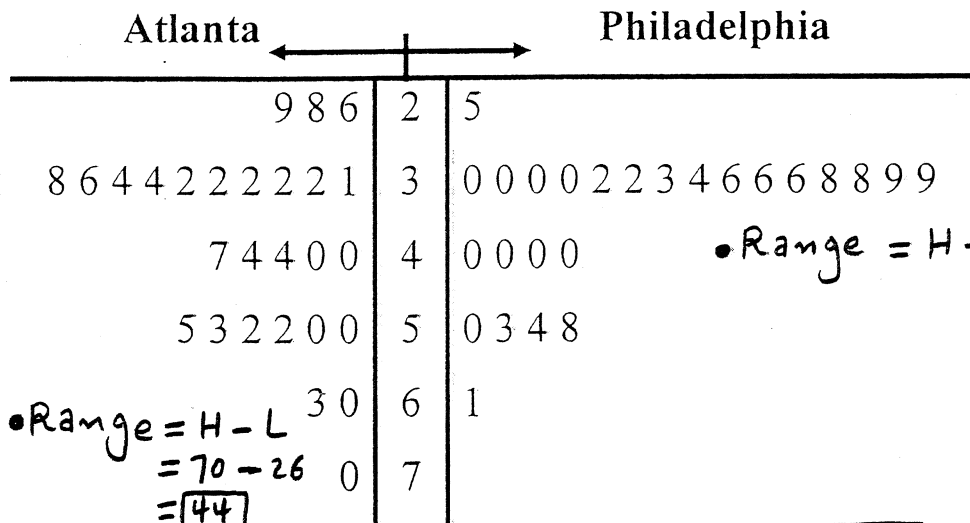
The number of stories in two selected samples of tall buildings in Atlanta and Philadelphia are shown.

Construct a back-to-back stem and leaf Plot, and compare the distributions.

Atlanta					Philadelphia				
55	70	44	36	40	61	40	38	32	30
63	40	44	34	38	58	40	40	25	30
60	47	52	32	32	54	40	36	30	30
50	53	32	28	31	53	39	36	34	33
52	32	34	32	50	50	38	36	39	32
26	29								

Solution

Arrange the data from L to H



• Range = H - L
 = 70 - 26
 = **44**

• Range = H - L = 61 - 25 = **36**

Compare the distributions. The buildings in Atlanta have a large variation in the number of stories per building.

A
L
S
A
A
D
I

Example:

If the stem and leaf plot of the data is

1	5
2	3 4
3	4 7

What is the raw data?

← ما هي البيانات الخام

Raw data is

15 , 23 , 24 , 34 , 37

On a Pareto chart, frequencies should be represented on the.....y-axis

The graph should be used to show the relationship between the parts and the whole is ..

* الرسم البياني الذي يوضح العلاقة بين الأجزاء والكل

- (a) Histogram (b) Pareto chart (c) pie graph (d) ogive

In a frequency distribution, the number of classes should be between: 5 , 20

* في التوزيع التكراري يتراوح عدد الفئات بين 5 و 20

- (a) 10 and 20 (b) 5 and 25 (c) 2 and 20 (d) 5 and 20

Another name for the Ogive is

المسمى الآخر لـ ogive

- (a) Histogram (b) freq. polygon (c) cumulative freq. graph

Find the boundaries for 8.6 - 8.8 ?

Lower boundary = 8.55

Upper boundary = 8.85

Chapter Quiz

Determine whether each statement is true or false. .

If the Statement is false , explain why.

- ① In the construction of a frequency distribution, it is a good idea to have overlapping class limits, such as 10-20, 20-30 , 30 – 40. (×)
-

- ② Histograms can be drawn by using vertical or horizontal bars. (×)
-

- ③ It is not important to keep the width of each class the same in a frequency distribution. (×)
-

- ④ Frequency distributions can aid the researcher in drawing charts and graphs. (✓)
-

- ⑤ The type of graph used to represent data is determined by the type of data collected and by the researcher's purpose. (✓)
-

6. In construction of a frequency polygon, the class limits are used for the x - axis. (x)
-

7. Data collected over a period of time can be graphed by using a pie graph. (x)
-

Select the best answer:

8. What is another name for the ogive?

- a. Histogram.
b. Frequency polygon
c. Cumulative frequency graph
d. Pareto chart
-

9. What are the boundaries for 8.6 - 8.87

- a. 8 - 9 b. 8.5 - 8.9 c. 8.55 - 8.85 d. 8.65 - 8.75
-

10. What graph should be used to show the relationship between the parts and the whole?

- a. Histogram b. Pie graph c. Pareto chart d. Ogive
-

11. Except for rounding errors , relative frequencies should add up to what sum?

- a. 0 b. 1 c. 50 d. 100

Complete these statements with the best answers.

- 12) The three types of frequency distributions are categorical, ungrouped and grouped.

- 13) In a frequency distribution, the number of classes should be between 5 and 20.

- 14) Data such as blood types (A, B, AB, O) can be organized into a (n) categorical. Frequency distribution.

- 15) Data collected over a period of time can be graphed using a (n) time series graph.

- 16) A statistical device used in exploratory data analysis that is a combination of a frequency distribution and a histogram is called a (n) Steam and leaf.

- 17) On a Pareto chart, the frequencies should be represented on the y- axis

- A pareto chart is useful for which of the following purposes?
 - A) Representing relative frequencies of categories in a specific year
 - B) Representing the cumulative frequencies of the data
 - C) Representing the frequencies of the data, sorted from largest to smallest**
 - D) Representing the frequencies of a data category over a period of several years

In pareto chart :

the data sorted from:

Largest to smallest.

- In a categorical frequency distribution, the number of observations in a class is called a(an) ...

A) interval. **B) frequency.** C) midpoint. D) category.

* من التوزيع التكراري النوعي
عدد المشاهدات من الفئة يساوي التكرار ← frequency

- The graph that should be used to show the relationship between the whole and the parts is called ...

A) time series graph. B) Pareto chart. C) frequency ploygon. **D) pie graph.**

* الرسم البياني الذي يوضح العلاقة بين الكل والأجزاء يسمى
→ Pie graph.

- In a frequency distribution, if the relative frequencies are 0.1, Y , 0.38 and 0.22, then the relative frequency Y is ...

A) 0.30 B) 0.60 C) 0.22 D) 0.10

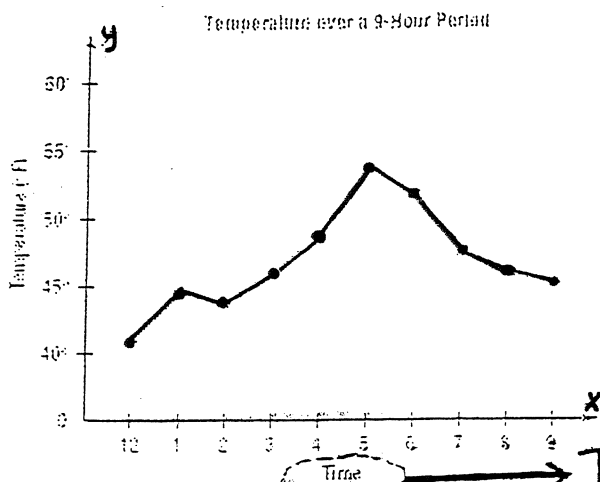
$$\Sigma \text{ relative frequencies} = 1$$

$$0.1 + y + 0.38 + 0.22 = 1$$

$$y + 0.70 = 1 \Rightarrow y = 1 - 0.70$$

$$y = \underline{\underline{0.30}}$$

- This graph is an example of



A) Ogive B) Time series graph C) Frequency Polygon D) Pie graph

- The data set that is collected over a period of time can be best represented by a (an) ...

A) time series graph B) histogram C) pie graph D) ogive

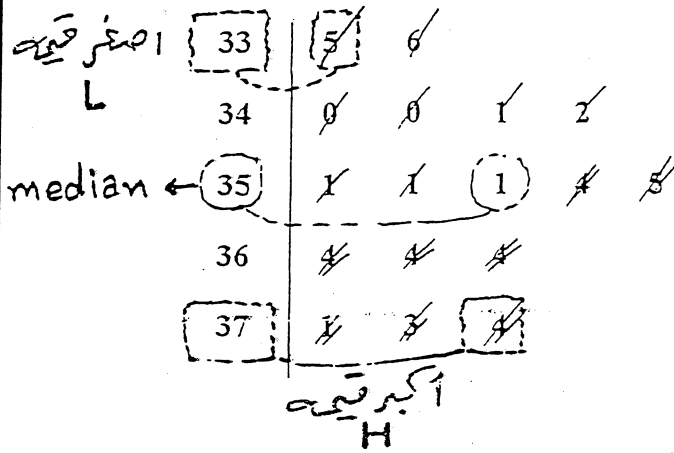
* البيانات التي تجمع خلال دوره زمني
تمثل بيانياً بالسلاسل الزمنية.

- The monthly incomes of eight computer operators are \$1,950, \$1,775, \$2,060, \$1,840, \$1,795, \$1,890, \$1,925 and \$1,810. What are these ungrouped values called?

A) Class limits. B) Class frequency. C) Raw data. D) Class boundaries

* الدخول الشهري لثمانى تمثل بيانات خام
Raw data

Use the following stem-leaf plot to answer the following three questions:



- The range value of the raw data set for the above stem and leaf plot is ...
A) 43 **B) 39** C) 44 D) 42

$$\text{Range} = H - L = 374 - 335 = \underline{\underline{39}}$$

- The raw data set for the above stem and leaf plot is called ...
A) bimodal. B) unimodal. C) trimodal. D) multimodal.

* العددان 364 و 351
كلتا منهما مكرر ثلاث مرات
∴ البيانات السابقة ثنائيه المنوال ← **bimodal**

- The median value of the raw data for the above stem and leaf plot is ...
A) 35 **B) 351** C) 51 D) 1

* الوسيط هو القيمة التي تتوسط البيانات

$$\therefore \text{median} = 351$$

What is the stem and leaf of 45?

(a) stem = 5 leaf = 4

* رقم الأعداد هو * باقى العدد هو *

(b) stem = 4 leaf = 5

What is the stem and leaf of 127?

(a) stem = 1 leaf = 27

(b) stem = 12 leaf = 7

(c) stem = 27 leaf = 1

(d) stem = 7 leaf = 12

The.....is a method of organizing data and is a combination of sorting and graphing?

(a) pareto chart

(b) pie graph

(c) stem and leaf plot

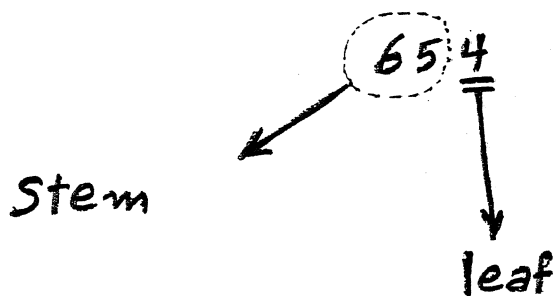
The stem part for the number 654 is ...

A) 6

B) 54

C) 4

D) 65



* رقم الأعداد هو * leaf 4
* باقى العدد هو * stem 65

In a pie graph, if the blood type A represents $\frac{9}{72}$ of the distribution, how many degrees would be needed to represent A?

A) 8°

B) 9°

C) 72°

D) 45°

$$** \text{ عدد الدرجات } = \text{ الكسر } \times 360^\circ$$

$$\text{number of degrees} = \frac{9}{72} \times 360^\circ = 45^\circ$$

A (An) ... is used to represent a data set that is collected over a period of time.

A) Pareto chart

(B) time series graph

C) ogive

D) pie graph

period of time \rightarrow time series graph

تعريف

Use the following to answer questions

The following table shows the distribution of the blood type for 70 students:

Classes	A	B	O	AB
Frequency	15	20	17	18

- The type of the frequency distribution is ... distribution.
 A) grouped frequency B) probability C) categorical frequency D) ungrouped frequency

The type is categorical frequency

الفئات classes ليست رقمية
 أي أنه الفئات عبارة عن كلمات أو رموز
 ∴ نوع التوزيع التكراري : نوعيه (وصفيه)

- The probability of selecting a student with AB blood type is ...
 A) 0.154 B) 0.348 C) 0.257 D) 0.27

$$P(AB) = \frac{\text{عدد الطلاب من نوع فصيلة الدم AB}}{\text{العدد الكلي للطلاب}} = \frac{18}{70} = 0.257$$

- The range value ...
 A) cannot be calculated B) is 3. C) is 5. D) is A - AB.

The range ; can not be calculated
 * لا يمكن حساب المدى لأنه البيانات نوعيه (وصفيه)

- In a pie graph, how many degrees would be needed to represent A?
 A) 72° B) 99.69° C) 77.14° D) 103.64°

The number of degrees
 for class A

$$= \frac{15}{70} \times 360 = 77.14^\circ$$

$$\frac{\text{تكرار الفئه A}}{\text{مجموع التكرار}} \times 360^\circ$$

Use the following table to answer the following five questions:

Ages	Number of Students
20.2 – 29.3	16
29.4 – 38.5	25
38.6 – 47.7	51
47.8 – 56.9	79
57.0 – 66.1	21
66.2 – 75.3	8

● What is the age midpoint for the third class?

- A) 33.95 B) 52.35 C) 61.55 **D) 43.15**

* منتصف الفئة الثالثة
جمع طرفي الفئة
2

$$\text{midpoint for third class} = \frac{38.6 + 47.7}{2} = 43.15$$

● What is the class boundary for the fourth class limit?

- A) 47.30 -- 57.40 B) 48.30 -- 56.40 **C) 47.75 -- 56.95** D) 47.85 -- 56.85

$$\begin{aligned} \text{class limit} &= 47.8 - 56.9 \\ \text{class boundary} &= 47.75 - 56.95 \end{aligned}$$

الفئة الرابعة

الفاصله فوزونه (بعد رقم واحد في class limit)

* نضع الرقم 5 من نهايه upper limit

* نضع الرقم 5 من الرقم الاخير من lower limit ثم نضع الرقم 5.

- أسم الجدول
- What is the name of the table?
 - A) **Grouped frequency distribution.**
 - B) Cumulative frequency distribution.
 - C) Categorical frequency distribution.
 - D) Ungrouped frequency distribution.
 - What is the relative frequency of students whose ages within the second class?
 - A) 25.5%
 - B) 0.255
 - C) 12.5%
 - D) **0.125**

relative frequency
التكرار النسبي للفئة الثانية

$$= \frac{\text{تكرار الفئة}}{\text{مجموع التكرارات}} = \frac{25}{200} = 0.125$$

- The data can be represented the best using ...
 - A) pie graph.
 - B) Ogive.
 - C) **histogram**
 - D) bar chart.

In a frequency distribution, if the relative frequencies are 0.20, 0.28, X and 0.16, then the relative frequency X is

- A) 0.63 B) 0.36 C) 0.46 D) 0.64

* مجموع $P(x)$ لابد أنه يكون 1

$$\sum P(x) = 1$$

$$0.20 + 0.28 + X + 0.16 = 1$$

$$X = 1 - 0.20 - 0.28 - 0.16$$

$$X = 0.36$$

The following table represents the favorite car make for a group of students

Classes	Frequency
Toyota	10
Nissan	11
Chevrolet	9
GMC	13
Honda	14

الفئة الثالثة

- The cumulative frequency for the third class is ...

A) 30 B) 48 C) 13 D) 43

* التكرار المتجمع المناظر للفئة الثالثة :

$$= 10 + 11 + 9 = \underline{\underline{30}}$$

- The sample size is ...

A) 5 B) 60 C) 57 D) 18

sample size: $n = \sum f = 10 + 11 + 9 + 13 + 14$
 $= \underline{\underline{57}}$

- The percentage of students who like Chevrolet is ...

A) 19.23% B) 83.33% C) 25% D) 15.79%

$$P = \frac{9}{57} \times 100\% = 15.79\%$$

- If a pie graph is used to represent the data, the degree for the Honda brand would be ...

A) 55.38 B) 15.38 C) 0.15 D) 88.42

$$\text{Number of degree} = \frac{14}{57} \times 360^\circ = \underline{\underline{88.42}}$$

- The type of data is ...

A) nominal. B) ordinal. C) discrete. D) continuous.

* نوع البيانات من الجدول أعلاه
وهو عبارة عن ماركات السيارات.

- The most appropriate measure of central tendency for this data is the ...

A) mean B) mode C) midrange D) range

* المقياس المناسب
من مقاييس النزعة المركزية لهذه البيانات الأسمية هو المنوال

- The continuous data can be organized into a table that is called ...
- A) grouped frequency distribution. C) categorical frequency distribution.
B) ungrouped frequency distribution. D) ordinal frequency distribution.

- The class width for the class limit 4.1 - 9 is ... $\rightarrow 4.1 - 9.0$ *وزنه الفاصله
- A) 6.55 B) 5.9 C) 4.9 D) 5

class boundaries : $4.05 - 9.05$

$$\text{class width} = 9.05 - 4.05 = 5$$

- What is the most appropriate measure of central tendency for the following data set?
Ali, Saeed, Ahmed, Saeed, Ali, Ali
- A) Mean. B) Mode. C) Median. D) Midrange.

* المقاييس المناسبة من مقاييس النزعة المركزية
للبيانات الأسمية هو الموال \leftarrow Mode

- The most appropriate graph for categorical data is ...
- A) pie graph. B) histogram. C) time series graph. D) stem and leaf plot.

* التمثيل البياني المناسب للبيانات النوعية (الوصفية)
هو القطاعات الدائرية \leftarrow pie graph

- The nominal data can be organized into a table that is called ...
- A) ungrouped frequency distribution. C) grouped frequency distribution.
B) nominal frequency distribution. D) categorical frequency distribution.

* البيانات الأسمية تنظم بجدول يسمى التوزيع التكراري النوعي

The number of patients in the waiting rooms within a hospital at a specific time are given by the following frequency distribution.

Number of patient	1	2	3	4	5	Total
Frequency	4	5	?	8	4	30

Answer the following five questions

- The percentage of the number waiting rooms that have 2 patients is ...
 A) 16.67% B) 12.50% C) 36.67% D) 83.33%

$$P = \frac{5}{30} \times 100\% = 16.67\%$$

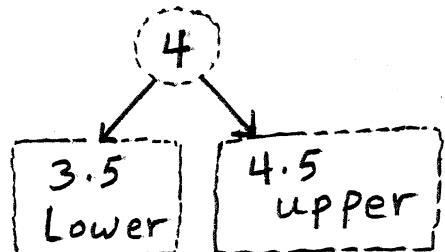
- The data can be best represented graphically by ...
 A) bar chart B) frequency polygon C) time series graph D) histogram

* الفئات عبارة عن ارقام منفصلة 1, 2, 3, 4, 5
 ∴ أفضل طريقة لتمثيل البيانات بيانياً
 هي طريقة الأعمدة ← bar chart

- The missing frequency for the third class is ...
 A) 7 B) 6 C) 3 D) 9

$$\text{The missing } f. = 30 - (4 + 5 + 8 + 4) = 9$$

- The lower class boundary for the fourth class is ...
 A) 5.5 B) 4.5 C) 2.5 D) 3.5



- The sample size is ...
 A) 30 B) 5 C) 20 D) 32

sample size

$$= \sum f = 30$$

* حجم العينة هو مجموع التكرارات.

تعريفات مهمة

- The pie graph: is a circle that is divided into sections according to the percentage of frequencies in each category of the distribution.

- When data are collected in the original form, they called raw data.
بيانات خام

- Ungrouped frequency distributions: is used for discrete data.

- Grouped frequency distribution: is used for continuous data.

- Histogram, frequency polygon or ogive graphs are used to represent continuous data graphically.

- Bar graph: are used mostly to represent discrete and ordinal data graphically.
* طريقه الأعمدة :
تستخدم لتمثيل البيانات المنفصلة والترتيبيه بيانياً .

- Pie graph and Pareto chart are used mostly to represent nominal data graphically.
* طريقه ال : باى ، باريتو
تستخدم لتمثيل البيانات الأسميه بيانياً .

- Stem and leave plot: is a combination of sorting and graphing. It retains the actual data while showing them graphically.

- The histogram displays the continuous data that are organized in a grouped frequency distribution by using vertical bars of various heights to represent the frequencies.

- The frequency polygon displays the continuous data that are organized in a grouped frequency distribution by using lines that connect points plotted for the frequencies at the midpoints of the classes.

← خطوط مستقيمة تصل بين النقاط
التي تمثل التكرارات المناظرة لمتصفات الفئات.

- The cumulative frequency graph or ogive represents the cumulative frequencies for the classes in a grouped frequency distribution.

- The bar charts displays the data by using vertical bars of various heights to represent the frequencies of discrete or categorical variables.

- Categorical frequency distributions are used for data that can be placed in specific categories, such as nominal or ordinal level data.

Given the following distribution:

Class boundaries Ages	Frequencies number of students
13.5 – 18.5	4
18.5 – 23.5	9
23.5 – 28.5	12
28.5 – 33.5	15
33.5 – 38.9	17

1. Number of students where age is less than 23.5 is = $4 + 9$

- (a) 4 (b) 9 (c) 13 (d) 5

2. Number of students where age is less than 33.5 is: = $4 + 9 + 12 + 15$

- (a) 15 (b) 57 (c) 40 (d) 25

3. Except for rounding errors, relative frequencies should add up to what sum?

- (a) 0 (b) 1 (c) 50 (d) 100

* باستبعاد اخطاء التقريب
يكون مجموع التكرار النسبي = 1

4. If class limits 23.4 – 28.4 the class width is.. $28.45 - 23.35 = \frac{5.1}{1}$

- (a) 5 (b) 2.5 (c) 5.1 (d) 6

5. In a pie graph if the blood type O was 36% of the distribution.

How many degrees would be needed to represent type O?

- (a) 129° (b) 29.6° (c) 129.6° (d) 360°

Solution

$$\begin{aligned} \text{Degrees} &= \frac{F}{n} \times 360^\circ \\ &= \frac{36}{100} \times 360^\circ \\ &= 129.6^\circ \end{aligned}$$

وأخيراً
أدعو الله أن يتقبل هذا العمل
ويكون فيه النفع للجميع....
جمال السعدي