الفصل الثاني

عرض البيانات الإحصائية

2-1Definition of data

Data is defined as a set of observations collected during the study of a phenomenon. Data can be:

1. **Quantitative data (numerical data):** such as the heights and weights of a group of students.
2. **Qualitative data (descriptive data):** such as skin color and gender.

There are two main types of data:

1. **Qualitative Data (Descriptive Data):** These are data that describe individuals or the community, such as hair color, eye color, skin color, or students' grades in a subject. Qualitative data can be further divided into:
    - **Nominal Data:** Data that are categorized without any order or ranking, such as cities or marital status.
    - **Ordinal Data:** Data that have a meaningful order or ranking, such as students' grades or military ranks.

2. **Quantitative Data (Numerical Data):** These are data that measure individuals or the community using numerical scales, such as students' heights measured in centimeters, students' weights measured in kilograms,. Quantitative data can be further divided into:
    - **Discrete Data:** Quantitative data that take specific whole number values and do not include fractions, such as the number of students who passed.
    - **Continuous Data:** Quantitative data that can take any value, including fractions and decimals, such as height.

2-2Data Presentation

statistical data can be presented in two main ways:

1. **Tabular Presentation:** Data are presented using frequency tables.
2. **Graphical Presentation:** Data are presented using graphs.

2-2-1Tabular Presentation of Qualitative Data

To create frequency tables for qualitative data, we first sort the attributes, then count the frequency of each attribute using tally marks, and finally calculate the frequency for each attribute. The above method can be illustrated using the following example:

Example 1:

The following data represent the social status of ten individuals, as shown in the following table-:

Single	Divorced	Single	Married	Divorced
Divorced	Widowed	Widowed	Married	Married

Task: Construct the frequency table for these data.

Category	Tally	Frequency
Single	//	2
Divorced	///	3
Married	///	3
Widowed	//	2
Total	10	10

Example 2:

Suppose you have the following data representing the grades of 35 students

Very Good	Acceptable	Excellent	Good	Excellent	Acceptable	Fail
Excellent	Good	Average	Excellent	Acceptable	Good	Excellent
Excellent	Very Good	Good	Fail	Very Good	Good	Acceptable
Good	Average	Very Good	Good	Very Good	Fail	Acceptable
Very Good	Very Good	Acceptable	Excellent	Very Good	Acceptable	Fail

Task: Construct the frequency table for these data(Homework)

Tabular Presentation of Quantitative Data

To construct a frequency table for quantitative

1. **Arrange the data:** Sort the data either in ascending order (from smallest to largest) or descending order (from largest to smallest), depending on the student's choice.

2. **Calculate the range of the data:** It is the difference between the largest and smallest values, according to the following formula:

$$\text{Range} = \text{Maximum Value} - \text{Minimum Value}$$

3. **Determine the number of classes:** The number of classes should be appropriate, calculated using the proper formula.

$$M=1+3.322\log(n)$$

4. **Calculate the class width (**Le=R/M): The class width must be calculated correctly.

5 **Select the starting point of the first class:** This usually represents the smallest value in the data or a slightly smaller value to facilitate calculations. Then, each class is determined by adding the class width to the beginning of the previous class. The end of each class is defined as follows:

1. **For discrete data:** The end of the class = beginning of the class + class width − 1
2. **For continuous data:** There is no need to write the end of the class explicitly, as the end of each class is the beginning of the next class.

6 **Construct the table:** Place tally marks to represent frequencies. The following example will illustrate all the points mentioned above.

Example 3:

The following table represents the final grades of 30 students in the Statistics course

38	69	61	44	65	42	57	67	93	51
76	39	83	74	67	84	24	50	54	50
52	92	52	66	36	70	52	87	54	85

Task: Construct the frequency table for these data

1/The table was arranged in ascending order.

24	36	38	39	42	44	50	50	51	52
52	52	54	54	57	61	65	66	67	67
69	70	74	76	83	84	85	87	92	93

2/ Calculating the Range

R = Max – Min

R = 93-24=69

3/ Calculating the number of classes

m = 1 + 3.322 log (n)

1+ 3.322 (30)= 5.907= 6

4/ Calculating the class width

Le=R/M , Le=69/6=111.5 =12

rounding Note: When rounding a number, if the digit to be rounded is 5 or greater, increase the preceding digit by 1 (for example, 12 rounds up to 13). If the digit is 4 or less, leave the number unchanged (for example, 11 remains 11).

5/ we notice that the smallest value is 24, which means that the first class starts at 24. The remaining classes go up to 84. Since the data are discrete, the class limits are as follows: 24, 35, 47, 59, 71, 83, 95, 72, 65, 48. Their class width is 36

Classes	Tally	Frequency
24-35	/	1
36-47	/////	5
48-59	///// ////	9
60-71	///// //	7
72-83	///	3
84-95	/////	5
Total		30

Example 4:

The following table represents the final grades of 16 students in the Statistics course

19	15.6	22	10	13	10.2
20	14	11	12.5	18	13.8

Task: Construct the frequency table for these data(Homework)

Translation of data charts/graphs

Charts include two types:

- charts for ungrouped data
- charts for grouped data

The graphical representations of ungrouped data include:

1. Pictorial diagrams.
2. Graphical charts, which include bar charts, pie charts, and line charts

The graphical representations of grouped data (tables) include:

1. Histogram
2. Frequency polygon
3. Frequency curve
4. Cumulative curves

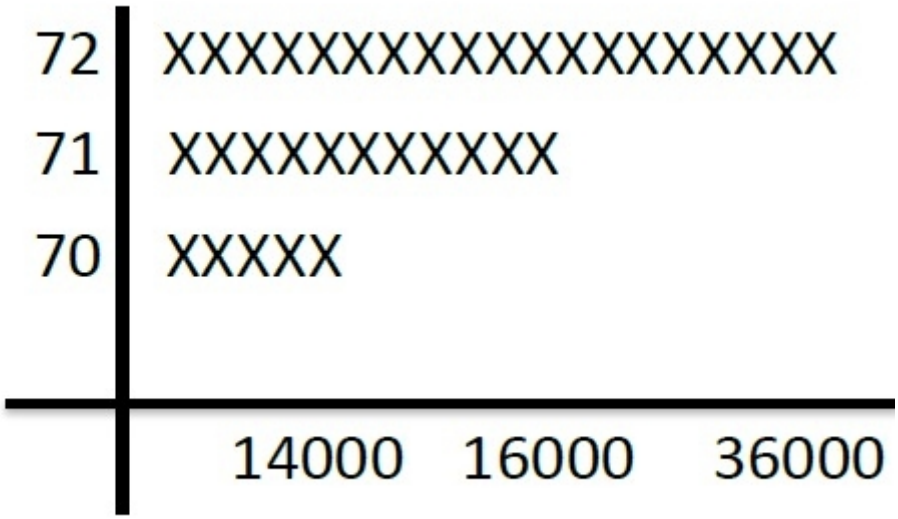
2-2-2Graphical presentations include:

1. **Pictorial charts**: Pictorial charts are very important for conveying information clearly to non-scientific and non-specialized audiences. In this method, data are represented using images or symbols that help the viewer easily understand the meaning of the data.

Example: Suppose that the number of children enrolled in kindergartens in Iraq over three years is given in the following table.

Year	Number
1970	14000
1971	36000
1972	16000

Required: Present these data using the pictorial chart method



2. **simple bar charts:**

These are rectangles with equal base widths but varying heights depending on the values associated with each bar. In this method, two axes are drawn.

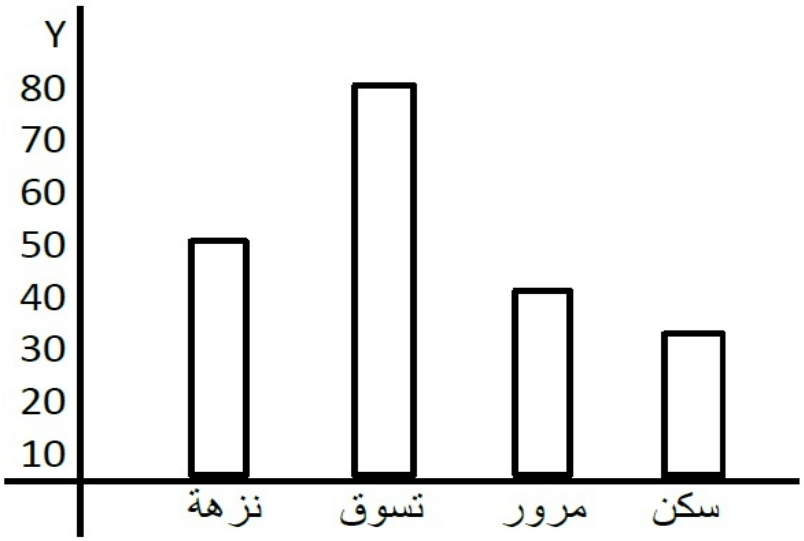
NOT

the horizontal axis (x-axis) represents the variable, and the vertical axis (y-axis) represents the corresponding value of that variable.

Example: A researcher studied a commercial center and selected a sample of 211 individuals who visited this center during a specific period.
The findings are shown in the following table.

Number of classes	Reasons for the increase
50	Picnic
80	*Shopping*
40	Traffic
30	Housing
200	Total

Required: Present these data using the simple bar chart method



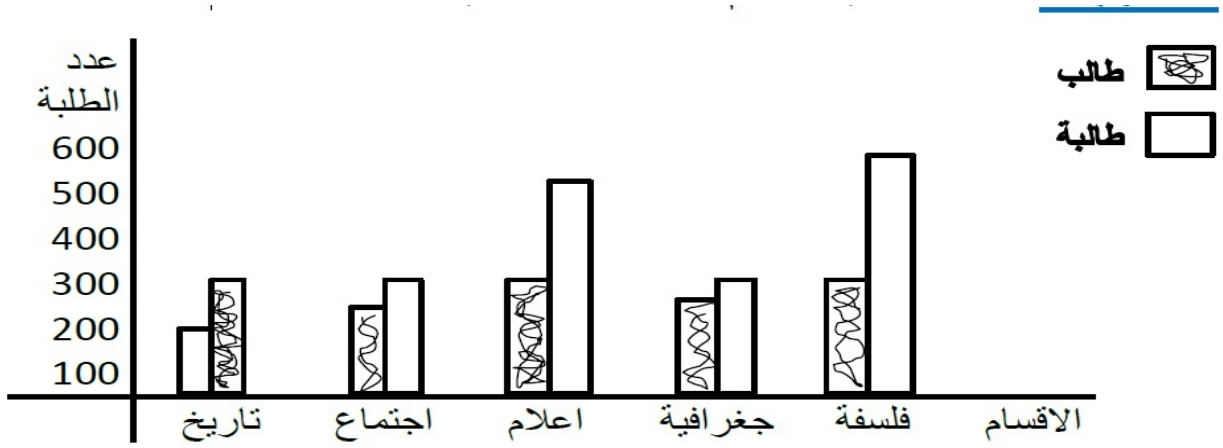
3 Adjacent Vertical Columns:

This method is similar to the simple vertical columns method, but a number of adjacent columns are drawn, each representing one of the variable values.

Example: The following table represents the number of students in the departments of the College of Arts, University of Mosul.

Department	History	Sociology	Media	Geography	Philosophy
student	300	250	300	250	300
Female student	200	300	500	300	600
Total	500	550	800	550	900

Display/present this data using the method of **adjacent vertical columns**.

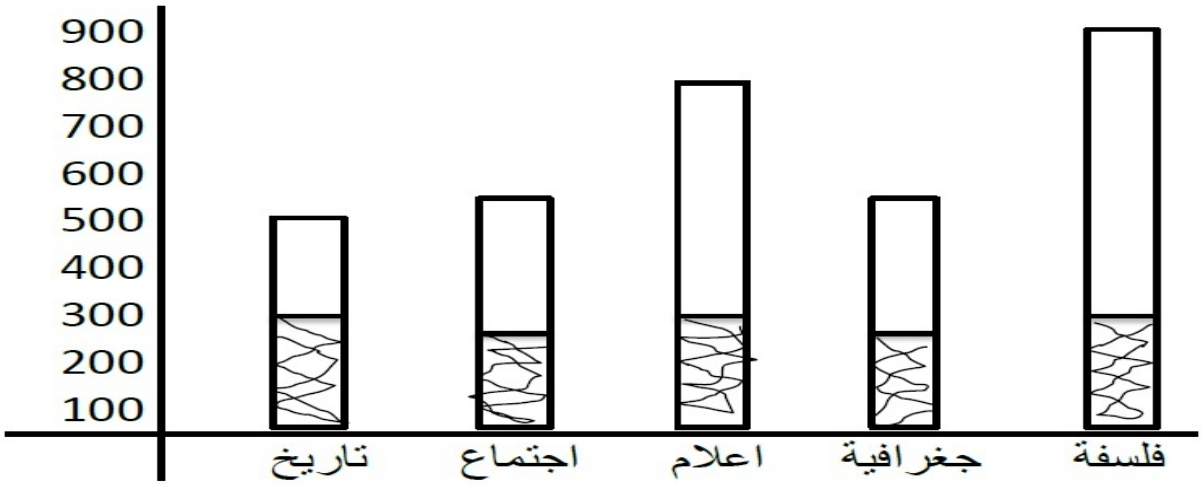


4. Stacked Vertical Columns:

This method is similar to the simple vertical columns method, but a column representing the first value of the variable is drawn, and then columns for the remaining variable values are placed on top of it. The end of the first column becomes the beginning of the second column.

Using the data from the previous example, the data can be represented using the **stacked columns method**.

Example: Based on the previous example's data, these data will be represented using the **stacked vertical columns method**.



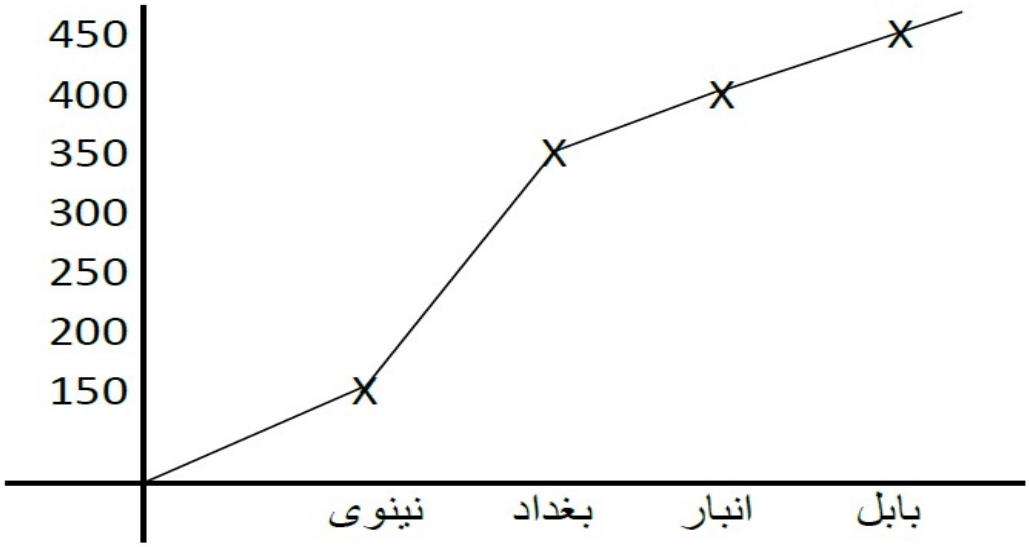
5.Broken Line (Line Graph):

In this method, the **x-axis** represents the variable (years), while the **y-axis** represents the value of the variable. Points are plotted for each value on the x and y axes, and then these points are connected with a **broken line**.

Example: The following table shows the production per dunum (unit area) of wheat in kilograms for all the governorates of the country

Governorate	Kg
Nineveh	150
Baghdad	350
Anbar	400
Babil	450

Display these data using a **line graph** (broken line graph)



6.Circle (Pie Chart):

It is a regular circle in which statistical data are displayed. The sum of the individual values is represented with in the area of the circle, and no part of the circle should remain empty. The angle of each sector is calculated using the following formula:

Angle of the sector (for the part) =(Actual frequency of the part/ Sum of frequencies360)

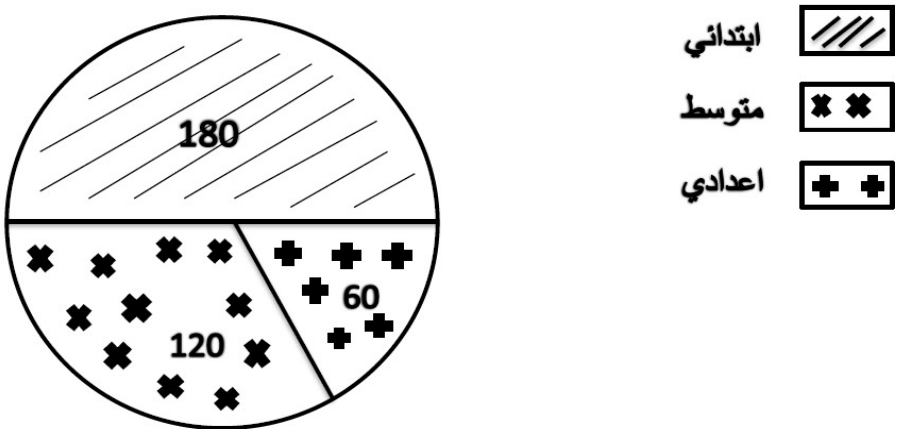
Example: Display the following data, which represent the number of students in a certain area, using pie charts

Stage	Number
Primary	9000
Intermediate	6000
Preparatory	3000
Total	18000

$$p = \frac{9000}{18000} \times 360 = 180$$

$$I = \frac{6000}{18000} \times 360 = 120$$

$$Pr = \frac{3000}{18000} \times 360 = 60$$



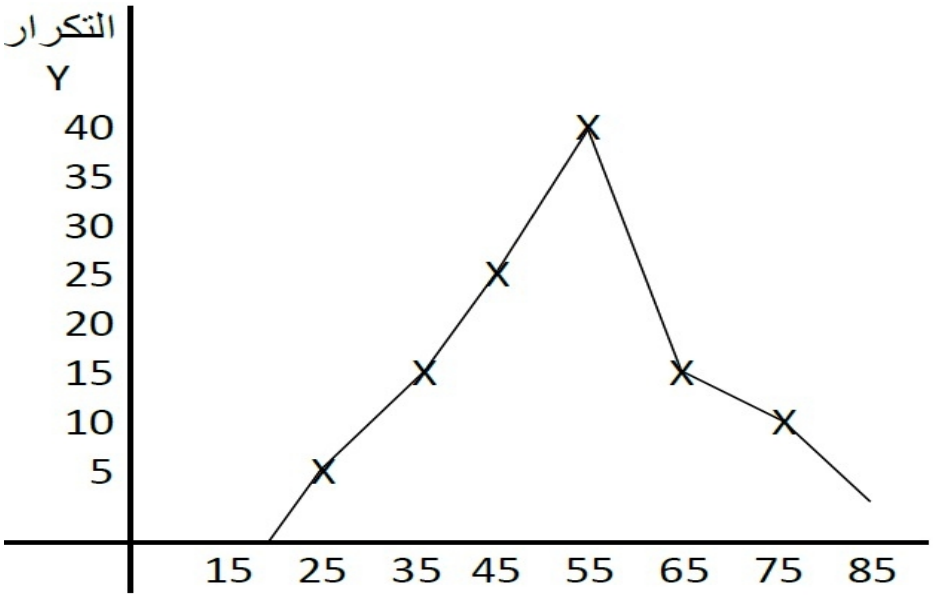
2-3The form of grouped data:

It is a representation of the simple frequency table, where the frequencies are plotted on the vertical axis and the class midpoints on the horizontal axis. The points are then connected with straight, broken lines, and afterwards the ends of the polygon are connected to the horizontal axis.

Example:

Given the following frequency distribution of the yield of one type of dates, draw the frequency polygon

Class midpoint	frequency	Class
25	5	-20
30	15	-30
45	25	-40
55	40	-50
65	15	-60
75	10	70-80



Class midpoint = (Highest value + Lowest value) / 2

$$20-30=\frac{20+30}{2}=25$$

$$30-40=\frac{30+40}{2}=35$$

$$40-50=\frac{40+50}{2}=45$$

$$50-60=\frac{50+60}{2}=55$$

$$60-70=\frac{60+70}{2}=65$$

$$70-80=\frac{70+80}{2}=75$$