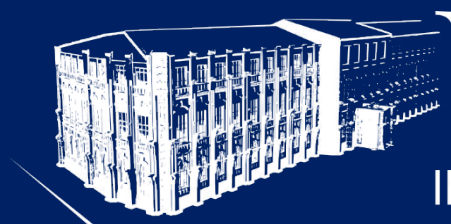


WESTMINSTER

INTERNATIONAL UNIVERSITY IN TASHKENT

An Accredited Institution of the University of Westminster (UK)

# Module Aims:



- To foster in students confidence to cope with the processing and analyzing of quantitative information.

- To provide an appreciation of numerical and statistical concepts relevant to the business environment.

# Learning outcomes:



- apply numerical skills to business and/or engineering problems
- present statistical data in a variety of formats, including electronic means
- apply basic rules of algebra and calculus
- using spreadsheets summarize numerical data into averages and deviations and apply them to a variety of business problems.

# In brief, you will learn how ...:



- To appreciate benefit of numerical data for businesses
- To make decisions based on the numerical data
- To interpret and represent numerical data in a most appropriate way depending on your aims
- To solve statistics and calculus problems using various quantitative methods
- Note: You can find out more about module content in module syllabus and 12-week teaching schedule.

# Teaching methods:



- 1-hour online lecture each week (online)
- 2-hour tutorial each week (offline)
- 1-hour workshop each week (offline)

You will learn the theory and its application

# Assessment methods:



**Two assessments (or components):**

- In-class test (30%+10%).
  - 30% goes to an in-class test in Teaching Week 6
  - 10% goes to weekly online mini-quizzes
- Final exam (60%) in Final exam week
  - True/false
  - Theory description
  - Problem solving
  - Open ended questions
  - Multiple choice

# LECTURE 1
# DATA & DATA REPRESENTATION

**Temur Makhkamov**
**Indira Khadjieva**
**QM Module Leaders**
tmakhkamov@wiut.uz
i.khadjieva@wiut.uz

**Office hours: by appointment**
**Room IB 205**
**EXT: 546**

# Lecture outline



## DATA

- the meaning and types of data
- sources of data
- the scales of measurements for data

## DATA REPRESENTATION TECHNIQUES AND TOOLS

- analyze the quantitative and qualitative data;
- display data in the form of table;
- display data in the form of graph.

# What is data? (1)



- **Data** –
- the facts and figures that are collected, analyzed and summarized.

*Examples:* data about people, countries, employees
nature, universities, number of products sold, costs, prices,
movies, cars, hospitals, registration numbers, tax codes etc

# What is data? (2)



- Data may be obtained through already existing-sources or through statistical studies.

  1. *already existing-source:*

  Salaries, sales, advertising costs, inventory levels can be disclosed from a company,

  2. *from a statistical study:*

  an experiment, a questionnaire, a survey, etc

# Primary and Secondary data



- **Primary data** – the data that are obtained as a result of conducting a questionnaire, a survey, an interview, an observation, etc.

  **Examples:**_____

- **Secondary data** – the data that come from existing sources. Government institutions, healthcare facilities, Internet and others can provide a great deal of information in a ready-to-estimate format.

  **Examples:**_____

# Questions:



What data is more costly (expensive):

primary or secondary?

What data is more reliable (trustworthy):

primary or secondary?

# Statistical data



**Table 1** A data set for 8 US Colleges & Universities

| College/ University | Enroll-ment | Type | Control | Tuition($) | | Room & Board ($) |
|---|---|---|---|---|---|---|
| | | | | Resident | Nonres-t | |
| Univ. of Alabama | 23,838 | Coed | Public | 5,278 | 15,294 | 5,380 |
| Westminster College | 953 | Coed | Private | 15,030 | 15,030 | 6,140 |
| Univ. of Colorado | 31,399 | Coed | Public | 5,643 | 23,539 | 8,300 |
| Columbia College | 1,446 | Women | Private | 20,302 | 20,302 | 6,022 |
| Texas A&M Univ. | 45,380 | Coed | Public | 6,966 | 15,216 | 7,660 |
| St. John's Univ. | 2,044 | Men | Private | 24,924 | 24,924 | 6,496 |
| Yale Univ. | 11,416 | Coed | Private | 33,030 | 33,030 | 10,020 |
| Meredith College | 2,139 | Women | Private | 21,200 | 21,200 | 5,940 |

*Source*: "Accredited US Senior Colleges and Universities" *The 2007 Information Please Almanac*. http://www.infoplease.com/edu/colleges

Q: What are the components of the statistical table?

# Components of the tabular data



- **Element** – the entity or item on which data are collected.

  *Examples:* Westminster College, Yale Univ., etc

- **Variable** – a characteristic of interest for an element.

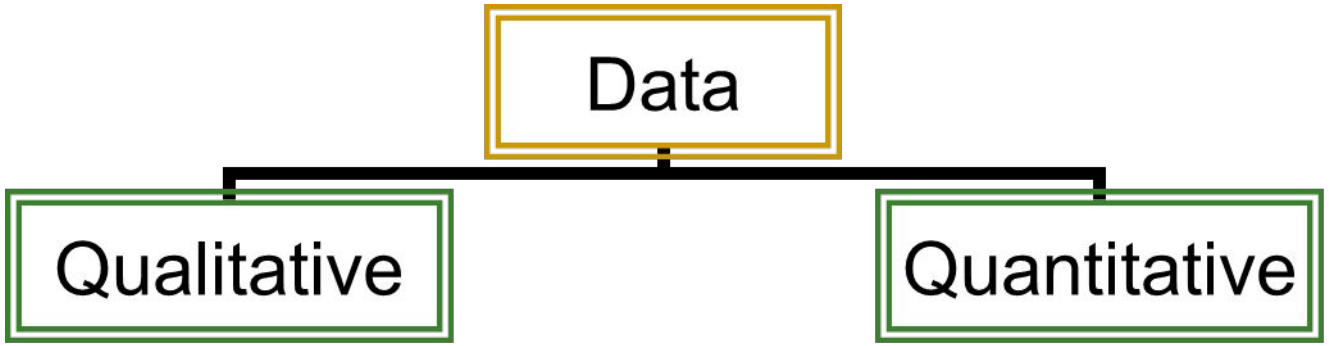
  *Examples:* Enrollment, type, etc

- **Observation** – a set of measurements collected for a particular element.

  *Examples:* 953, coed, public, $6,140, etc

# Main types of data






- **Qualitative data** provide labels or names for variables. They can be nonnumeric descriptions or numeric codes.

  *Examples:* Coed, Public, etc

- **Quantitative data** show an amount of variables. They indicate either “how much” or “how many” of something.

  *Examples:* 953 students, $6,140 for Room & Boarding, etc

# Question:



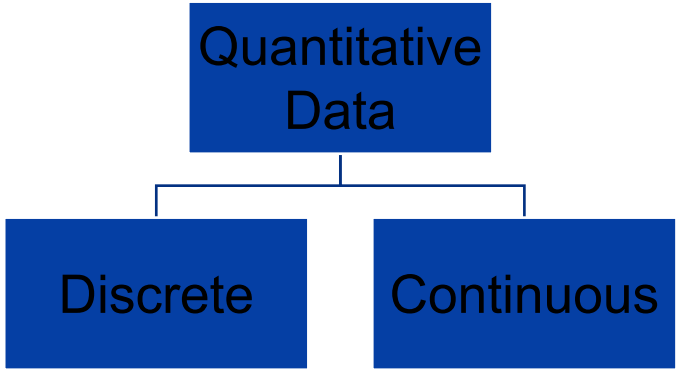
- Consider **this room** as an element.

Are its variables such as,

Names of students quantitative or qualitative?
Mode of students quantitative or qualitative?
Number of students quantitative or qualitative?

# Quantitative Data






- **Discrete data –** the data obtained as a result of counting.

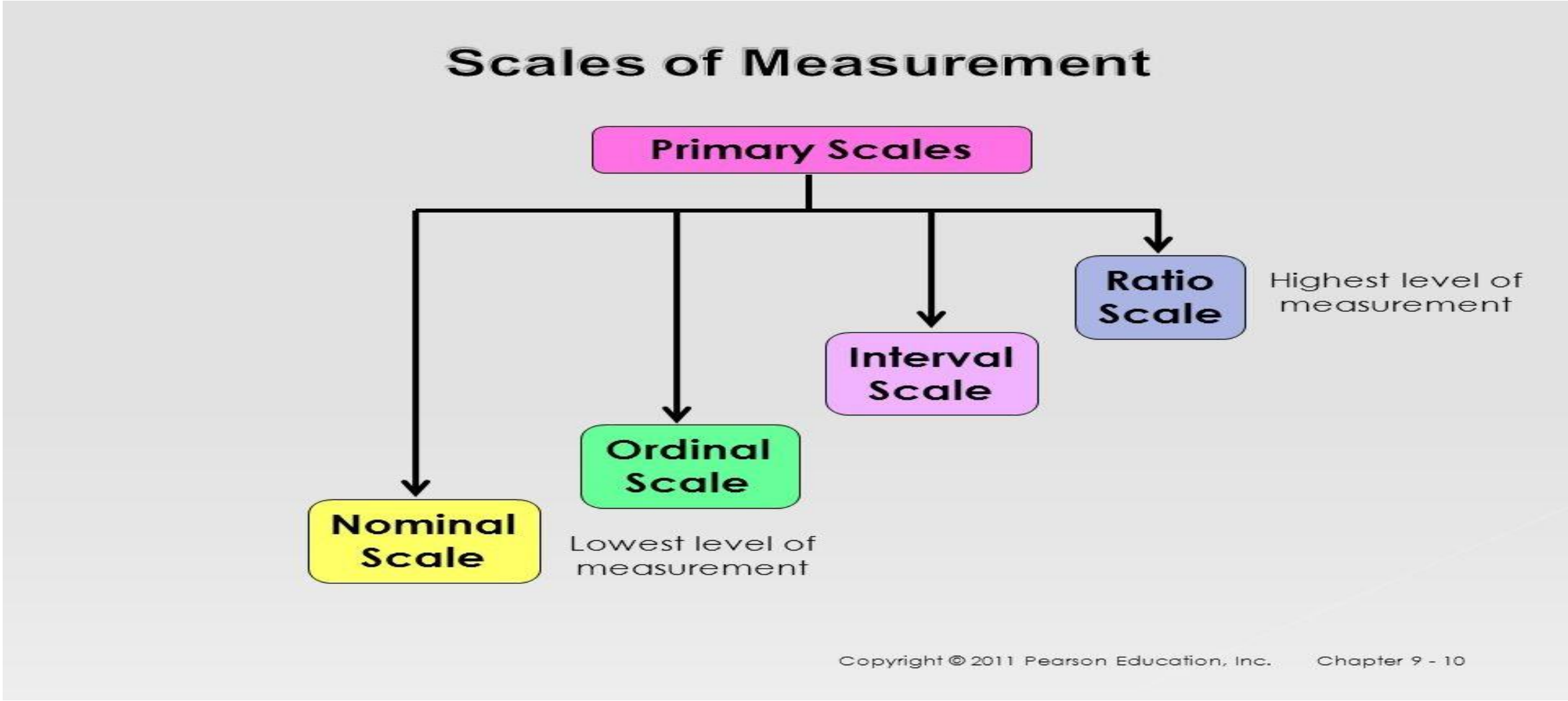
  *Examples: Number of enrolled students: 500, 1000, 2458, etc.*

- **Continuous data –** the data that can take any value within a continuum, limited only by the precision of the measurement instrument.

  *Examples: Length or height of some object: 5 cm, 5.35 cm,*

# Scale of Measurement





# SM for Qualitative Data (1)



- **Nominal scale** – a scale of measurement that uses name or label to define a characteristic of an element.


# NOMINAL

****Nominal sounds like name****

## Notes

- Lowest level of measurement
- Discrete categories
- No natural order
- Categorical or dichotomous
- May be referred to a qualitative or categorical

## Examples

- Gender
  - 0 = Female
  - 1 = Male

Dichotomous

- Group Membership
  - 1 = Experimental
  - 2 = Placebo
  - 3 = Routine

Categorical

- Marital status, colour, religion, type of car …

# SM for Qualitative Data (2)



- **Ordinal scale** – a scale of measurement that is nominal and allows ranking or ordering the data according to some criteria.


# ORDINAL

****Ordinal sounds like order****

## Notes

- **Ordered categories**
- Relative rankings
- Unknown distance between rankings
- Zero is arbitrary

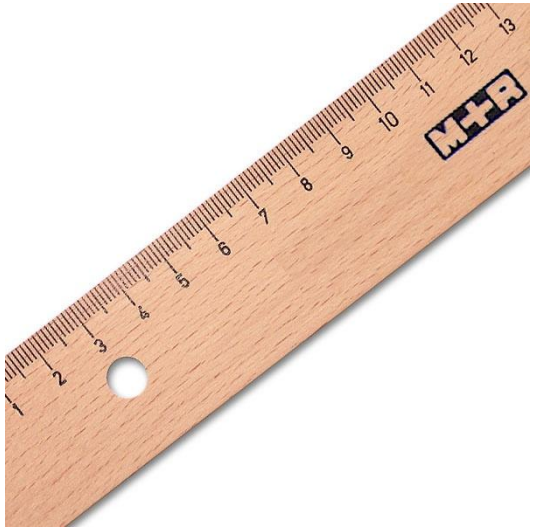
## Examples

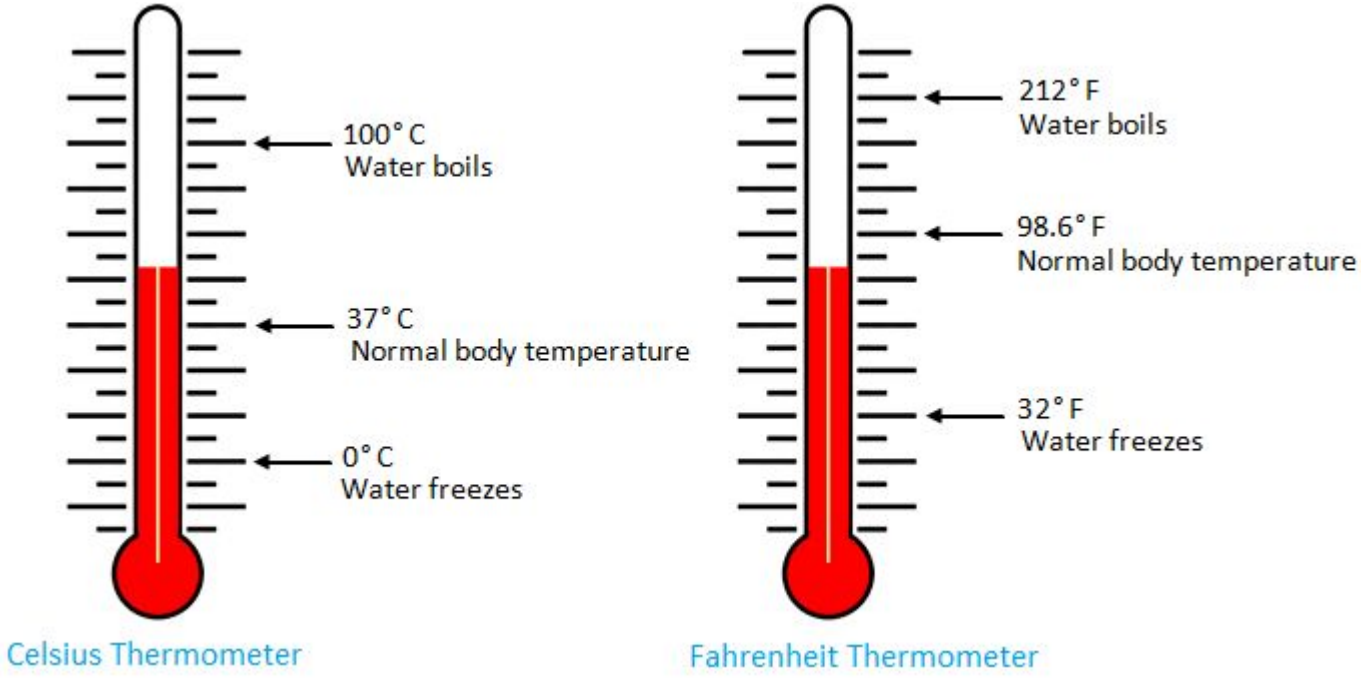
- Likert scales
- Socioeconomic status
  - 1 = Low
  - 2 = Middle
  - 3 = High
- Size
  - 1 = Small
  - 2 = Medium
  - 3 = Large
- Size, ranking of favorite sports, class rankings, wellness rankings

# SM for Quantitative Data (1)



- **Interval scale –** a scale of measurement that is ordinal and intervals between data can be used to compare variable observations.





# Interval

## Notes

- Ordered categories
- Equal distance
  - Can measure differences
- **Zero is arbitrary**
  - Temperature
    - Celsius or Fahrenheit
  - Elevation
  - Time

## Possible Measures

- All ordinal tests
- Mean
- Standard deviation
- Addition and subtraction
  - *Cannot multiply or divide*

# SM for Quantitative Data (2)



- **Ratio scale –** a scale of measurement that is interval and allows considering the ratio of two data values.


# Ratio

## Notes

- Precise, Ordered, Exact
- Equal intervals
- **Natural zero**
  - Weight
  - Time
  - Degrees Kelvin

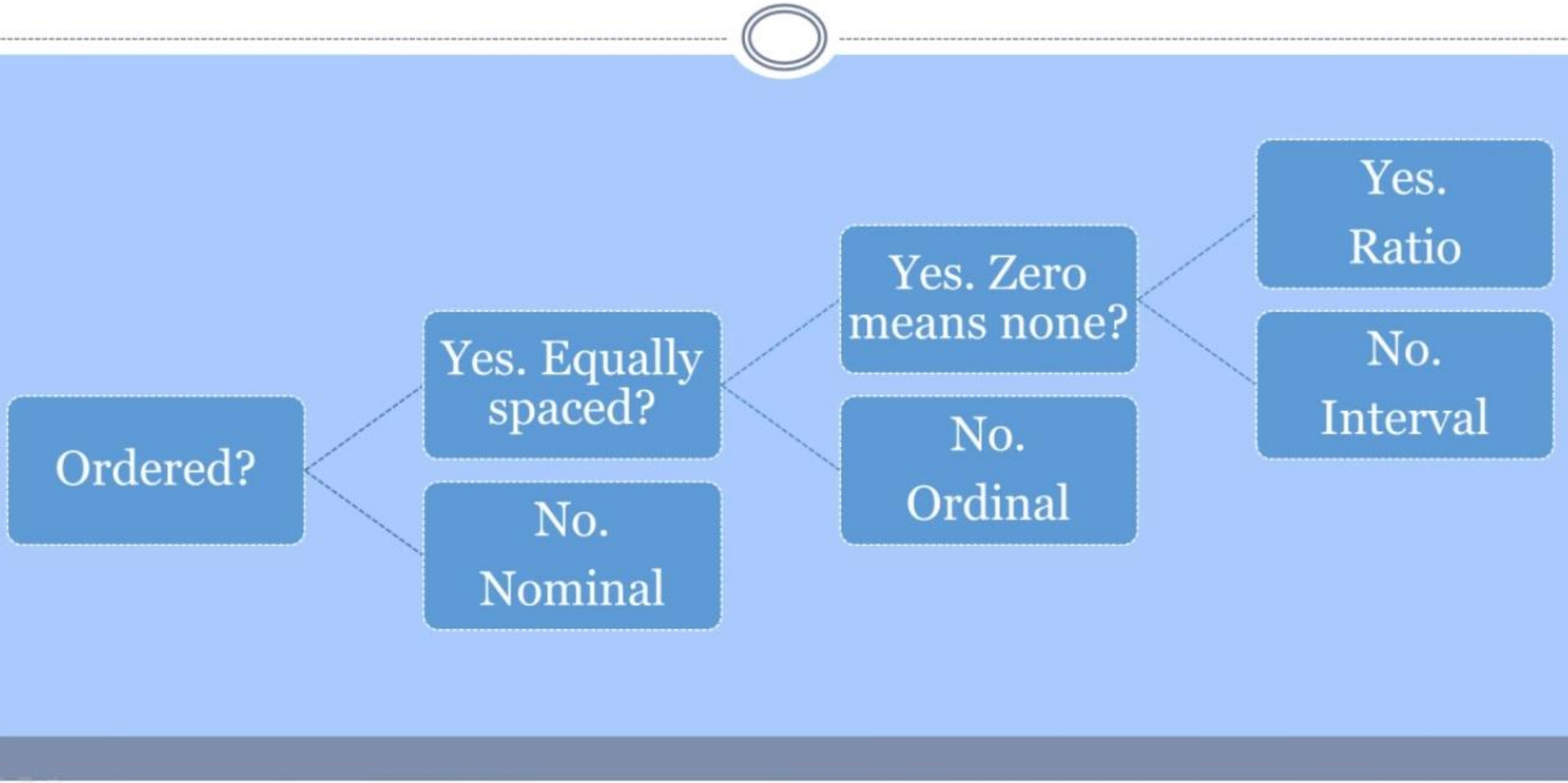
## Possible Measures

- **All** operations are possible
  - Descriptive and inferential statistics
- **Can** make comparisons
  - An 8 kg baby is twice as heavy as a 4 kg baby
- **Can** add, subtract, multiply, divide (ratios)

# Characteristics of Levels of Measurement

| | Nominal | Ordinal | Interval | Ratio |
|---|---|---|---|---|
| Labeled | ☑ Yes | ☑ Yes | ☑ Yes | ☑ Yes |
| Ordered | ☒ No | ☑ Yes | ☑ Yes | ☑ Yes |
| Known difference | ☒ No | ☒ No | ☑ Yes | ☑ Yes |
| Zero is arbitrary | N/A | ☑ Yes | ☑ Yes | ☒ No |
| Zero means none | N/A | ☒ No | ☒ No | ☑ Yes |

# Level of Measurement Decision Tree



# Raw data



- **Raw data** – the data that has not been processed (analyzed, categorized, put in a table) yet.

*Example:*

Number of students (total 100), who attended 12 lectures: 100, 98, 85, 76, 64, 55, 76, 87, 96, 98, 99 & 100

# Aggregate data

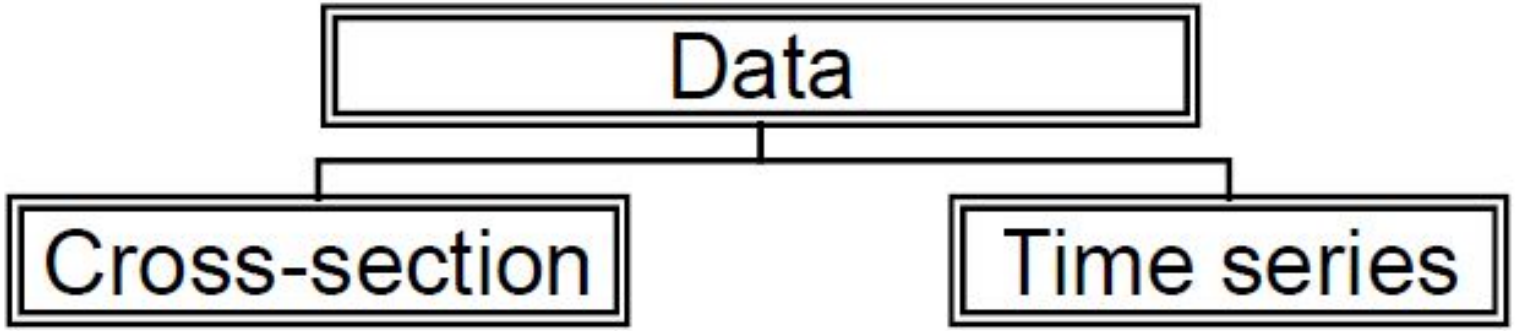


- **Aggregate data –** the data that has already been processed to serve one's goal.

*Example*:

On four lectures, the attendance of students was lower than 80 and on other eight lectures it was greater or equal to 80.

(the raw data above have been analyzed).




**Cross-section data** – data collected at the same point in time or based on the same period of time.

*Example*:

Numbers of different models of automobiles produced by GM Uzbekistan in 2020.

**Time series data** – data that consist of observations collected at regular intervals over time.

*Example*:

Number of automobiles produced by GM Uzbekistan during the period from 2010 to 2020.

# Population and Sample



- **Population** – a collection of all elements of interest in a particular study.
- **Sample** – a subset of the population

*Example:*

All University students vs CIFS students

CIFS students vs 3CIFS1 group

**Note:** Data about a large group of elements are difficult to collect due to various restrictions, therefore only a small part of the group is considered.

# PART II. Data representation tools and techniques

# Section I Qualitative data:



- **Case 1**. Research conducted on 50 individuals' choice on GM Uzbekistan automobiles.

**Table 1**. Data from a sample of 50 new-car purchases

| | | | | |
|---|---|---|---|---|
| Nexia | Cobalt | Nexia | Nexia | Spark |
| Matiz | Spark | Nexia | Matiz | Cobalt |
| Nexia | Cobalt | Nexia | Matiz | Spark |
| Matiz | Spark | Malibu | Cobalt | Cobalt |
| Matiz | Matiz | Malibu | Matiz | Malibu |
| Cobalt | Malibu | Matiz | Cobalt | Matiz |
| Matiz | Nexia | Malibu | Malibu | Spark |
| Nexia | Spark | Malibu | Matiz | Matiz |
| Malibu | Matiz | Matiz | Nexia | Cobalt |
| Nexia | Spark | Nexia | Spark | Spark |

# Tabular Methods:

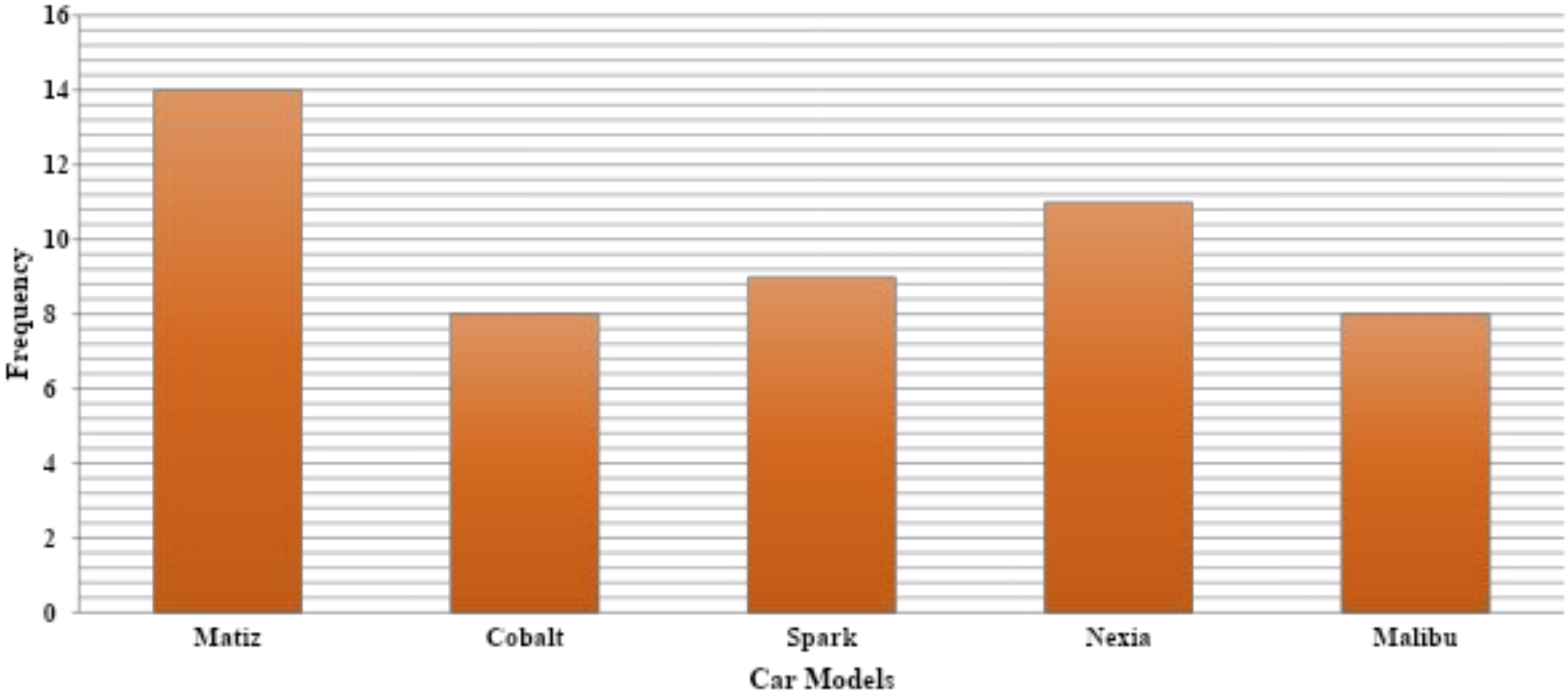


- **Frequency and Relative frequency tables**

| Automobile purchased | Tally | Frequency | Relative frequency |
|---|---|---|---|
| Matiz | 𝍸 𝍸 \|\|\|\| | 14 | 0.28 |
| Cobalt | 𝍸 \|\|\| | 8 | 0.16 |
| Spark | 𝍸 \|\|\|\| | 9 | 0.18 |
| Nexia | 𝍸 𝍸 \| | 11 | 0.22 |
| Malibu | 𝍸 \|\|\| | 8 | 0.16 |
| Total | | 50 | 1.00 |

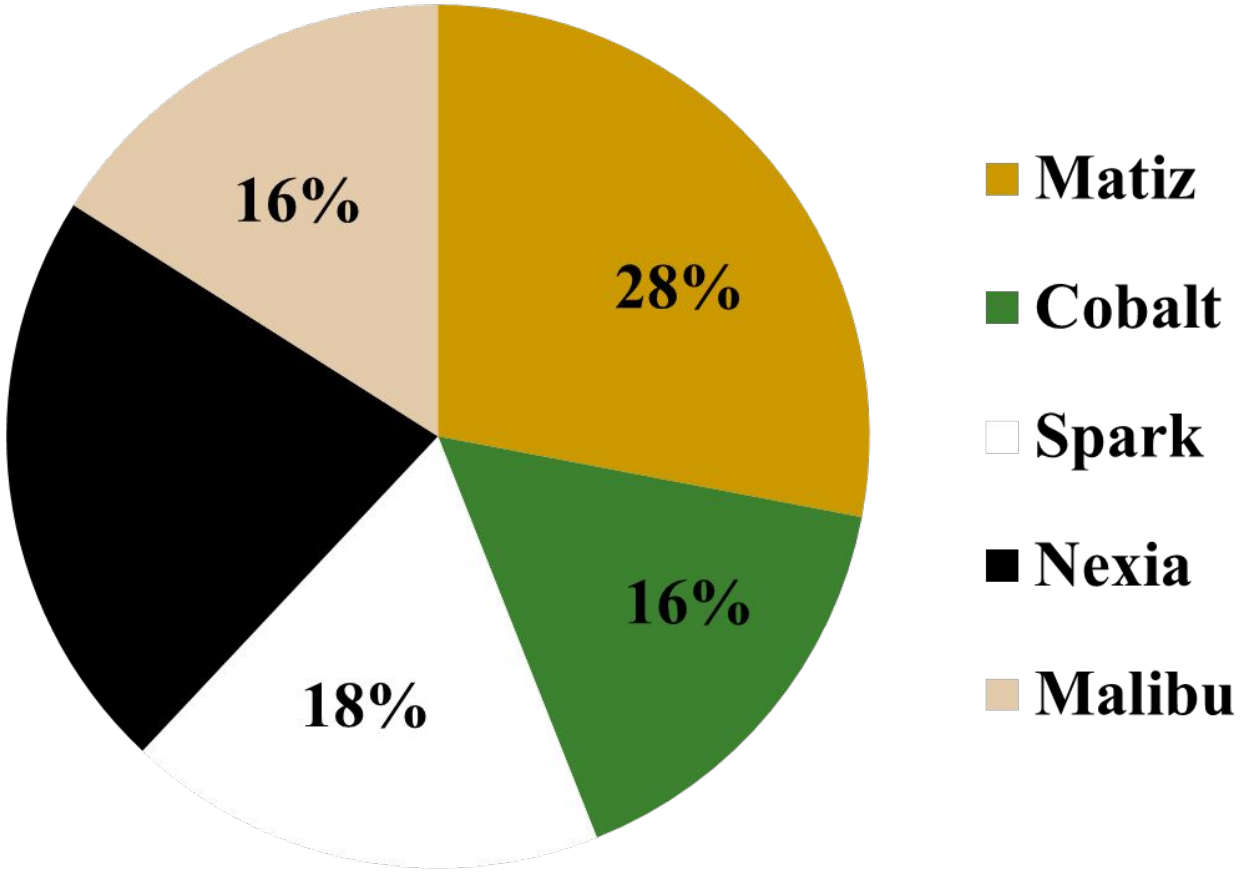
# Graphical Method: Bar graph





# Graphical Method: Pie chart





# Quantitative data: Discrete



- **Case 2**. The store sold the following numbers of refrigerators on 30 different days. Analyze and present the data in tabular and graphical forms.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 5 | 0 | 15 | 1 | 23 | 6 | 5 | 18 | 8 | 10 |
| 2 | 10 | 6 | 0 | 0 | 11 | 2 | 13 | 6 | 3 |
| 19 | 7 | 12 | 1 | 5 | 16 | 0 | 14 | 4 | 8 |

# Tabular Methods:



## Frequency, relative and cumulative frequency table

Range = 23 – 0 = 23; Group width = 23:5 = 4.6 ≈ 5;

Thus, make the group width = 5 for convenience.

| Number of car sold | Tally | Frequency | Cumulative frequency | Relative frequency |
|---|---|---|---|---|
| 0-4 | \|\|\|\|\| \|\|\|\|\| | 10 | 10 | 0.33 |
| 5-9 | \|\|\|\|\| \|\|\|\| | 9 | 19 | 0.3 |
| 10-14 | \|\|\|\|\| \| | 6 | 25 | 0.2 |
| 15-19 | \|\|\|\| | 4 | 29 | 0.13 |
| 20-24 | \| | 1 | 30 | 0.03 |
| Total | | 30 | | 1.00 |

# Tabular Method:



## Stem-and-Leaf diagram

**A store sold the following numbers of refrigerators on 30 days:**

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 5 | 0 | 15 | 1 | 23 | 6 | 5 | 18 | 8 | 10 |
| 2 | 10 | 6 | 0 | 0 | 11 | 2 | 13 | 6 | 3 |
| 19 | 7 | 12 | 1 | 5 | 16 | 0 | 14 | 4 | 8 |

**1-Method**

| | |
|---|---|
| 0 | 5016582600263715048 |
| 1 | 5800139264 |
| 2 | 3 |

**Stem = 10, Leaf = 1**

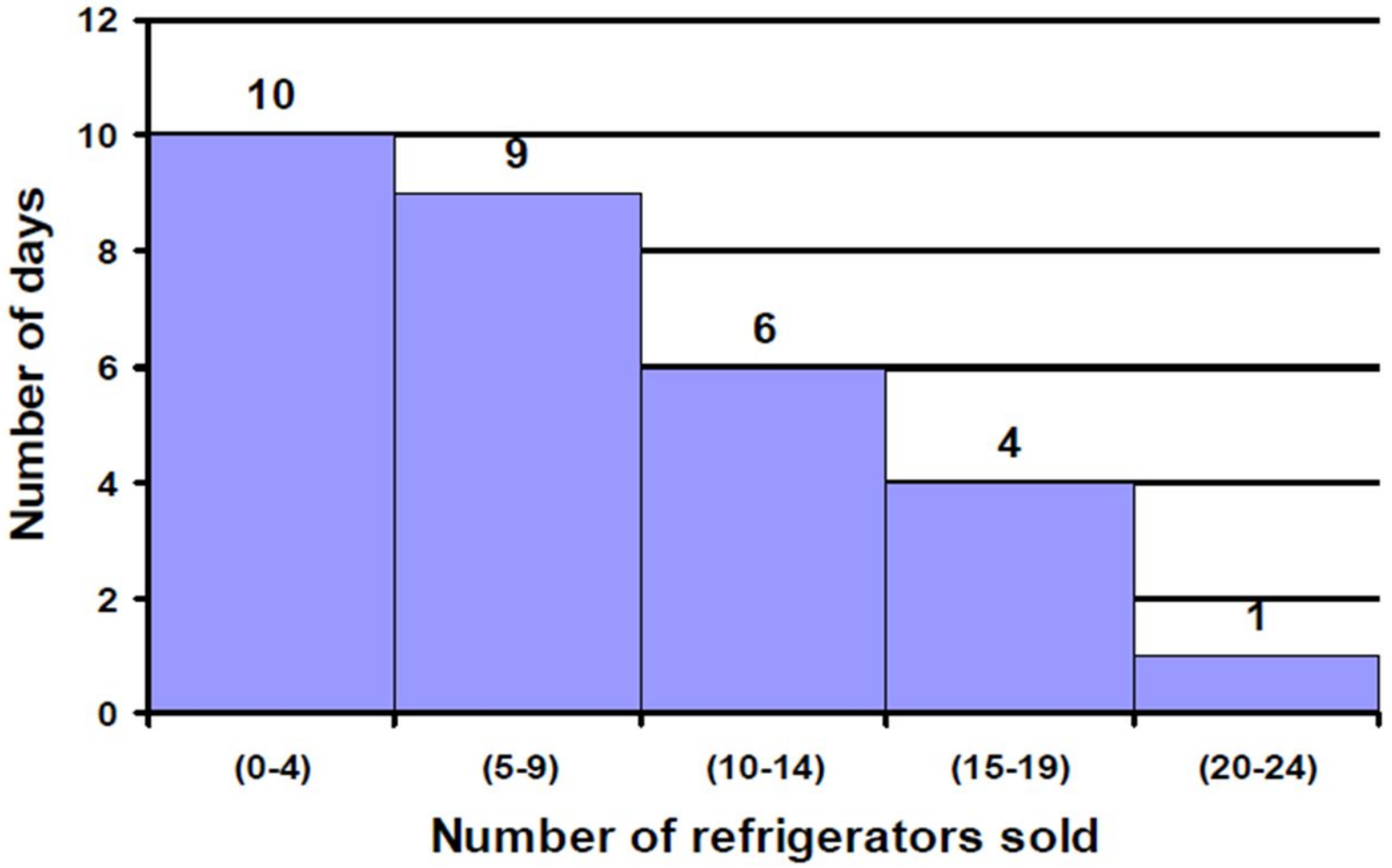
**2-Method**

| | |
|---|---|
| 0 | 0120023104 |
| 0 | 56586758 |
| 1 | 001324 |
| 1 | 5896 |
| 2 | 3 |

# Graphical Method: Histogram



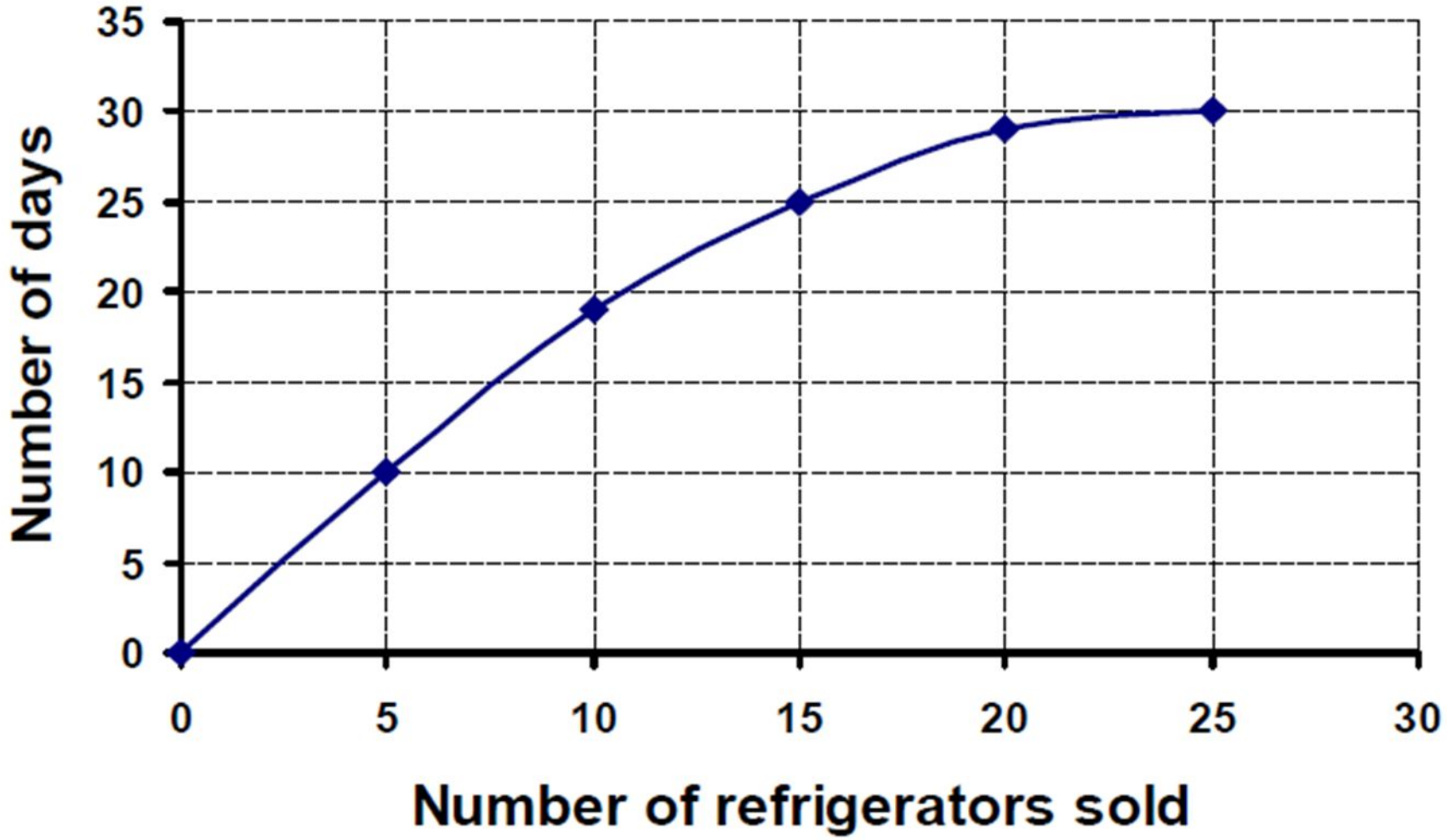
## Histogram



# Graphical Method



## Cumulative frequency



# Quantitative data: Time series



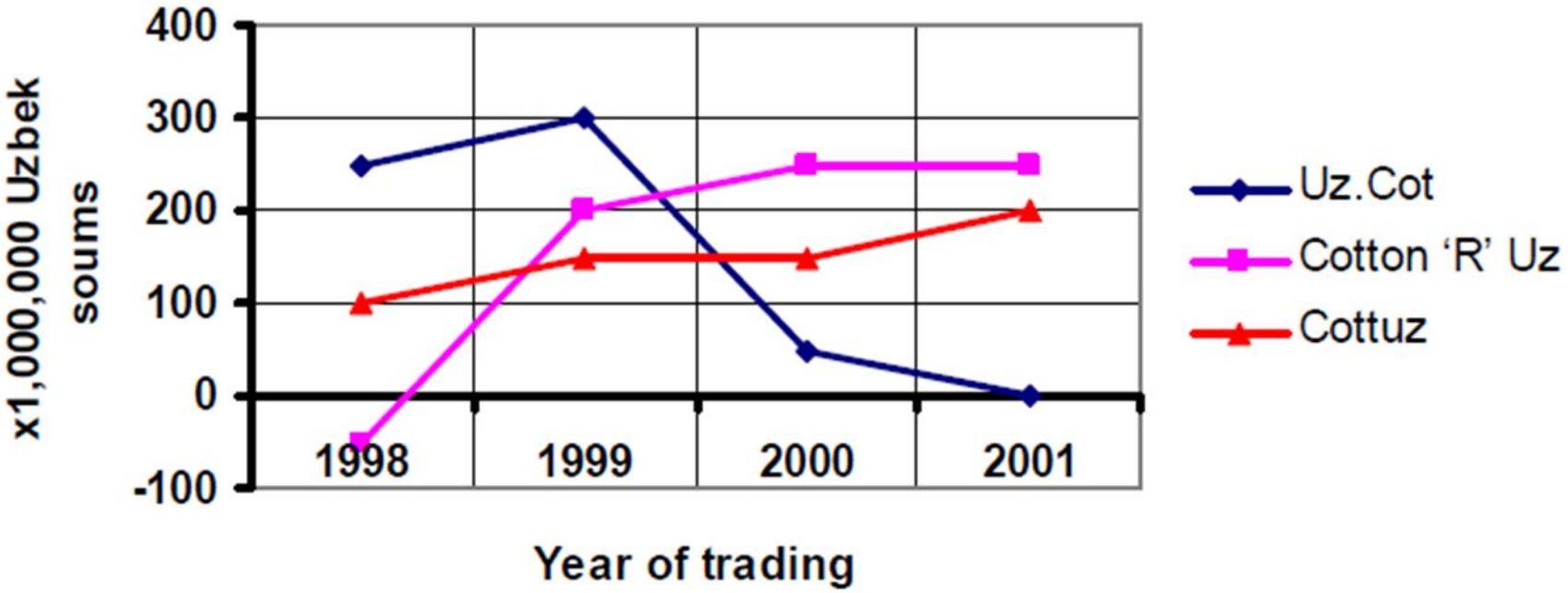
Case 3. the following table shows the profit made by three cotton companies over four years. Display this data graphically

| Name/Year | 1998 | 1999 | 2000 | 2001 |
|---|---|---|---|---|
| Uz.Cot | 250 | 300 | 50 | 0 |
| Cotton 'R' Uz | – 50 | 200 | 250 | 250 |
| Cottuz | 100 | 150 | 150 | 200 |

# Quantitative data: Time series



## Times series graph (line graph)



# Quantitative data: Time series



Case 4:

The company XYZ produces three types of products (A, B, and C). The total sales of the Product A in 1999, 2000 and 2001 were £40,000, £45,000 and £50,000, of the Product B were £30,000, £40,000 and £50,000 and of the Product C were £50,000, £55,000 and £60,000 respectively. Construct a table for this data and illustrate it with a help of bar chart.

# Tabular form



| Product/Year | 1999 | 2000 | 2001 |
|---|---|---|---|
| A | 40 | 45 | 50 |
| B | 30 | 60 | 40 |
| C | 50 | 55 | 60 |

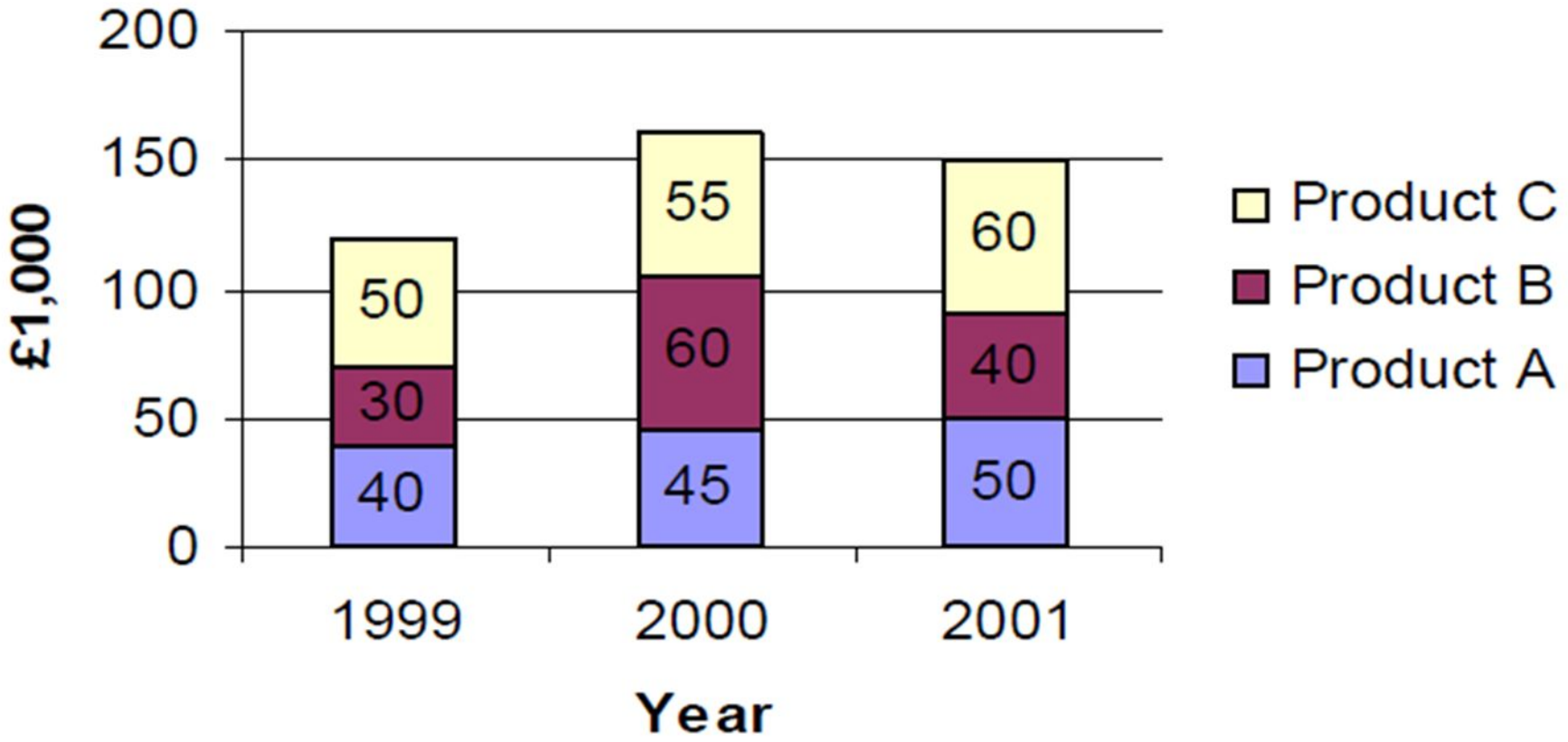
or

| Year/Product | A | B | C |
|---|---|---|---|
| 1999 | 40 | 30 | 50 |
| 2000 | 45 | 60 | 55 |
| 2001 | 50 | 40 | 60 |

# Graphical form

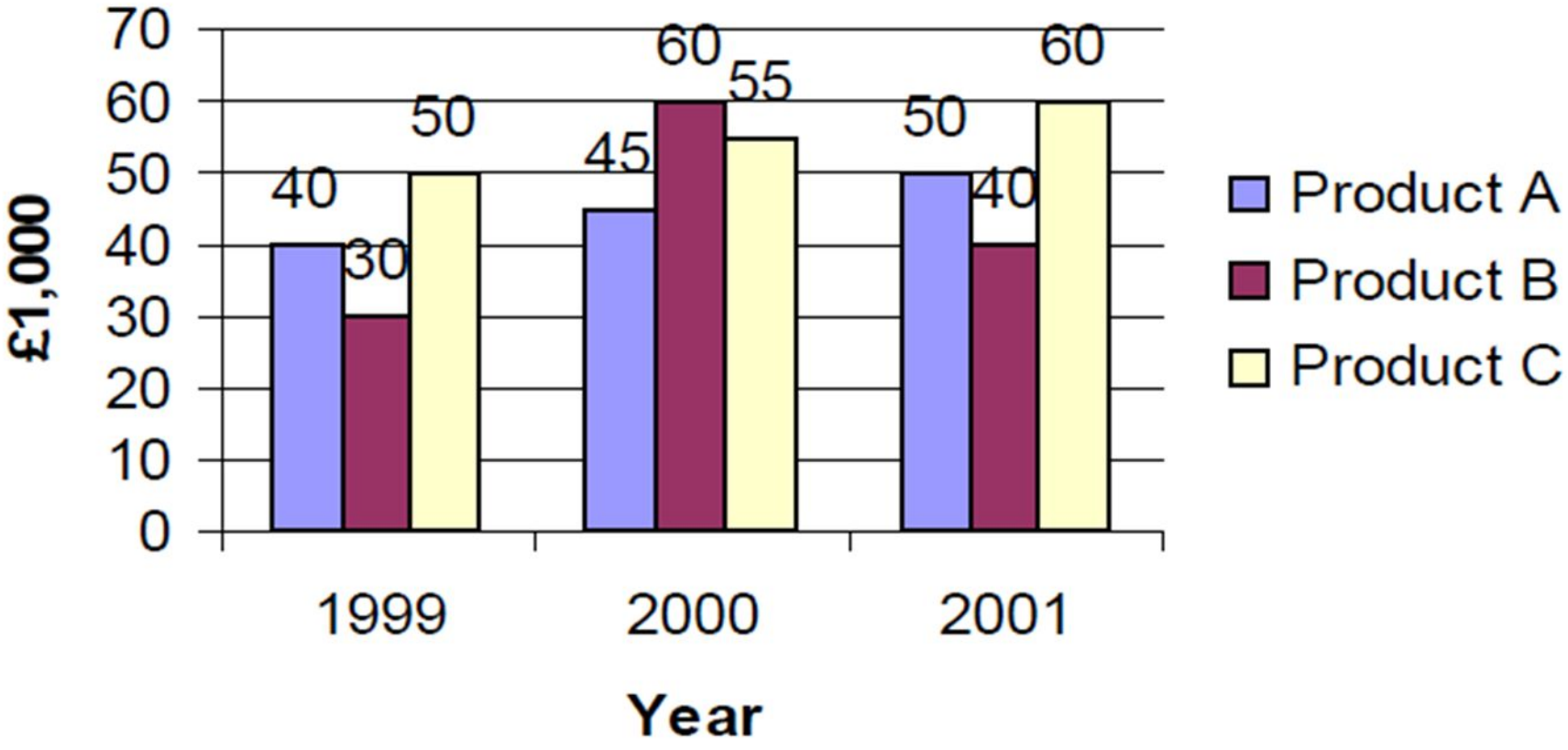


## Component bar graph



# Graphical form

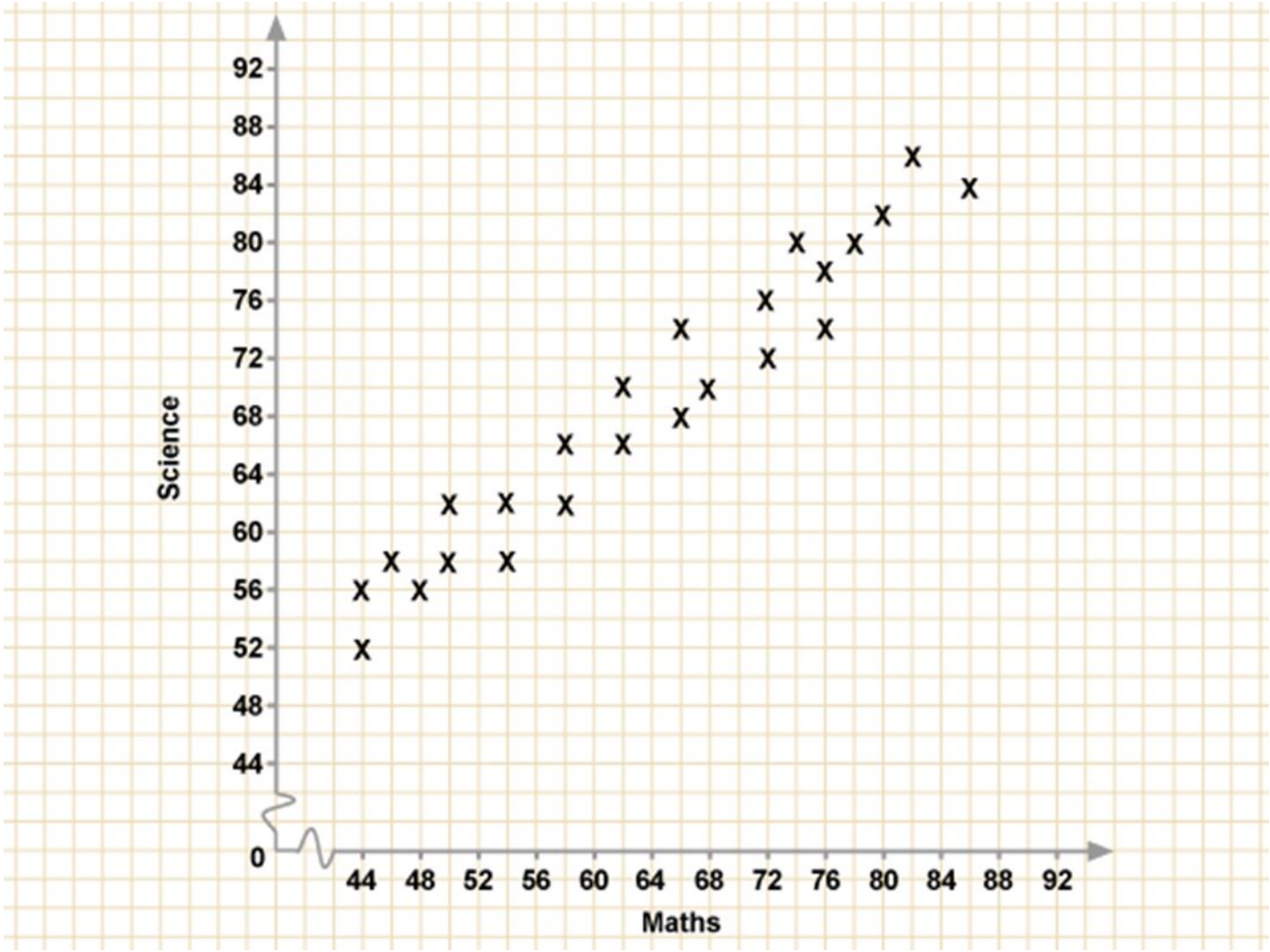
## Multiple bar graph



# Graphical Method



## Scatter graph



# Concluding remarks:



Today, you learnt:

a) The components of statistical table;
b) The main types of data;
c) The scales of measurement of the data
d) analyze statistical data;
e) use tabular methods to display data
f) use graphical methods to display data

# Essential readings (Part 1)



- Jon Curwin…, “Quantitative Methods…”, Chapters 1-2
- Glyn Burton…, “Quantitative Methods…”, Chapter 1
- Richard Thomas, “Quantitative Methods…”, Chapter 1.1
- Mik Wisniewski…, “Foundation Quantitative…”, Chapter 3
- Clare Morris, “Quantitative Approaches…”, Chapter 3

# Essential readings (Part 2):



- Jon Curwin…, “Quantitative methods…”, Chapter 4
- Glyn Burton…, “Quantitative methods…”, Chapter 1
- Richard Thomas, “Quantitative methods…”, Chapter 1.2-1.4
- Mik Wisniewski…, “Foundation Quantitative…”, Chapters 5-6
- Clare Morris, “Quantitative Approaches…”, Chapter 5
- Louise Swift “Quantitative methods…”, Chapter DD1.