I- A 9th-grade teacher *collects* the grades of his students in Arabic test per 30.

> 29 ; 12 ; 20 ; 22 ; 28 ; 15 ; 30 ; 25
> 24 ; 13 ; 25 ; 17 ; 23 ; 19 ; 29 ; 28
> 20 ; 19 ; 17 ; 21 ; 15 ; 21 ; 15 ; 23

1) How can you describe the *organization* of the above data?
   a) Ordered   b) Random   c) Simple
2) Which *objects* (people, animals …) are *under study*?
   ………………………………………………………………………………………………
3) What is the *aspect* that is being *studied*?
   ………………………………………………………………………………………………
4) What is the *number* of *subjects* under study?
   ………………………………………………………………………………………………
5) What is the highest mark collected? …………………………………………………..
6) Which is the most frequent mark? …………………………………………………….
7) How many students took 15? …………………………………………………………

II- A 9th-grade teacher *collects* the grades of his students in a math test per 30 and organize them in the following way:

Grades	12	13	15	17	20	22	25	28
Number of students	1	2	2	1	4	5	3	7

1) How can you describe the *organization* of the above data?
   a) Ordered   b) Random   c) Simple
2) Which *objects* (people, animals …) are *under study*?
   ………………………………………………………………………………………………….
3) What is the *aspect* that is being *studied*?
   ………………………………………………………………………………………………….
4) What is the *number* of *subjects* under study?
   ………………………………………………………………………………………………….
5) What is the highest mark collected? …………………………………………………..
6) Which is the most frequent mark? …………………………………………………….
7) How many students took 15? …………………………………………………………

Conclusions:
   a) Which of the above ways of collecting data was more readable? …………………….
   b) If the second way is called the statistical way, then
      i. What does statistics do? ……………………………………………………………
      ii. How is it useful? ……………………………………………………………………
      iii. Where would you think that we can use statistics? …………………………………...

✓ **Introduction:** Statistics which is a synonym for "<u>numerical facts</u>" is a relatively new branch of mathematics that aims at:

1) Collect data.
2) Classifying, summarizing and organizing data.
3) Reading data in a more efficient way.
4) Representing data in different forms (pie graph, histogram, bar graph …).
5) Interpret and find relation among data.
6) Draw out conclusions from a given set of data.

✓ **Usages:** almost every field of studies benefit from organizing statistical surveys.

  👀 _**Educational field**_: Teachers evaluating students over a specified set of objectives (factorize, develop, graph, conduct a geometric proof…).

  👀 _**Entertainment**_: A TV show analyzing a football match (goals scored, shoots on target percentage of ball possession…)

  👀 _**Industry**_: A company introducing a new product to the market (study of profit per month)

✓ **Statistical vocabulary:**

- _**Population:**_ is the set of observed elements having a common property.
  ➢ The set of students in a class.
  ➢ The set of teams in a certain league.

- _**Size:**_ is the total number of elements in a population.
- _**Character (values, variable) (x_i):**_ is the common property of the population under study.
  ➢ The height, weight, grades…. of an individual in a set of population.
  ➢ The color of eyes, gender (male, female), behavior of an individual in a population.
    ▪ Types of characters:
    1) _**Quantitative:**_ a character is said to be quantitative if it can be measured.
       ✓ Length, number of children, number of books read ….
    2) _**Qualitative:**_ a character is said to be qualitative if it cannot be measured.
       ✓ The color of eyes, gender (male, female), behavior, scent, taste, shapes….

- _**Frequency(n_i):**_ is the number of times a character is observed.

👀 _Note that_: The total frequency or size is the sum of all frequencies and it is denoted by N.

- _**Relative frequency(R.f):**_ is the ratio of the frequency(n) to the size (N) of an object.

$$\textbf{\textit{In symbols: }} R.f = \frac{n}{N}$$

👀 _Note that_: The relative frequency is a number strictly included between 0 and 1

$$\textbf{\textit{In symbols: }} 0 < R.f < 1$$

✓ **Measure of central tendency:**

- **_Range(R):_** is the difference between the highest and lowest observed values for a quantitative character.
- **_Mode:_** is the variable that admits the highest frequency.
  - Having two modes is called **"bimodal"**.
  - Having more than two modes is called **"multimodal"**.
- **_Mean (Average):_** is of two main types
  - ➢ **_Arithmetic Mean:_**
    Def: is the ratio of the sum of values to the size of the population.
    Formula: $\overline{X} = \dfrac{\sum x_i}{N}$
  - ➢ **_Weighted mean:_**
    Def: is the ratio of the sum of product of values by their frequencies to the size.
    Formula: $\overline{X} = \dfrac{\sum x_i \cdot n_i}{N}$
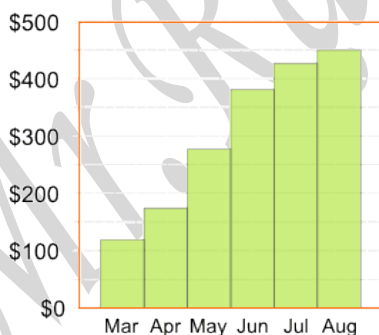
✓ Cumulative frequency:

- Cumulative means "how much so far".

  Think of the word _"**accumulate**"_ which means to gather together.
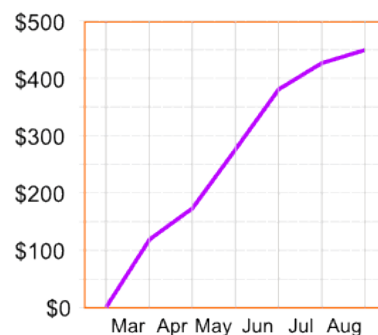
- Types of cumulative frequencies:
  - ➢ **_Increasing cumulative frequency (ICf)_**

    You can make cumulative graphs if you want.

Increasing Cumulative Histogram   Increasing Cumulative Line Graph

  - ➢ **_Decreasing cumulative frequency (DCf)_**

    To have cumulative totals, just **add up the values as you go**.

Histograms vs Bar Graphs

Bar Graphs are good when your data is in **categories** (such as "Comedy", "Drama", etc).
But when you have continuous data (such as a person's height) then use a Histogram.
It is best to leave gaps between the bars of a Bar Graph, so it doesn't look like a Histogram

👀 Using a calculator:

    ↝ How can we use the calculator to find some statistical indicators?

CASIO $fx-991ES$	CASIO $fx-991ES$

1- Mode2
2- Shift/clr/1/=
3- Enter data:
   i)   Variable/shift/,/frequency/m+
4- To find:
   a) **Mean**: press shift/2/1/=
   b) **Standard deviation**: press shift/2/2/=
	1- Mode/3/stat/1:1-var
	2- On
	3- Shift/mode
	4- Down
	5- 4:stat/1:on
	6- Same as step 1.
	7- Fill data.
	8- Ac
	9- Shift/1
	10- Press: 4: var
	    Choose one you want to cal.
	     i. Mean: $\bar{x}$
	     ii. Standard deviation.

Project

Choose one of the following surveys, and then answer the related questions given at the end

S-1: Study the distribution of tourists among five Lebanese cities.
S-2: Study the number of goals scored in the last ten games for your best five football teams.
S-3: Study the number of points scored by five players in a basketball match between your two
    favorite teams.
S-4: Study the number of hours you spend preparing for each of five of your school
    assignments.

For each statistical survey specify (on your own):
   1) The population and the size of the chosen sample space.
   2) The variable under study (character) and its type (Qualitative or Quantitative)
   3) The range of the data if possible.
   4) The highest and the lowest values among the specified data.
   5) The average (mean) of your data.