General Revision:

The following raw data represents the marks of 11th grade students in math per 20 in a certain test:

9	13	15	13	12	7	17	9	18	13
18	12	13	11	17	13	13	7	9	11
9	15	15	9	12					

1) Statistical vocabulary:
   a. Specify the:
      i. Population: ……
      ii. Statistical unit: ……..
   b. Indicate the variable (character) under study and specify its nature.
      …….
   c. List two modalities of the variable under study: …….
   d. Choose one of the following terms that best describe the variable:
      i. Continuous: means the variables are in interval form (classes) such as height, weight…
      ii. Discrete: means the variables are whole numbers such as number of brothers …
2) Rearrange the given raw data in a statistical table of increasing cumulative frequency.

Mark (x_i)	7	9	11	12	13				Total
Frequency (n_i)	2	4	2	3					
I.C.F									

3) Specify the significance of the numbers: $n_i = \{1,4 \& 2\}$ ……
4) Determine and interpret each of the following
   a. Statistical indicators: indicate certain aspects about a given series
      i. Mode, M_o: ……
      ii. Range, R:……...
   b. Central tendencies: such values shows the central value of a quantitative statistical series
      i. Median, M_e: …….
      ii. Mean, $\bar{x} = \dfrac{\sum_{i=1}^{k} n_i x_i}{N}$: ……
5) The standard deviation is given by: $\sigma = \sqrt{V} = \sqrt{\dfrac{1}{N}\sum_{i=1}^{k}\left(n_i {x_i}^2\right) - \left(\bar{x}\right)^2}$
   a. Use your calculator to calculate σ: …….
   b. Explain the obtained value. ……
   c. Find the percentage of students whose mark is in the interval $\left[\bar{x}-\sigma, \bar{x}+\sigma\right]$
      ……..
6) Is it true that average of the class is 12.28 with an error of 3.62? ……..

In the above example the modalities where ***discrete***, that is to say single valued numerals.
Now, we are interested in ***continuous*** data, a data is continuous if the modalities are listed as ***interval*** (***classes***).

1- Converting a range of discrete data into classes or intervals:

To divide discrete data into classes of equal size we consider the following factors:

- ✓ The number of intervals.
- ✓ The size or amplitude of the class.
- ✓ The upper and lower limits of the interval.

Consider the following data:

9	13	15	13	12	15	17	9	18	13
18	12	13	11	17	13	13	7	9	11
15	15	15	9	12					

1) Regroup the above data in classes of amplitude 3:

Classes (C_i)	[7;10[	[10;13[		

2) If the center of a class(interval) is its midpoint then complete the following table:

Classes (C_i)	[7;10[	[10;13[			Total
Frequency (n_i)	5				
Center x_i					
R.F					
I.C.F					
D.C.F					

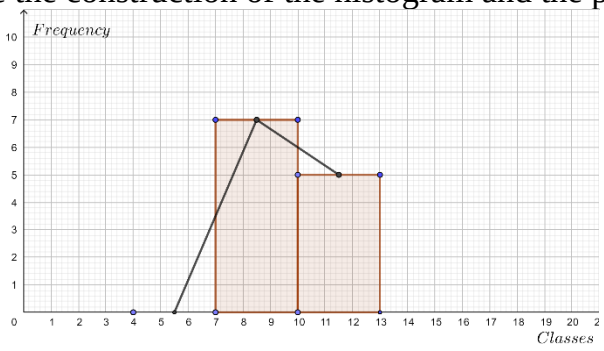
3) Prove that the median class is [13;16[……

4) Determine the modal class: ……

5) If the average of a statistical series $(X_i; n_i)$ is: $\bar{x} = \frac{1}{N}\sum_{i=1}^{k} n_i c_i = \frac{1}{N}(n_1c_1 + n_2c_2 + \ldots\ldots + n_kc_k)$,

then approximate the mean mark of the given students: ……..

6) Complete the construction of the histogram and the polygon of frequencies



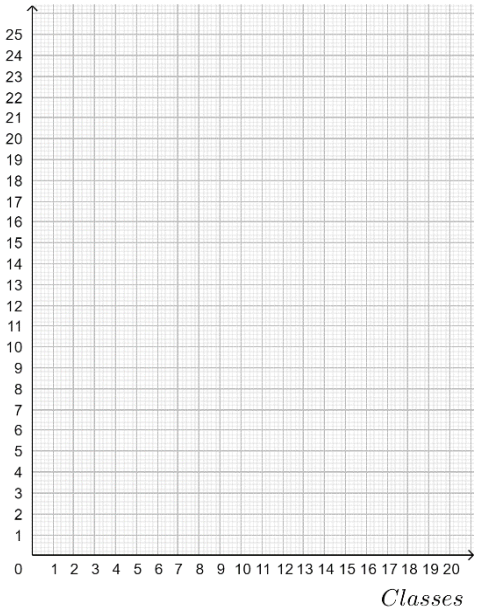
7) The following points are on the polygons of:

I.C.F: $A(7;0), B(10;7), C(13,12), D(\ ;\ ) \& E(\ ;\ )$

D.C.F: $F(7;25), G(10;18), H(13,\ ), I(\ ;\ ) \& E(\ ;\ )$

a. Use table to complete the above points

b. Construct the polygons of I.C.F &I.D.F on same graph:



The median is the « abscissa » of the point of intersection of the polygon of I.C.F(or D.C.F) with the straight line o of equation y = N/2 , where N is the total frequency .

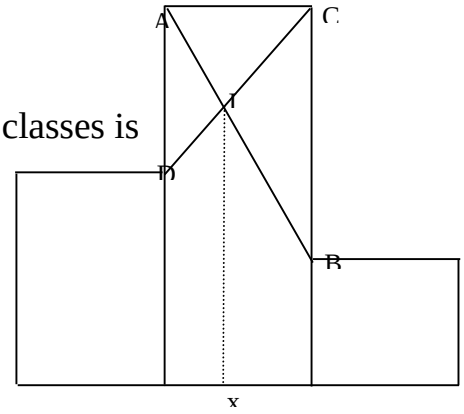
OR: The median m is the "abscissa" of the point of intersection of the polygons of the increasing and the decreasing cumulative frequencies .

OR: We take from I.C.F the segment that contains the point M(m;N/2) , C , M and D are collinear then $\frac{y_M - y_C}{y_D - y_C} = \frac{x_M - x_C}{x_D - x_C}$ we get $x_M = m \approx 13.2$

8) Mode – Modal class

The mode M_o , is the value of the variable having the **highest frequency** , that is most **the individuals of the population have this value .**

An approximation of the mode of a series grouped in classes is the **center of the modal class** .



Mode Graphically

The abscissa of the point of intersection
Of segments [AB] and [CD] the diagonals
Of the rectangle relative to modal class .

Question1: Given the data: 1 1 2 3 3 3 4 4 4 5 5 5 5 8 9 .

1- Calculate the mode and the median of this data.
2- Group this data in classes of amplitude 2.
3- Calculate the mode and the median .Conclude.

Question 2: A bookstore announced the following distribution of books sold last week according to prices:

1-Complete the table showing centers, increasing and decreasing cumulative frequencies:

Price (in 1000 LL)	[5; 10[	[10; 15[	[15; 20[	[20; 25[	[25; 30]
Frequency	40	70	48	56	46

2- Calculate the average price of a book sold.
3- What is the percentage of books sold at a price of at least15 000 LL?
4- Draw the polygon of I C f and that of decreasing cumulative frequencies.
5- Estimate graphically the median price. What does the answer signify?

Recall That - Range, mean, variance , standard deviation .

Classes	$[x_1 ; x_2[$	$[x_2 ; x_3[$	$[x_3 ; x_4[$	.....	$[x_k ; x_{k+1}]$
Frequencies	n_1	n_2	n_3	.....	n_k
Centers	c_1	c_2	c_3	.....	c_k

- Mean: $\bar{x} = \frac{1}{N}(n_1c_1 + n_2c_2 + ... + n_kc_k)$.
- Range: $e = x_{k+1} - x_1$.
- Variance: $V = \frac{1}{N}(n_1{c_1}^2 + n_2{c_2}^2 + ... + n_k{c_k}^2) - \bar{x}^2$
- Standard deviation: $\sigma = \sqrt{V}$.