The third axis

Tabular and Graphical Display of Quantitative Data

(Quantitative variables)

I- Tabular and Graphical Display of discrete Quantitative variables

1- **Frequency table of Discrete Quantitative variable**

Suppos that you have collected some discrete data. It will be difficult to get a ‘’feel’’ for the distribution of the data just by looking at it in list form. It may be worthwhile constructing a frequency table.

The frequency of a value is the number of observations taking that value (xi).

A frequency table is a list of possible values (the first collumn) and their frequencies (the second collumn)

Example : the following data represents the number of rooms for 20 houses

1،2،1،3،2،4،5،6،6،4،2،3،5،1،2،3،3،6،3،3

- Record this data in a frequency table.

-What is the statistical population, statistical unit, statistical phenomenon of this data.

- Solution : frequency table represents the number of rooms for 20 houses

Number of rooms (xi)	Frequency (ni)
1	3
2	4
3	6
4	2
5	2
6	3
Sum	20

statistical population : houses

statistical unit : house

statistical phenomenon : number of

2- **Relative Frequencies and percentages**:

Relative Frequencies: In addition to the absolute frequencies (counts), relative frequencies are often included in the table. These are the proportions or percentages of observations in value, which can help in comparing different values of discrete variable (xi).

$$fi=ni/\Sigma ni \quad / \quad \Sigma fi=1$$

Percentages: Expressing the data in percentages is useful when you want to compare the relative sizes of values of discrete variable (xi).

$$fi(\%) = (fi=ni/\Sigma ni) *100 \; / \; \Sigma fi \; (\%) =100$$

Example : frequency table represents the number of rooms for 20 houses (calculate relative frequency and percentages)

Number of rooms (xi)	Frequency (ni)	Relative frequency (fi)	Penrsantage (fi
1	3	0.15	15
2	4	0.2	20
3	6	0.3	30
4	2	0.1	10
5	2	0.1	10
6	3	0.15	15
Sum	20	1	100

3- **Cumulative frequency (ascending / descending)** :

Cumulative frequency, often referred to as cumulative frequency distribution , is a statistical concept used to analyze and summarize data. It represents the total frequency of values less than or equal to a given value in a dataset.

In ascending cumulative frequency, you start with the smallest data values and add up the frequencies as you move through the dataset in ascending (increasing) order. It begins with the smallest value and ends with the largest.
Ascending cumulative frequency is useful for understanding how many data points fall below or equal to a given value. It helps in calculating percentiles, identifying the median, quartiles, and other important statistics.

In descending cumulative frequency, you start with the largest data values and work your way down through the dataset in descending (decreasing) order. It begins with the largest value and ends with the smallest.
Descending cumulative frequency can be useful when you want to know how many data points are above or equal to a given value. It provides insights into the upper end of the data distribution.

Example : frequency table represents the number of rooms for 20 houses (calculate Cumulative Ascending and descending frequency

Number of rooms (xi)	Frequency (ni)	CAF	CDF
1	3	3	20
2	4	7	17
3	6	13	13
4	2	15	7
5	2	17	5
6	3	20	3
Sum	20	//////////////	//////////////

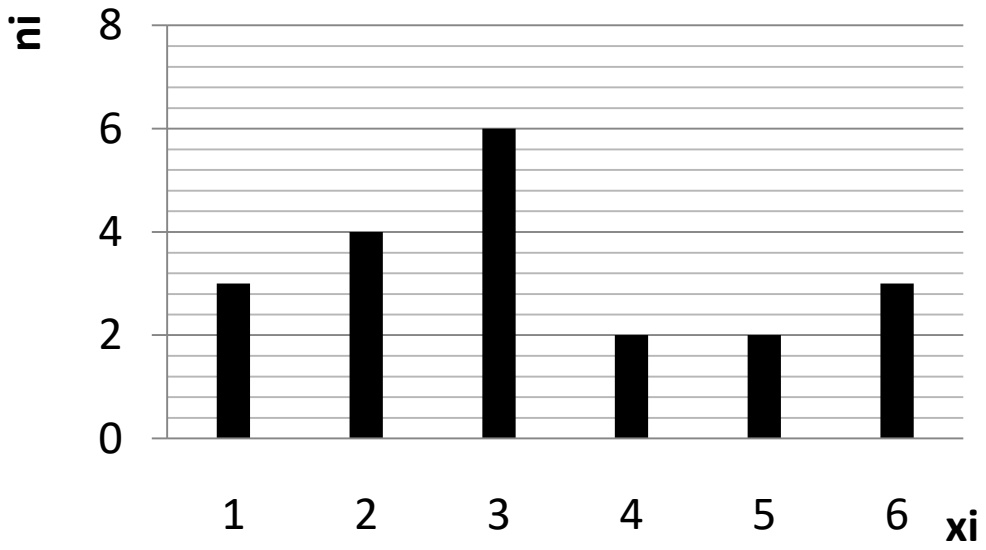
You want to know the following:
-The number of houses with fewer than or equal to 3 rooms : 13 houses (CAF)
-The number of houses with 4 rooms or more: 7 rooms (CDF)

4- **Graphical display of Discrete Quantitative variable**

A- **Bar chart**: consists of bars corresponding to each of the possible values, whose heights are equal to frequencies.

Example : we use the same previous example to represent the bar chart.

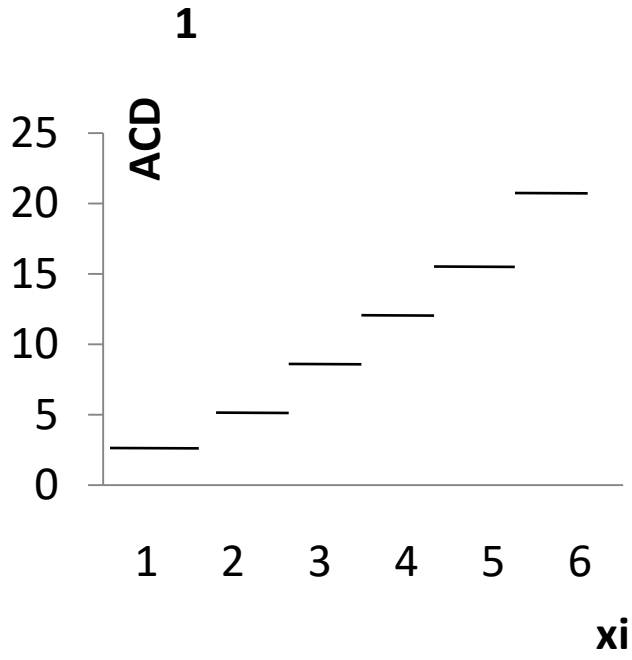
- Bar Chart represente the number of rooms for 20 houses

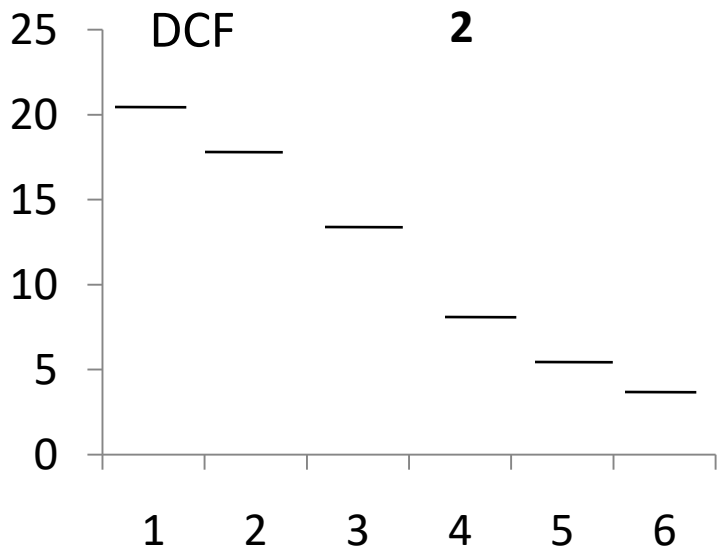


B- **Graphical representation of Cumulative frequency (ascending / descending)**

When represented graphically, ascending/ descending cumulative frequency is typically plotted with data values on the x-axis and cumulative frequencies on the y-axis. This creates a cumulative frequency , which shows the increasing/decreasing straight segments based on the cumulative ascending / descending recurrent value.

Example : we use the same previous example to represent Ascending (1) and Descending (2).





II- Tabular and Graphical Display of Continuous Quantitative variables

1- **Frequency table of Continuous Quantitative variable**

When the variable is continuous, we do not look at the frequency of each value, but groupe the value in the class intervals (classes).

The length of the intervals can be caculated using the STURGE formula:

$$k = \frac{E}{1 + 3.32 \log(n)}$$

- Where:

n: Sample size

E:range= the difference between the largest and smallest values

K: Interval lenght

- Example: Find the appropriate length for the interval to emplty the following data into continuous frequency table.

24	27	44	33	16	20	32	34	25	32
60	57	55	60	32	56	54	34	27	45
53	44	33	57	53	25	54	44	23	28
34	53	62	61	41	51	32	43	36	64
43	43	47	41	37	52	54	42	55	62

Solution:

$$K = E \ / \ (1+3.32 \log(n))$$

$$K = (64-16)/(1+3.32\log 50) = 6.67 \approx 7$$

The frequency table becomes as follows:

Class	ni
[16-23[	2
[23-30[	7
[30-37[	10
[37-44[	7
[44-51[	5
[51-58[	13
[58-65[	6
Sum	50

The data in this table has already been grouped for us into 7 classes.

Each of intevals from 16-23 , 23-30 and so on is called a class interval .

In this exemple: the class interval 16-22 has a frequency of 2. this means that 2 person age between 16-23 inclusive but we dont know their exact age.

- **Note:**

If the interval length are unequal, the heights of the rectangles are chosen so that the area of each rectangle equals the frequency . We calculate a new frequency called the modified frequency (hi) using the ratio method to the multiple of the lowest class length according to the following law:

$$hi=ni^{*}\ (k/ki)$$

Where:

hi: modified frequency

ni: normal frequency

k: the smallest length of the class

ki: length of the classes.

Important note: We adjust the frequency when the class lengths are unequal when:

- We draw the histogram.
- We determining the modal class and calculating the mode.

- Example: the following table illustrates the distribution of 100 workers based on daily wages.

modified frequency hi hi=ni* (k/ki)	Interval lenghts ki	ni frequency	Classes
5	k=5	5	25-20
7.5	10	15	35-25
20	5	20	40-35
8.5	15	25	55-40
7.5	20	30	75-55
5	k=5	5	80-75

2 -**Cumulative frequency (ascending / descending)** :

Cumulative frequency, often referred to as cumulative frequency distribution , is a statistical concept used to analyze and summarize data. It represents the total frequency of class less than to a given value in a dataset.

In ascending cumulative frequency, you start with the smallest data values and add up the frequencies as you move through the dataset in ascending (increasing) order. It begins with the smallest value and ends with the largest.

Ascending cumulative frequency is useful for understanding how many data points fall below to a given the upper limit of the class. It helps in calculating percentiles, identifying the median, quartiles, and other important statistics.

In descending cumulative frequency, you start with the largest data values and work your way down through the dataset in descending (decreasing) order. It begins with the largest value and ends with the smallest.

Descending cumulative frequency can be useful when you want to know how many data points are above to a given the lower limit of the class. It provides insights into the upper end of the data distribution.

Example: Calculate CAF / CDF of the following frequency table

CDF	CAF	ni	classes
96	5	5	20-10
91	14	9	30-20
82	27	13	40-30
69	42	15	60-40
54	62	20	80-60
34	79	17	90-80
17	90	11	100-90
6	76	6	120-100
		96	Sum

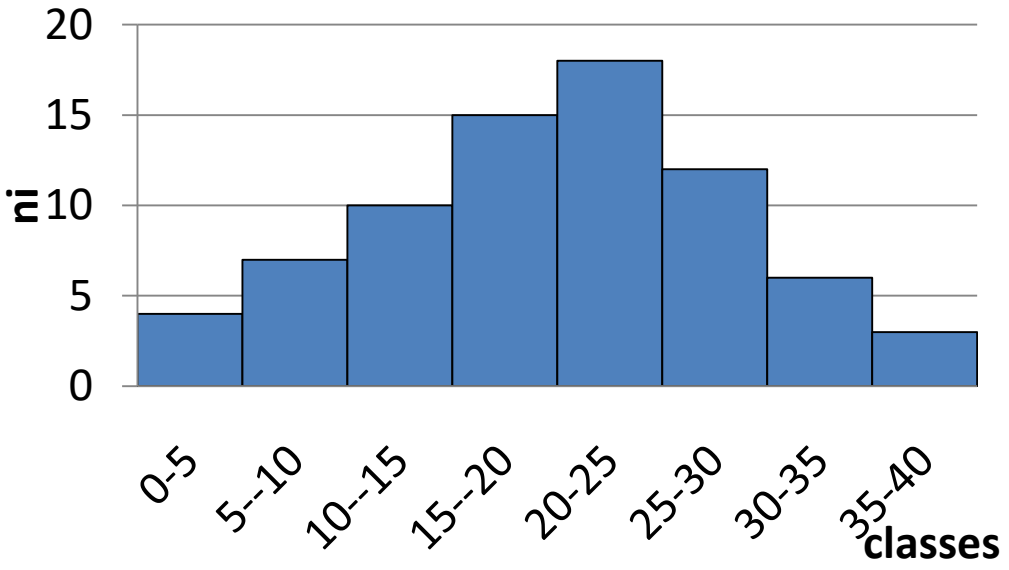
3- Graphical display of Continuous Quantitative variable

A- Histograms:

In statistics, data is often represented using a histogram. A histogram is constructed by dividing the data into a number of classes and then number in each class or frequency is represented by a verticle rectangle. The area of the retangle represents the frequency of each class.

Exemple : the table below gives the distribution of the phenomenon of being late to work for 75 workers in an institution

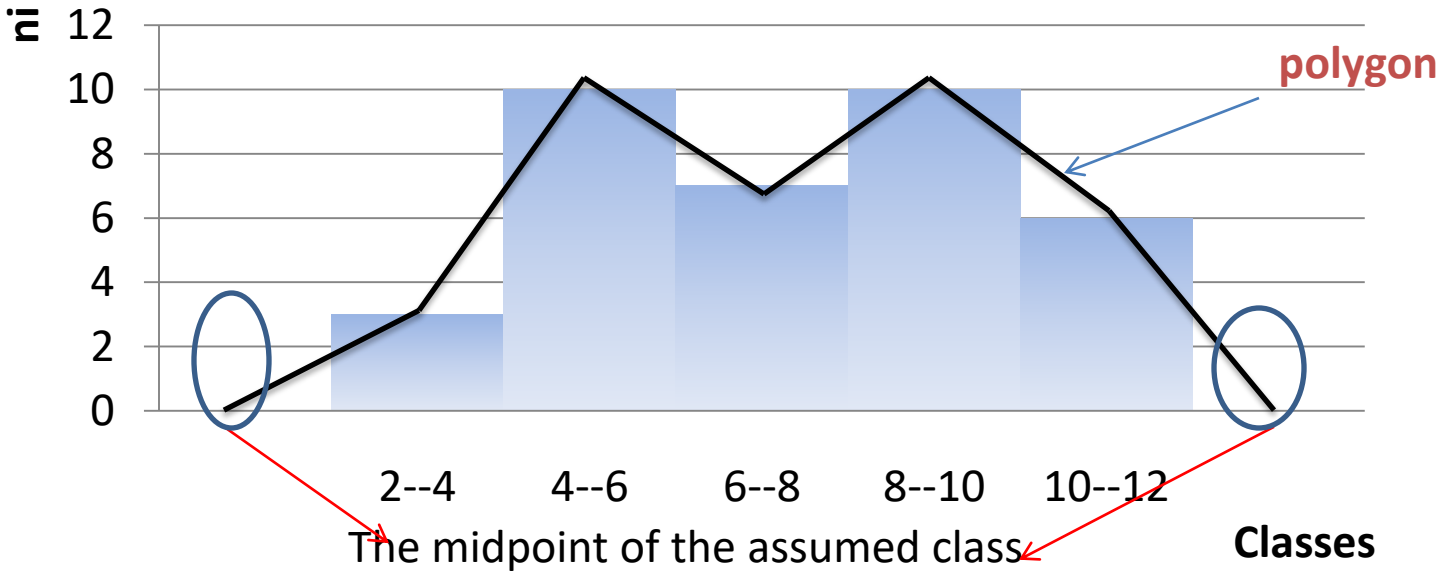
classes	ni
0-5	4
5-10	7
10-15	10
15-20	15
20-25	18
25-30	12
30-35	6
35-40	3



- **B- Polygon:**

the polygon is a closed geometric figure with straight sides. It is formed by connecting a serie of points , called vertices (class centrer and corresponding frequencies), with straight line segments.

Example of Graphic representation of the polygone



C- **Graphical representation of Cumulative frequency (ascending / descending)**

When represented graphically, ascending/ descending cumulative frequency is typically plotted with data values on the x-axis and cumulative frequencies on the y-axis. This creates a cumulative frequency , which shows the increasing/decreasing straight segments based on: the cumulative ascending / descending recurrent the upper limit of the class / the lower limit of the class.

Example : we use the same previous example to represent Ascending and Descending.