

DESCRIPTIVE STATISTICS

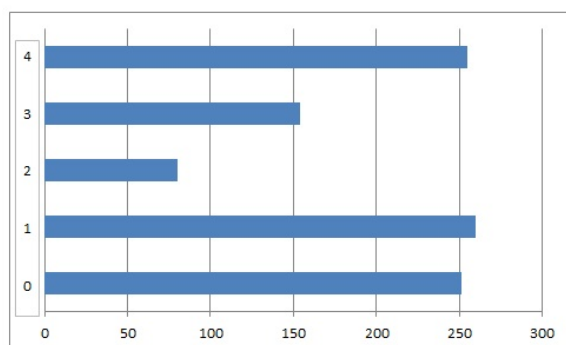
Some exercises

1: To the 1000 habitants of small town is asked to express an opinion on a new municipal service, using a scale from 0 to 4 (0 = poor, 4 = excellent). The responses obtained are summarized in the table below. Provide indices of position and variability adequate and a graphical representation of the appropriate answers.

judgemnt	0	1	2	3	4
frequence	251	260	80	154	255

Sol:

S_i	f_i	p_i	F_i	P_i
0	251	251/1000	251	251/1000
1	260	260/1000	511	511/1000
2	80	80/1000	591	591/1000
3	154	154/1000	735	735/1000
4	255	255/1000	1000	1000/1000



graphical representation

mean: $\bar{x} = 1.90$

median: $\hat{x} = 1$

variance: $\sigma^2 = 2.43$

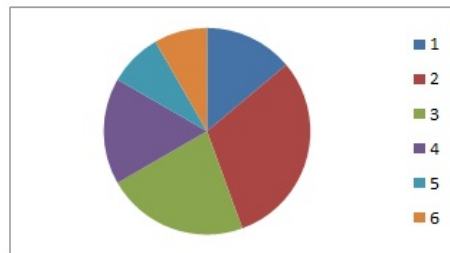
standard deviation $\sigma = 1.56$.

2: To a number of 36 individuals living in Turin is asked the number of times they went to see a movie in the last month. The 36 answers obtained are the following:

1, 3, 4, 2, 2, 4, 5, 5, 1, 1, 2, 3, 4, 3, 2, 6, 6, 1,
2, 2, 3, 2, 1, 3, 4, 2, 3, 3, 3, 5, 6, 4, 2, 2, 4, 2.

- Provide the relative frequencies and relative cumulative frequencies of the answers.
- Provide a Pie chart.
- Determine mean, mode, median and variance of the responses.
- The same survey was held in Tashkent, and the answers provided by 36 individuals gave an average value equal to 2.5 and a variance equal to 3.6. Compare the relative variability and comment.

Vani	f_i	p_i	P_i
1	5	5/36	5/36
2	11	11/36	16/36
3	8	8/36	24/36
4	6	6/36	30/36
5	3	3/36	33/36
6	3	3/36	36/36
Tot.	36	36/36	



Sol:

c)

$$\text{mean: } \bar{x} = 3$$

$$\text{median: } \hat{x} = 3$$

$$\text{mode: } \tilde{x} = 2$$

$$\sigma^2 = 2.11$$

d)

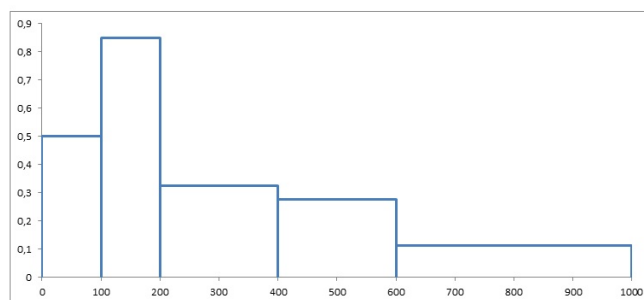
$$c.v.(TO) = \frac{\sqrt{2.11}}{3} = 0.48 \quad c.v.(TAS) = \frac{\sqrt{3.6}}{2.5} = 0.76$$

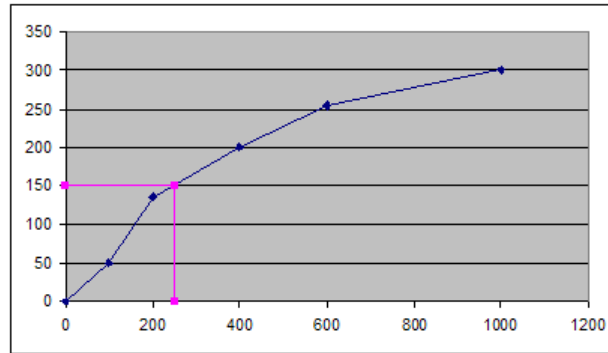
The variability is higher in Tashkent, both absolutely and relatively.

3: For 300 days are detected the overall consumption of electricity in a small mountain village. The data obtained, expressed in KW, are summarized in the table that follows. Provide a graphic representation, and find in an approximate way the mean, median and standard deviation of consumption.

class	[0, 100)	[100, 200)	[200, 400)	[400, 600)	[600, 1000]
freq.	50	85	65	55	45

Sol:





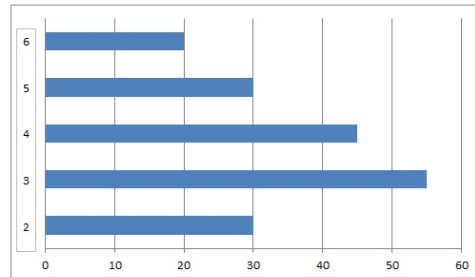
mean: $\bar{x} \simeq 327.5$
 median: $\tilde{x} \simeq 250$
 variance: $\sigma^2 \simeq 60870$
 standard deviation: $\sigma \simeq 247$

4: Given the following absolute frequencies (relative to the number of rooms of a population of 180 houses), determine the relative frequency and relative cumulative frequency, graph them, and determine the average, median, mode and standard deviation of the data.

rooms	2	3	4	5	6
freq.	30	55	45	30	20

Sol:

S_i	f_i	p_i	F_i	P_i
2	30	30/180	30	30/180
3	55	55/180	85	85/180
4	45	45/180	130	130/180
5	30	30/180	160	160/180
6	20	20/180	180	180/180



mean: $\bar{x} = 3.75$
 mode: $\tilde{x} = 3$
 median: $\hat{x} = 4$
 $\sigma^2 = 0.92$
 $\sigma = 0.96$

5: Given the following discrete quantitative data

$\{3, 5, 1, 1, 0, 2, 8, 2, 3, x\},$

with $x \in \mathbb{R}$, find an appropriate value for x such that

- a) $\bar{x} = 3$;
- b) $\hat{x} = 2$;
- c) $\hat{x} = 2.5$;
- d) $\hat{x} < \bar{x}$.

Given now $x = 1$,

- e) find the variance and the coefficient of variation of the data set;
- f) write the relative frequencies and the relative cumulated frequencies;
- g) graphically represent the distribution of data.

Sol:

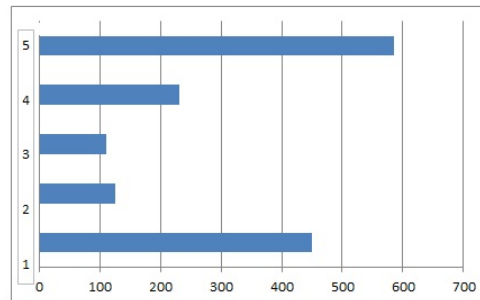
- a) $x = 5$
- b) $x = 2$
- c) $x = 3$
- d) $x = 1$
- e) $\bar{x} = 2.67, \sigma^2 = 5.56, c.v. = 0.88$

6: A data set of 1500 observations is summarized below. Provide indices of position and variability adequate and a graphical representation of the appropriate data.

measure	1	2	3	4	5
freq.	450	125	110	230	575

Sol:

S_i	f_i	p_i	F_i	P_i
1	450	450/1500	450	450/1500
2	125	125/1500	575	575/1500
3	110	110/1500	685	685/1500
4	230	230/1500	915	915/1500
5	585	575/1500	1500	1500/1500



mean: $\bar{x} = 3.22$

median: $\hat{x} = 4$

mode: $\hat{x} = 5$

$\sigma^2 = 2.96$

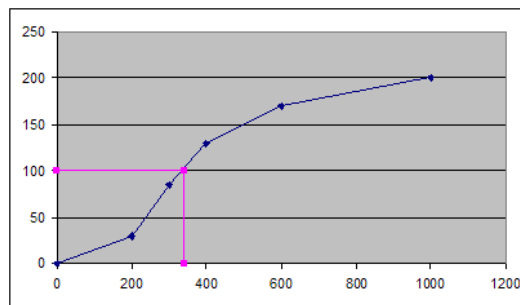
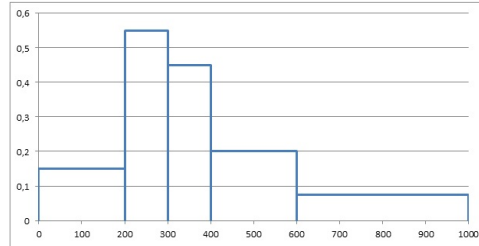
$\sigma = 1.72$

7: For 200 days the water consumption are detected in a small town. The obtained data, expressed in m^3 , are summarized in the table below. Represent the observed consumption graphically, and

find an approximation of the mean, median and standard deviation of consumption.

class	[0, 200)	[200, 300)	[300, 400)	[400, 600)	[600, 1000]
freq.	30	55	45	40	30

Sol:



mean: $\bar{x} \simeq 382.5$
 median: $\tilde{x} \simeq 334.4$
 $\sigma^2 \simeq 45950$, $\sigma \simeq 214$.

8: To 30 men has been asked how many books they have read in last month. The 30 responses received are as follows:

1, 0, 4, 2, 2, 4, 5, 0, 1, 1, 2, 3, 4, 3, 2,
 2, 2, 3, 0, 1, 3, 4, 0, 0, 3, 3, 5, 6, 4, 2.

- Provide the relative frequencies and relative cumulative frequencies of the answers.
- Provide a Pie chart.
- Determine mean, mode, median and variance of the responses.
- The same survey was held to 30 women, and the 30 answers gave an average value equal to 2.5 and a variance equal to 3. Compare the relative variability and comment.

Sol:

mean: $\bar{x} = 2.4$

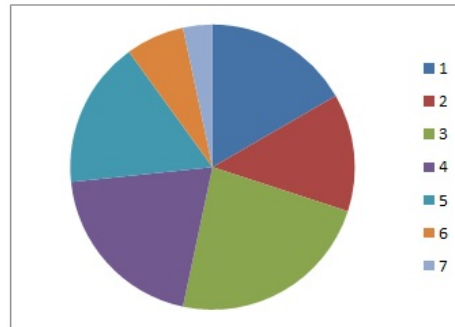
median: $\hat{x} = 2$

mode: $\hat{x} = 2$

$\sigma^2 = 2.64$, $\sigma = 1.62$.

$c.v.(inv) = 0.67$ $c.v.(est) = 0.69$ (almost the same relative variance)

Vani	f_i	p_i	P_i
0	5	5/30	5/30
1	4	4/30	9/30
2	7	7/30	16/30
3	6	6/30	22/30
4	5	5/30	27/30
5	2	2/30	29/30
6	1	1/30	30/30
Tot.	30		

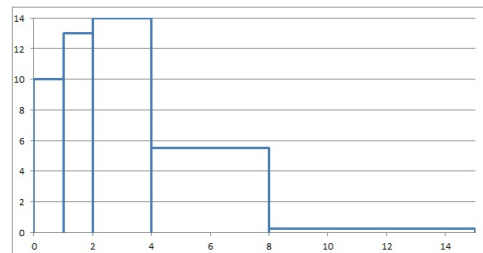


9: One hundred quantitative continuous data are grouped into classes as follows:

class	$[0, 1)$	$[1, 2)$	$[2, 4)$	$[4, 8)$	$[8, 15]$
freq.	10	13	28	22	27

- provide appropriate graphical representations;
- provide an approximation of mean, median and variance;

Sol:



$$\begin{aligned}
 \text{mean: } \bar{x} &\simeq 5.51 \\
 \text{median: } \hat{x} &\simeq 3.9 \\
 \sigma^2 &\simeq 16.1, \quad \sigma \simeq 4.
 \end{aligned}$$