



**GDAŃSK UNIVERSITY  
OF TECHNOLOGY**

FACULTY OF MANAGEMENT AND ECONOMICS

# **ESSENTIALS OF STATISTICS**

## **NO. II.**

### **MEASURES OF LOCATION**

**KAROLINA TURA, PHD**

**DEPARTMENT OF ECONOMIC SCIENCE**

# AGENDA

## 1. Measures of locations

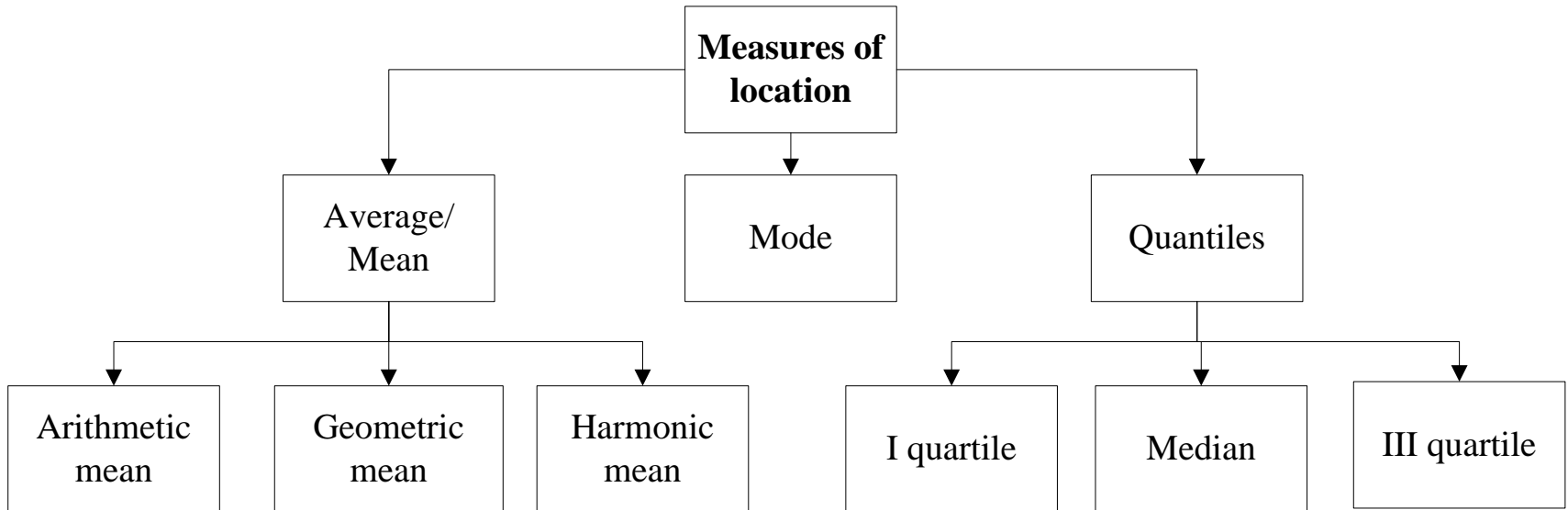
1. Average

2. Mode

3. Quantiles

## 2. Practice

# MEASURES OF LOCATION



# ARHITMETIC MEAN

Detailed form

$$\bar{x} = \frac{\sum_{i=1}^{n_i} x_i}{N}$$

Average value of ...

Frequency table

$$\bar{x} = \frac{\sum_{i=1}^k x_i n_i}{N}$$

Frequencies of each class interval

Midpoint of each class interval

Frequency table with equal interval class

$$\bar{x} = \frac{\sum_{i=1}^k \dot{x}_i n_i}{N}$$

$\bar{x}$

Sample mean

$\mu$

Population mean

# TASK 1. DETAILED FORM

15 students scored the points on the exam of Statistics:

1,2,3,4,4,5,5,5,6,6,6,6,7,8,8.

Find and interpret:

- a) mean,
- b) mode,
- c) median,
- d) I quartile,
- e) III quartile.

$$\bar{x} = \frac{\sum_{i=1}^{n_i} x_i}{n} = \frac{76}{15} \approx 5$$

# TASK 2. FREQUENCY TABLE

The number of hours (per week) which students spend on learning Statistics in 2014 is given in table.

Find and interpret:

- a) mean,
- b) mode,
- c) median.

$x$ Hours (per week)	$n$ Frequency
0	25
1	54
2	11
3	9
4	1

$$\bar{x} = \frac{\sum_{i=1}^k x_i n_i}{N} = \frac{107}{100} = 1.07$$

## HINT

	$x_i$	$n$	$x_i n$	$n_{cum}$
	0,00	25,00	0,00	25,00
	1,00	54,00	54,00	79,00
	2,00	11,00	22,00	90,00
	3,00	9,00	27,00	99,00
	4,00	1,00	4,00	100,00
Summary		100,00	107,00	393,00

# TASK 3. FREQUENCY TABLE WITH EQUAL CLASS INTERVALS

Observations of consumer credit borrowers were collected at the Bank X in 2008.

Results were grouped in a series of observations given in the table below.

Find and interpret:

- mean,
- mode,
- median.

Credit [zł] $x$	Credit borrowers $n$
0-5	300
5-10	470
10-15	693
15-20	328
20-25	120
25-30	35

## HINT

	$x_i$	$n$	$n_{cum}$	$\dot{x}$	$\dot{x}_i n$
	0-5	300,00	300,00	2,50	750,00
	5-10	470,00	770,00	7,50	3525,00
	10-15	693,00	1463,00	12,50	8662,50
	15-20	328,00	1791,00	17,50	5740,00
	20-25	120,00	1911,00	22,50	2700,00
	25-30	35,00	1946,00	27,50	962,50
Summary		1946,00	8181,00	90,00	22340,00

$$\bar{x} = \frac{\sum_{i=1}^k \dot{x}_i n_i}{N} = \frac{22340}{1946} \approx 11.48$$

# MODE

The value that occurs most often...

The level of variable which occurs with the greatest frequency...

The measurement that occurs most frequently in a data set ...

1. We have to put the values into ascending order

The lower bound of the range with the largest frequency

Width of the range with the largest frequency

Frequency of the range with largest frequency

$$D = x_d + h_d \frac{n_d - n_{d-1}}{(n_d - n_{d-1}) + (n_d - n_{d+1})}$$

Frequency table with equal interval class



# TASK 1. DETAILED FORM

15 students scored the points on the exam of Statistics:

1,2,3,4,4,5,5,5,6,6,6,6,7,8,8.

Find and interpret:

- a) mean,
- b) mode,
- c) median,
- d) I quartile,
- e) III quartile.

6

# TASK 2. FREQUENCY TABLE

The number of hours (per week) which students spend on learning Statistics in 2014 is given in table.

Find and interpret:

- a) mean,
- b) mode,
- c) median.

$x$ Hours (per week)	$n$ Frequency
0	25
1	54
2	11
3	9
4	1

1

## HINT

	$x_i$	$n$	$x_i n$	$n_{cum}$
	0,00	25,00	0,00	25,00
	1,00	54,00	54,00	79,00
	2,00	11,00	22,00	90,00
	3,00	9,00	27,00	99,00
	4,00	1,00	4,00	100,00
Summary		100,00	107,00	393,00

# TASK 3. FREQUENCY TABLE WITH EQUAL CLASS INTERVALS

Observations of consumer credit borrowers were collected at the Bank X in 2008.

Results were grouped in a series of observations given in the table below.

Find and interpret:

- a) mean,
- b) mode,
- c) median.

Credit [zł] $x$	Credit borrowers $n$
0-5	300
5-10	470
10-15	693
15-20	328
20-25	120
25-30	35

$$D = x_d + h_d \frac{n_d - n_{d-1}}{(n_d - n_{d-1}) + (n_d - n_{d+1})} =$$

$$= 10 + 5 * \frac{693 - 470}{(693 - 470) + (693 - 328)} \approx$$

$$\approx 11.9$$

## HINT

	$x_i$	$n$	$n_{cum}$	$\dot{x}$	$\dot{x}_i n$
	0-5	300,00	300,00	2,50	750,00
	5-10	470,00	770,00	7,50	3525,00
	10-15	693,00	1463,00	12,50	8662,50
	15-20	328,00	1791,00	17,50	5740,00
	20-25	120,00	1911,00	22,50	2700,00
	25-30	35,00	1946,00	27,50	962,50
Summary		1946,00	8181,00	90,00	22340,00

# MEDIAN

The median of a quantitative data set is the middle number when the measurements are arranged in ascending (or descending) order...

The midpoint of the distribution- the number such as half of observations are smaller and half are larger

For a median, 50% of the data are less than it, and 50% of the data are bigger than it

$$PosMe = \frac{n}{2} \quad N \text{ is even}$$

$$PosMe = \frac{n + 1}{2} \quad N \text{ is odd}$$

## Frequency table with equal class intervals

Lower bound of the median range

Cumulative frequency of the range before the median range

$$Me = x_{Me} + (PosMe - n_{cum, n-1}) \frac{h_{Me}}{n_{Me}}$$

$h_{Me}$  ← Width of the median range  
 $n_{Me}$  ← Frequency of the median range

# TASK 1. DETAILED FORM

15 students scored the points on the exam of Statistics:

1,2,3,4,4,5,5,5,6,6,6,6,7,8,8.

Find and interpret:

- a) mean,
- b) mode,
- c) median,
- d) I quartile,
- e) III quartile.

$$PosMe = \frac{n + 1}{2} = \frac{15 + 1}{2} = 8 \rightarrow Me = 5$$

# TASK 2. FREQUENCY TABLE

The number of hours (per week) which students spend on learning Statistics in 2014 is given in table.

Find and interpret:

- a) mean,
- b) mode,
- c) median.

$x$ Hours (per week)	$n$ Frequency
0	25
1	54
2	11
3	9
4	1

$$PosMe = \frac{n}{2} = \frac{100}{2} = 50 \rightarrow Me = 1$$

## HINT

	$x_i$	$n$	$x_i n$	$n_{cum}$
	0,00	25,00	0,00	25,00
	1,00	54,00	54,00	79,00
	2,00	11,00	22,00	90,00
	3,00	9,00	27,00	99,00
	4,00	1,00	4,00	100,00
Summary		100,00	107,00	393,00

# TASK 3. FREQUENCY TABLE WITH EQUAL CLASS INTERVALS

Observations of consumer credit borrowers were collected at the Bank X in 2008.

Results were grouped in a series of observations given in the table below.

Find and interpret:

- a) mean,
- b) mode,
- c) median.

Credit [zł] $x$	Credit borrowers $n$
0-5	300
5-10	470
10-15	693
15-20	328
20-25	120
25-30	35

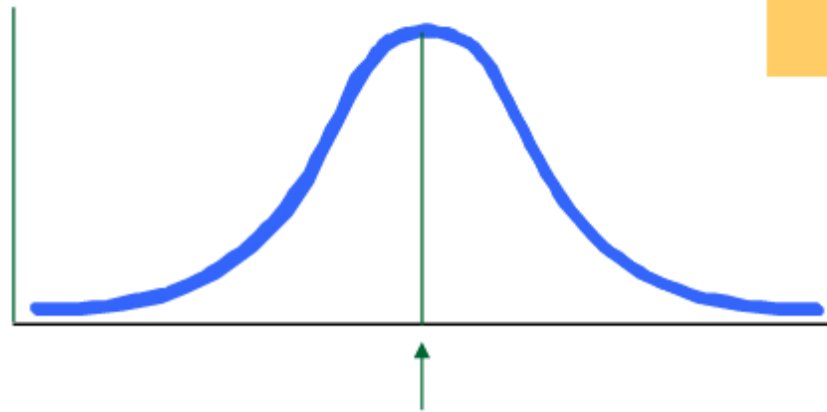
$$PosMe = \frac{n}{2} = \frac{1946}{2} = 973$$

$$Me = x_{Me} + (PosMe - n_{cum, n-1}) \frac{h_{Me}}{n_{Me}} = 10 + (973 - 770) \frac{5}{693} = 11.5$$

## HINT

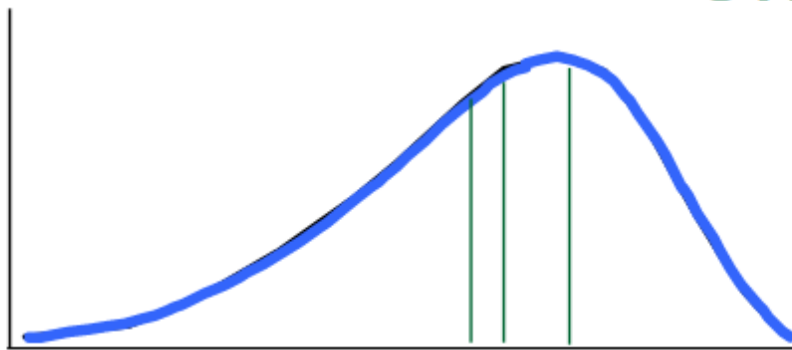
	$x_i$	$n$	$n_{cum}$	$\dot{x}$	$\dot{x}_i n$
	0-5	300,00	300,00	2,50	750,00
	5-10	470,00	770,00	7,50	3525,00
	10-15	693,00	1463,00	12,50	8662,50
	15-20	328,00	1791,00	17,50	5740,00
	20-25	120,00	1911,00	22,50	2700,00
	25-30	35,00	1946,00	27,50	962,50
Summary		1946,00	8181,00	90,00	22340,00

The mean is pulled toward the skew.



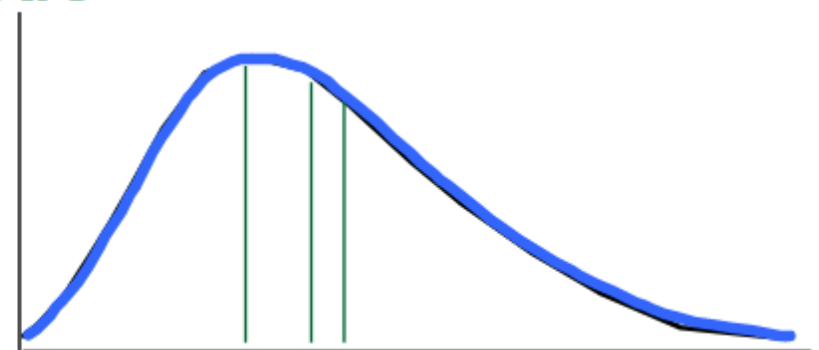
Mode = Mean = Median

**SYMMETRIC**



Mean ——— Median ——— Mode

**SKEWED LEFT**  
(negatively)



Mode ——— Median ——— Mean

**SKEWED RIGHT**  
(positively)



# QUANTILES

The I quartile of a quantitative data set is the 1/4 number when the measurements are arranged in ascending (or decsending) order...

For a I Quartile, 25% of the data are less than it, and 75% of the data are bigger than it

$$PosQ_1 = \frac{n}{4} \quad N \text{ is even}$$

$$PosQ_1 = \frac{n + 1}{4} \quad N \text{ is odd}$$

## Frequency table with equal class intervals

Lower bound of the quartile range

Cumulative frequency of the range before the quartile range

$$Q_1 = x_{Q_1} + (PosQ_1 - n_{cum, n-1}) \frac{h_{Q_1}}{n_{Q_1}}$$

Width of the quartile range

Frequency of the quartile range

# QUANTILES

The III quartile of a quantitative data set is the  $\frac{3}{4}$  number when the measurements are arranged in ascending (or decsending) order...

For a III Quartile, 75% of the data are less than it, and 35% of the data are bigger than it

$$PosQ_3 = \frac{3n}{4} \quad N \text{ is even}$$

$$PosQ_3 = \frac{3(n+1)}{4} \quad N \text{ is odd}$$

## Frequency table with equal class intervals

Lower bound of the quartile range

Cumulative frequency of the range before the quartile range

$$Q_3 = x_{Q_3} + (PosQ_3 - n_{cum, n-1}) \frac{h_{Q_3}}{n_{Q_3}}$$

Width of the quartile range

Frequency of the quartile range

# TASK 1. DETAILED FORM

15 students scored the points on the exam of Statistics:

1,2,3,4,4,5,5,5,6,6,6,6,7,8,8.

Find and interpret:

- a) mean,
- b) mode,
- c) median,
- d) I quartile,
- e) III quartile.

$$PosQ_1 = \frac{n+1}{4} = \frac{15+1}{4} = 4 \rightarrow Q_1 = 4$$

$$PosQ_3 = \frac{3(n+1)}{4} = \frac{3(15+1)}{4} = 12 \rightarrow Q_3 = 6$$

# TASK 2. FREQUENCY TABLE

The number of hours (per week) which students spend on learning Statistics in 2014 is given in table.

Find and interpret:

- mean,
- mode,
- median,
- I quartile,
- III quartile.

$x$ Hours (per week)	$n$ Frequency
0	25
1	54
2	11
3	9
4	1

$$PosQ_1 = \frac{n}{4} = \frac{100}{4} = 25 \rightarrow Q_1 = 0$$

$$PosQ_3 = \frac{3n}{4} = \frac{3 * 100}{4} = 75 \rightarrow Q_3 = 1$$

## HINT

	$x_i$	$n$	$x_i n$	$n_{cum}$
	0,00	25,00	0,00	25,00
	1,00	54,00	54,00	79,00
	2,00	11,00	22,00	90,00
	3,00	9,00	27,00	99,00
	4,00	1,00	4,00	100,00
Summary		100,00	107,00	393,00

# TASK 3. FREQUENCY TABLE WITH EQUAL CLASS INTERVALS

Observations of consumer credit borrowers were collected at the Bank X in 2008.

Results were grouped in a series of observations given in the table below.

Find and interpret:

- a) mean,
- b) mode,
- c) median
- d) I quartile,
- e) III quartile.

Credit [zł] $x$	Credit borrowers $n$
0-5	300
5-10	470
10-15	693
15-20	328
20-25	120
25-30	35

## HINT

	$x_i$	$n$	$n_{cum}$	$\dot{x}$	$\dot{x}_i n$
	0-5	300,00	300,00	2,50	750,00
	5-10	470,00	770,00	7,50	3525,00
	10-15	693,00	1463,00	12,50	8662,50
	15-20	328,00	1791,00	17,50	5740,00
	20-25	120,00	1911,00	22,50	2700,00
	25-30	35,00	1946,00	27,50	962,50
Summary		1946,00	8181,00	90,00	22340,00

$$PosQ_1 = \frac{n}{4} = \frac{1946}{4} = 486.5$$

$$Q_1 = x_{Q_1} + (PosQ_1 - n_{cum, n-1}) \frac{h_{Q_1}}{n_{Q_1}} =$$

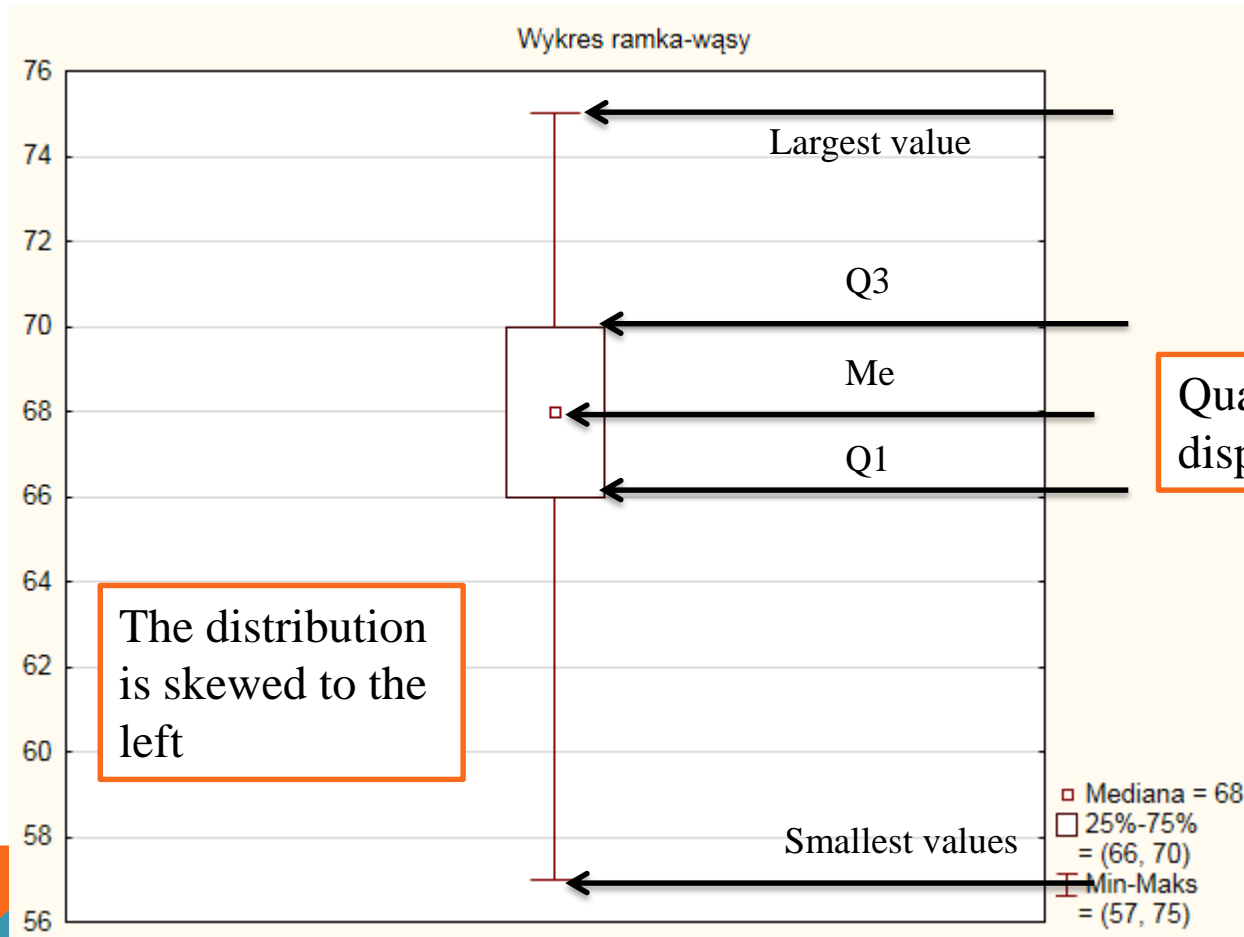
$$= 5 + (486.5 - 300) \frac{5}{470} \approx 6.98$$

$$PosQ_3 = \frac{3n}{4} = \frac{3 * 1946}{4} = 1459.5$$

$$Q_3 = x_{Q_3} + (PosQ_3 - n_{cum, n-1}) \frac{h_{Q_3}}{n_{Q_3}} =$$

$$= 10 + (1459.5 - 770) \frac{5}{693} \approx 14.97$$

# BOX WHISKER PLOTS



**STATISTICA**

## TASK 4.

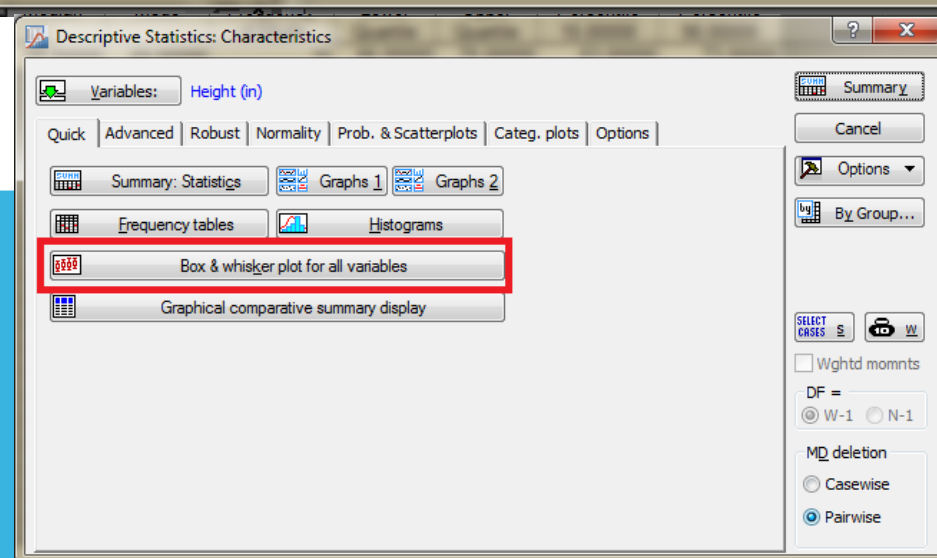
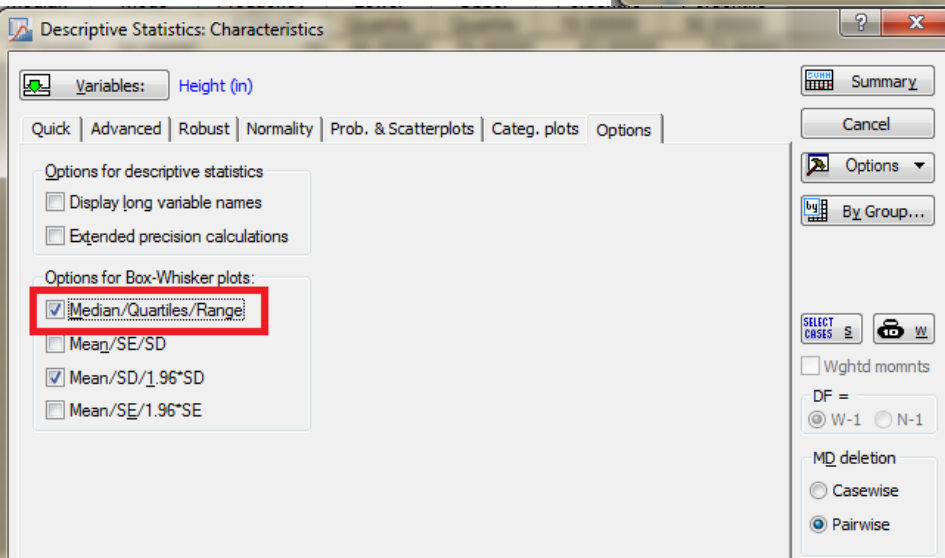
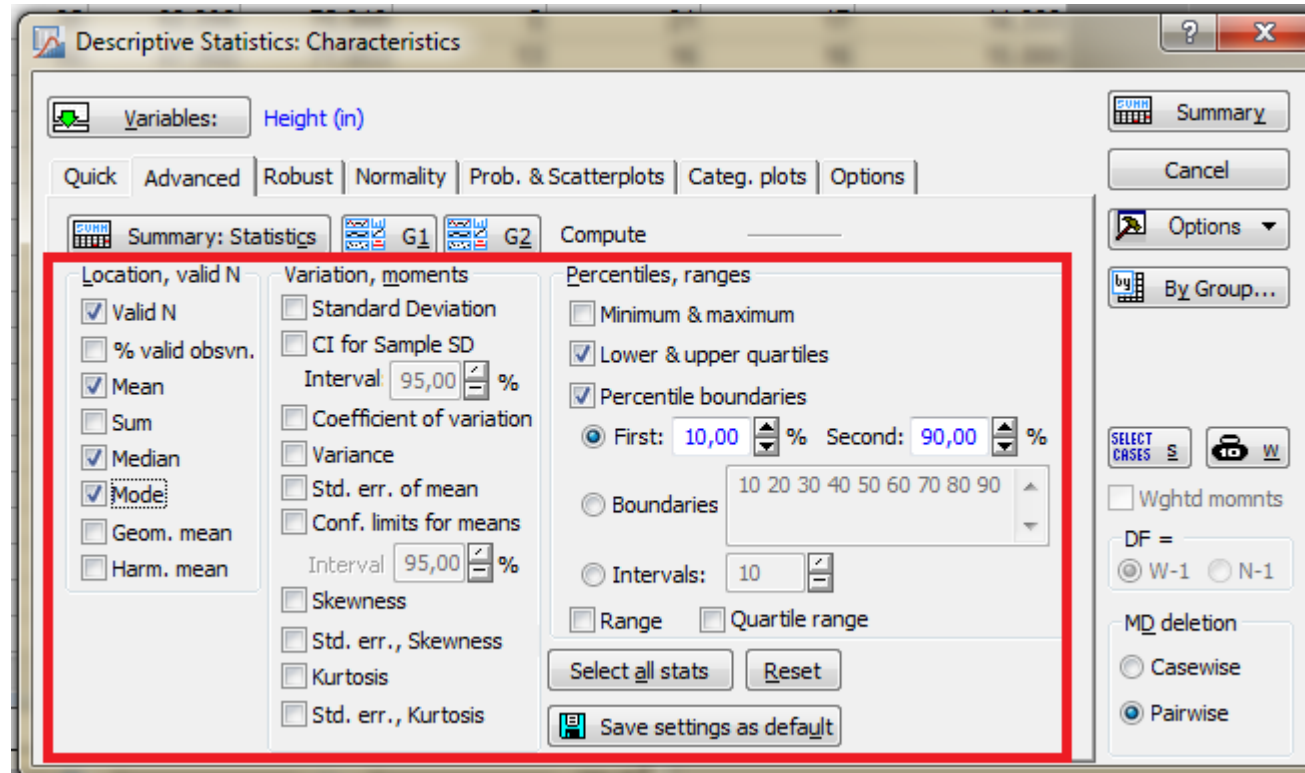
Scientists examined height of randomly selected men from the city of Gdynia. On the basis of the data contained in the file CharacteristicsHeight.sta perform an analysis. Create a box whisker plot.

Find and interpret:

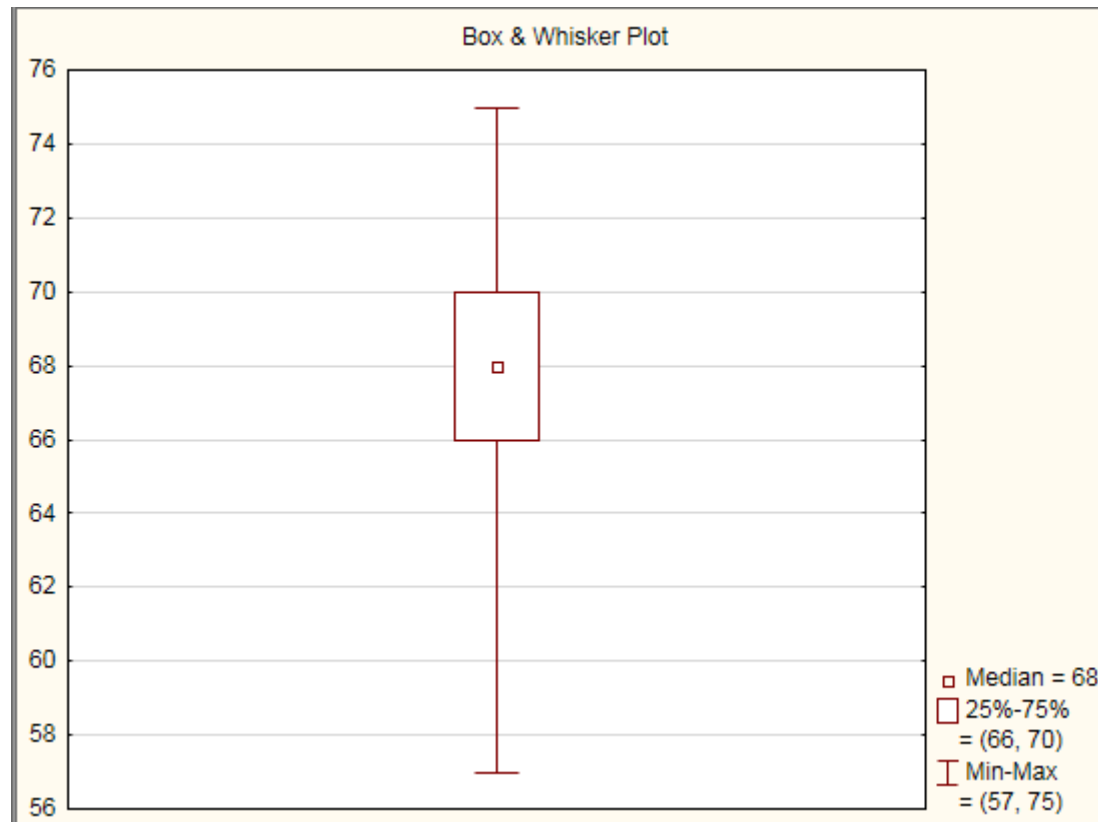
- a) mean,
- b) mode,
- c) median,
- d) I quartile,
- e) III quartile,
- f) I decile.



# HINT



# HINT



Variable	Descriptive Statistics (Characteristics)								
	Valid N	Mean	Median	Mode	Frequency of Mode	Lower Quartile	Upper Quartile	Percentile 10,00000	Percentile 90,00000
<b>Height (in)</b>	100	67,89000	68,00000	69,00000	15	66,00000	70,00000	63,00000	73,00000

# PREPARATION FOR THE NEXT CLASSES

McClave, J. T., Benson, P. G., Sincich, T. (2008) , *Statistics for Business & Economics*, Pearson Education Inc., New Jersey, p. 74-100;

**Thank you for your  
attention**



**GDAŃSK UNIVERSITY  
OF TECHNOLOGY**

FACULTY OF MANAGEMENT AND ECONOMICS