



**GDAŃSK UNIVERSITY  
OF TECHNOLOGY**

FACULTY OF MANAGEMENT AND ECONOMICS

# **ESSENTIALS OF STATISTICS**

## **NO. I.**

### **METHODS OF DESCRIBING SETS OF DATA**

**KAROLINA TURA, PHD**

**DEPARTMENT OF ECONOMIC SCIENCE**

# AGENDA

## 1. Theory

1. Basic concepts

2. Frequency tables

## 2. Practice

# BASIC CONCEPTS

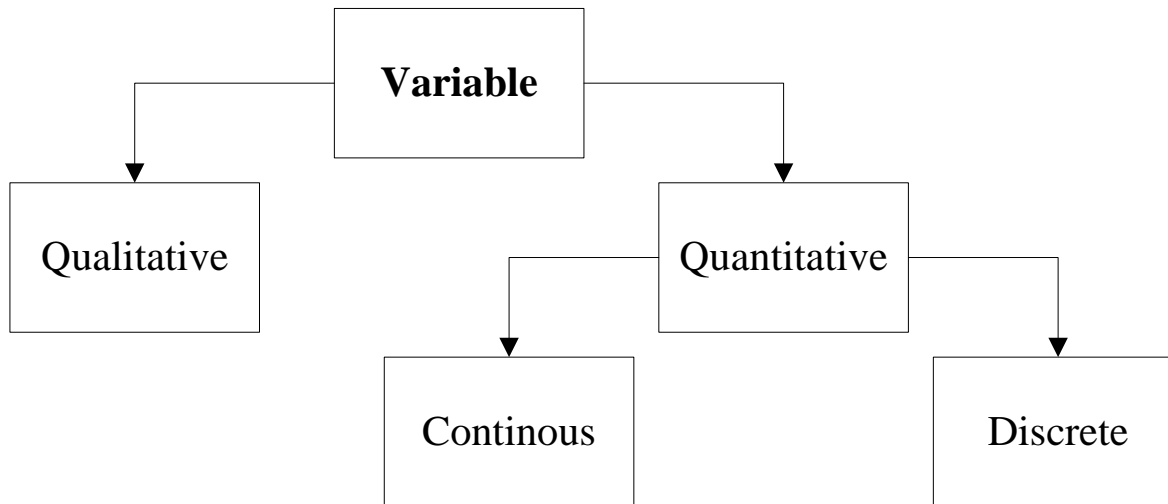
1. **Statistics-** the science of data. It involves collecting, classifying, summarising, analyzing and interpreting numerical information.
2. **Descriptive statistics-** utilizes numerical and graphical techniques to look for patterns in a data set, to summarize the information revealed in a data set and to present the information in a convenient form.

# BASIC CONCEPTS

1. **Population-** a set of units that we are interesting in studying
2. **Statistical unit-** an object upon which we collect the data
3. **Sample-** a subset of the units of a population

# BASIC CONCEPTS

1. **Variable**- characteristic or a property of an individual experimental unit



# EXAMPLE 1.

„Cola wars” is the popular term for the intense competition between Coca-Cola and Pepsi. Suppose, as part of Pepsi marketing campaign, 1 000 cola consumers are given a blind taste test. Each consumer is asked to state a preference for brand „Coca-Cola” or „Pepsi”.

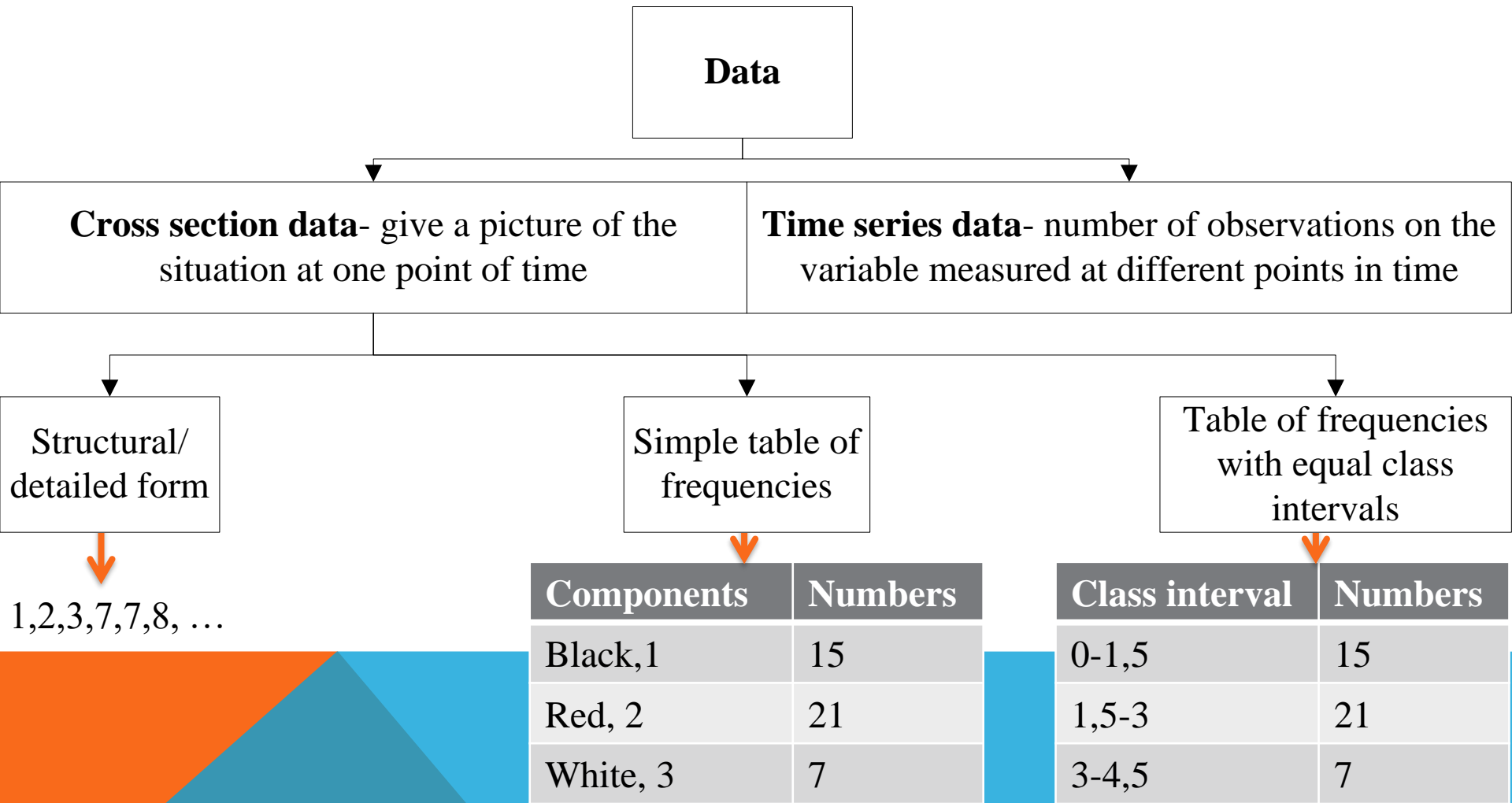
- a) Describe the population.
- b) Describe the statistical unit.
- c) Describe the sample.
- d) Describe the variable of interest.

## EXAMPLE 2.

The average age of viewers of „Dancing with the stars” is 50 years. Suppose a rival network decided to check this hypothesis. To test hypothesis, it samples 500 viewers of „Dancing with the stars” and determines the age of each.

- a) Describe the population.
- b) Describe the statistical unit.
- c) Describe the sample.
- d) Describe the variable of interest.

# BASIC CONCEPTS- DATA





# HOW TO CREATE A SIMPLE FREQUENCY TABLE?

## EXAMPLE

Show the proportion that fall into each class

Show the total number of individuals, obtained by cumulating frequencies

	Frequency	Relative frequency [%]	Cumulative frequency	Cumulative relative frequency [%]
cats	5	33%	5	-
dogs	3	20%	8	57%
fish	2	14%	10	66%
No pet	5	33%	15	100%

*Relative frequency = frequency / sum of frequencies \* 100%*

Do you have pets?

Dan	Maria	Inga	Steve	Jamie	Eve	John	Ping	Ivo	Chan	Carol	Ana	Basil	Mario	Alex
fish	cat	-	cat	-	-	dog	-	cat	dog	cat	fish	cat	dog	-

# HOW TO CREATE A FREQUENCY TABLE WITH EQUAL INTERVALS?

1. How to approximate number of class intervals?

Number of class intervals  $\rightarrow k = \left\lfloor \sqrt{n} \right\rfloor$

Floor of the number

Total frequency

*Rule of thumb*

2. How to approximate class width?

Largest value    Smallest value

Class width  $\rightarrow h = \frac{x_{\max} - x_{\min}}{k}$

Number of class intervals

# HOW TO CREATE A FREQUENCY TABLE WITH EQUAL INTERVALS?

## EXAMPLE

Suppose that, given the job of collecting and summarising relevant data on the firm's calls, you record the duration of a sample of 30 long distance calls placed in a given week. The results are shown in table. Create a frequency distribution with equal class intervals.

11,8	3,7	16,6	13,5	4,8	8,3
8,9	9,1	7,7	4,3	12,1	6,1
10,2	8	11,4	6,8	9,6	18,5
15,3	12,3	8,5	15,9	18,7	11,7
6,2	11,2	10,4	7,2	5,5	14,5

$$k = [\sqrt{30}] = 5$$

$$h = \frac{x_{\max} - x_{\min}}{k} = (18,7 - 3,7) / 5 = 3$$

Range	Frequency	Relative frequency [% ]	Cumulative frequency	Cumulative relative frequency [% ]
<3,7;6,7)	6	20	6	
<6,7;9,7)	9	30	15	50
<9,7;12,7)	8	26,7	23	76,7
<12,7;15,7)	4	13,3	27	90
<15,7;18,7)	3	10	30	100

**STATISTICA**

# SIMPLE FREQUENCY TABLES EXAMPLE

Statistics>Basic Statistics/Tables>Frequency tables

The image displays three sequential screenshots of the SPSS software interface, illustrating the steps to generate a frequency table for a specific variable.

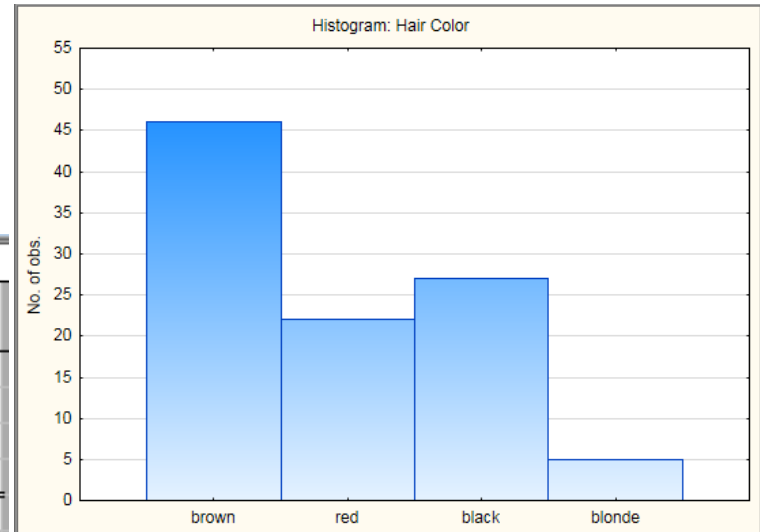
**First Screenshot: Basic Statistics and Tables: Characteristics**  
This dialog box shows the 'Quick' tab with a list of statistical options. The 'Frequency tables' option is highlighted with a red box. Other options include Descriptive statistics, Correlation matrices, and various t-tests.

**Second Screenshot: Frequency Tables: Characteristics**  
This dialog box is titled 'Frequency Tables: Characteristics' and shows the 'Variables:' field set to 'none'. The 'Display options for frequency tables' section is expanded, and the following options are checked and highlighted with a red box:  
 Cumulative frequencies  
 Percentages (relative frequencies)  
 Cumulative percentages  
Other options like '100% minus cumulative percentages', 'Logit transformed proportions', and 'Normal expected frequencies' are unchecked.

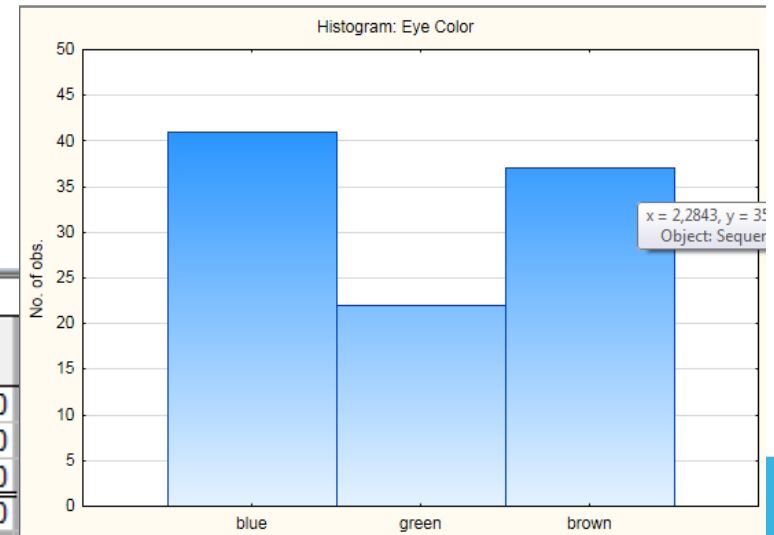
**Third Screenshot: Frequency Tables: Characteristics**  
This dialog box is also titled 'Frequency Tables: Characteristics' but with the 'Variables:' field set to 'Eye Color'. The 'Summary:' section is expanded, and the 'Histograms' option is highlighted with a red box. Other options include 'Descriptive statistics' and '3D histograms, bivariate distributions'.

# EXAMPLE

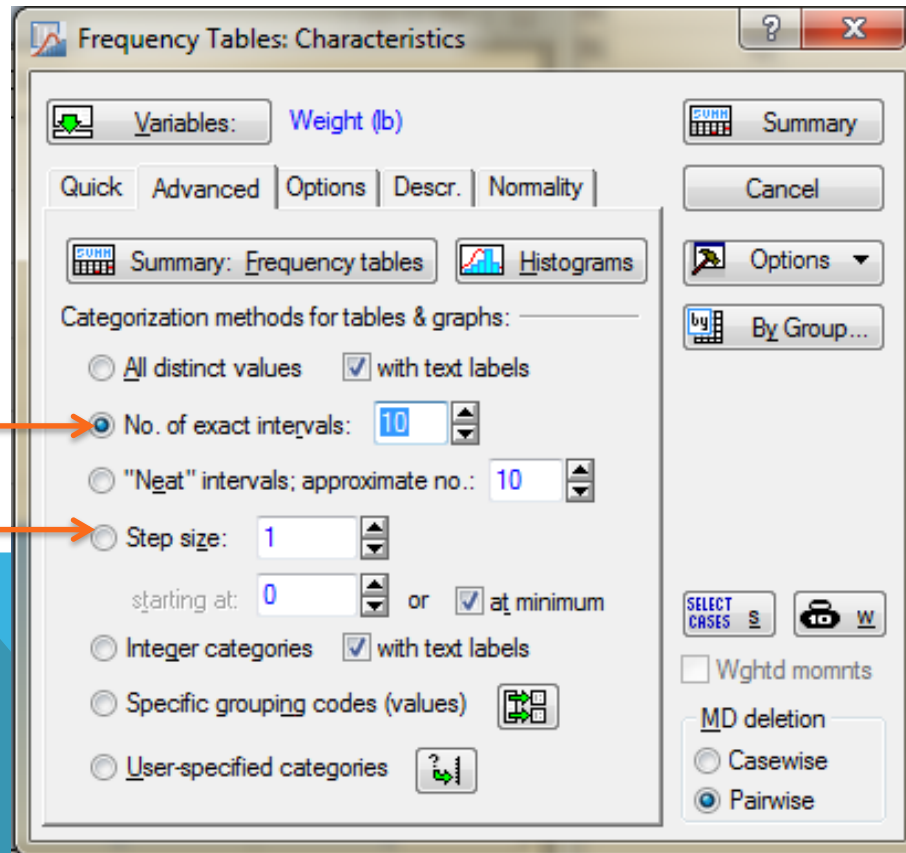
Frequency table: Hair Color (Characteristics)				
Category	Count	Cumulative Count	Percent	Cumulative Percent
<b>brown</b>	46	46	46,00000	46,0000
red	22	68	22,00000	68,0000
black	27	95	27,00000	95,0000
blonde	5	100	5,00000	100,0000
Missing	0	100	0,00000	100,0000



Frequency table: Eye Color (Characteristics)				
Category	Count	Cumulative Count	Percent	Cumulative Percent
<b>blue</b>	41	41	41,00000	41,0000
green	22	63	22,00000	63,0000
brown	37	100	37,00000	100,0000
Missing	0	100	0,00000	100,0000



# FREQUENCY TABLE WITH EQUAL CLASS INTERVALS EXAMPLE



Number of class intervals

Class width

# EXAMPLE

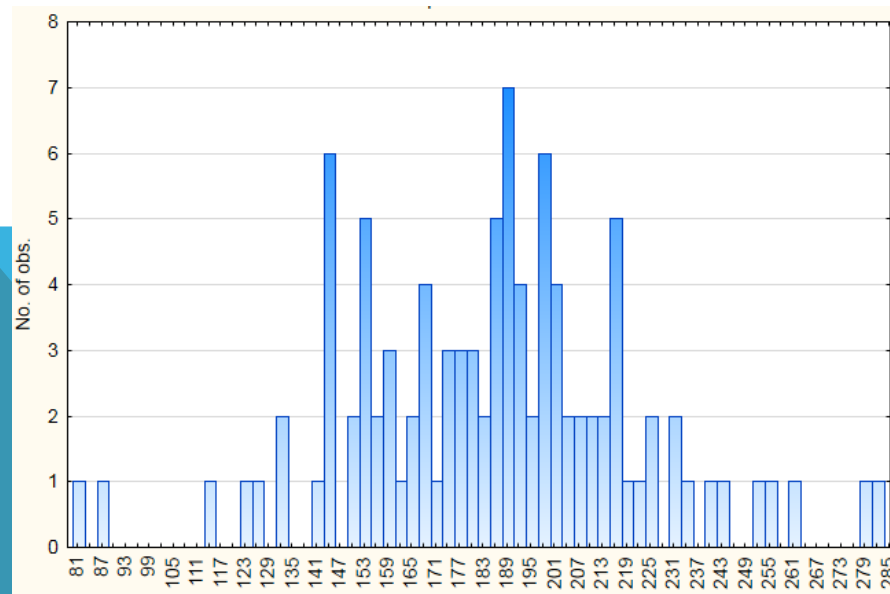
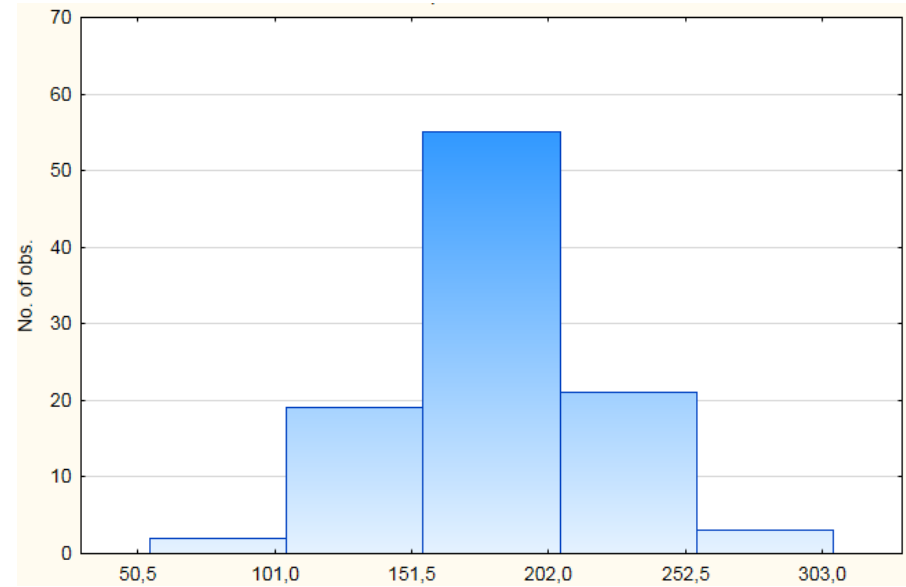
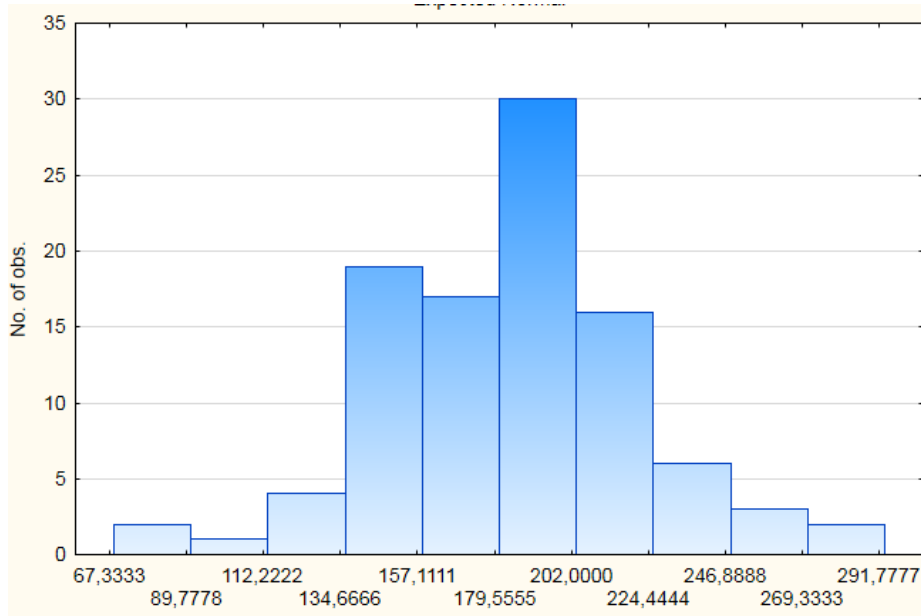
		Frequency table: Weight (lb) (Characteristics)			
From	To	Count	Cumulative Count	Percent	Cumulative Percent
68,7777777777778	<x<=91,2222222222222	2	2	2,00000	2,0000
91,2222222222222	<x<=113,6666666666667	1	3	1,00000	3,0000
113,6666666666667	<x<=136,1111111111111	4	7	4,00000	7,0000
136,1111111111111	<x<=158,5555555555556	19	26	19,00000	26,0000
158,5555555555556	<x<=181	17	43	17,00000	43,0000
181	<x<=203,4444444444444	30	73	30,00000	73,0000
203,4444444444444	<x<=225,8888888888889	16	89	16,00000	89,0000
225,8888888888889	<x<=248,3333333333333	6	95	6,00000	95,0000
248,3333333333333	<x<=270,7777777777778	3	98	3,00000	98,0000
270,7777777777778	<x<=293,2222222222222	2	100	2,00000	100,0000
Missing		0	100	0,00000	100,0000

		Frequency table: Weight (lb) (Characteristics)			
From	To	Count	Cumulative Count	Percent	Cumulative Percent
54,75	<x<=105,25	2	2	2,00000	2,0000
105,25	<x<=155,75	19	21	19,00000	21,0000
155,75	<x<=206,25	55	76	55,00000	76,0000
206,25	<x<=256,75	21	97	21,00000	97,0000
256,75	<x<=307,25	3	100	3,00000	100,0000
Missing		0	100	0,00000	100,0000

		Frequency table: Weight (lb) (Characteristics)			
From	To	Count	Cumulative Count	Percent	Cumulative Percent
80	<=x<83	1	1	1,000000	1,0000
83	<=x<86	0	1	0,000000	1,0000
86	<=x<89	1	2	1,000000	2,0000
89	<=x<92	0	2	0,000000	2,0000
92	<=x<95	0	2	0,000000	2,0000
95	<=x<98	0	2	0,000000	2,0000
98	<=x<101	0	2	0,000000	2,0000
101	<=x<104	0	2	0,000000	2,0000
104	<=x<107	0	2	0,000000	2,0000
107	<=x<110	0	2	0,000000	2,0000
110	<=x<113	0	2	0,000000	2,0000
113	<=x<116	1	3	1,000000	3,0000
116	<=x<119	0	3	0,000000	3,0000
119	<=x<122	0	3	0,000000	3,0000
122	<=x<125	1	4	1,000000	4,0000
125	<=x<128	1	5	1,000000	5,0000
128	<=x<131	0	5	0,000000	5,0000
131	<=x<134	2	7	2,000000	7,0000
134	<=x<137	0	7	0,000000	7,0000



# EXAMPLE



# PREPARATION FOR THE NEXT CLASSES

McClave, J. T., Benson, P. G., Sincich, T. (2008) , *Statistics for Business & Economics*, Pearson Education Inc., New Jersey, p. 63-74;

**Thank you for your  
attention**



**GDAŃSK UNIVERSITY  
OF TECHNOLOGY**

FACULTY OF MANAGEMENT AND ECONOMICS