Mathematical Statistics Problem Sets 1&2
Index Numbers. Descriptive Statistics

1. The dynamics of gross wages in 2012-2017 in Poland may be summarized by the following series of chained simple indices:

Year	2013	2014	2015	2016	2017
previous year =100	103.6	103.3	103.7	103.4	105.5

   Find the change between 2013 and 2017 (a simple index number with a base of 2013). Find the average rate of change between 2013 and 2017. Assuming the same average rate of change will apply to subsequent years, approximate the value of the average gross wage in 2018, if the average gross wage in 2017 was equal to 4528 PLN.

2. The dynamics of the number of emigrants from Poland (stock) in the years 2004-2017 may be summarized by the following chain of simple index numbers, with 2004 as the base year:

Year	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
index	1.45	1.95	2.27	2.21	2.1	2.0	2.06	2.13	2.2	2.32	2.4	2.52	2.54

   Calculate the average rate of change of the migrant stock for the years 2004-2017. Calculate the chain index numbers. In which year was the stock of migrants the lowest, and in which year the highest? In which year was the annual rate of growth the lowest, and in which year the highest? Knowing that the stock of migrants in 2017 was estimated at 2.540 million, predict the number of migrants in 2018 (extrapolating the average rate of change for 2004-2017).

3. The following table provides monthly price chain index numbers for February, March and April, as well as the value of sales for February and April for two articles:

	Sales value		Price chain index		
Article	II	IV	II	III	IV
A	100	150	1.2	0.9	0.9
B	200	220	1.0	1.1	1.2

   Calculate the aggregate index numbers (using the Lasyperes, Paasche and Fisher formulas) for the prices and quantities, as well as an index number for the values to compare the sales in February and April. Interpret the results. What had a greater impact on the sales value between February and April: price or quantity changes? What was the impact of price changes on the aggregate sales value between January and February?

4. A group of 20 amateurs took part in a 20 km street race. The results obtained (in minutes) were:

$$111, 131, 132, 135, 139, 141, 148, 149, 149, 152, 153, 157, 158, 175, 176, 181, 184, 186, 190, 190.$$

   Illustrate the race times by means of a histogram of

   (a) bins of width equal to 10, starting from 110;

(b) bins of width equal to 10, starting from 105;

(c) 4 bins of equal width.

Discuss the differences.

5. The series:

$$2, 2, 2, 2, 2.5, 2.5, 2.5, 3, 3, 3, 3, 3, 3, 3, 3.5, 3.5, 3.5, 3.5, 4, 4, 4, 4, 4.5, 5, 5$$

describes the exam grades of a group of Mathematical Statistics students.

(a) Plot the bar chart for student grades.

(b) Find the mean, median, standard deviation and interquartile range of the sample. Interpret the values of the statistics.

6. A group of students were interrogated on Mardi Gras to determine the number of cakes they had eaten. The results of the survey are summarized in the following table:

number of cakes eaten	0	1	2	3	4	5	6
number of female students	5	10	7	2	1	0	0
number of male students	2	5	6	8	3	0	1

Compare the Mardi Gras customs of male and female students (central tendency, variation and asymmetry).

Visualize the results

(a) by means of histograms,

(b) by means of box plots.

7. An analysis of the earnings of a group of engineers yielded the following results

earnings (in thousands of $)	Empirical CDF
2.4	0
2.8	0.1
3.4	0.19
3.8	0.32
4.2	0.5
4.6	0.75
5.0	0.9
5.4	0.96
5.8	1

Visualize the distribution of earnings by means of a histogram. Calculate the mean, median, mode, variance, interquartile range and interpret the results.