

Statistics and Exploratory Data Analysis

Marcin Chlebus, Ewa Cukrowska-Torzewska

Laboratory 3: Measures of dispersion and shape

Exercises for students

Exercise 1:

Use data on White House staff salaries for 2010 ('White_house.csv').

- a) Summarize the distribution of salaries using appropriate measures of dispersion. Which groups of workers (based on a position) exhibit the greatest variation in salaries?
- b) Calculate the mean, median and mode of the salaries. What can you say about the shape of the distribution of the salaries? Verify your answer and calculate relevant measures of shape.
- c) Check your data for the presence of outliers using the three rules we covered during the class.

Exercise 2:

Use data on fruit consumption per person by country for the years 1961-2013 ('Fruit_consumption.csv').

- a) Identify countries that exhibit the maximum and minimum fruit consumption in 2013.
- b) For each country calculate variance and standard deviation of fruit consumption.
- c) What can you say about the distribution of fruit consumption in the United States?