Econometrics (R) class 2021/2022
Exercise sheet 1

Exercise

The file "maddison.csv" contains Real GDP per capita in 2011US$ for 172 countries from the research project Maddison Project Database (2018).

1) Give the frequency table of the variables "Entity" and "Year". For which period do we have data on all the countries?

2) Give the summary of the descriptive statistics for the whole dataset. What does the difference between the mean and the median GDP tell you?

3) Draw the boxplots of the variable "GDP". Are there any outliers? How can you detect the highest outliers, without putting the whole dataset on screen?

4) Draw the scatterplot with "Year" on the x axis and "GDP" on the y axis. What do you observe?

5) Create the subset France of all GDP-values for France.

6) Draw the scatterplot with "Year" on the x axis and "GDP" on the y axis for France only. What do you observe?

7) Zoom in on the graph by just putting the evolution between 1950 and 2016. Draw a blue line through the points of the graph.

8) Repeat questions 5) and 7) for Luxembourg.

9) Create the subset comp containing only the values for the years 1950 and 2016.

10) Test if GDP is on average significantly higher in 2016 than it was in 1950.

11) Analyze the GDP evolution in Burundi.

12) Order the countries in a decreasing order with respect to their GDP in 2016.

13) Create the variable Russia by recoding the variable Entity so that Russia becomes 1 and all other countries 0 and add it to the datafile.

14) Draw a histogram of the variable GDP. Do you think that it follows a normal distribution?

15) Erase the last line of the dataset by two different methods.