# Quantitative methods of business research. Seminar 1.

**Name: Mitchin Daniil**

The goal of the first seminar is to get familiar with RStudio interface and to learn how to process the data there. You will familiarize yourself with R functions which allow you to check if model assumptions hold and test different hypothesis.

We will use a pre-made SPSS dataset. It is real and interesting data about business environment and enterprise performance in East European countries that was originally collected by the European Bank for Reconstruction and Development (EBRD) and the World Bank Group (the World Bank).

For the seminar we will follow the protocol below, and it will be similar for the home task (although the datafile and specific questions will be slightly different).

**Seminar 1 protocol**

**Download the file** BEEPS\_12\_inclass.sav **from MS Teams seminar group (see in Files) and copy it to your PC.**

Open the file (**File -> Import Dataset -> SPSS ->** Find the file in your folders). Name the R object **beeps12** and get the first impression about it.

Now, we will explore the data and variable views.

**Q1. Getting acquainted with the data.**

1. How many variables are there?

14 country 1, country 2

1. Companies from which countries are selected?

2012

1. When was this data collected?
2. What does variable ***own*** tell you about? What is its type? What is the measurement level?

**Q2. Analyzing for outliers and building graphs**

Let’s analyze if there are any outliers in your data. For variable ***employ*** build the boxplot.

What can you say from the chart?

What are the mean and the standard deviation for the variable employ?

If any outlier appeared, please for simplicity exclude it (or them) from the dataset.

How many observations left?

What are the mean and the standard deviation for the variable ***employ*** now? (To check it use descriptive statistics function)

**Q3. Normality check.**

Let’s now test the assumption for future tests – the normality assumption.

Let’s first do that using graph. Build a Q-Q plot.

What can you say about the normality assumption?

Evaluate Skewness and Kurtosis for the variable ***employ***. Interpret the results below:

Another way to test normality is to use Kolmogorov Smirnov test. Do this again for the variable ***employ***.

1. What are the null and alternative hypotheses for this test?
2. Interpret the results of the test. Does normality assumption hold? Why?

Let’s now separately test hypotheses about different countries. To do that you need to split file into groups by country name.

**Q4. Use an appropriate test and check if the mean number of full-time employees working in Country 1 in 2012 is equal to 34 people. How about Country 2?**

First describe variable ***employ*** for two different countries (check for mean, median, standard deviation).

1. What are the null and alternative hypotheses for this test?
2. Which test was used?

c) Interpret the results. Do not forget the meaningful explanation (not only the statistical one!)

**Q5. Use an appropriate test and test if the average number of full-time employees working in Country 1 in 2012 and the similar number in Country 2 are equal.**

1. What are the null and alternative hypotheses for this test?
2. Which test was used?

c) Interpret the results. Do not forget the meaningful explanation (not only the statistical one!)

**Q6. Use an appropriate test and test if the average number of full-time employees working in 2012 doesn't differ for companies with different ownership.**

1. What are the null and alternative hypotheses for this test?
2. Which test was used?

c) Interpret the results. Do not forget the meaningful explanation (not only the statistical one!)

**Q7. Creating new variables**

Let’s modify some variables to make them more useful for analysis. Add up the percentage of exported goods directly and indirectly (**q14b** and **q14c** variables) to create variable named **percexport.**

Let’s now convert the continuous variable **percexport** into the categorical variable **exporter**.

• If a company’s **percexport** score is 50 or lower, then that score will be transformed to a value of 1 on the new exporter variable (i.e., company will be placed in the low exporting category)

• If a company’s **percexport** score is 51 or higher, then that score will be transformed a value of 2 on the new exporter variable (i.e., company will be placed in the low exporting category)

What is the share of low exporters out of total number?

**Q8. Checking on total sales**

Now let’s check ***total sales*** variable: is it normally distributed, are there any outliers, think of some tests that can be interesting for sales comparison in this dataset (2 different tests).

**Report on normality check**

**Report on outliers**

**Keep the script as your working file for further references and to perform your homework.**

**#Home assignment have to work data of my countries. and Repeat the same task at the seminar**