



4th Lab Exercise

Familiarizing with R

For the needs of the specific exercise, data from Eurostat are used and more precisely the 2011 Census Hub

Based on particular queries and for your convenience, an Excel file was created (EurostatCensus2011Data.xlsx) containing 3 sheets:

- In the first sheet, data for every country's population are presented categorized by age groups
- In the second sheet, data for every country's population are presented categorized by occupation
- In the third sheet, data for every country's non-local population citizenship is presented categorized by continent of origin

Based on the data of the specific file, please:

1. Calculate the total population of each country.
2. Calculate the total population of the European Union (EU) and the total population per age group.
3. Using a **pie chart** represent the total age distribution of the EU population.
4. Using **barplot** represent the occupation distribution of the EU population
 - i. As a percentage (%)
 - ii. In absolute values
5. Using a **pie chart** represent the percentage of non-local EU population per continent of origin.

Remarks

- In order to read the contents of an Excel file, a useful function is `read_xlsx` from the package `readxl`. Indeed, to read the first sheet of an Excel file (*MyFile.xlsx*) as a dataframe, use the following command:

```
data=as.data.frame(read_xlsx("MyFile.xlsx",sheet=1))
```
- In order to output your plots in a pdf file, a useful function is `pdf` (internal function). Hence, to store all the plots into a single pdf file (*Output.pdf*), prior to plotting use the command `pdf("Output.pdf")` and after plotting use the command `dev.off()` to actually create the pdf file.