

Assignment

Dear students,

Thank you for your great participation on Thursday and Friday.

The first assignment has the objective of gaining knowledge and familiarity with the process of visual explorative analysis using R.

1. Part (70%)

In the first part of the assignment you will solve all 15 exercises that are in the [presentation](#) and save your final results in a [markdown](#) file. The link to the first exercise is <https://slides.com/sandravizmad/rggplot2#/1/15> and to the last one is <https://slides.com/sandravizmad/rggplot2#/5/27>

Even though the whole code is linked in the presentation I will put all the gist code links here together again for you.

Installations

<https://gist.githubusercontent.com/sandravizmad/7d1cec9d4f079b4fa447d5cd53f0942b/raw/9550d70839a0520fdcdc9b7ab7034977f27c6991/R-%2520Installations>

Themes

<https://gist.githubusercontent.com/sandravizmad/ab4fdc1eb81225f6345638866083a0a9/raw/50c36ae2c27992c0ec80d73f126bfdbc107d7ea3/R-%2520Themes>

Distributions

<https://gist.githubusercontent.com/sandravizmad/903ca78eb57f9a3f29d746b658b467f3/raw/61141bb10f1f0c2ece71fba13ff9260d00a2ae5a/R-%2520Distributions>

Relationship

<https://gist.githubusercontent.com/sandravizmad/787655fc1ee7232e463081cf5375ff79/raw/deb38a9a004ef8832666b9cb7c70f3d146567dbd/R-%2520Explorative%2520analysis>

Time series

<https://gist.githubusercontent.com/sandravizmad/4bcf825bcd8c7046442d3023a9885962/raw/bbc8499b36f97282e602f774f4548a66a495321d/R%2520Time%2520series>

Advanced

<https://gist.githubusercontent.com/sandravizmad/cc8d7a229609188845a1d1d21c935c68/raw/a7092dbe438ea1e08a6cf62d78a34c5657eb035b/R%2520-%2520Advanced%2520viz>

If you click in the presentation on the exercises text you will find some additional information that will help you understand the context.

2. Part (30%)

Select a dataset from the following links or choose your own one.

<https://data.oecd.org/> or

<https://www.makeovermonday.co.uk/data/>

Then you have to upload your dataset to a R script. You do this by using the following code.

```
adults <- read.csv("https://gist.githubusercontent.com/sandravizmad/8273490c96bf54a345cd119fb458bc04/raw/4234e3d287a4d24f6331a663b3dacc04ab67620/Adults")
```

In case the file you are using is on your computer, just make sure it is saved in the same folder as your script.

After this I want you to explore the data in the way I presented to you on Friday. Remember the most important objective is to find patterns in distributions, relationship or times series analysis.

You decide on the charts and how many charts you want to create (but not more than 5). I want that you give also a short description of your analysis similar to the one in this example: [World Data Visualization Challenge - Analysis](#)

The submission should be **one final markdown file** that includes part 1 and part 2.

The time needed for the assignment shouldn't be longer than 10 hours.

The deadline for this assignment is the 16/11/2020.

If you have any questions feel free to contact me in slack.

Enjoy,

Sandra