

Homework Assignment 7

Maria Kunevich

2022-11-10

Assignment Description

Download the `gapminder` data set into R `Markdown` session.

1. Perform quick EDA and pick up two variables you want to explore in more depth (for example, life expectancy and gdp).
2. Prepare the data set that includes only variables of your interest in a suitable format for analysis (use `dplyr` package and `tidyr` when necessary).
3. Explore two variables and how they are associated with each other (correlation analysis). Include the description of assumptions for correlation analysis, concluding with the type of analysis you choose. Create necessary visualisations to support your analysis (a histogram, a scatter plot, etc) and include your interpretation of the graphs.
4. Create a report in R `Markdown` with the following sections:
 - Introduction (brief description of the data set and variables)
 - Questions for EDA
 - Description of data cleaning and transformation
 - Description of correlation analysis (steps for visualisation, checking assumption for correlation analysis, performing correlation analysis)
 - Data interpretation and conclusions

Each section is worth of 1 point, the whole analysis - 5 points.