

COMPUTER PRACTICE 1

DATA EXPLORATION AND VISUALISATION

OBJECTIVES

- Obtain practical skills of exploratory data analysis provided by self-developed scripts using Python libraries

GUIDELINES

1. Recall theoretical concepts of data exploration and visualisation, provided in topics “Data Types” and “Data Exploration and Visualization”.
2. Create Python Notebook for scripting and analysis running. You can use Jupyter Notebook or Google Colab as work environment.
3. Load the data set provided for the Computer Practice 1 in the course and read through its description.
4. Identify the categorical and continues variables in data set.
5. Calculate and analyse appropriate descriptive statistics for all continues variables.
The report should represent the values and analysis of at least 8 appropriate statistics.
6. Calculate and analyse appropriate descriptive statistics for all categorical variables.
The report should represent the values and analysis of all appropriate statistics selected by you.
7. Visualize variable values using the appropriate techniques:
 - Histograms
 - Box-plots
 - Bar charts
 - Line charts

The report should include all titled and numbered graphs with results of your analysis of each graph with your findings and conclusions.
8. Analyse relationship between 2-3 pairs of variables using appropriate techniques:
 - Scatter plots
 - Pearson correlation coefficients’ values
 - Cross-tabulation

The report should include all titled and numbered graphs, values of Pearson correlation for each of pares and the results of your analysis, reflections regarding the form and strength of relationship. Please make your decision statistically sound.
9. Upload your report (.pdf) and Python Notebook (.ipynb) with all your scripts used for analysis and generated results to the learning environment.