

Visual Data Science

Lab Part 1

A report (at least 5 A4 pages) has to be submitted via TUWEL until November 18th, 2020.

You can decide whether you would like to write the report in German or in English.

In the report, please discuss the following questions:

1. Describe your results and findings of task 1. Use at least one screenshots to outline how you solved the visual data analysis.

- a. Why did you select the specific combination of nutrition elements?
- b. How did you find out how many clusters need to be formed to describe the data best?
- c. Which Python functionality did you use for the statistical analysis and which visualization type did you employ?
- d. Which aspects of the task (e.g., identifying the number of clusters) where easier to solve by using statistical analysis, and which by using visualization?

2. Describe your results and findings of task 2. Use at least one screenshots to outline how you solved the visual data analysis.

- a. Describe the correlations you found in the data.
- b. Which Python functionalities did you employ for the statistical analysis?
- c. Which visualization type did you use?
- d. Do you think that correlations can be missed when either using only statistical analysis or when using only visualization?

3. Describe your results and findings of task 3. Use at least one screenshots to outline how you solved the visual data analysis.

- a. Which groups did you use and why?
- b. Describe the differences among the three groups.
- c. Which Python functionalities did you employ for the statistical analysis and which visualization type did you use?
- d. Was it easier to do the comparison with statistical analysis only, or by employing visualization?

4. Summarize what you learnt when solving the three tasks.

- a. Which tasks were easier to solve when also using visualizations, and which were not?
- b. What do you think are the main challenges / problems when using only statistical analysis, or only visualizations?