Task 2: Individual Data Visualization Project

Prof. Dr. Tillmann Schwörer

1 Project Overview

"Graphical excellence is the well-designed presentation of interesting data - a matter of substance, of statistics, and of design ... It consists of complex ideas communicated with clarity, precision, and efficiency." (Edward Tufte)

- **Task**: In spirit of this quote, put yourself into the shoes of a data journalist and tell a visual story about a topic or question of your choice.
- **Submissions**: Submit your project on Moodle, or submit a link to a public GitHub repository. Submission deadline is 10.01.2023, 23:59.

2 Structure of the project

- **Goal**
  - Describe the goal of your visualization project: limit yourself to 1 central topic or question that you want to address (it may well have several sub-aspects).
  - Explain what motivates your project, e.g.: Why did you choose it? What makes it interesting and/or important?
- **Data Acquisition**
  - Use a real-world dataset to provide insights on your topic, or to answer your question.
  - You can assemble a data set using data from e.g. statistical offices, APIs, or via web scraping. You can also use a ready-made data set from a public source. But note the further explanations in section "Grading".
  - Provide information on relevant aspects of the data. In particular such information that relates to the reliability/authenticity of your data, or information that is needed to understand its contents: Where did you obtain the data? Is this the original data source? How/when was the data collected? What are relevant definitions?
- **Data Preprocessing**
  - If necessary, clean, combine, enrich your data set.
  - Explain what your preprocessing steps are, and why they are important.
- **Visual Story**
  - This is the main part of the project: Tell an insightful visual story on your topic or question.
  - Carefully choose a handful of plots (~ 4-6), to provide a broad picture.
  - Refine your plots to make them effective, visually pleasing, and scientifically sound.
  - Design and finetune the plots such that the main messages or patterns are intuitively understandable.
  - Interactive features can be helpful, but be aware that you are partially losing control over what the user sees.
  - Always represent the data truthfully, and mind confirmation or publication bias.
  - The plots should be designed such that they speak for themselves. However, you still need to tell your story in words: What is in your eye the relevant pattern that emerges from the plot? What does it fit with your other findings and your overall story?

- **Conclusion**
  - Summarise your main findings.
  - Critically reflect on limitations of your project.
  - Provide ideas on related aspects or questions for future research.

3 Rules

- **Language**: English or German.
- **Programming Language**: Python
- **Individual work**: You are allowed to exchange ideas with other students, but this is not a teamwork exercise.
- **Copying and Citing**: You may use all the code from the lectures. Copying and adapting from other sources is allowed in small quantities. Cite all relevant resources on which your project is based, or from which you draw inspiration.
- **Reproducibility**: Make sure that your project is fully reproducible.

4 Grading

- **Storytelling (60 points)**:
  - Coherent story: topic/question, chosen plots, and explanations fit nicely together
  - Interesting: do the plots provide interesting, potentially surprising information?
  - Effective: are the plots effective, i.e. intuitively understandable? Are the plots optimized for the human eye and mind?
  - Visually pleasing: does it look appealing?
  - Scientific: are data choice, cleaning, and visual representation truthful?
- **Documentation (25 points)**:
  - Reproduciblilty
  - Clear textual explanation through the entire project
  - Remove all aspects that do not strictly fit into your story
  - Clean code; no debug output or error messages, or commented out code.
- **Innovativeness and complexity** (15 points + 10 bonus points)
  - Innovative data, topic, or insights
  - Innovative visualization techniques (not for its own sake, but serving the goal of creating effective, visually pleasing, and scientifically sound visualizations)
  - Complex data acquisition and cleaning
  - Particularly innovative or complex projects may be awarded up to 10 bonus points.