Group project 1

GRA4157

October 10, 2024

Data exploration and visualization

In this project you will explore and visualize a dataset of choice. Define one or more questions you aim to answer using the selected dataset, which should be answered during your presentation. You are free to choose the dataset and the questions yourself. You will work in groups of 1-3 students. Your efforts will be presented in a 10-20 minute presentation, depending on the group size. Every group member needs to contribute equally to the presentation, but it is fine if each group member has responsibility for different tasks within the project. You will receive feedback on the following:

You will receive feedback on the following

- The question you are trying to answer.
- How is the data presented and did you find an interesting dataset?
- How was it collected? scraping? public api?
- The raw data should be presented in an easily readable format. Table, map(?), statistics?
- You should do some calculations/interpretations on the data. If you e.g. have data on position and time, you can compute velocity. If, for example, you have sales per day, you can compute sales per year. Try to be creative!
- Examples: Temperature data (over time), elevation data, population data, GDP data.

The group presentations will be held Friday 18th of October. The presentations are not graded, but you will use material from the presentation to write part of a report for the final assignment/exam in the course. Below are some templates from last year which you can use as examples or build upon. The the code for the templates are located under /GRA4157/lectures/08-project-and-intro-to-ml/templates/

The templates below are located under /GRA4157/lectures/06-visualization-project/templates/

Template 1

We have previously worked with the Oslo city bike dataset. In the Bicycle template you will find examples on how to plot markers for each bike station, create lines between each bike station, as well as creating a heat map. You can use this as inspiration to visualize many different aspects of the trips for all stations or for a single station.

Template 2

Strava is an Internet service to track physical exercise that incorporates features of social networks. Each activity is tracked via a .gpx file that can be read into python. In this case, you can track your own data to visualize on a map, and to add specific information to the map that is not included in the standard strava map (e.g. velocity, elevation, heart rate). It is possible to track your own activities with the Strava app or use an exercise watch to do analysis on. Alternatively, several pro (and other) athletes upload their exercises to their profiles, so you can choose to display data from their activities (e.g. `https://www.strava.com/pros/laurenstendam`)