Group project 1

GRA4157

October 10, 2025

Data exploration and visualization

In this project, you will explore and visualize a data set of your choice. Define one or more questions you aim to answer using the selected dataset, which should be addressed during your presentation. You are free to choose both the data set and the questions yourself.

You will work in groups of 1 to 3 students. Your efforts will be presented in an 8–12 minute presentation, depending on the size of the group. Each group member needs to contribute equally to the presentation, but it is acceptable if each group member takes responsibility for different tasks within the project.

Feedback criteria

You will receive feedback on the following points:

- The question you are trying to answer.
- How the data is presented, and whether you found an interesting dataset.
- How the data was collected (e.g. scraping, public API, etc.).
- Presentation of the raw data in an easily readable format (e.g., table, map, statistics).
- Calculations or interpretations based on the data (e.g., computing velocity from position and time, or annual sales from daily sales).

Be creative in your approach! Example datasets include temperature data over time, elevation data, population data, or GDP data.

Presentation details

The group presentations will be held on **Friday, October 17th**. The presentations are not graded, but you will use the material from your presentation to write part of a report for the final assignment or exam in the course.

Templates and examples

Below are some templates from last year that you can use as examples or build upon. The code for the templates is located under:

`/GRA4157/lectures/08-project-and-intro-to-ml/templates/`

Additional visualization templates can be found under:

`/GRA4157/lectures/06-visualization-project/templates/`

Template 1: Oslo city bike dataset

This template demonstrates how to plot markers for each bike station, create lines between stations, and generate heat maps. You can use it as inspiration to visualize various aspects of the trips for all stations or for a single station.

Template 2: Strava activity data

Strava is an internet service that tracks physical exercise and integrates social networking features. Each activity is stored as a `.gpx` file, which can be read into Python. You can visualize your own data on a map and enrich it with additional metrics such as velocity, elevation, or heart rate.

It is possible to track your own activities using the Strava app or an exercise watch. Alternatively, you can use publicly available data from professional athletes (e.g., https://www.strava.com/pros/laurenstendam).