ICT for Smart Mobility

Exercise – Data collection and analysis - Examples

Consider the 3 datasets in the folder: "Exercises/1_Data_collection"

- Dataset1.zip
- Dataset2.xml
- Dataset3.csv

Analyze the files in the dataset (**Python** is suggested).

Task A

These datasets are related to transport data collection. What kind of data do they contain?

Associate the 3 datasets above to the 3 following data sources and descriptions:
- **Survey.** Data regarding how people commute to go to work, obtained by interviewing a sample of the population.
- **Point traffic sensor data.** Count measurements of different inductive loops in a city, measured in a 5-minute time window.
- **General Transit Feed Specification (GTFS) data.** Public transportation schedules and associated geographic information according to the GTFS format.

Task B

Analyze the content of the **survey dataset**.
- How many people in total use the metro to go to work in 2021?
- Are there more females or males that commute to work in the survey?
- Which category (gender+age class) uses more motorcycle/moped?

Task C

Analyze the content of the **Point traffic sensor data**.
- Which is the road with the highest flow (vehicles/h)?
- Which is the lowest average flow speed (km/h)?
- Locate through the coordinates in a map the position of the inductive loop further to the east.

Task D

Analyze the content of the **GTFS dataset**.
- How many stops are there in the dataset?
- Identify the names of the routes that have service on Sundays.
- Where is located the stop/s with the latest arrival time (arrival_time)?