Exercise 2 - Work Plan
194.192 Experiment Design and Execution

Group 19

November 2025

Questions

Question number 8 was selected. Based on the overarching theme of comparing European rail and air transport networks, and after refining the initial ideas, the following research questions will be addressed:

1. How do rail travel times compare to air travel times between 10 to 12 biggest European cities by population? This requires supplementing the dataset with external demographic data.
2. Are there any routes on which rail leads to shorter journey times than air travel?
3. How can we estimate the typical travel time between the city's central rail station and its nearest major airport, and how should this time be incorporated into the total air travel duration?
4. Which European city is the most well-connected in terms of achieving the shortest travel times (including indirect connections) to other major cities? This will require defining connectivity metrics and using shortest-path algorithms (e.g., Floyd–Warshall).

Dataset

The primary dataset used in this project was obtained from 4TU.ResearchData, an international open research data repository managed by the technical universities in the Netherlands. Our dataset, titled *"Mapping the Market Potential for Air–Rail Integration"*, provides travel times and connection structures relevant to European air and rail transport. It is publicly available under the following DOI link: Dataset Link. To address specific research questions, additional data may be required, including Population data to determine the largest European cities and Airport location and access-time estimates (e.g., average travel time from city centre to the major airport).

Plan to answer the questions

The research questions will be answered through data analysis of the provided dataset and, where necessary, additional external data sources. The work will begin with preprocessing and structuring the dataset in order to address missing values, inconsistencies and required transformations. Afterwards, travel times for both air and rail connections will be calculated, with air travel durations adjusted to include estimated airport access and processing times. To assess connectivity between cities, graph-based methods will be applied, making use of shortest-path algorithms to compare the efficiency of indirect and direct connections. The findings will then be presented through visualisations such as charts, network diagrams, and isochrones, enabling an intuitive interpretation of relative travel times and connectivity across Europe.

Work division between group members

The preprocessing of the dataset will be done by all members of the group. The preprocessing will include exploring the data and descriptive analysis of the dataset, dealing with outliers and missing values, as well as the dataset manipulation to achieve desired format. Once the dataset is clean and ready for further use, each member of the group will take one of the four proposed questions and answer it. All group members will be in charge of reviewing all other members' solutions, and not just their own.