



Navigating Zurich

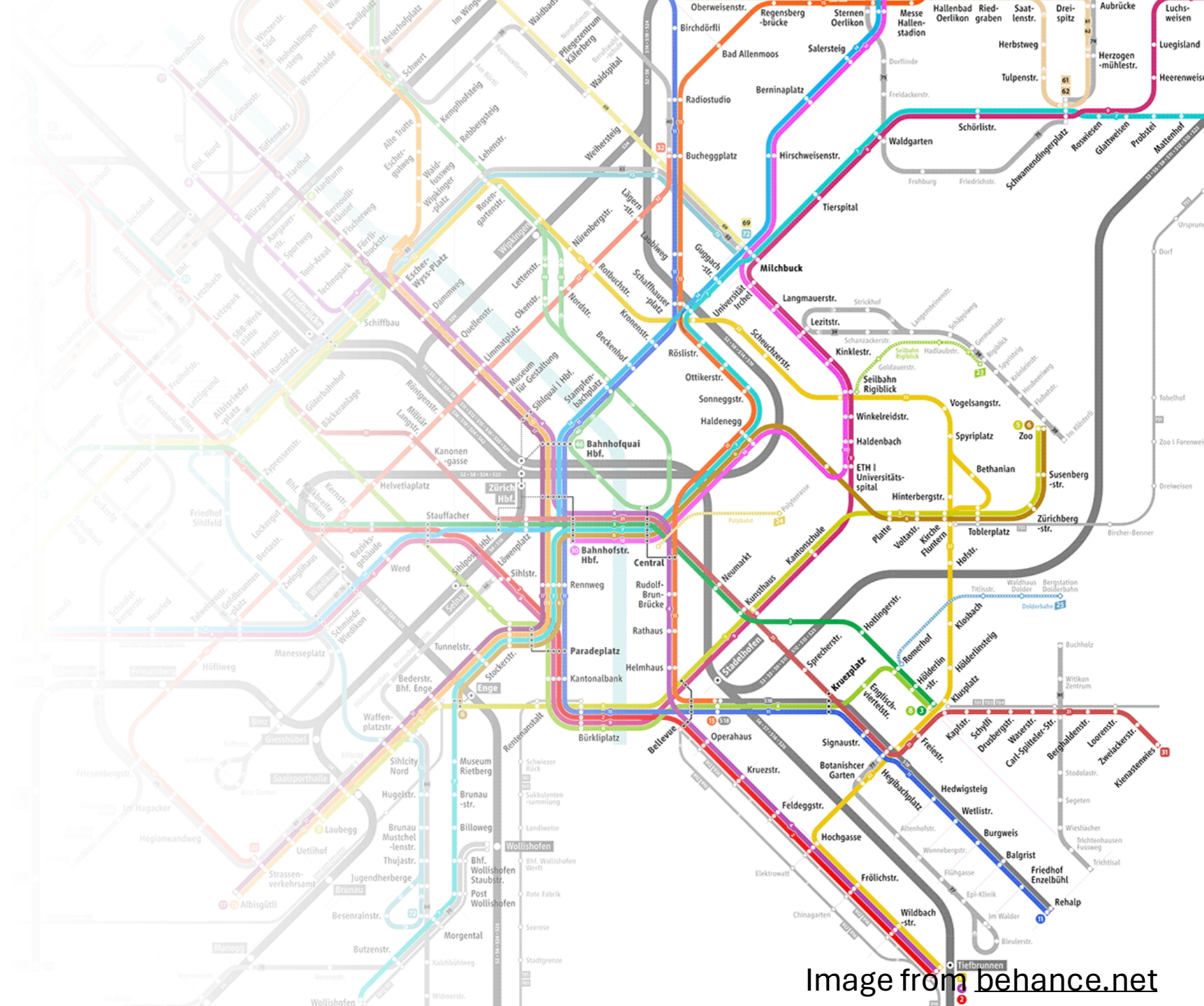
A Comprehensive Analysis of Urban Traffic Dynamics

Project Presentation by

Pascal Sager, Luca Zhao, Weijia
Zhong, Xiaohan Zhu

Outline

- Introduction
 - Research Questions
 - Dataset
- Geospatial Analysis
- Documentation
- Discussion & Limitations





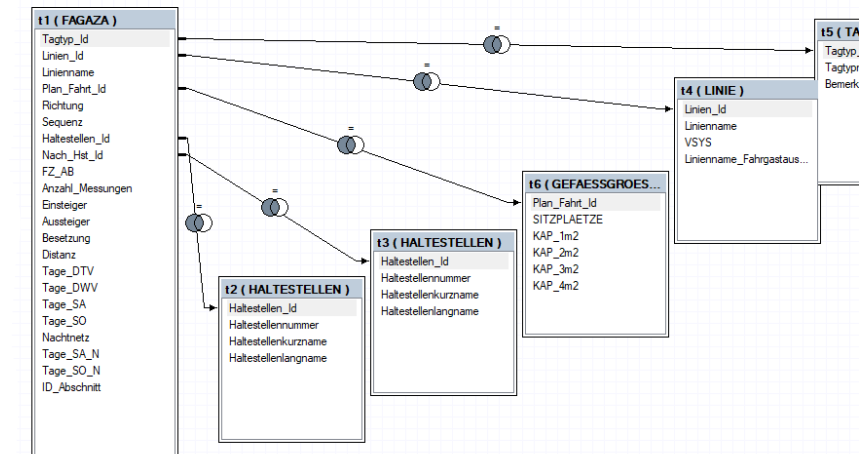
Research Questions

We analyze public transportation in the city of Zurich and focus on:

- The **development** of Zurich's public transportation system over time
- The **utilization** intensity of Zurich's public transportation infrastructure
- The **spatiotemporal distribution** of passengers
- Analysis of the **interplay between diverse factors**, including spatial location, weekday versus weekend patterns, and academic calendar
- **Prediction of seat availability** on public transit

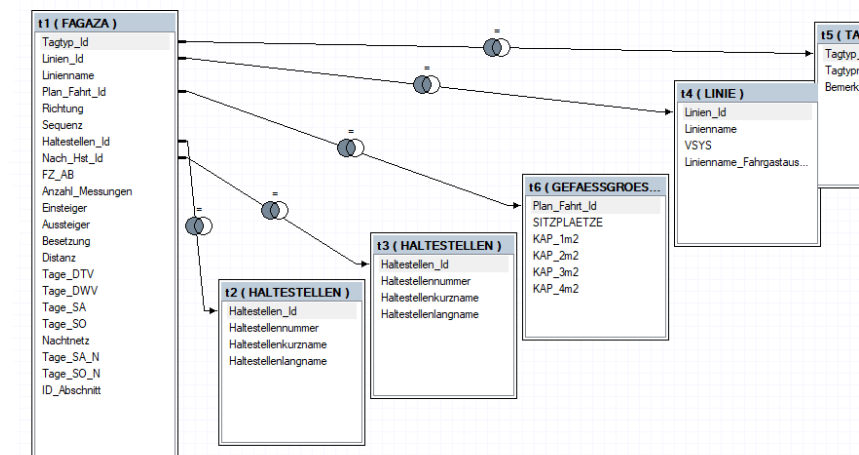
Data

- Dataset containing annual **passenger boarding, alighting, and vehicle occupancy** numbers of Zurich's transport authority (VBZ)
 - 6 tables
 - 1M+ entries
 - 39 features (columns)
- Enriched with external **GPS** coordinates of stops



Data Quality

- The overall quality of the data is decent, there are some inconsistencies across different years.
- Missing values are present.



Data Cleaning

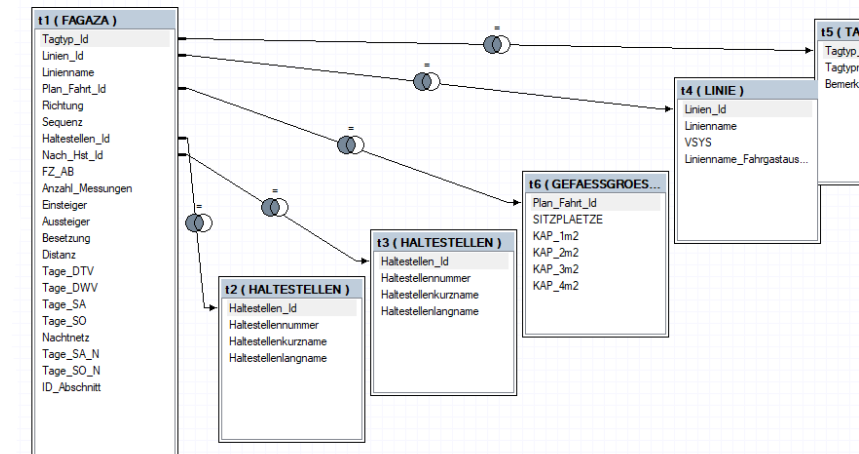
- Item non-response, about 0.3% ~ 0.4% for each year.

We simply removed the rows because:

1. No measurements were provided
2. The percentage of missing value is lower than 1%

- Incomplete data about the GPS coordinates

We used the library GeopPy to retrieve the coordinates and manually correct generated data.





Development over time

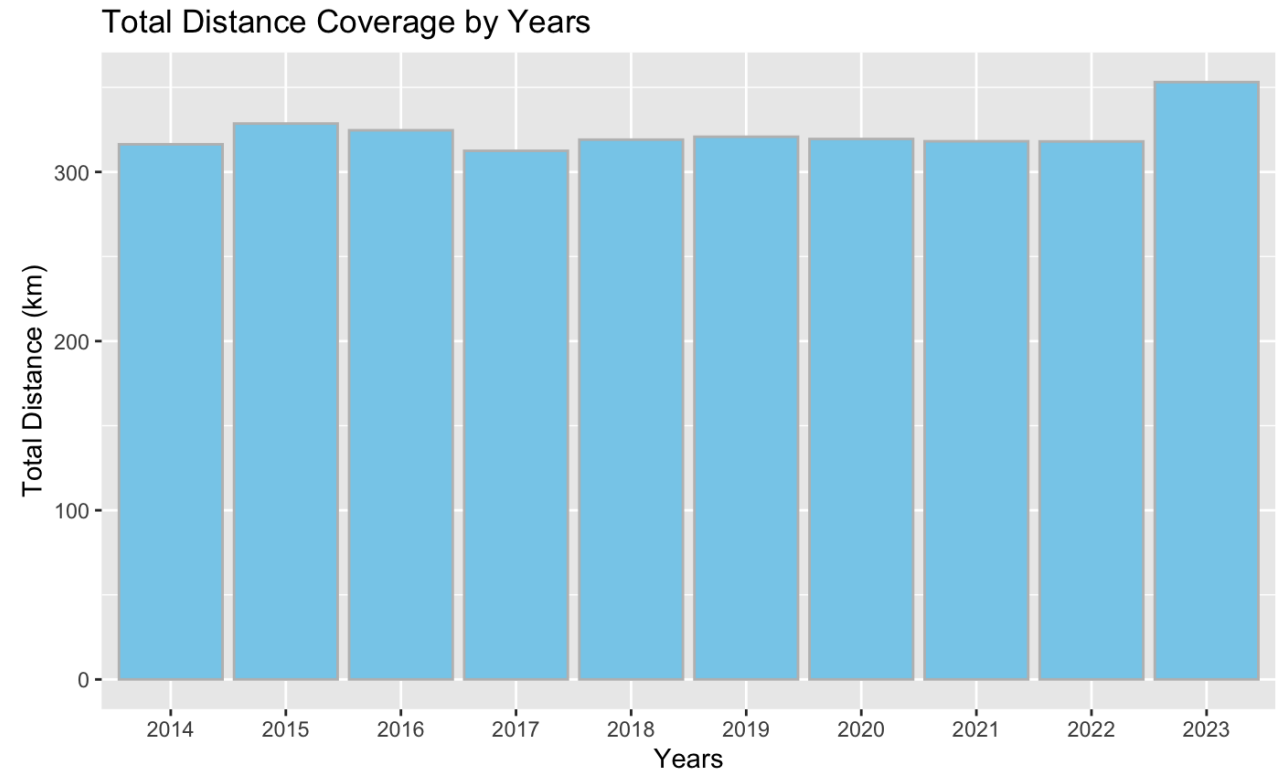
The Development of VBZ by Spatial Coverage (km)

Method

- Summing up the distance in between each unique stops

Findings

- Little changed happened across year 2014 to 2023.
- Some development happened in year 2013 (what are those?)



Passenger Volume Change across the years (mio)

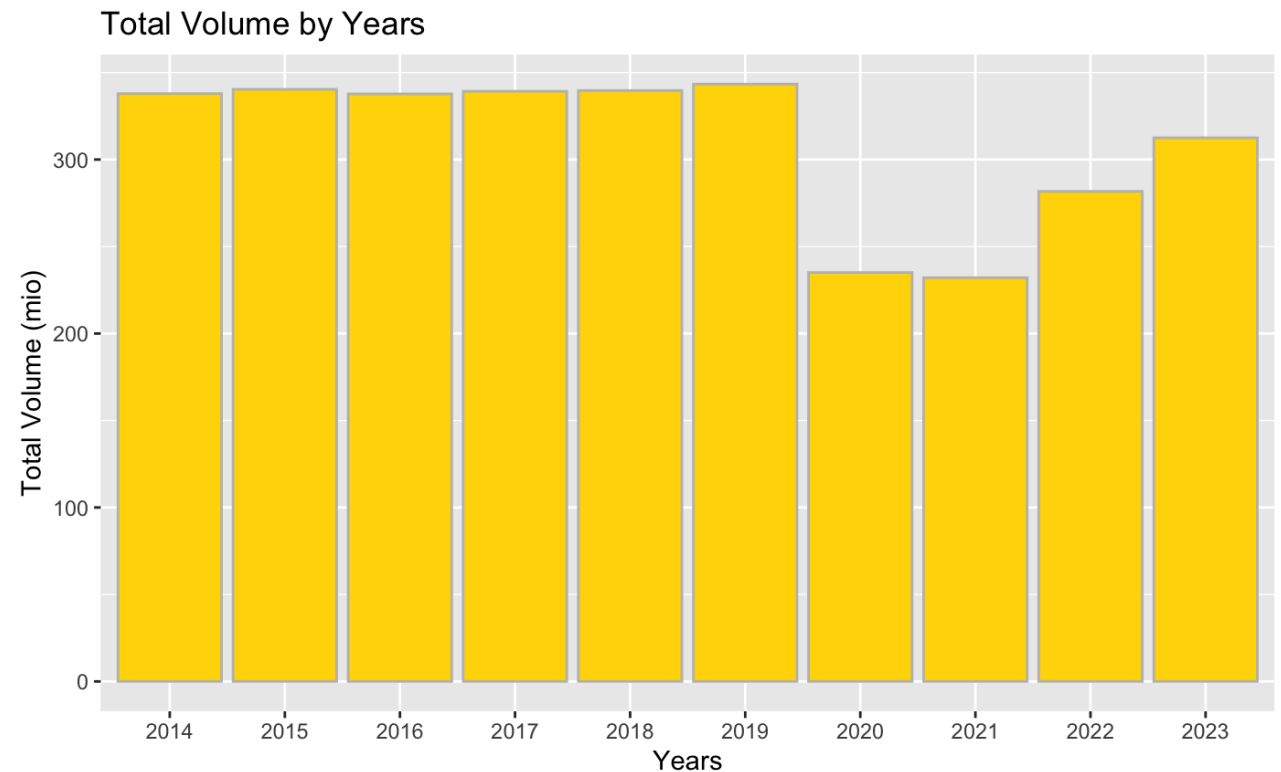
Method

- The number of passengers that get aboard for unique rides per day * number of days

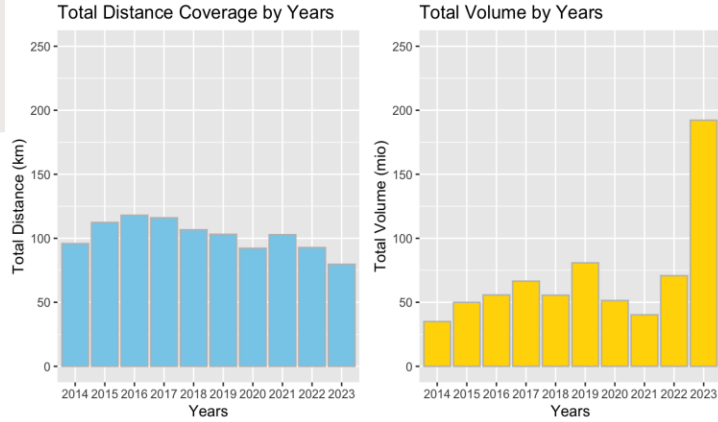
Findings

- Significant drop during COVID
- Slowly gets back 2022-2023, still not fully back to the level pre-COVID
- Why? Do people get more used to cars? Can cross-validate with vehicle traffic volume

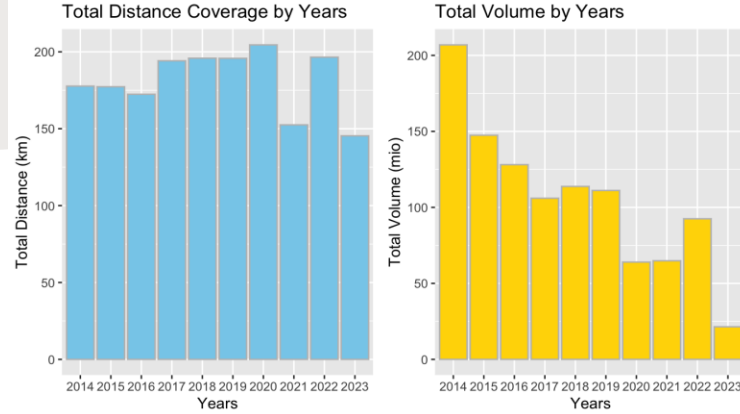
Also examine the changes in sub-categories



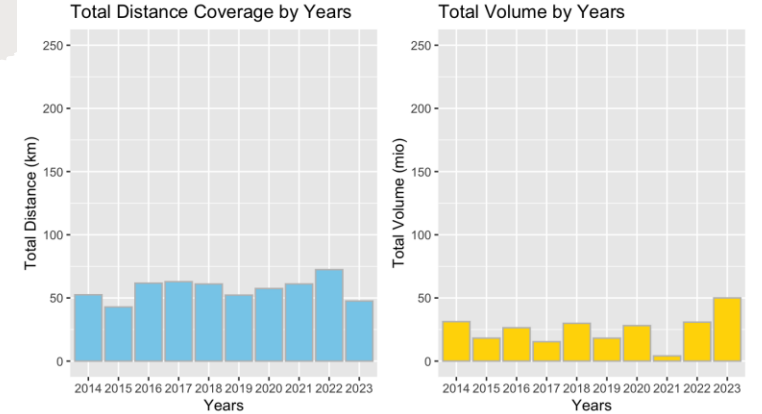
Change in Tramways



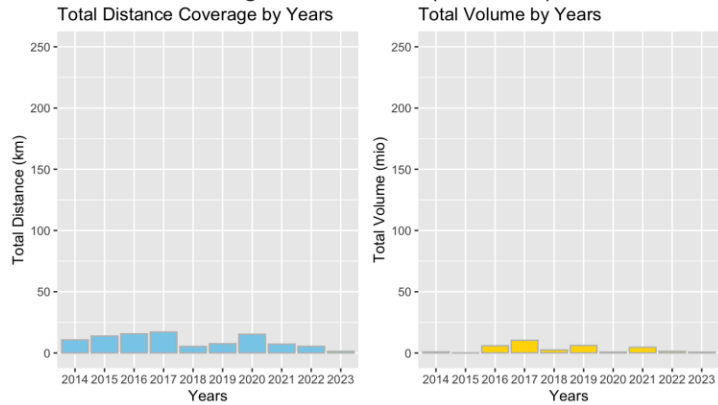
Change in Bus



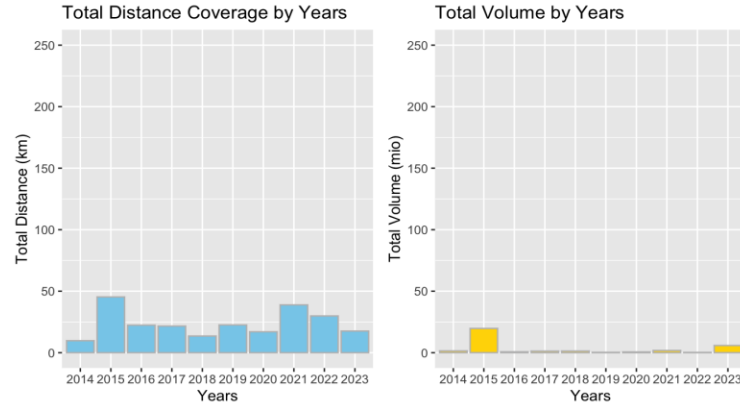
Change in Trolleybus



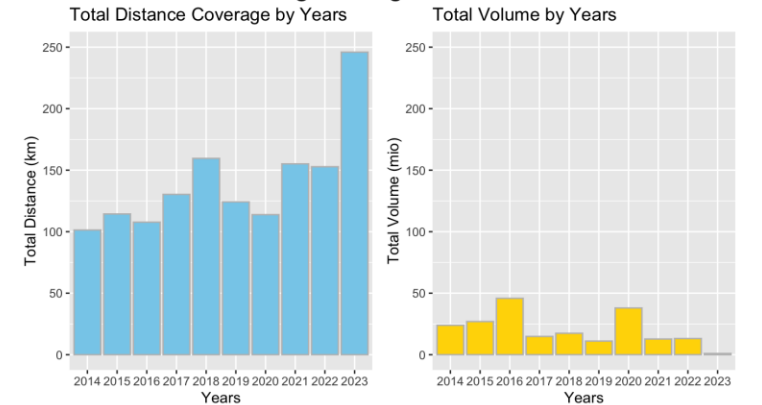
Change in Cablecar (Seilbahn)



Change in Forchbahn



Change in Night network



Increase in Tram vs decrease in bus is interesting.
Layout to be changed later

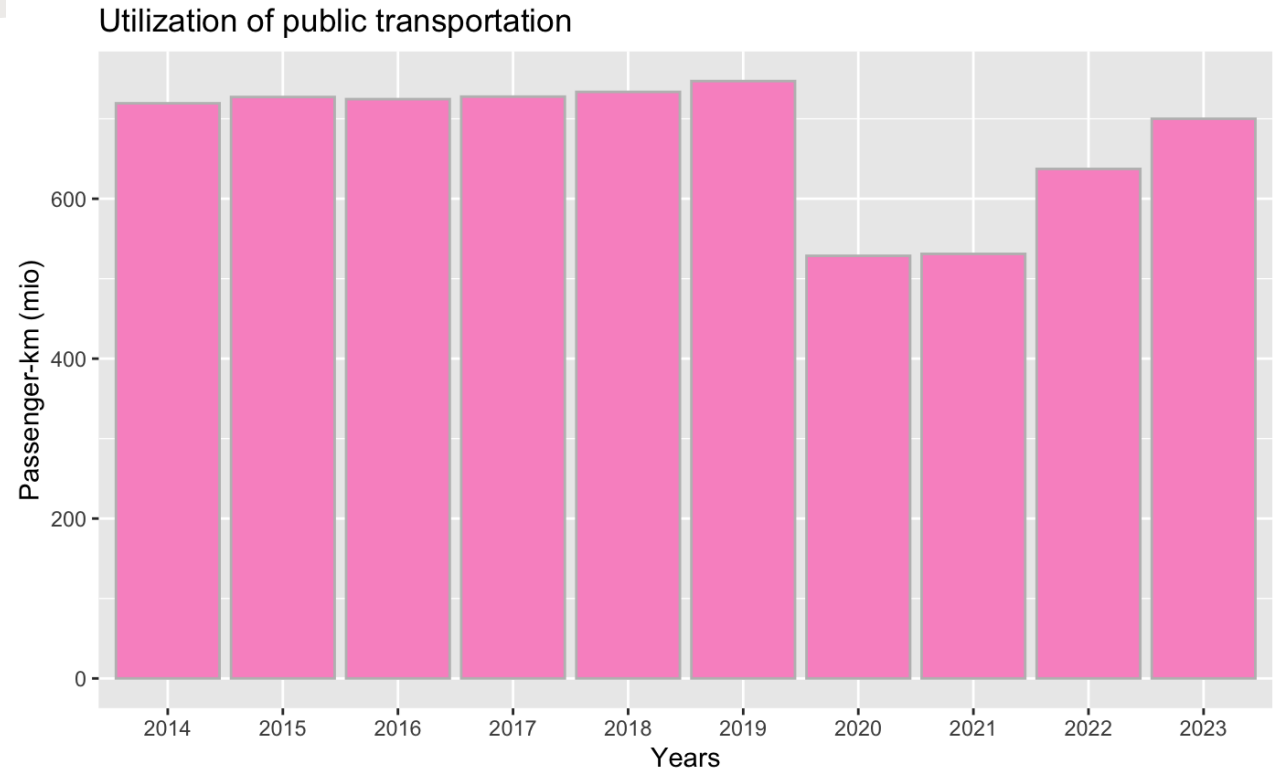
Utilization of Urban Transportation (total)

Method

- By years, compute the transport of passengers over kilometers.

Findings

- Very similar trend to passenger volume



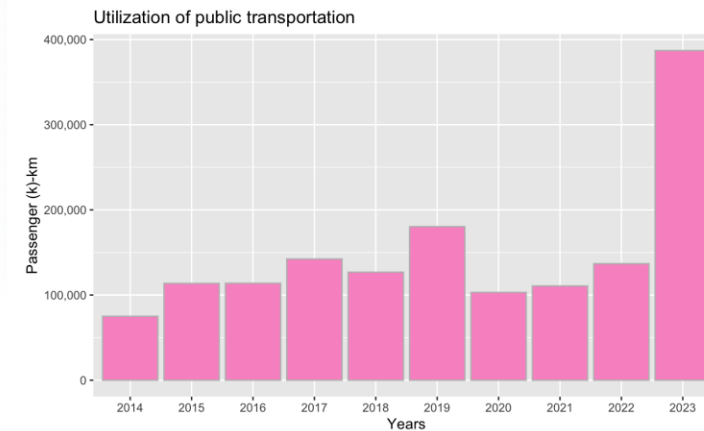
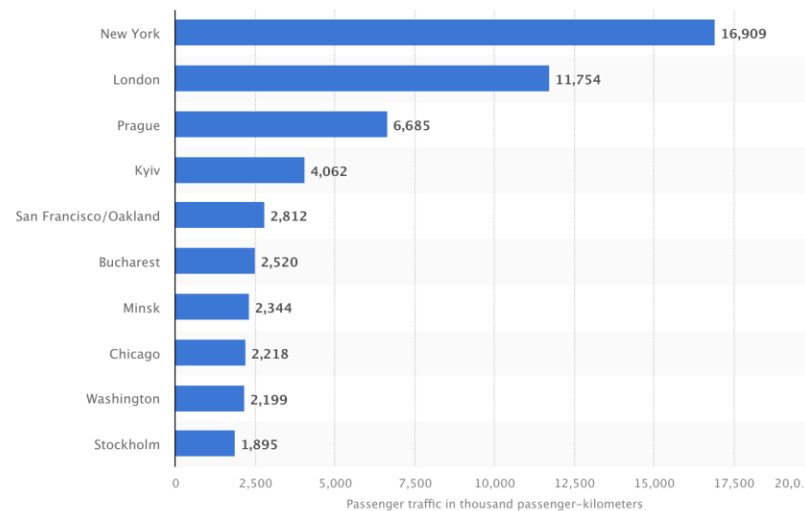
Utilization of Urban Transportation (tram)

Method

- Compare mobility of tram system in Zurich area to metros in other city in the world, year 2019 (source: statista)

Findings

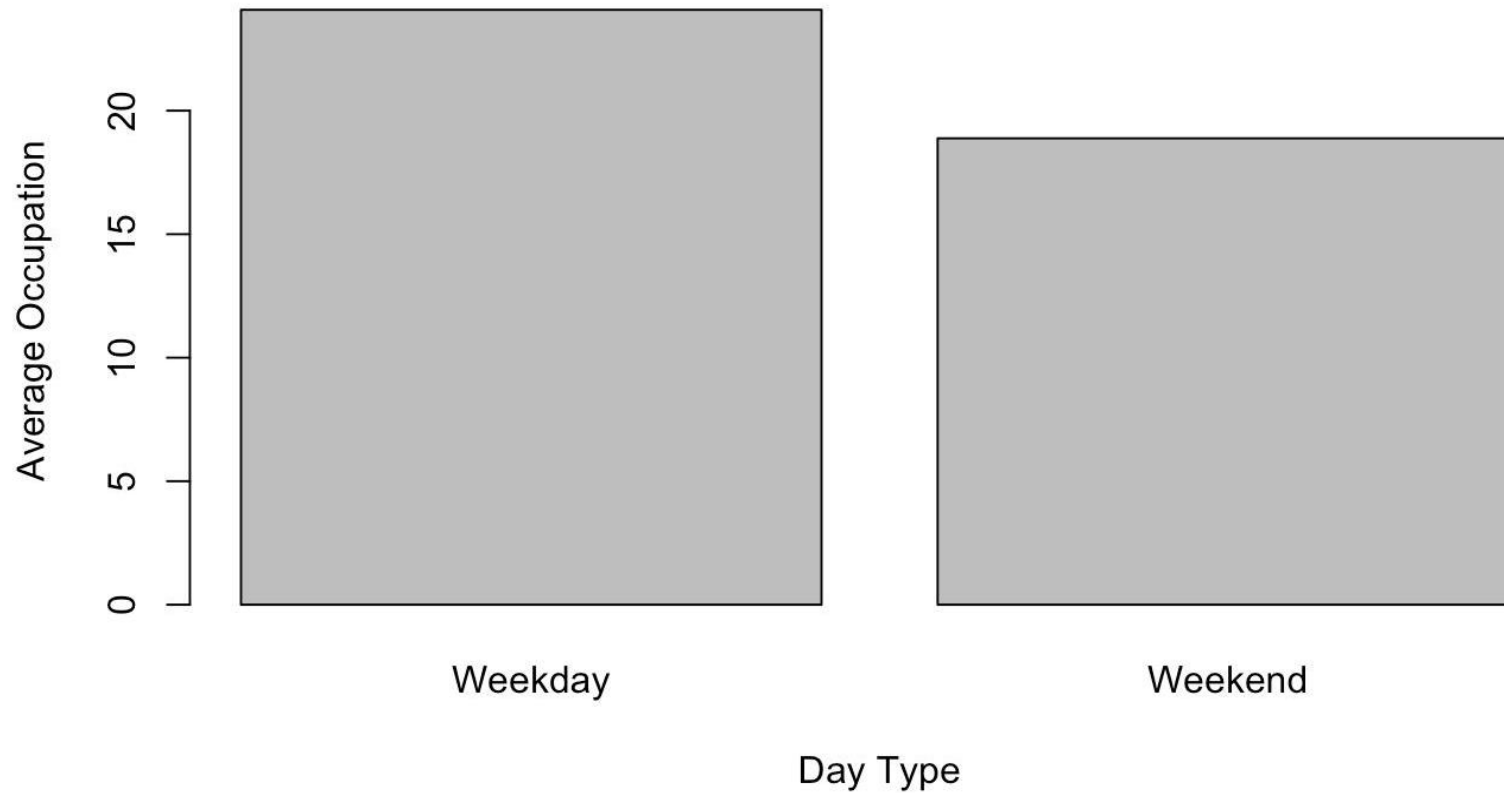
- The numbers seems off. What's wrong?





Utilization intensity

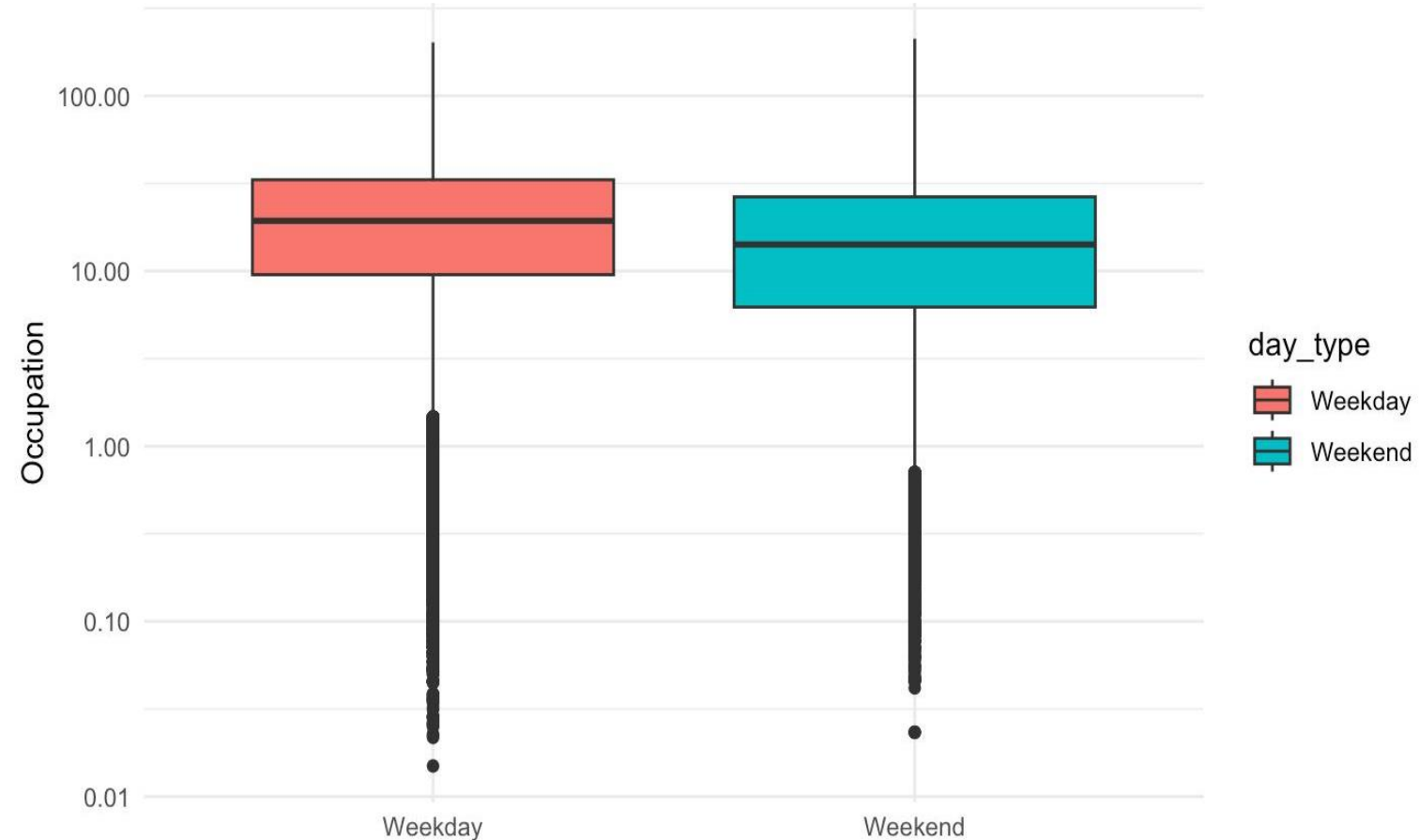
Average Occupation Weekday vs Weekend



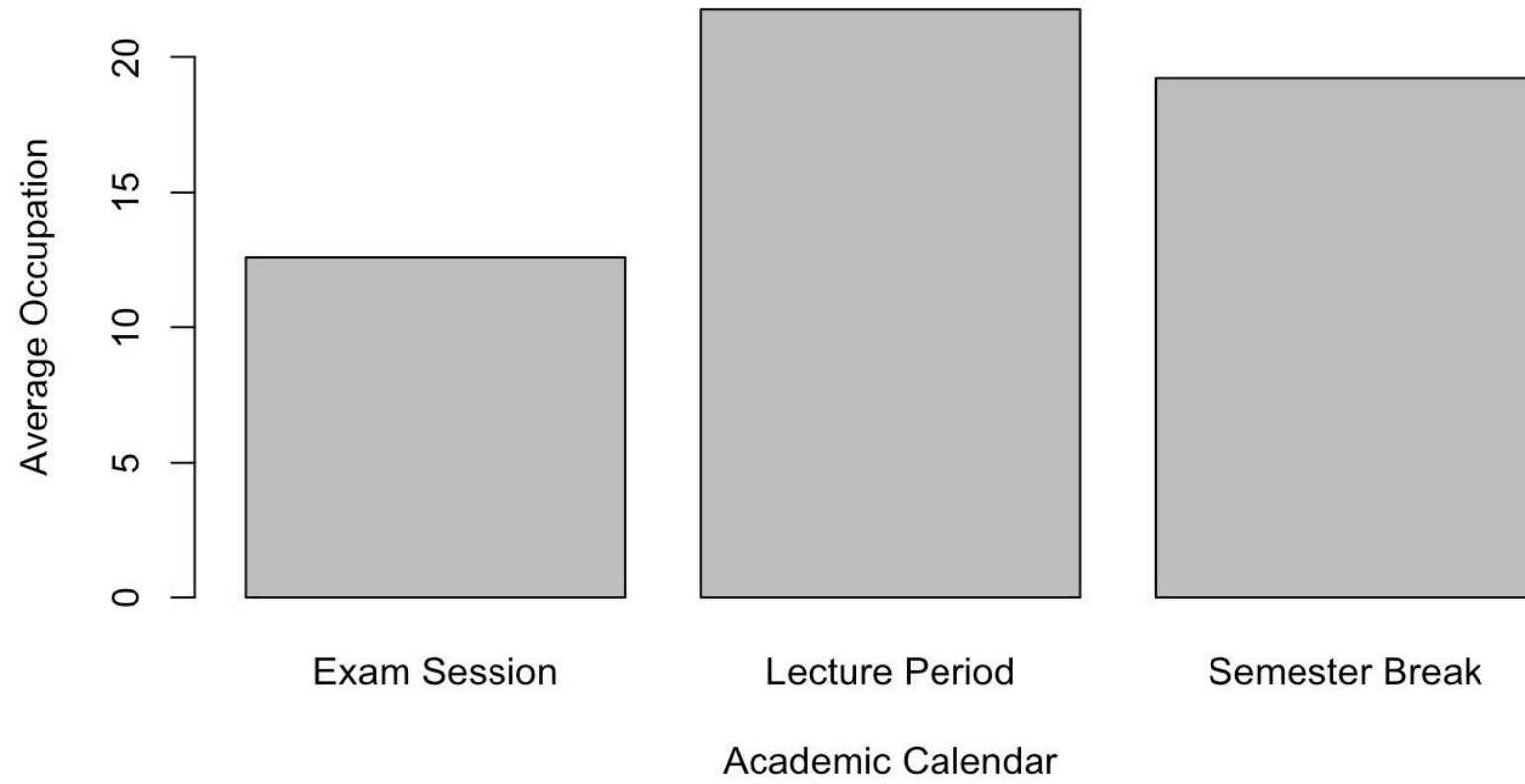
Box Plot comparison

Findings:

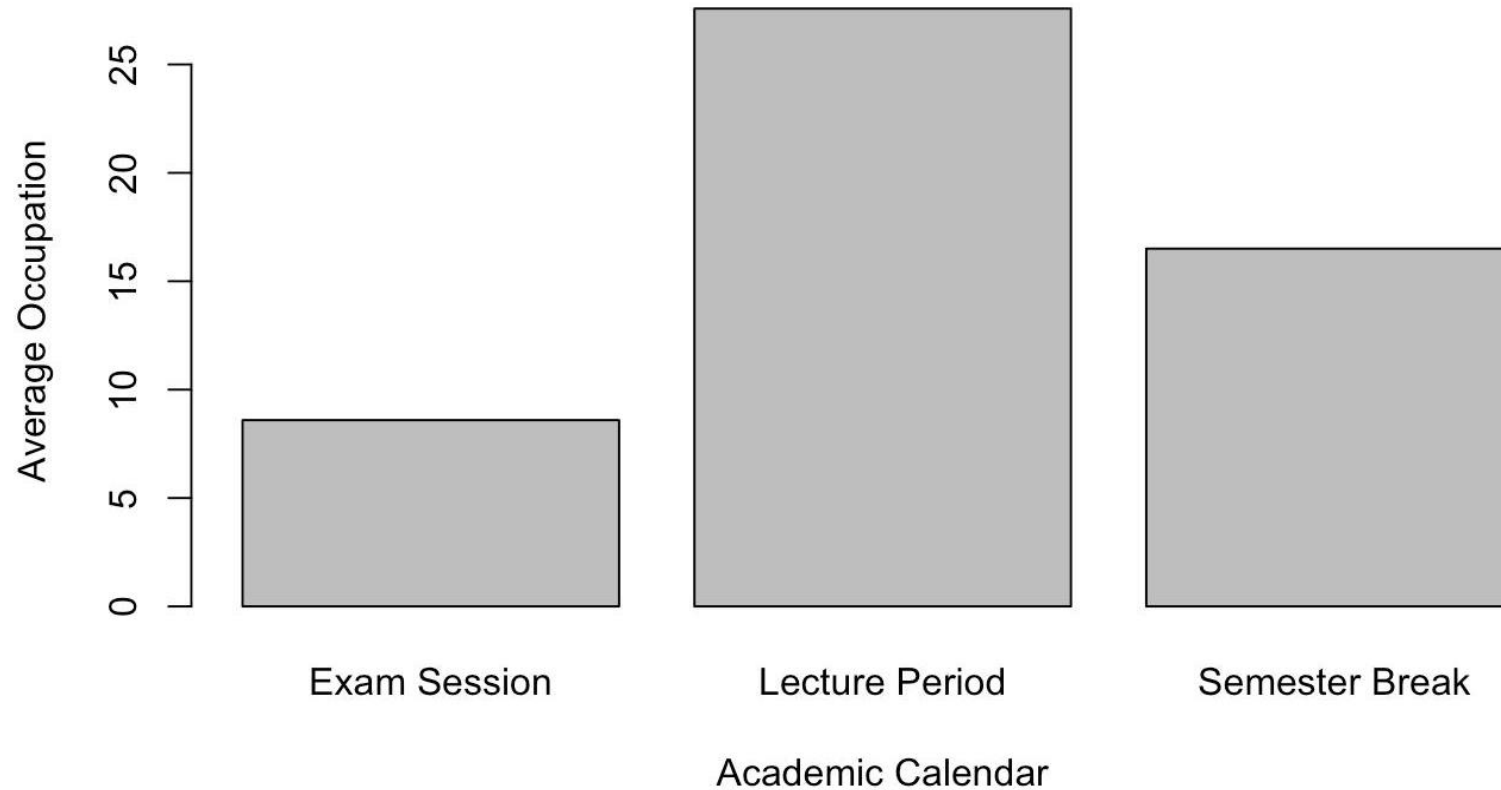
- The IQR shows a slightly wider spread during weekends, indicating more variabilities of passenger volume
- High whiskers on both day types, meaning that many routes have very low occupancy.



Academic Period: Average Occupation of All Stops



Academic Period: Average Occupation of University Stops



Comparison

Exam Session:

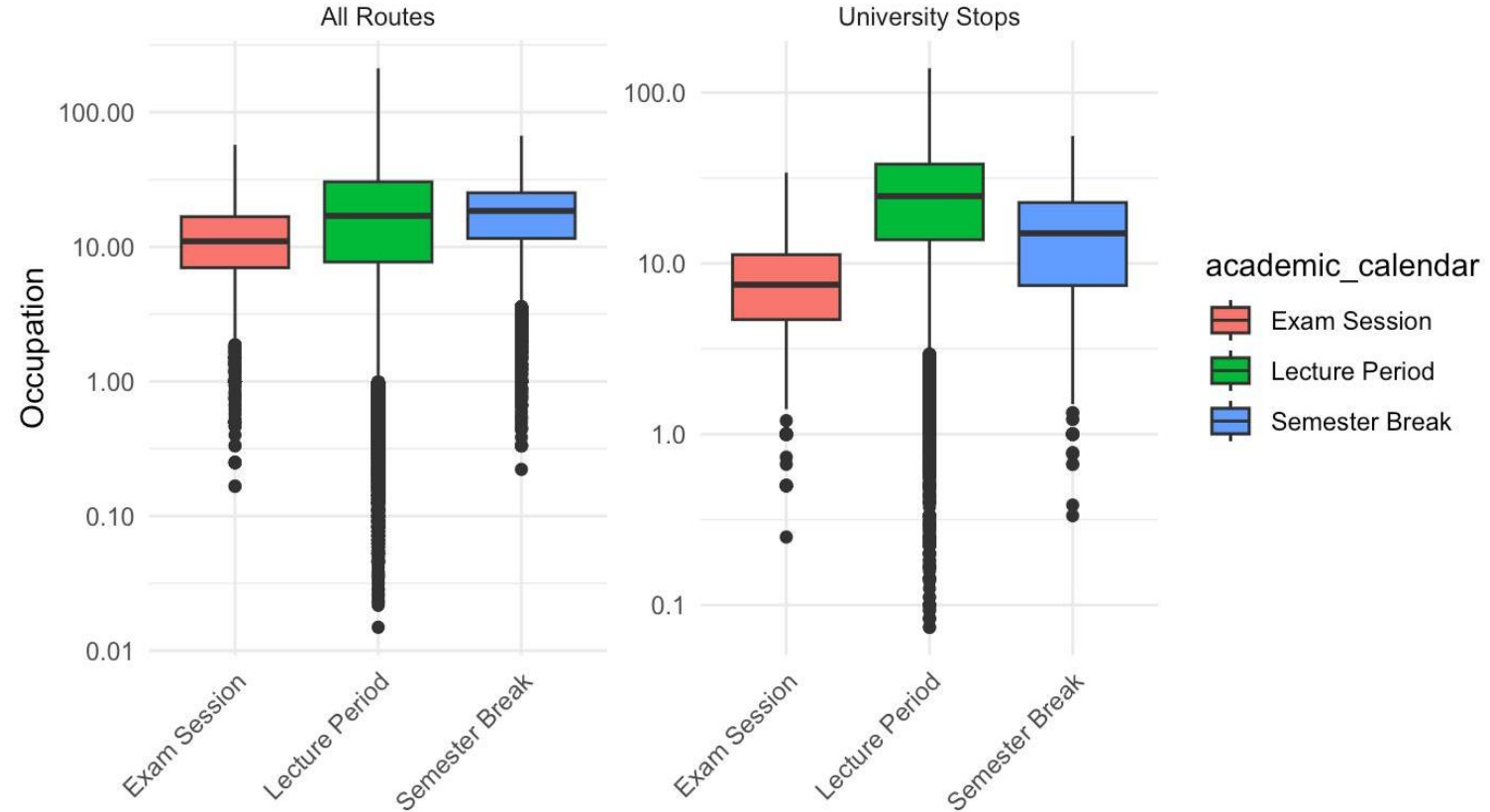
University stops' median occupation is lower compared to all routes, suggesting fewer passengers are traveling near university stops during exam sessions.

Lecture Period:

Compared to other time frames, during lecture period the occupation is the highest for university stops.

Semester break:

Overall highest occupation, but not for university stops.





Spatial temporal distribution

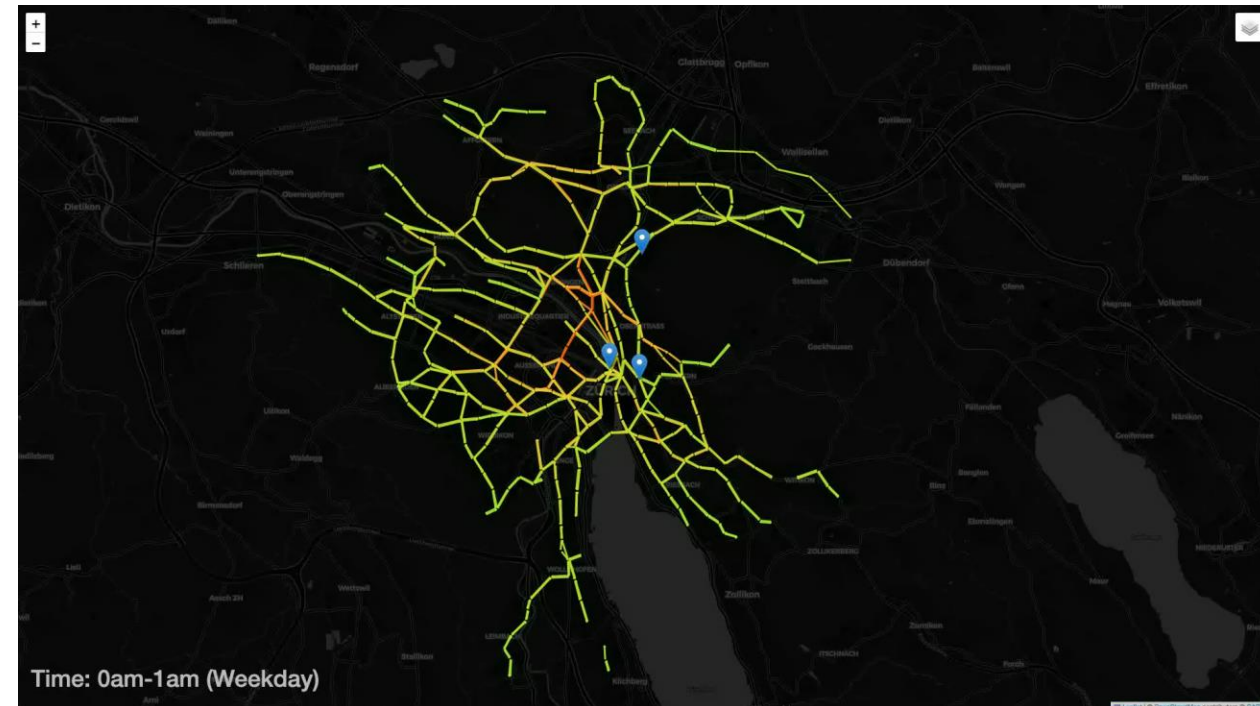
Geospatial Analysis

Method

- Merge data with GPS coordinates
- Aggregate data per hour
- Draw arrows on top of the map of Zurich
- Color arrows according to seat occupation

Findings

- Weekday: A lot of traffic around 8 am and 5 pm towards / away from the main station
- Irchel: Students arrive ~8 am and leave ~6 pm
- Weekend traffic constantly increases until a peak is reached around ~5 pm



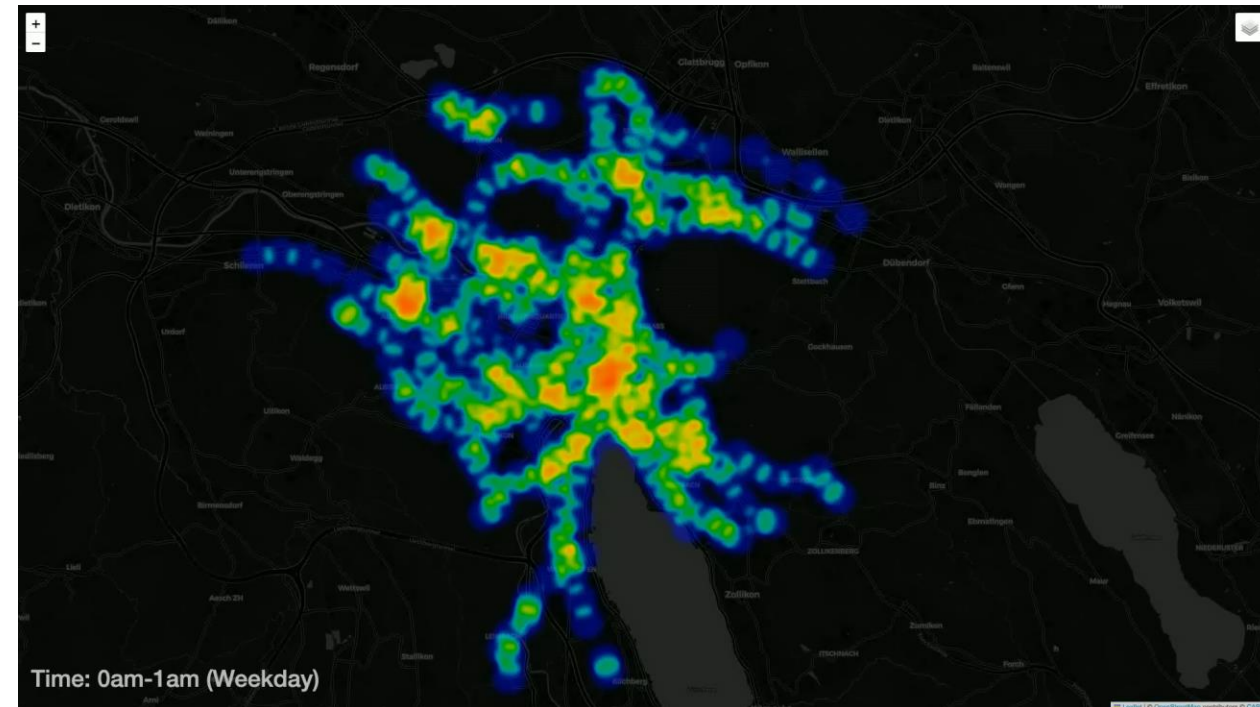
Geospatial Analysis

Method

- Merge data with GPS coordinates
- Aggregate data per hour
- Sum passengers boarding and alighting at stops
- Plot heatmap on top of city of Zurich

Findings

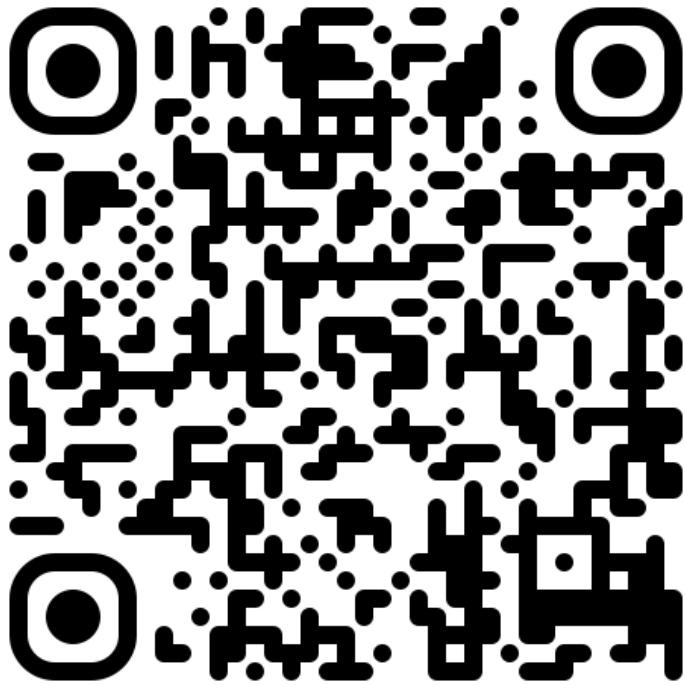
- Busy areas: Main station, Alstätten, Paradeplatz, Bellevue





Prediction:
number of available seats

Documentation



<https://sagerpascal.github.io/uzh-data-science-project/>

Navigating Zurich: A Comprehensive Analysis of Urban Traffic Dynamics

Pascal Sager*, Luca Zhao*, Weijia Zhong*, Xiaohan Zhu*

Zurich University

Project Work: Introduction to Data Science

*Indicates Equal Contribution

[Project Proposal](#) [Code](#)

