

Proposal Foot & Bike Traffic in Zurich

Julie Tschanz, Philipp Wyss, Damian Brühlhart, Mike Krähenbühl

What is the general topic ?

Renting public bikes has become more and more popular these past few years and to approach the rising demand one would want to discover the pattern of bike renting to better fit the offer to the demand.

What data will be used

Traffic Data - City of Zurich

Data (Source): [Daten der automatischen Fussgänger- und Velozählung - Viertelstundenwerte](#)

The data were collected by the civil engineering office of the city of Zürich and recense the foot and bike traffic at different geographic points throughout the city. The data set contains the following variables :

- **zählgerät** : ID
- **Freundschlüssel Zählstelle** : ID of the location where the measurement was taken
- **Datum** : Date in the format YYYY-MM-DD
- **Velo_in** : Number of bikes arriving in the station
- **Velo_out** : Number of bikes getting out of the station
- **Fuss_in** : Number of pedestrians arriving in the station
- **Fuss_out** : Number of pedestrians getting out of the station
- **Koordinate_Ost** : West coordinates (longitude)
- **Koordinate_Nord** : North coordinates (latitude)

More readings about how this data collection is being conducted:

[Automatische Zählungen des Fussverkehrs](#)

[Automatische Zählungen des Veloverkehrs](#)

Weather Data - City of Zurich

- [Hourly meteodata](#)

The data set includes hourly values from 1992, divided into annual files. The variables are :

- **Datum** : Date
- **Standort** : Location
- **Parameter** : Parameter
- **Intervall** : Interval
- **Einheit** : Unit
- **Wert** : Value
- **Status** : Status

It is a long table holding following measurements (**Unit**):

Unit

| Unit |
|------------------------|
| air pressure |
| precipitation duration |
| global radiation |
| temperature |
| relative humidity |
| wind direction |

For example after successful import `weather.head()` gives you following export:

| Datum | Standort | Parameter | Intervall | Einheit | Wert | Status |
|-----------------------|-------------------------|-----------|-----------|---------|--------|--------------|
| 2023-01-01T00:00+0100 | Zch_Stampfenbachstrasse | T | h1 | °C | 11.57 | provisorisch |
| 2023-01-01T00:00+0100 | Zch_Stampfenbachstrasse | Hr | h1 | %Hr | 72.29 | provisorisch |
| 2023-01-01T00:00+0100 | Zch_Stampfenbachstrasse | p | h1 | hPa | 971.62 | provisorisch |
| 2023-01-01T00:00+0100 | Zch_Stampfenbachstrasse | RainDur | h1 | min | 0.00 | provisorisch |
| 2023-01-01T00:00+0100 | Zch_Stampfenbachstrasse | StrGlo | h1 | W/m2 | 0.01 | provisorisch |

Population Data - City of Zurich

- <https://opendata.swiss/de/dataset/bevolkerung-seit-1901>
- The data set contains the variation of population of the city of Zürich (Wirtschaftliche Wohnbevölkerung der Stadt Zürich nach Jahr, seit 1901.). The latest update is from the 8th of February 2024.
- The dataset holds following attributes:
- **StichtagDatJahr** : Year
 - **AnzBestWir** : Economic resident population of the city of Zurich by year

How will the data be processed ?

The different data sets will be aggregated based on date and exploratory analysis will be done to have an overview of foot and bike traffic during different periods of the year and in different weather conditions.

The number of bikes (maybe also pedestrians) at a given time of the day (or during a whole day) will be set as the outcome variables and the weather as well as the population will be set as the predictor variables.

What questions will we try to answer with the project?

- **Question #1** : How does weather, season and population affect foot and bike traffic in Zurich City and is it possible to predict?
Why does it matters : This could help public transport planners, city planners, street food vendors or for example city bike rental companies to improve their ressource allocation problem.
2. When do city bike rental locations need to be ready with high capacity depending on conditions (weather, season, ...)?

What techniques will be used?

- Exploratory analysis
- Correlation-analysis
- Lasso Regression for variable selection and variable analysis
- Visualization of geo data (if possible)
- Model for prediction of traffic (GLM, random forest, ...)
- Maybe pose as classification problem