Applying Traffic & Demographic Database for City Planning

Project Status & Data Modeling

Team: ACID

Database Management Project, HSLU 2025

Presenting Our Case

The 'Why'

- **Hypothetical Client:** City of Zurich Urban Planning Office
- **Problem:** Increasing population density vs. limited road infrastructure
- **Mission:** Identify "Stress Zones" where population growth outpaces traffic capacity

Analysis Hypothesis: The 'Stress Index'

To quantify the 'stress' on urban districts, we propose the following hypothesis based on linked data:

Data Linkage & Key Metrics

With our data now unified by `DistrictID`, we can analyze key growth indicators:

- **Traffic Volume Growth %** (from Table A)
- **Population Growth %** (from Table B)

The Insight: Identifying District Bottlenecks

The relationship between these growth metrics helps us categorize districts:

- **Commuter Hubs:** Districts where Traffic Growth > Population Growth
- **Residential Zones:** Districts where Population Growth > Traffic Growth

The Raw Data Landscape

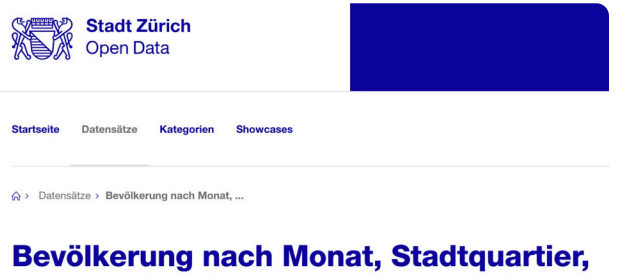
2 + 1 Data Sources

Source A: Traffic Data

Time-stamped car counts on Zurich streets since 2012.

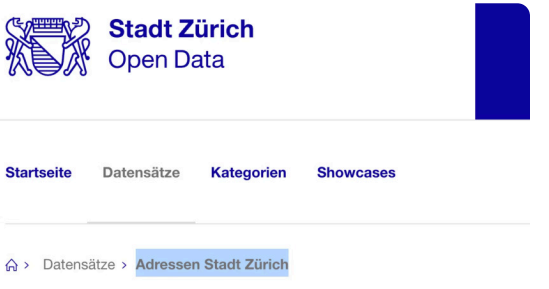
Over 2M rows of hourly sensor readings.

Source B: Population Data



Monthly demographic aggregates for various districts/kreise.

Source X: Geo-Address Data



Connects traffic and population data by street and district.

Deep Dive: The 'Missing Link' Solution

The Problem & The Solution

The Missing Link

Traffic data (Source A) lacks the "Quarter_name" needed to directly join with Population data (Source B).

External Dataset

The `str_stadtquartier.csv` dataset provides the precise street-to-district mapping required.

Visualizing the Connection

This external dataset enables us to add the "Quarter_name" to the traffic data, finally allowing a join with the population data.

population_raw.csv

Column	Example Value	English Translation
StichtagDatJahr	1998	Reference date year
StichtagDatMM	1	Reference date month (number)
StichtagDatMonat	Januar	Reference date month (name)
StichtagDat	1998-01-31	Reference date
SexCd	1	Sex code
SexLang	männlich	Sex (male)
AlterV20ueber80Sort_noDM	1	Age group sort order
AlterV20ueber80Cd_noDM	1	Age group code
AlterV20ueber80Kurz_noDM	0-19	Age group short (0-19)
HerkunftCd	1	Origin code
HerkunftLang	Schweizer*in	Origin (Swiss)
KreisCd	2	District code
KreisLang	Kreis 2	District 2
QuarCd	021	Quarter code
QuarLang	Wollishofen	Quarter name (Wollishofen)
DatenstandCd	V	Data status code
DatenstandLang	Veröffentlicht	Data status (Published)
AnzBestWir	950	Population count

traffic_data_cleaned.csv

Column(translated)	Example Value
measurement_site_id	Z001M001
measurement_site_name	Unknown
counting_site_id	Z001
counting_site_name	Seestrasse (Strandbad Wollishofen)
position_description	Unknown
east_coordinate	2683009.89
north_coordinate	1243936.2
direction	outbound
signal_id	789
signal_name	Badanstalt Wollishofen
num_detectors	1
timestamp	2013-01-01T00:00:00
vehicle_count	224.0
vehicle_count_status	Measured

str_stadtquartier_raw_for_join.csv

Column	Example Value	English Translation
adresse	Zähringerstrasse 43	Address
anzahl_fla_projektiert		Number of units (planned)
anzahl_fla_real	2	Number of units (realized)
flaeche_projektiert		Area planned (m²)
flaeche_real	449	Area realized (m²)
flaeche_total	449	Total area (m²)
gwr_egid	140003	Federal building ID
hausnummer	43	House number
lokalisationsname	Zähringerstrasse	Street name
objectid	1	Object ID
stadtkreis	1	City district
statistisches_quartier	Rathaus	Statistical quarter (Rathaus)

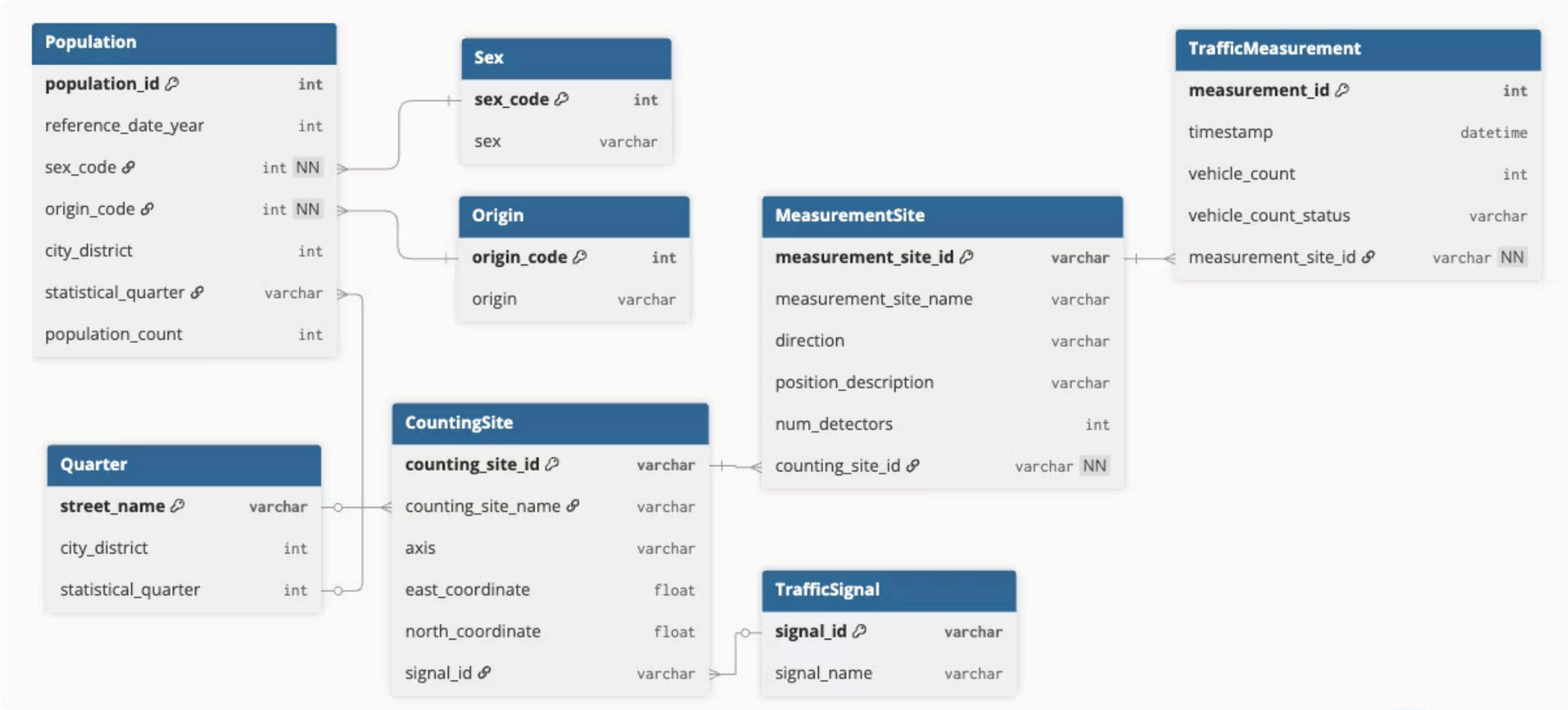
Conceptual Data Model

Entity-Relationship Diagram (3NF achieved without transitive dependency)

- **Entities:**

  Identified core objects: TrafficMeasurement, CountingSite, District, PopulationCount, Street, and an implicit mapping entity for Street-District relationship.

- **The Challenge:**
  - Too many no-use variables, hard to decide which to keep/drop in the beginning.
  - Identify exact structures of the model
  - Team communication on complex structure



Challenges & Future Steps

- **Data Cleaning/Processing:**

  Inconsistent street name formats (e.g., "St." vs. "Strasse"), and missing street & district connections, require extensive data cleaning and standardization efforts.

- **Data Integration:**

  Normalize and integrate three data sources (traffic, population, and addresses) using SQL. Transform into our data model.

- **Timeframe Mismatch:**

  Aggregated hourly traffic data to align with less frequent monthly population data, presenting a challenge in data harmonization.

- **Next Step:**

  Develop Analysis