GymCPH Access Data

Case

Magnus Berg Sletfjerding
March 21st, 2025

The Idea

The Solution

Experiment overview 👮

Summary

The Idea

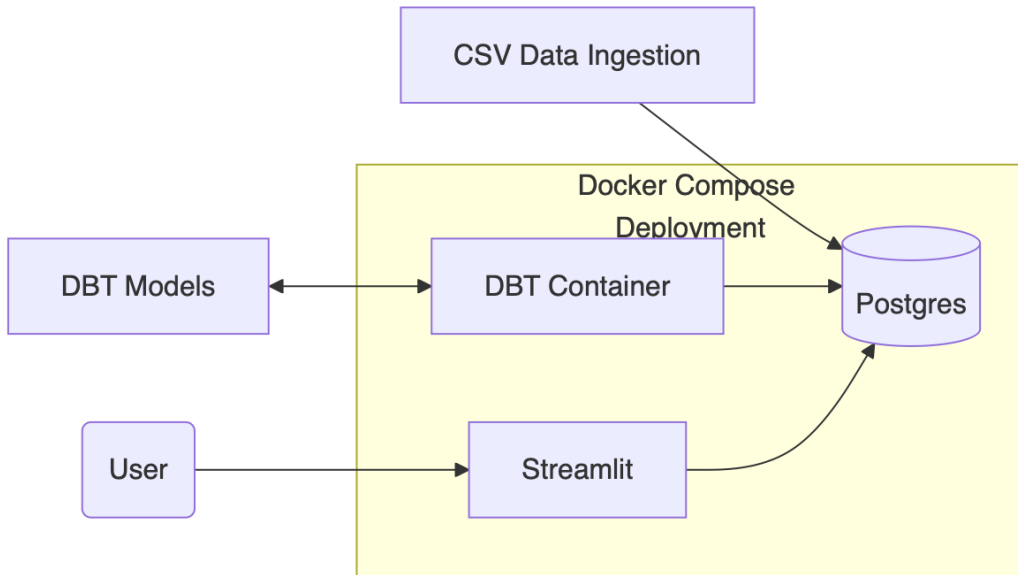## Find data for a CPH based gym and analyse it 🏸

GymCPH Intro

1. GymCPH is a medium sized gym in the Greater Copenhagen Area
2. They have asked us to find recommendations for how to improve their offering
3. Their main goal is customer satisfaction and their current revenue allows them to operate comfortably.
4. Owned by a non-profit they do not seek profits directly but would like to retain members.
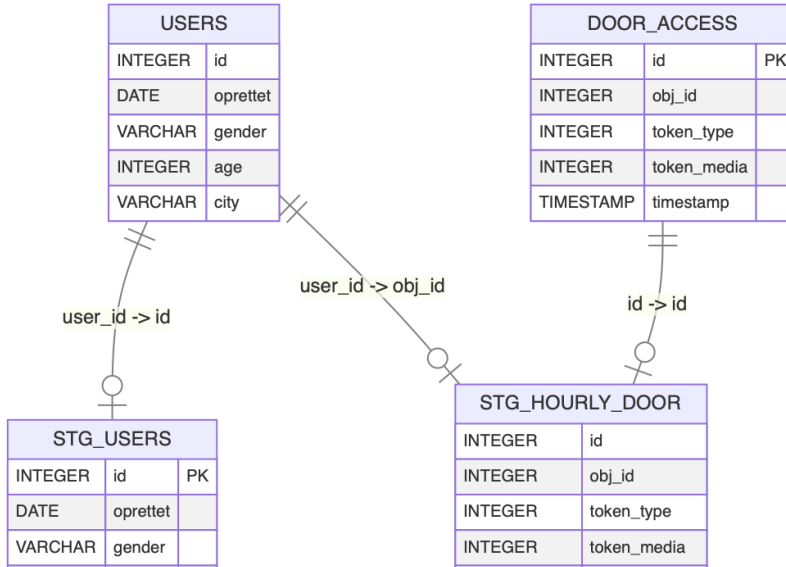
Data available

1. Data from door access system
2. Membership database dump (anonymised) 🇪🇺
3. All public data

The Solution

USERS

INTEGER	id
DATE	oprettet
VARCHAR	gender
INTEGER	age
VARCHAR	city

DOOR_ACCESS

INTEGER	id	PK
INTEGER	obj_id	
INTEGER	token_type	
INTEGER	token_media	
TIMESTAMP	timestamp	

user_id -> obj_id

id -> id

user_id -> id

STG_USERS

INTEGER	id	PK
DATE	oprettet	
VARCHAR	gender	

STG_HOURLY_DOOR

INTEGER	id
INTEGER	obj_id
INTEGER	token_type
INTEGER	token_media

Experiment overview 🧐✈️

Actions

1. First I uploaded the Data into Snowflake, and began running SQL scripts on it.
2. However, once I tried to use the DBT Labs connection I ran into trouble. ➡️
3. At this point I was rather pressed on time, so I said *bye Snowflake* and swapped to my trusty postgres 🐃

Blocker

Discover	Data products	Providers	My requests

Requested: 8 hours ago

1. Runs locally
2. No login required
3. No quarantine time required to pull
   data

Summary

1. Add weather data to check correlation over time
2. Segment by age and find if there are some times that are more/less popular with age.

What I would look into for scalability

- Snowpipe setup for streaming data into it
- Airbyte type ELT job
- Connecting to internet to be able to process larger datasets
- Figure out how to load HDF4 files from MODIS to do Geo Based things