Source

insideairbnb.com

Object Type: Gzipped CSV Files

Interface: Webscraping using Python

Data Warehouse

PostgreSQL hosted in Supabase

Bronze Schema

ETL Script

Raw Data

Object Type: Tables

Load Type: Full Load

Insertion Type: Truncate and Insert

Transformations: Each row of data saved as JSONB format row in table

Silver Schema

ETL Script

Processed Data

Object Type: Tables

Load Type: Full Load

Insertion Type: Truncate and Insert

Transformations: Data Cleaned and Split into Dimension and Fact Tables (Star Schema).

Gold Schema

Visualization Ready Data

Object Type: Materialized Views

Transformations: Joins and Aggregations using the Silver Tables

Target

Streamlit Webapp

Object Type: Python based Streamlit App

Interface: Streamlit Cloud

All ETL Scripts Deployed in a Docker Container and automated using GitHub Actions