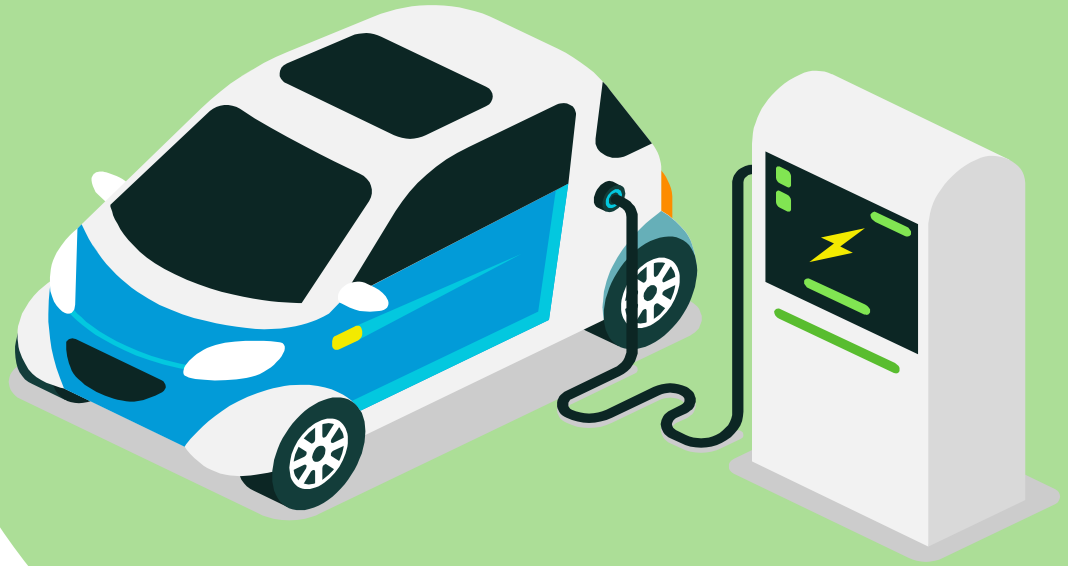
Uof T Data Analytics Project 3 – Data Engineering: Electric Vehicles in Washington State

Team:
Lucas Perez
Sultan Raheem
Mahind Rao
Rachel Schoen



WHY?

Create a database to answer:

How does the presence and accessibility of electric vehicle (EV) charging infrastructure impact the adoption of EVs and contribute to reducing carbon emissions across different regions?

Ethical Considerations

Bias & Fairness

Benefit & Harm

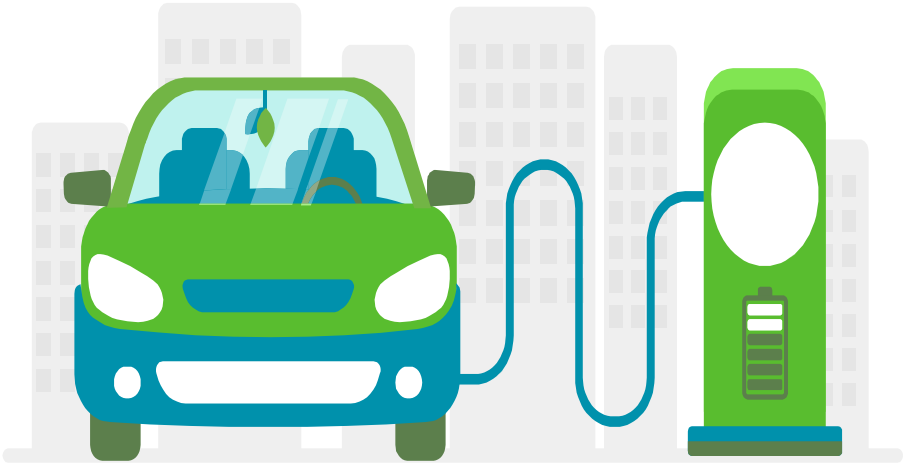
Data Ownership & Governance

Transparency & Accountability

Privacy & Protection

Public Interest & Impact

SQL Database



Structured Data

Complex Queries

Data Integrity & ACID Compliance

Scalability for Structured Data

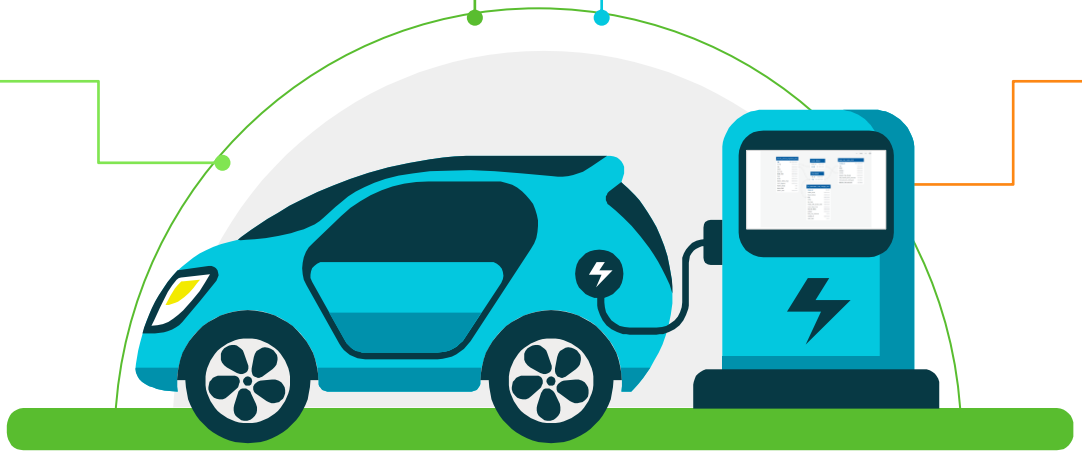
Maturity & Ecosystem

Data Base Design

ghgp_data_carbon_2022 is linked to the **county_dataset** and **city_dataset** to relate carbon emissions data to specific counties and cities.

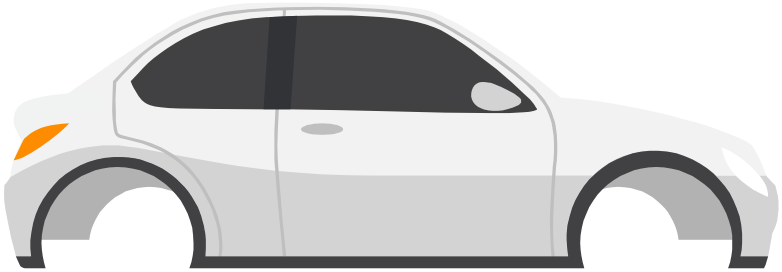
Electric_Vehicle_Population_Data is linked to the **county_dataset** and **city_dataset** to associate vehicle data with specific locations.

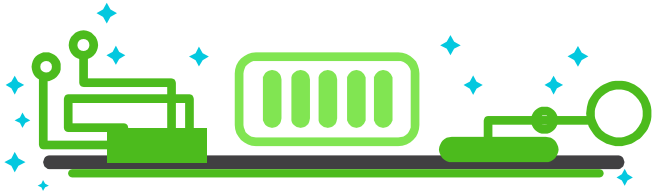
Five Datasets: ghgp_data_carbon_2022, Electric_Vehicle_Population_Data, EV_Alternative_Fuel_Charging_Stat, county_dataset and **city_dataset**



EV_Alternative_Fuel_Charging_Stat is linked to the **city_dataset** to associate charging station data with specific cities.

Data & Delivery





01 Data Preparation

Datasets are CSV files taken from reputable sources. Cleaned CSV on Excel.

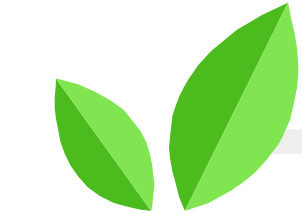
02 Data Validation using Pydantics

Pydantics chosen for its automated testing.

03 Reading the Data & displaying it

Used SQLalchemy to read data & display it.

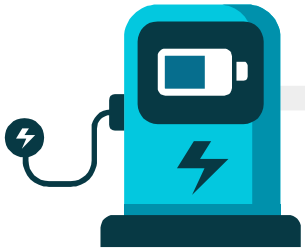
Data Validation Using Pydantics



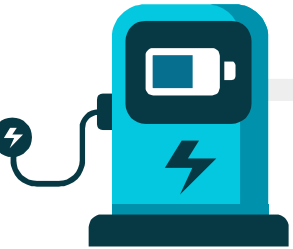
Initializing the Pydantic Workspace

Deploy Pydantic

Validate all Datasets



Read the Data and Display it
SQLalchemy



Convert Postgres database to SQLite

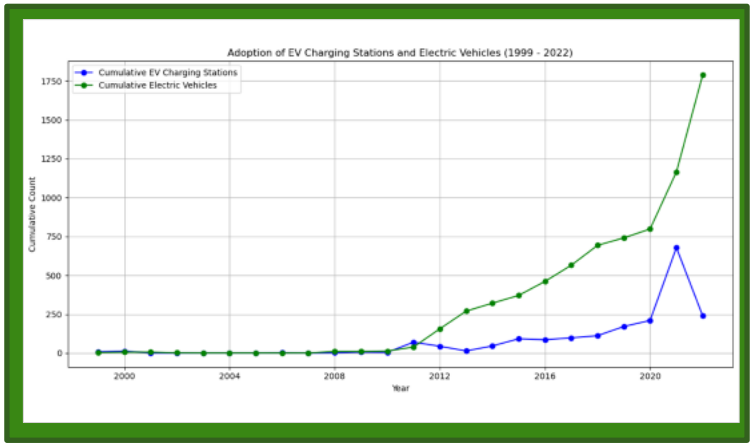
Create Engine

Create a link

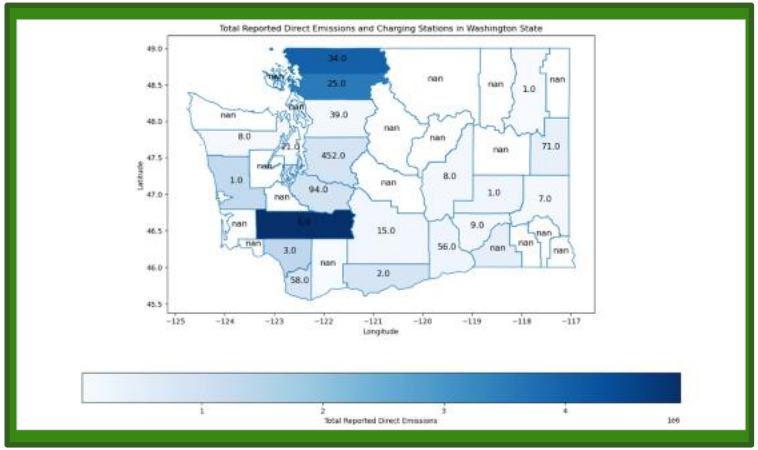
Transform to Pandas data frame

Data Base Interactions & Visualizations

How does the availability and accessibility of EV charging stations correlate with the adoption rates of electric vehicles in various counties?

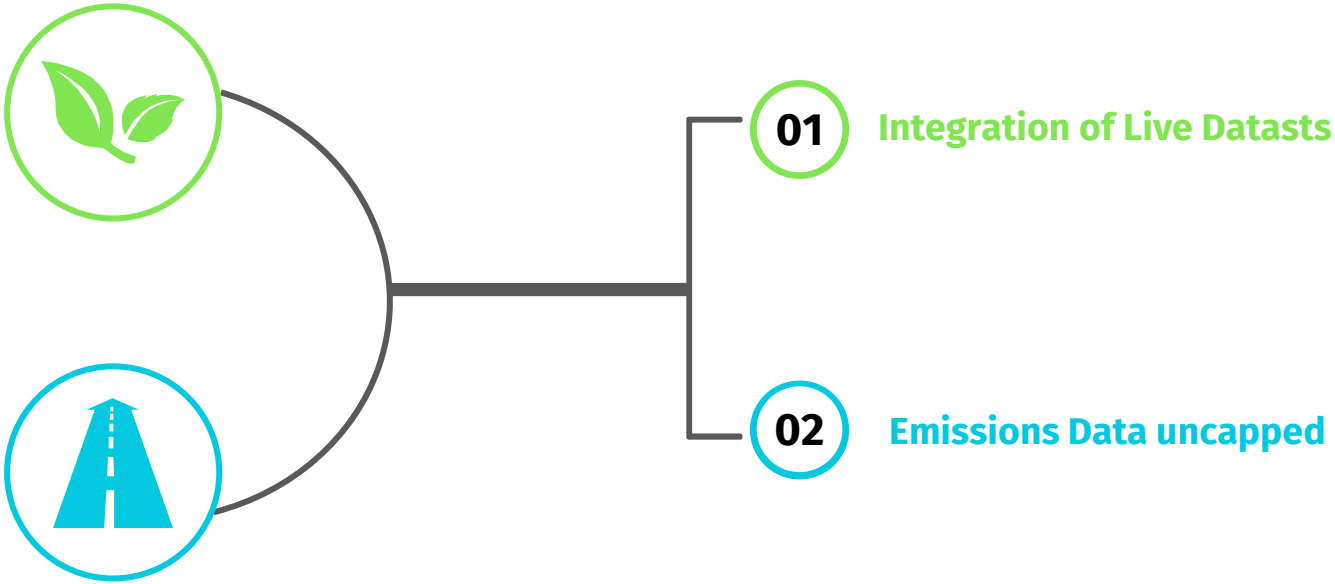


What is the relationship between the increase in EV charging stations and the reduction in industrial emissions by facilities within those regions?

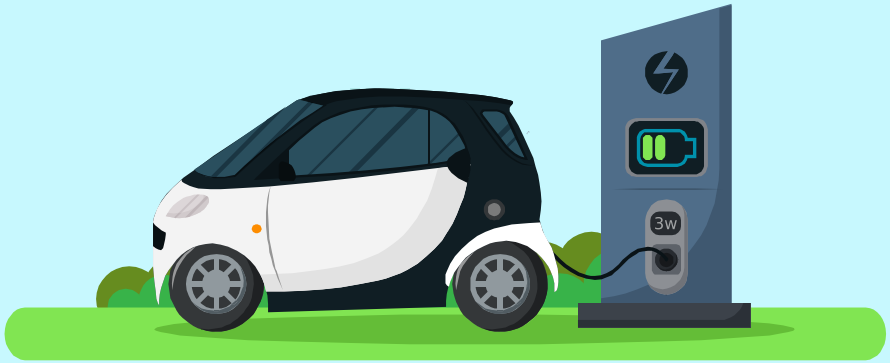


*Assumption: Emissions are derived from different industries' output and are related to automobiles

Closing Remarks



Questions?



References

https://www.kaggle.com/datasets/saketpradhan/electric-and-alternative-fuel-charging-stations

https://catalog.data.gov/dataset/electric-vehicle-population-data/resource/8abbca4e-2f1c-4249-989f-b96201dd4710?inner_span=True

https://www.epa.gov/ghgreporting/data-sets

https://www.bigendiandata.com/2017-06-27-Mapping_in_Jupyter/

https://stackoverflow.com/questions/55636285/how-to-fix-valueerror-zero-size-array-to-reduction-operation-fmin-which-has-no

https://chatgpt.com/share/23883159-f752-479a-a899-5be44adbeaea

https://github.com/matplotlib/matplotlib/issues/16284

https://realpython.com/python-map-function/#understanding-map

https://chatgpt.com/share/2b7f03b5-3f03-408d-84cb-0e37741ca0ba

https://chatgpt.com/share/2b7f03b5-3f03-408d-84cb-0e37741ca0ba

https://jcutrer.com/python/learn-geopandas-plotting-usmaps

https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.loc.html

https://www.earthdatascience.org/courses/use-data-open-source-python/use-time-series-data-in-python/date-time-types-in-pandas-python/customize-dates-matplotlib-plots-python/#:~:text=Using%20the%20DateFormatter%20module%20from,number%20with%20lower%20case%20m

https://stackoverflow.com/questions/62299949/month-starting-date-and-ending-date-between-a-range-of-date-in-python

https://www.recurrentauto.com/research/washington-electric-vehicles