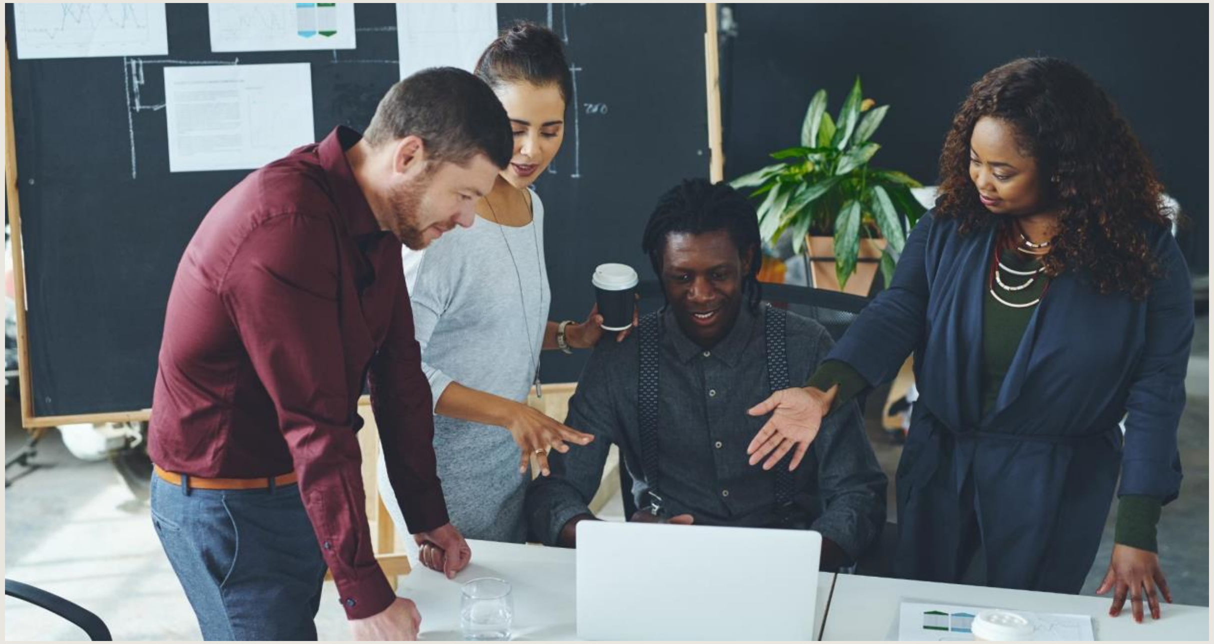
EVSA
(CHARGING) INFRASTRUCTURE-RELIABILITY AND USAGE

TEAM MEMBERS:



Peter Minor

Shammi Pereira

Cassandra Gomez

Colin McNatt

PROJECT BACKGROUND:

Public satisfaction with EVSE chargers has declined year after year for the past 5 years.*

- Common issues include:
  - Accessing charge stations (full stations)
  - Hardware malfunctions (cords, plugs, display)
  - Software problems (charge cards denied, communication with vehicle)

The adoption of EVs (BEV and PHEV) depends on user perception of available charging which is a contributing factor to range anxiety a major reason consumers do not purchase EVs.

*JD Powers Public Satisfaction with EVSE Chargers (https://www.jdpower.com/business/electric-vehicle-experience-evx-public-charging-study).

INTERESTING QUESTIONS:

- Develop a definition of EVSE performance and reliability based on charging errors, availability and utilization.
- What <u>patterns</u> of <u>best</u> and <u>worst</u> performance of EVSE chargers can be identified by data mining?
- How can public policy be targeted to improve consumer satisfaction with Public EVSE chargers?

PRIOR WORK:

Source	Summary
Gamage et. Al. (2023). "Reliability of Corridor DC Fast Chargers and the Prevalence of no-Charge Events". University of California, Davis, Electric Vehicle Research Center.	Focus on transport corridors (interstate highways) and EVSE chargers. Utilizes EV Watts database. Defines performance factors utilizing EVSE charger flags and wait times.
Energetics (2023). "EV Watts Bi-Annual Activity Update". Energetics consulting.	Summary of national public EVSE charger performance and availability by type of charger (level 2 or level 3)
Idaho National Laboratories (2015). "What can be Learned from the EV Project to Inform Others who are interested in Similar Studies", Idaho National Laboratories (2015).	Interesting questions to be asked about the state of performance of EVSE infrastructure. Paper written before the creation of EV Watts Database.
ClearResult (2024). "Electric Vehicle Charging Station Reliability" EV WATTS White Paper Series.	Defines Charger reliability and provides nationwide summaries of charger reliability.
Energetics (2023). "Multi-Unit Dwelling Utilization". Energetics Consulting.	Analysis of EVSE reliability and utilization in multifamiliy venues.

SUMMARY OF PRIOR WORK:

- Recent studies have focused on specific segments of the EVSE charger network such as transport corridors or individual aspects of the network such as:
  - Specific venues (multi-family housing chargers)
  - Specific use cases (fleets and transport corridors)
  - General and national reliability and performance
- Studies have also defined accessibility and performance of EVSE charger utilization in the EV Watts database, but have only reported broad measures of reliability, nationally.

DATA SETS TO BE USED:

- EV Watts Public EVSE Data Set includes:
  - EVSE Charger Table (PK-EVSE_id)
  - Connector Table (PK-connector_id, FK-EVSE_id)
  - Session Table (PK-Session_id, FK-EVSE_ID)
  - Vehicle Table (PK-Vehical_id)
  - Vehicle Trips Table (PK-Trip_ID, FK-Session_ID)
- Other datasets as needed (e.g., census data on metropolitan areas)
- Data set is publicly available from the US Department of Energy https://livewire.energy.gov/ds/evwatts/evwatts.public
- The Dataset has been downloaded, and segments have been loaded and are currently being reviewed for completeness and coverage.

DATA LOCATION AND COORDINATION:

- The data are being stored on a Microsoft OneDrive which all team members have access to.
- Each team member is responsible for downloading their own copy of the data.
- Code for reading in and cleaning the data will be stored on GitHub and each team member will be able to "pull" the latest code for cleaning the data from the team repo.

DATA CLEANING & PREPROCESSING:

- **Rename column names** to make them clearer
- Determine **missing/NaN values** and remove rows with missing values, remove columns with many missing values, impute a mean or median for the missing values or fill in the missing values with inference-based values
- Handle **duplicates**
- Manage **noisy data** through binning, regression and/or clustering techniques
- **Convert categorical values** to numerical values
- **Normalize data** in range 0 to 1 using Manhattan Distance and Simple Matching Distance
- Plot boxplots for each column to determine **outliers** and remove them
- Plot histograms for each column to determine **probability distribution** and **modality** of the dataset
- Plot a correlation matrix to determine **positive and/or negative correlation** between the columns

DATA INTEGRATION:

- The core of the EV Watts Database is the EV Charger Table. For this table to be useful, it must be joined with other tables, such as sessions and vehicle trips
- In most cases, the relationship between the charger table and other tables is a one-to-many relationship.
- The tables will have to be joined carefully on their primary and foreign keys to be sure the dataset size does not grow to an unusable size. Sampling or reduced time frames may be used to reduce the size of the data set.
- Look up tables for charger error flags will have to be integrated.

DATA INTEGRATION CONTINUED:

- The EV Watts Data set does not include detailed location information, such as GPS coordinates.
- Instead, the database indicates the metropolitan statistical area of the charger.
- The Census bureau maintains data on population and business and can be integrated into the analysis as needed (https://www.census.gov/geographies/reference-maps/2020/geo/cbsa.html).

DATA TRANSFORMATIONS:

A major task in this research will be creating an indicator of EVSE quality.

We propose to create an **index of quality**, which will combine several features such as number of charger error codes, access (based on availability of the charger), charge level (watts), and venue convenience.

DATA AGGREGATIONS:

- The data set are highly detailed. There can be dozens of charging sessions per charger per day, and there are 365 days in a year. There are 3.5 years of data.
- At the same time, certain variables, such as location of the charger, are somewhat aggregated (metropolitan statistical areas), to maintain confidentiality of the data.
- As a result, the data will have to be aggregated over time and space.
- This will be a help in the final analysis, but the initial task will require a lot of computer processing power.
- Data samples will be taken to develop code before running it on the entire data set.

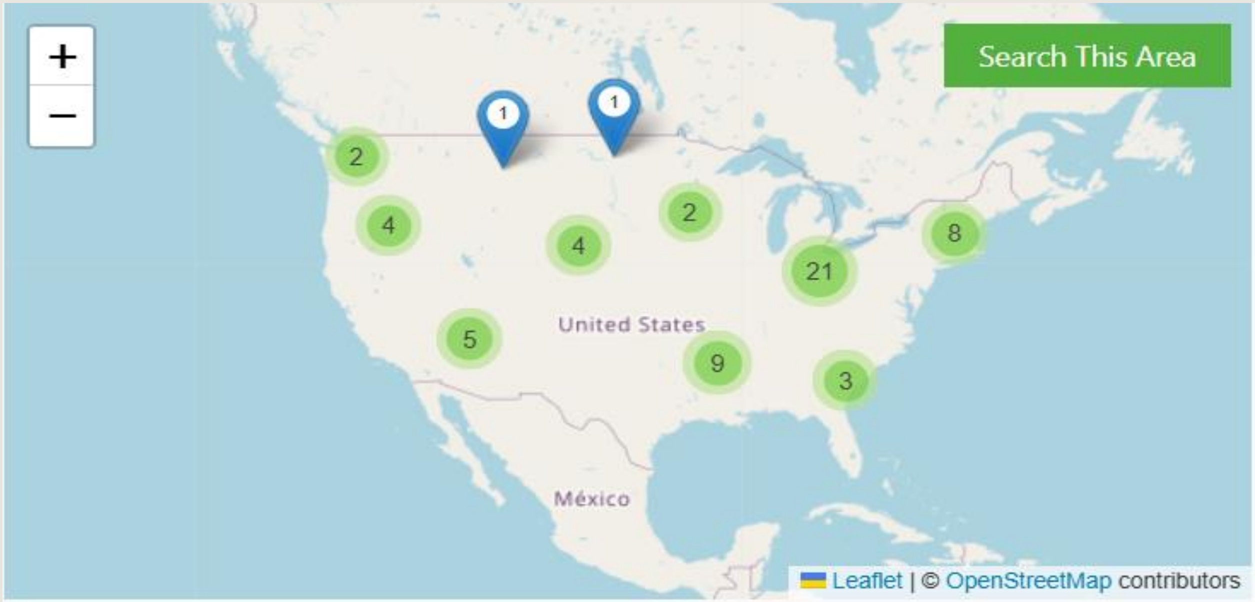
TOOLS:

- Principal Tools:
  - Python (Jupyter Notebooks)
  - Pandas (in Python)
  - NumPy (in Python)
  - Sci-Kit Learn (aka, SKLearn, Data mining and statistical tools, in Python)
  - Matplotlib and Seaborn (Graphing, in Python)
  - Plotyly and Plotly Dash (creating a dashboard, if needed, in Python)
  - GitHub
  - MS OneDrive (MS Word, Excel, PowerPoint)

EVALUATION:

DATA INTEGRITY

Data Map



Green numbers indicate multiple datasets close to each other. Zoom in to see individual datasets, represented in blue.

EV Watts Public Database:

How sound is the data?

Who performed the study and why?

This study was conducted by Energetics and various partners in the hopes of providing a clear image on user practices and habits with EV usage.

Data Range and Reasoning

The data covered has to do with Electric Vehicles, the Service Equipment for said vehicles, and overall data between the two. This gives a comprehensive insight as to user patterns and possible correlations.

Scalability and Integrity

The important part of our data is how many data objects were included with the initial study. The amount seems to be about 1,132 vehicles. This was collected over roughly 4 years from 2019 to 2022.

EVALUATION CONTINUED:

HOW WILL WE UTILIZE OUR DATASETS?

EVSE Quality Assessment:

Range Between Charging stations

We look to the duration between charge sessions and how much driving vehicles did before charging again. This helps us gather an overall image of usage patterns. It could also help in understanding where new stations could be placed in the future.

Patterns of Quality Differences

By observing the differences between charge stations and their individual power output as well as their locations, we can derive a potential pattern for lower or higher outputs. We can use this to discern if the quality of the charge was sufficient based on the frequency or duration based on the time between one charge to the next.

Public Opinion and EVSE Chargers

We can look toward specifically low use and low output EVSE to understand usage patterns for these generally lower quality machines. With this, we can observe where some injections of more power output may be utilized to perhaps increase public opinion and in turn, usage of these stations.

EVALUATION OF RESULTS:

As previously discussed, a primary measure of performance must be derived from the EV Watts Data Base.

As a first measure for evaluating the accuracy of these measures, we will compare our measures to broad datasets, such as JD Powers consumer satisfaction surveys, to ascertain any correlation (e.g., if our study says all EV chargers work as intended, and JD powers says low satisfaction, we will have to reconsider our measure).

We will also compare summarize our detailed regional, state and venue data to the broad national summaries of EVSE reliability.

EVALUATION CONTINUED:

IMPACT

We will be looking at the potential impacts of our work for public policy.

To assess impacts, we will review key results for interesting features of interest to policy makers, such as weak and strong areas of the EVSE public charging infrastructure.