CSC 805 - Data Visualization
Visualization Project - Phase 3

Jijeong Lee, Parth Panchal, Mark Kim

November 26, 2023

Contents

1 Project Report

1.1 System Architecture

1.2 Dataset Description

The data set used in this visualization project consisted of:

1. Washington State Electric Vehicle Population Data; and

2. Alternative Fuel Stations By State (Updated 10-12-2023).

1.2.1 Washington State Electric Vehicle Population Data

The Washington State electric vehicle data contains 143596 records with fields that are described below (irrelevant fields omitted). A screenshot of the overview of the data is provided in Figure 1.

VIN the last 10 digits of the vehicle ID number;

Location the location where the vehicle was registered by:

- county,
- city,
- state,
- postal code,
- latitude,
- longitude, and
- 2020 census tract;

Model Year the model year of the vehicle;

Make the make of the vehicle;

Model the model of the vehicle;

Type the electric vehicle type (BEV or PHEV);

Electric Range the electric range in miles;

Base MSRP the base MSRP of the vehicle;

Legislature District the legistlative district of the vehicle; and

Electric Utility the electric utility of the location of the vehicle.

1.2.2 Alternative Fuel Stations By State

The Alternative Fuel Station data contains 66259 records with the fields described below (irrelevant fields omitted). A screenshot of the overview of the data is provided in Figure 2.

Station Name the name of the station;

Location the location of the charging station by:

- street address,
- city,
- state,

Figure 1 (screenshot):

| Overview | Alerts 16 | Reproduction |

Dataset statistics

Number of variables	17
Number of observations	143596
Missing cells	204567
Missing cells (%)	8.4%
Duplicate rows	0
Duplicate rows (%)	0.0%
Total size in memory	18.6 MiB
Average record size in memory	136.0 B

Variable types

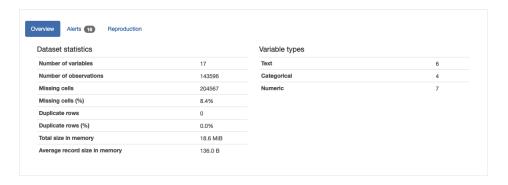|	
Text	6
Categorical	4
Numeric	7

Figure 1: WA State EV Data Overview

- postal code,
- latitude, and
- longitude;

EVSE Num the number of charging outlets by type:

- Level 1,
- Level 2, and
- DC Fast;

EV Network the charging network that the station belongs to; and

Open Date when the station opened.

Figure 2 (screenshot):

| Overview | Alerts 24 | Reproduction |

Dataset statistics

Number of variables	26
Number of observations	66259
Missing cells	349708
Missing cells (%)	20.3%
Duplicate rows	0
Duplicate rows (%)	0.0%
Total size in memory	13.1 MiB
Average record size in memory	208.0 B

Variable types

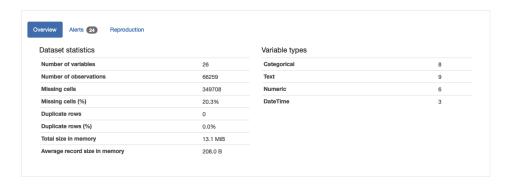|	
Categorical	8
Text	9
Numeric	6
DateTime	3

Figure 2: Charging Station Data Overview

Although we initially planned on using vehicle population for the entire US, we finally decided to abandon that effort for two reasons:

1. the sheer size of the dataset (over 5 Gb) would require compute resources that we did not have; and

2. the integration of the datasets was requiring many lookups and conversions that were rate-limited (geoid, county, and/or zip-code conversions, etc).

Each state provided data in formats that were not standardized with each other, and location data was stored in ways that made data integration near impossible with the time frame that we had. One example of this was where registration data for vehicles was being stored as congressional districts and/or county name. Although we found a free API capable of making these conversions, the request rates are throttled to one request per second. With this conversion model, we could make only 86400 requests per day, so with over 1 million records, this was not viable.

1.3   System Description

1.4   Screenshots

2   Demo Video

Demo Video

3   Project Source Code

EV data visualization project source code