Business Analytics Final Group Project:
Optimizing EV Charging Station Locations Across Washington State

Caterina Almazan (cga2133), Devyani Vij (dv2485), Kimmi Woods (khw2120)

Links to Collab Notebooks:

- BA-Final-Project.ipynb
- OPT-MODEL-BA-Final-Project.ipynb

Introduction to EV market:

The Electric Vehicle (EV) market has shown notable presence and potential growth as consumers become increasingly inclined to adopt more environmentally-conscious practices in their day-to-day lives and weigh options that may allow them to save money in the long term.
To fully capture the potential value of the market, companies may be looking to benefit from the implementation of optimization and analytics as a way to increase sales and appeal to new and existing customers, and government policy makers would better be able to allocate their resources to more efficiently incentivize this switch from fossil fuel vehicles.

Defining the Problem:

EV sales are continuing to rise, and will reach record levels by the end of 2024; however, the rate of EV sales growth is also slowing, at least temporarily. Worldwide sales of Plug-in hybrid electric vehicles (PHEV) will rise a little more than 20% this year compared to last year, according to a recent report from the International Energy Agency. That's a significant rise but still less than the 35% global increase between 2022 and 2023 [1].

One large barrier for the traction of EVs and PHEVs is a lack of public infrastructure to support these vehicles, namely, the charging stations. A lack of or limited access to chargers remains the biggest complaint among these car owners. There are about 144,000 public EV chargers in the United States, according to the Department of Energy. About 42,000 of those are in California. States like Mississippi and Montana, while far less populous, have only a few hundred. Away from home, charging your EV costs more than using home-installed chargers, sometimes twice as much [2]. According to the U.S. Department of Energy, most drivers of EVs and PHEVs charge their vehicles overnight at home using AC Level 1 or AC Level 2 charging equipment [3]; however, for many people – particularly those in apartment buildings and other non-home owners – charging at home is simply not an option. Additionally, some EV owners may just find it too difficult to visit a charging station that is too far away and inconveniently located for them.

Taking this into account, we decided to focus on applying clustering analysis and optimization techniques to the EV industry, specifically in the State of Washington, which is currently making plans to implement major investments in EV charging infrastructure and addressing barriers to EV ownership: Washington saw a huge 43% increase in the total number of EVs and PHEVs registrations. That figure rose from 118,050 in January 2023 to 168,850 by the year's end, according to Department of Licensing data [4]. Atlas Policy predicts that "governments, electric utilities, and private investors" have and will continue funneling large investments into creating these charging stations to support the growing demand for EVs [5].

Based on analytics of EV registration, locations of current charging stations, and general demographics of certain areas, we can provide important recommendations to all three of these station funders regarding which existing zip codes / clusters of EV cars that have larger than average distances from the nearest chargers, so that these underserved areas may be better targeted for future charging stations that can be built in those regions of Washington. We also aim to provide analysis as to the current utilization rate of each station cluster to provide further recommendations on which charging stations are being underutilized. Finally, honing in on the largest city of Seattle, Washington, we do less of a theoretical analysis and get to concrete recommendations for locations that either current charging stations should increase infrastructure, or temporarily closed stations should reopen.

Data Description & EDA:

For our project, we utilized two datasets: we gathered EV population data from a Kaggle dataset ("Population") showing registered EV information through Washington State Department of Licensing [6], as well as data on EV charging station location from the Joint Office of Energy and Transportation [7]. In our code, we split the charging station dataset ("Stations") into two dataframes for currently available stations ("Available_Stations") and temporarily unavailable stations ("Temp_Unavailable_Stations").

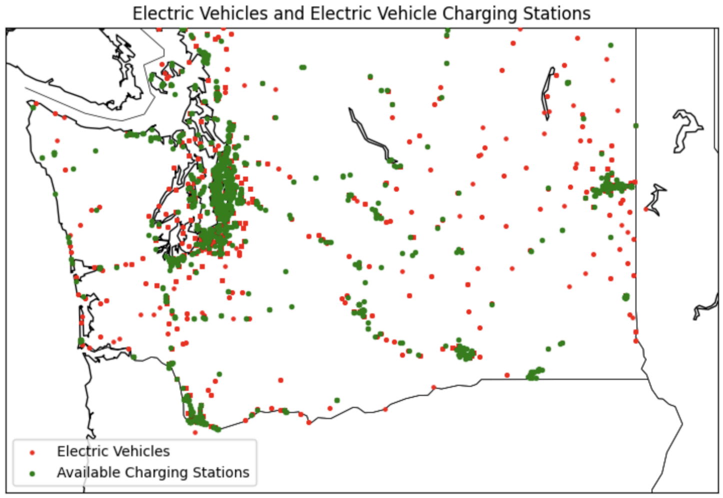
Both of these sets needed lots of cleaning and filling in for missing values. We deleted columns which we thought were unnecessary to our project. Columns we utilized the most were Latitude, Longitude, and zip code from both Population and Available_Stations; these show the coordinates and postal codes for the locations of the registered EV cars and for the current available charging stations, respectively. We also used EV Level2 EVSE Num and EV DC Fast Count columns from Available_Stations which shows the number of Level 2 chargers and Direct Current Fast Charging (DCFC) chargers available at each station, respectively. Level 2 chargers can charge a BEV to 80 percent from empty in 4-10 hours while DCFC equipment can charge a BEV to 80 percent in just 20 minutes to 1 hour [8].

Additionally, we noticed some discrepancies in the data: In the EV population data, the electric ranges for 124 of the cars (make/model/year) with missing data. We utilized ChatGPT to help us provide estimates for these 124 models, and merge them with our existing Population dataset.

Looking at the charging stations data set, the data source website stated that information on the private stations is not always published online or in the data download, but may be tracked only in the backend Station Locator database. Due to this mention of the incompleteness of the private station data, we made the decision to only consider public charging stations. We also made the decision to solely focus on BEVs, rather than also including PHEVs. This is because PHEVs can always have the option to run on gasoline, and we aim to center our analysis on the need for and impact of Washington's charging stations; thus, we dropped the PHEVs from our dataset.

Data Modeling, Model Results, & Analysis

For the initial step in our analysis, we created a geographic visualization to understand the spatial distribution of electric cars and charging stations. By plotting the latitude and longitude coordinates of both cars and stations on a map using Basemap in Python, we could visually inspect their locations relative to each other. This visualization helped us gain insights into the overall coverage of charging infrastructure and the distribution of electric vehicles, which are crucial factors in assessing range anxiety and planning for future infrastructure development

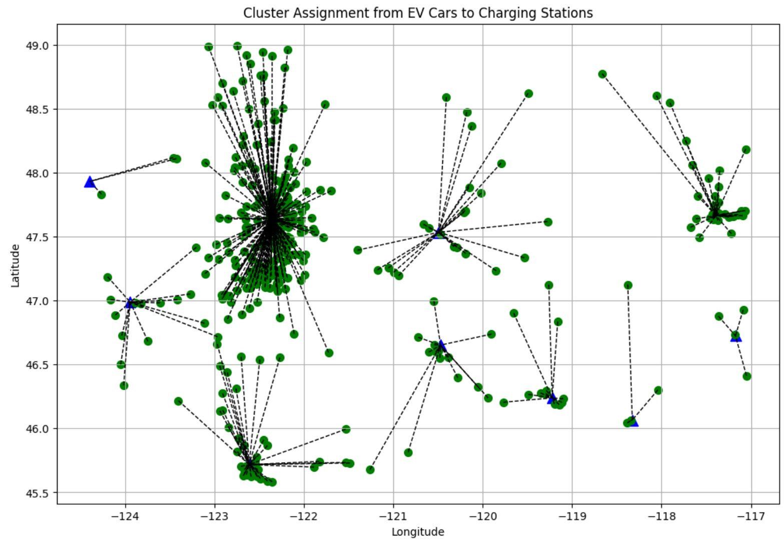


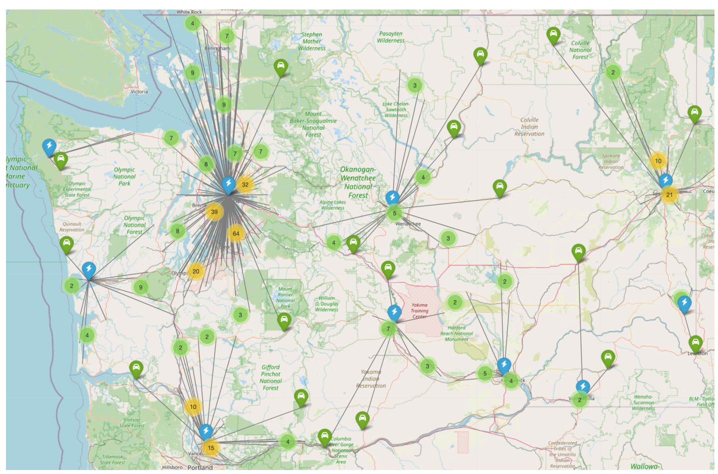
I. Clustering

Due to the vast amount of data, particularly the number of EV registration data points in our Population dataset, we decided the best way to assign cars to available charging stations was to form clusters, as there are too many cars to work with individually. Initially, we experimented with forming and assigning clusters of EVs to clusters of charging stations via K-Means clustering; however, we found this algorithm was a bit restrictive and inflexible with our amount of data, and also yielded a k value that was much too low (4) for the number of charging stations clusters we expected to see formed. We then utilized DBSCAN (Density-Based Spatial Clustering of Applications with Noise) to cluster the charging stations based on their geographical coordinates. Unlike with K-Means, DBSCAN can discover clusters of arbitrary shapes and does not require the number of clusters to be specified in advance. When creating the clusters of EVs from the Population dataset, we used HDBSCAN, an extension of DBSCAN which is more flexible with larger datasets.

From this method, we determined the number of clusters for available charging stations to be 10 and the number of EV car clusters to be 371 (Info on the exact sizes of each of these clusters can be seen in our Collab code). We then assigned each of the 371 car clusters to the closest charging station cluster, in addition to determining the euclidean distance between them (All these assignments and distances are printed in our code).

To better visualize this cluster assignment from EV Cars to Charging Stations, we created a scatter plot; additionally, due to the vast number of cars within each cluster, we thought an interactive map of these assignments and placement of clusters would be the most beneficial and creative. Below is a screenshot of this interactive map, but we highly recommend exploring the clusters and interactive features in our code.

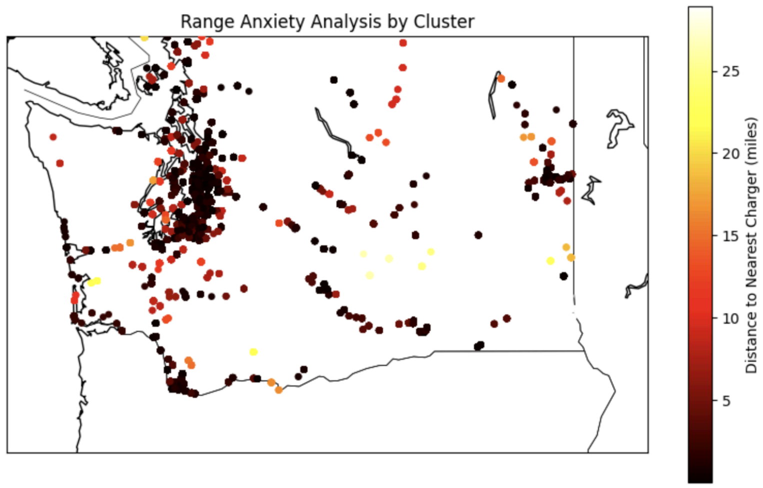




II. Range Anxiety

After creating the clusters of stations and cars, we performed an analysis on range anxiety. Range anxiety analysis is crucial for understanding the limitations and challenges of EVs in terms of their driving range and access to charging infrastructure. Range anxiety directly affects consumer confidence in EVs. By analyzing this factor, we can better understand the barriers to EV adoption and develop strategies to overcome them as it helps identify areas where charging infrastructure is lacking or where it needs to be expanded. This information is vital for planning the placement of new charging stations to maximize coverage and convenience for current and future EV owners.

In order to study range anxiety, we wanted to calculate the distance between the centroid of each car cluster and the closest charging station. We calculated this using Manhattan distance. We used Manhattan distance because it is a good approximation of the driving distance between two points, especially at small scales. We found that each car cluster is an average of 1.18 miles from the closest charging station. We next visualized the distribution of cars in the dataset and the distances to the nearest charging stations, which provided insight into areas where range anxiety might be a concern.

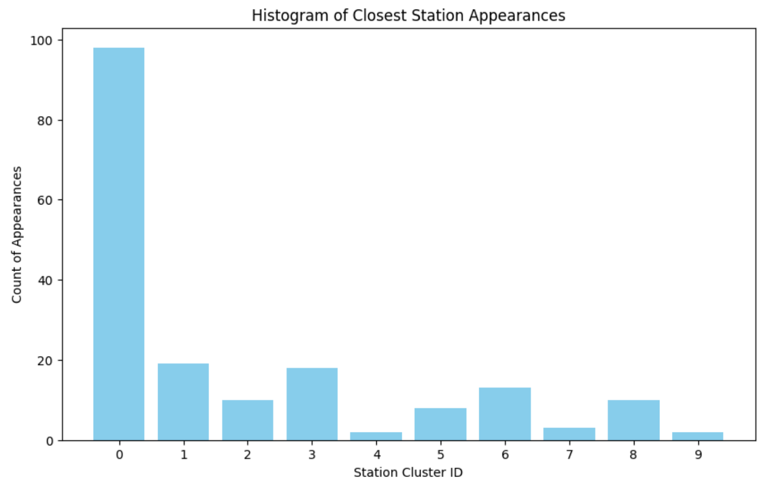


To further figure out where range anxiety may be a problem, we calculated which clusters have a distance to the nearest station that is significantly higher than average, using t-tests. We then collected the postal codes of the cars in those clusters (shown in code). This list of postal codes can more precisely inform policymakers about where to construct new charging stations.

Using these significant postal codes, we determined their corresponding car cluster ID numbers to better understand which regions of Washington on the interactive map these postal codes corresponded to. For each significant car cluster, we examined its closest station and created a histogram displaying the distribution and frequency of the closest charging station cluster for each of the statistically significant EV clusters.

From examining this histogram, the list of postal codes that are significantly further than average from their charging station clusters, and the interactive map, we recommend that Washington policymakers target charging station developments within these significant postal code regions, but also particularly in the southeast region (station clusters 4 & 7) and northwest corner (station cluster 9) of Washington State. Simply because these regions are less populous than regions serviced by charging station cluster 0, near the greater-Seattle area, does not mean EVs in less dense regions of Washington must also travel longer than average distances to reach available public charging stations.

Policy makers can prioritize these areas in the future to build new EV charging stations to promote sustainability across the State of Washington. While this is a recommendation set in a greater context, we believe this high-level approach of evaluating which EV clusters have larger travel distances to charging stations which deviate from the average is important knowledge for policy makers, as it serves as the foundational steps of making charging station accessibility increasingly equitable; from this point of departure, policy makers can begin further investigation and allocation of management and resources at the State-level to execute these impactful changes for citizens in less populated or underrepresented areas of Washington.



III. Station Utilization

We conducted a station utilization analysis to assess the efficiency of charging station clusters. We made several assumptions, including cars visiting stations in their assigned station cluster, charging only during a 16-hour window, and getting 100 miles of charge every 4 days (driving an average of 25 miles a day). Using these assumptions, we calculated the number of cars each station cluster could handle in a day based on the types of chargers available. From research, we found that a Level 2 charger can charge 3 cars to 100 miles in a 12-hour period, and a DC fast charger can charge 24 cars to 100 miles in a 12-hour period[8]. We then determined the daily charging needs of each car cluster and the number of cars that would visit each station cluster per day. Dividing the latter by the former gave us the utilization percentage for each station cluster. The results indicate most clusters are overutilized, with all but 2 clusters having over 90% utilization. Clusters 3, 8, and 9 exceed 300% utilization, suggesting potential issues with overloading. Cluster 5 has 70.5% utilization, and cluster 4 has 48.4% utilization. This analysis can help in optimizing charging station placement and capacity to meet the demand more efficiently. We suggest that the stations in clusters 3, 8, and 9 should increase the number of chargers they have. The stations in clusters 4 and 5 have a sufficient number of chargers.

IV. Optimization Model

Due to having such a large dataset, in order to still give specific recommendations for optimizing EV charging stations, the optimization model we ran was just on subsets of zip codes that define larger cities in Washington. We assigned a "Proximity Utilization" to each Available Station in this area, which gives a metric for how efficiently each charging station is being utilized based on the demand from EVs within a threshold distance (we used 40 km which is 25 miles, the average miles driven a day from our assumptions in the previous section) and the station's capacity to serve them. After estimating demand from assumptions listed in the previous section and using the number of ports and how long they could be used every day to estimate a station's capacity, utilization was calculated as the ratio of total daily demand by all nearby EVs to the total daily capacity of the station. It essentially measures what percentage of the station's capacity is expected to be used by the EVs in the immediate area if each EV follows the estimated demand pattern. We did also include a pricing penalty in the model, which applied a reduction factor to the utilization if the station charges for the charging service. This could be seen as an adjustment to reflect potentially lower usage when users are required to pay, as opposed to free charging stations.

For each of the 3 cities we observed, our recommendations are based on two outputted visualizations - a utility map and a heatmap for demand. Stations that are deep red in the heatmap and yellow in the utility map mean they experience high demand, but that they have the adequate capacity and infrastructure to service this demand.

On the other hand, The purple dots in the top plot that are also in that deep red zone are severely lacking the charging port access needed to service the high population of customers in the area, and are identified as “underutilized”. Therefore, we are recommending that those stations who are already familiar to customers install more Level 2 or DCFC charging ports to account for the high demand in their area.

Given each city we look at is only a subset of zip codes from the dataset, and there very possibly could have been different utilization scores and demand assigned to each station by expanding that range and providing for more data for the stations and cars on the outskirts of the city, we will not get too into the color analysis on the outer points given those points have incomplete information.

i. Seattle Metropolitan Area

See Appendix 1.1 for Seattle’s EV Charging Stations Utilization Map and Appendix 1.2 for Seattle’s Heatmap of EV Accessibility at Charging Stations.

While there are 83 stations we would recommend more ports at, the most pressing stations would be: ['505 First Garage', 'Northgate Mall - Tesla Supercharger', 'Lighthouse Apartments', 'Broadstone Strata', 'Tilt49 Realty', 'Union Square - Tesla Supercharger', 'Seattle, WA - Union Street (Wall Connector) - Tesla Destination', '531611 : Amazon Phase VI - Bigness', 'Troy Block', '2+U', 'The Waverly', 'Stadium Place', '531613 - 1000-1100 Dexter Avenue N', '325 Eastlake', '1700 Seventh', 'Slalom Hawk Tower', 'Komo Plaza Garage', '431603 : AVA Capitol Hill', 'Ballard Blocks - Tesla Supercharger', 'Pike Place Market', 'Hill 7 Station', 'Columbia Center Tower', '531620 : Centre 425/Everest', '531604 : 600 Broadway', '531603: 701 Pike Street (Seattle), LLC', 'Residence Inn by Marriott - Tesla Destination', '531616 - Marion & Minor Garage'].

ii. Spokane

We recommend more level 2 or 3 ports at the following locations: ['Liberty Lake, WA - Tesla Supercharger', 'TA Petro - Tesla Supercharger', 'Walmart 2539 (Spokane Valley, WA)', '803 N Post St Parking', '9000 N Division St', 'AgWest Farm Credit', '201 W North River Dr Garage', 'Parkade Plaza Parking Garage', 'Steam Plant Parking', 'Kroger-Fred Meyer 657 (Spokane, WA)', 'Washington State Department of Ecology - Public Parking Lot', 'Goodwill Industries Outlet Store', 'Spokane Public Library - Central', 'SpringHill Suites by Marriott - Tesla Destination', 'SFCC', 'Oxford Suites Spokane Valley - Tesla Destination', 'Hilton Garden Inn Spokane Airport - Tesla Destination']

iii. Tacoma

Note - we adjusted the threshold to 15 km for Tacoma since it is a much smaller city and to show more distinction in the heat map.

We recommend more level 2 or 3 ports at the following locations: ['Tacoma Power Lot D EV Charging', 'Tacoma Mall - Tesla Supercharger', 'Evergreen Shopping Center - Tesla Supercharger', 'The Lex', 'City of Tacoma - Solid Waste Management', '531608 : UW Tacoma C17 Garage', '923 Commerce Street', 'Brewery Lofts EVC.', 'Convention Center', '531609 : Pierce County Annex', 'Brewery Block TAC', 'Tacoma Mall', 'City of Tacoma - Center for Urban Waters', '531617 : Court C Parking Level - Charging Stations', 'Central WW Treatment Plant', 'Yakima Garage', 'Safeway Tacoma', 'Walmart 4137 - Tacoma, WA']

Conclusion

In our analysis of electric vehicle data in Washington state, we wanted to highlight the areas that are in need of further infrastructure. Improving this will encourage more people to switch to electric vehicles and enhance equity in access to electric vehicles for different demographics. Our recommendations are tailored towards policymakers in Washington and at the Department of Energy.

Due to the vast size of this data, breaking the EVs and charging stations into clusters allows policymakers to better make actionable steps to increasing availability and equitable access of maintenance of EVs. Creating these cluster assignments allowed us to break down the various regions of Washington State to perform our analysis. We were able to execute this effectively using innovative algorithms such as DBSCAN and HDBSCAN that were powerful enough to harness this large amount of data.

Our analysis of range anxiety highlights an easy way to facilitate wider electric vehicle (EV) adoption. By identifying areas with limited access to charging infrastructure, policymakers can strategically plan for new charging stations, improving equity in EV access and promoting environmental sustainability. Our findings suggest targeting station developments in specific regions, notably the southeast and northwest corners of Washington State, to enhance charging convenience for EV owners in less populous areas. These recommendations lay the groundwork for policymakers to take actionable steps towards a more sustainable transportation future in Washington.

Our analysis of station utilization revealed significant differences in charging station utilization across station clusters. Clusters 3, 8, and 9 are experiencing high demand, exceeding 300% utilization, while clusters 4 and 5 are underutilized at 48.4% and 70.5%, respectively. To address these disparities, we recommend expanding charging infrastructure in overutilized clusters and exploring smart charging solutions to optimize utilization and promote equitable access to EV charging by placing charging stations in underserved communities.

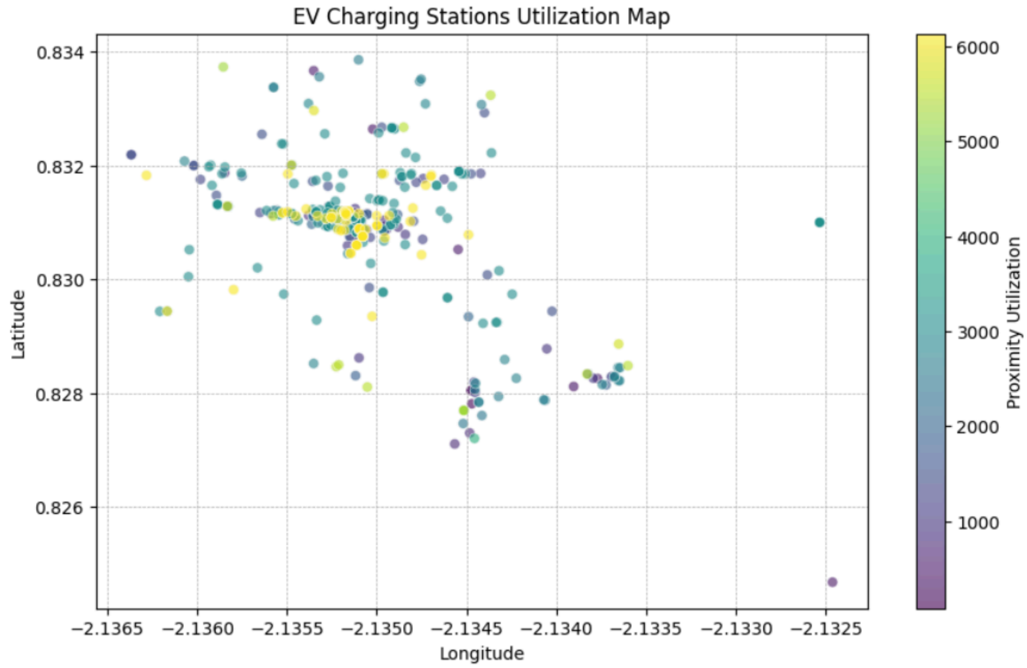
In our optimization model, we were able to identify EV stations that are already open in Washington State and are being underutilized in areas of high demand due to a lack of infrastructure. In the three largest cities, Seattle, Spokane, and Tacoma, we identified a list of these stations where more Level 2 or Level 3 chargers need to be installed in order to account for high density of EVs in those areas. A way we would've liked to take this analysis further is, now that we've found stations that customers would benefit from more charging ports at, to look at these stations from a business owner point of view and consider the trade off between the increased demand able to be serviced via an additional port vs. the cost of actually installing this new port. For the time being, this is not entirely feasible, though, as most of our stations dataset either had missing values for the cost of using their chargers, or were free due to the way that government is still providing alternate ways for EV charging stations to be beneficial to a business owner through incentives like tax benefits or direct funding for public charging stations.

Additionally, based off of how the stations dataset was organized, we wanted to predict if opening certain temporarily unavailable stations would be beneficial to those businesses by seeing if they improved the aggregate utilization of a subset of stations through our optimization model.

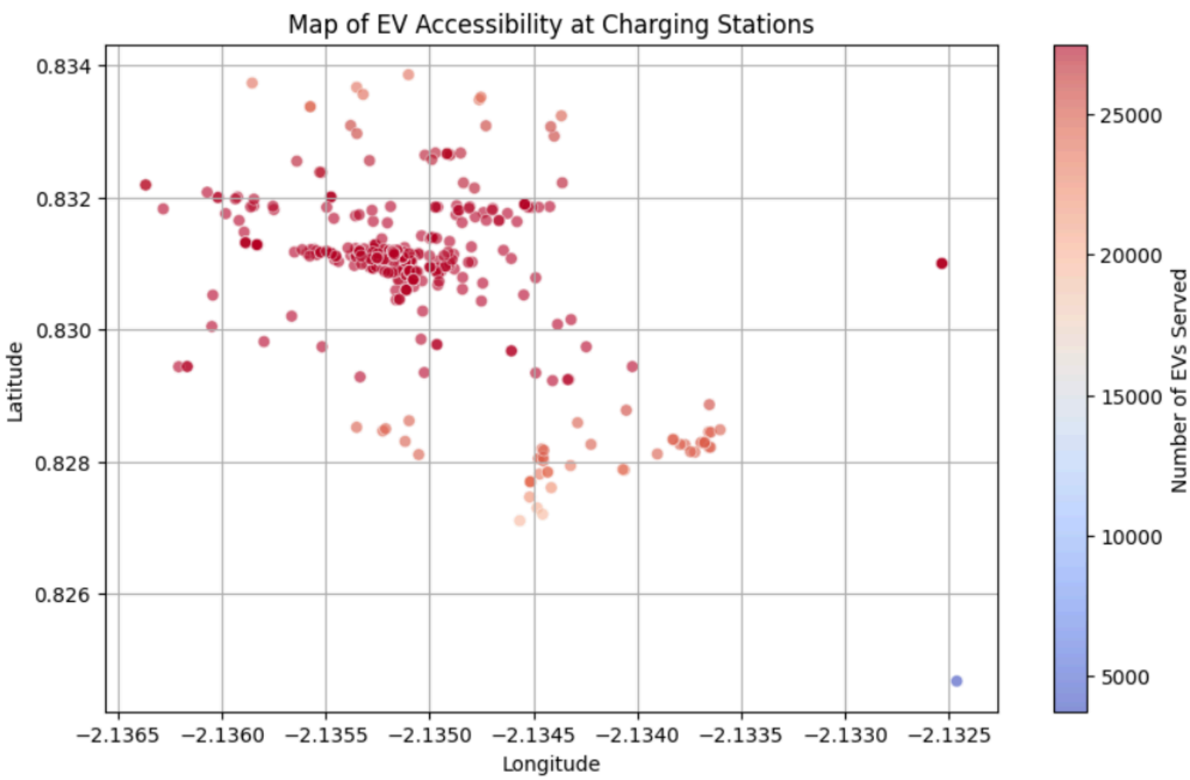
Overall, if we could be able to continue investigating this topic, we would be interested in exploring demographic information concerning the population of each region of Washington State: For example, average income per cluster, exact population size, levels of education, etc. These metrics would better allow us to make a more holistic evaluation and recommendation of optimal placement of EV charging stations across the state while ensuring that no region is underrepresented due to their demographic factors.

Appendix

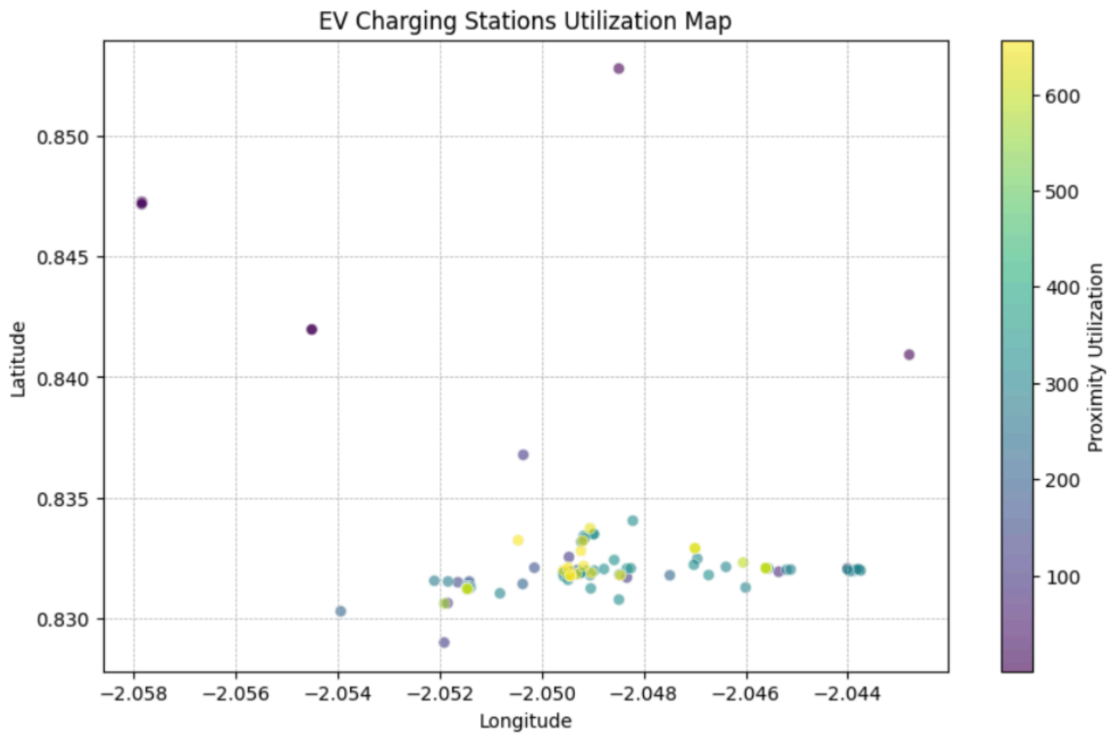
Appendix 1.1 Seattle Utilization Map



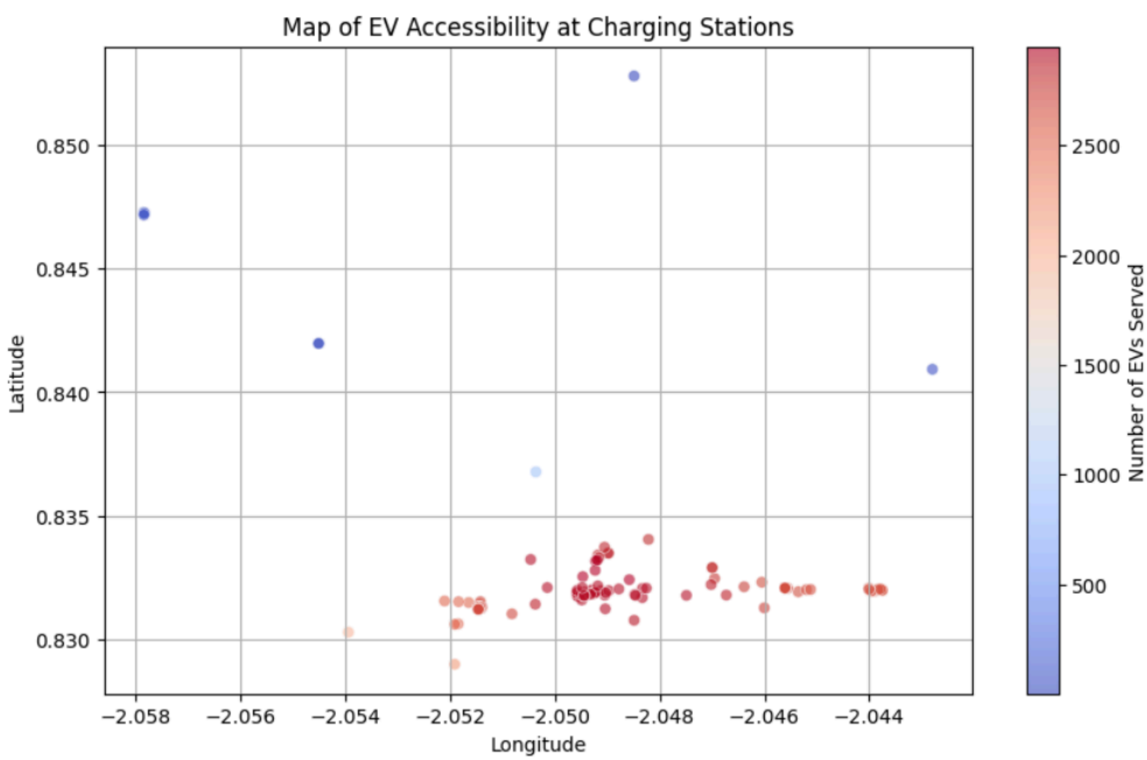
Appendix 1.2 Seattle Customer Demand Heatmap



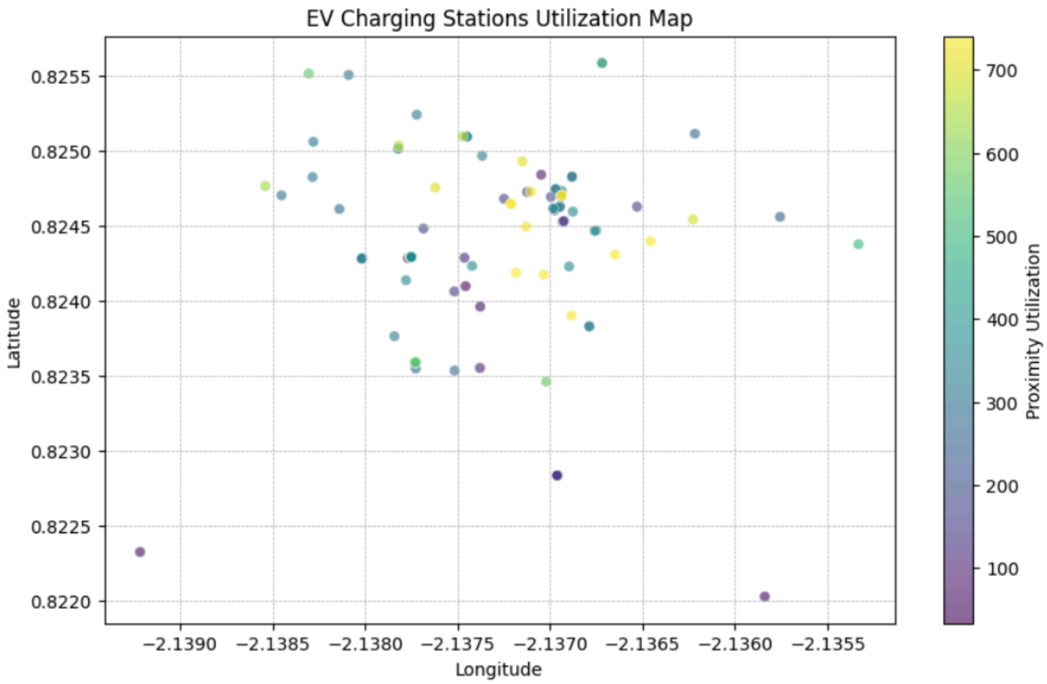
Appendix 2.1 Spokane Utilization Map



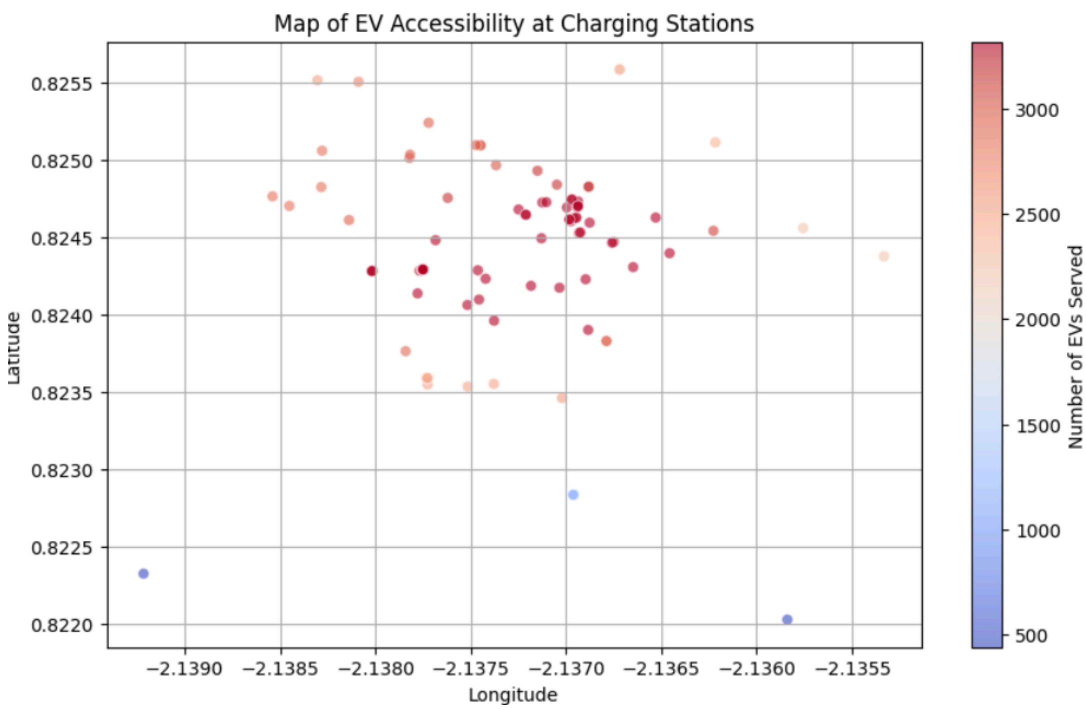
Appendix 2.2 Spokane Customer Demand Heatmap



Appendix 3.1 Tacoma Utilization Map



Appendix 3.2 Tacoma Customer Demand Heatmap



References

[1] Valdes-Dapena P. No, The electric vehicle market isn't crashing. It just needs recharging. | CNN Business. CNN. 2024. Available from: https://www.cnn.com/2024/05/01/cars/electric-vehicle-market-sales-tesla-ford/index.html

[2] Valdes-Dapena P. Why EV charging is still such a pain | CNN Business. CNN. 2023. Available from: https://www.cnn.com/2023/11/04/business/why-ev-charging-is-still-such-a-pain/index.html

[3] Alternative Fuels Data Center: Charging Electric Vehicles at Home. afdc.energy.gov. U.S. Department of Energy. Available from: https://afdc.energy.gov/fuels/electricity-charging-home#:~:text=Most%20drivers%20of%20electric%20vehicles

[4] A Record Year for Electric Vehicles and Plug-In Hybrids in Washington. Washington State Department of Ecology. Available from: https://ecology.wa.gov/blog/april-2024/a-record-year-for-electric-vehicles-and-plug-in-hybrids-in-washington#:~:text=Washington%20also%20saw%20a%20huge,increase%20in%20a%20single%20year.

[5] N Nigro. Investment In Publicly Accessible EV Charging In The United States. Atlas Public Policy; 2023. Available from: https://atlaspolicy.com/wp-content/uploads/2023/05/Investment-in-Publicly-Accessible-EV-Charging.pdf#:~:text=In%20response%2C%20significant%20funding%20from%20governments%2C%20electric%20utilities%2C,years%2C%20reaching%20nearly%20%2420%20billion%20by%20early%202023.

[6] Electric Vehicle Population Dataset. www.kaggle.com. Available from: https://www.kaggle.com/datasets/yashusinghal/electric-vehicle-population-dataset

[7] Electric Vehicle Charging Stations. Joint Office of Energy and Transportation. Available from: https://driveelectric.gov/stations

[8] Electric Vehicle Charging Speeds. US Department of Transportation. www.transportation.gov. Available from: https://www.transportation.gov/rural/ev/toolkit/ev-basics/charging-speeds#:~:text=Level%202,PHEV%20in%201%2D2%20hours.