**IBM Capstone Project**

Optimal Location for EV Service in the Washington State.

**Introduction**

Due to the growing population of electric and hybrid vehicles in the US, the new market emerging in the field of auto service. Followed by the Tesla Motors, world top carmakers offering more electric vehicles in their model line which constantly takes a significant portion of the car market in the US.

The growing number of electric vehicles collaterally boosts the EV charging market, as more consumers decide to purchase electric cars as higher demand for charging stations. On the other hand, the wider access to charging stations at specific locations stimulates consumers to choose an electric vehicle.

It is obvious that any vehicle with electric drivetrain requires maintenance and replacement services distinct from traditional gasoline vehicle, thus, a significant number of EV on the roads creates a market for “non-traditional” auto service.

Based on market opportunity analysis, potential auto service business planning to step into the EV market and provide maintenance and replacement services in the state of Washington, however choosing the most efficient deployment locations for EV service facilities that would maximize profits requires location-based data analysis and the estimated market growth. Following data analysis will help to visualize the potential EV service market as well as clustering potential customers into groups and define the most efficient deployment locations that would satisfy the EV owner’s needs.

**Data**

Following data required to perform the analysis:

1. List of zip codes that covers the state of Washington (WA) including geolocation and county.
2. The geographical location of currently registered EV vehicles in WA will be obtained directly from (data.wa.gov), official state government site.
3. The geographical location of EV charging stations in WA will be obtained from the National Renewable Energy Laboratory (NREL) through an API.

**Methodology**

The process will take 3 stages:

1. Cleaning and data preparation.
2. Transformation of the EV data and merging with EV charging station data, population, and population density data.
3. Applying K-clustering.
4. Visualizing data with folium.

After data was cleaned and prepared for analysis, we created a list of 586 zip codes located in the state of Washington. The data frame accompanied by associated city, total population, population density, and geo-coordinates (latitude and longitude).



In the next step, we are visualizing the data by applying coordinates on the map of the Washington state.

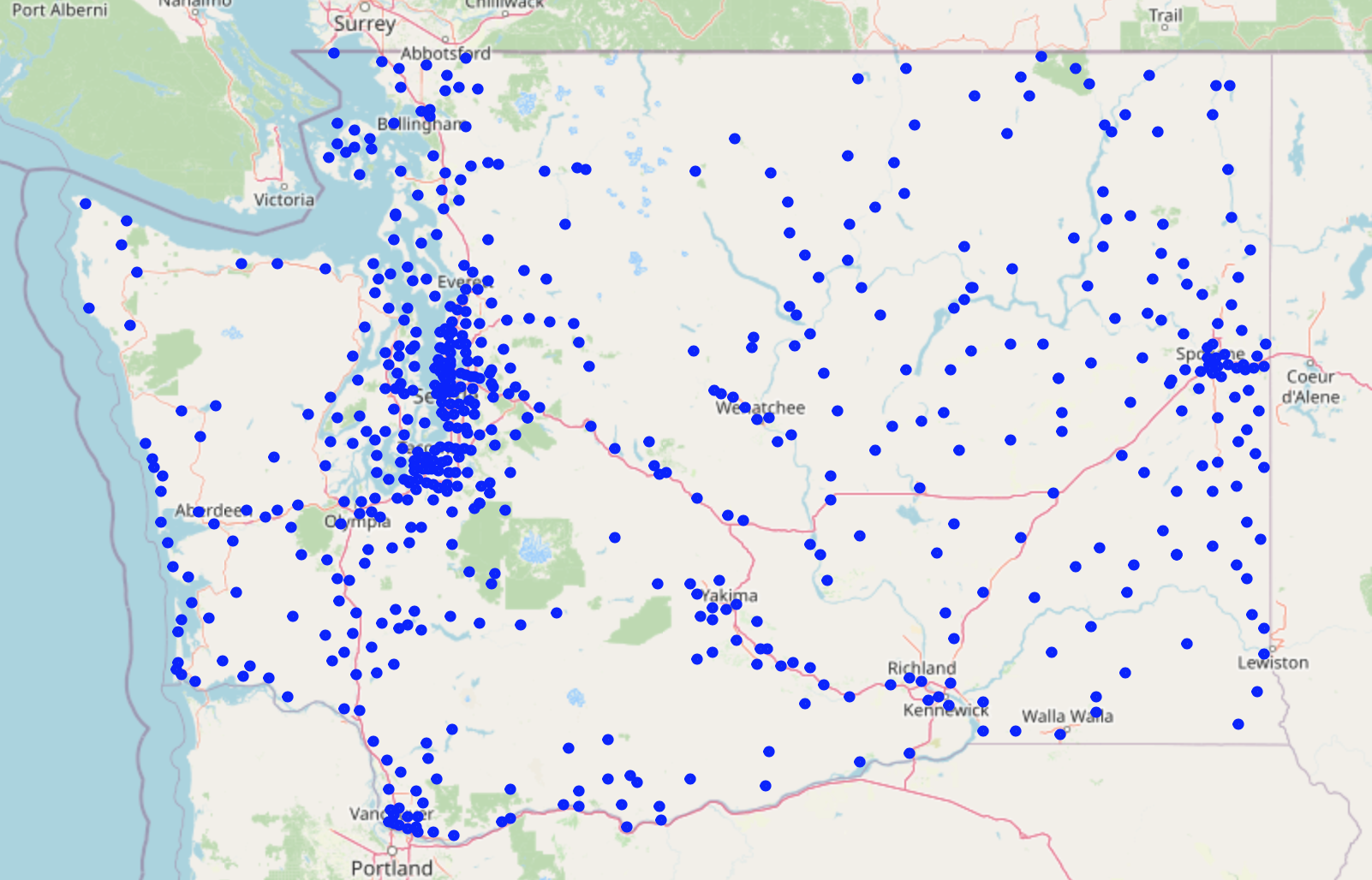


Figure 2. Visualized centers of WA zip codes

At this stage, we use API to including the data from the US Department of Energy that contain location data for each EV charging station in the US and Canada. After cleaning the data we have located every publicly available EV charging station in WA with level 2 and fast charging. In general, the API provides data for all EV charging levels (1, 2, Fast Charging) however, we filtered data for only level 2 and Fast Charging.

Figure 3. EV charging station location data 

In addition to zip code coordinates, demographical data and charging station data, we are including EV populational data in the WA. The data is available as a dataset of 39491 rows where each row represents one specific EV. The data was cleared and grouped by each Zip code with the number of EV and EVC along with population and population density within each zip code.

After dropping zip codes with 10 or fewer electric vehicles from dataframe, we have 285 zip codes.

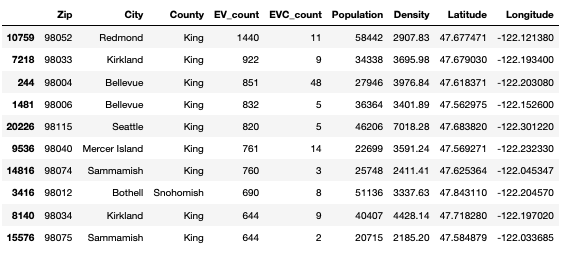


Figure 2. EV, EV charging grouped by zip code

**Cluster Analysis**

To identify groups (clusters) with similar parameters we applying the K-means clustering algorithm. To prepare data for clustering we replaced original parameters with a new parameter, "Number of EV per 1000 population", "Association between EVCS and Population Density", "Number of EVCS per 100 EV".

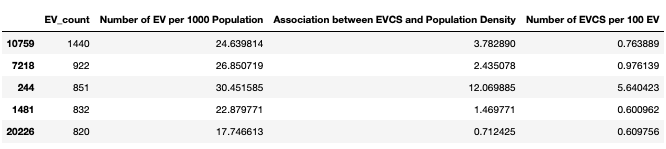


Figure 3. Modified dataframe

Consequently, the elbow method is used to identify the optimal number of clusters.

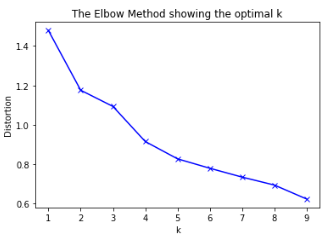


Figure 4. Clusters

The graph shows that the optimal number of clusters is 4.

The following map created based on results of K-clustering and illustrates different clusters by the color of the marker and the size of the marker is proportional to a quantity of EV in each zip code.

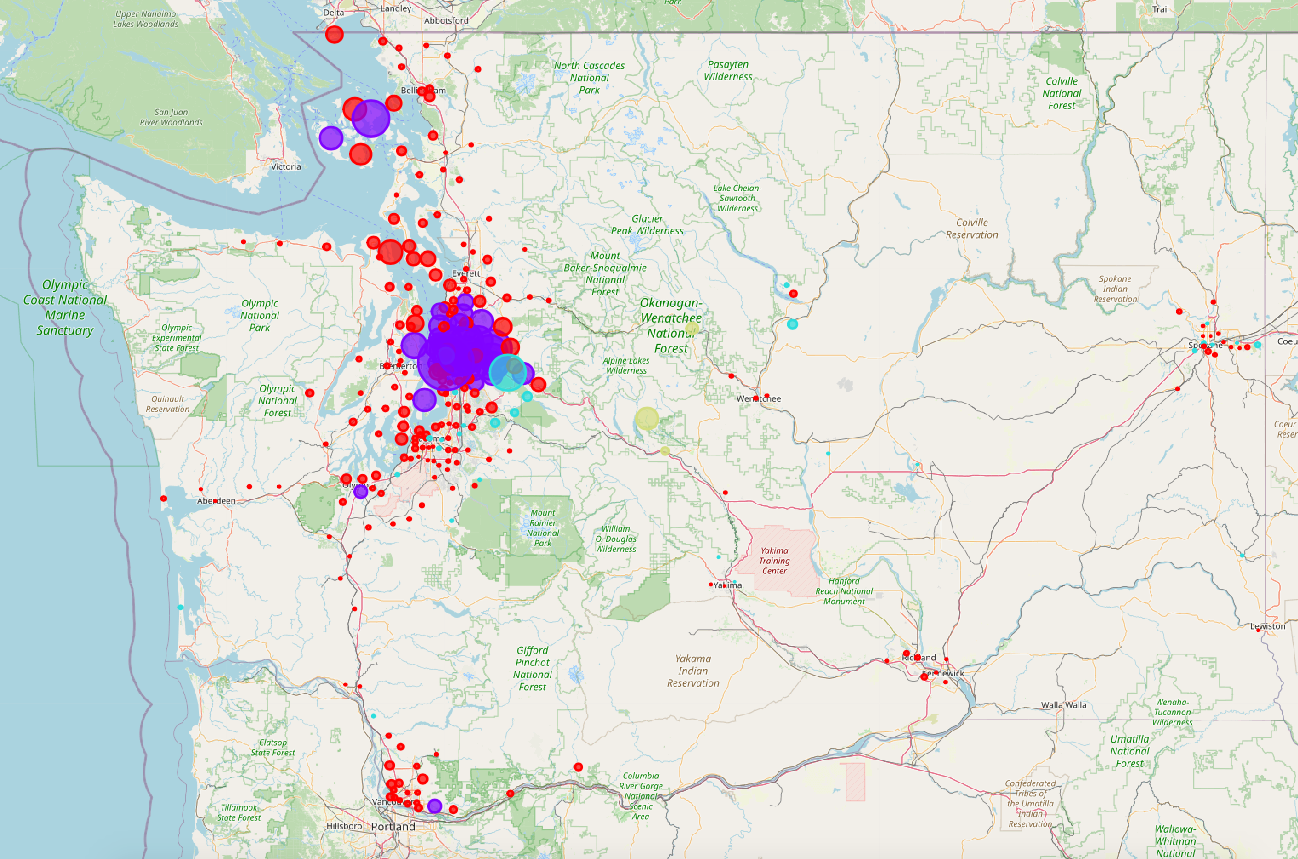


Figure 5. Washington state map with clustered zip codes.

On the scatter plot we can observe dark red circles representing the centroid of each cluster. The dark blue dots are the cluster 1, concentrating toward the upper right corner of the graph while, yellow dots are the cluster 4 which represents the majority of zip codes but have a smaller count of EV and EV per 1000 population.

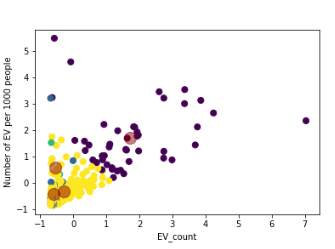


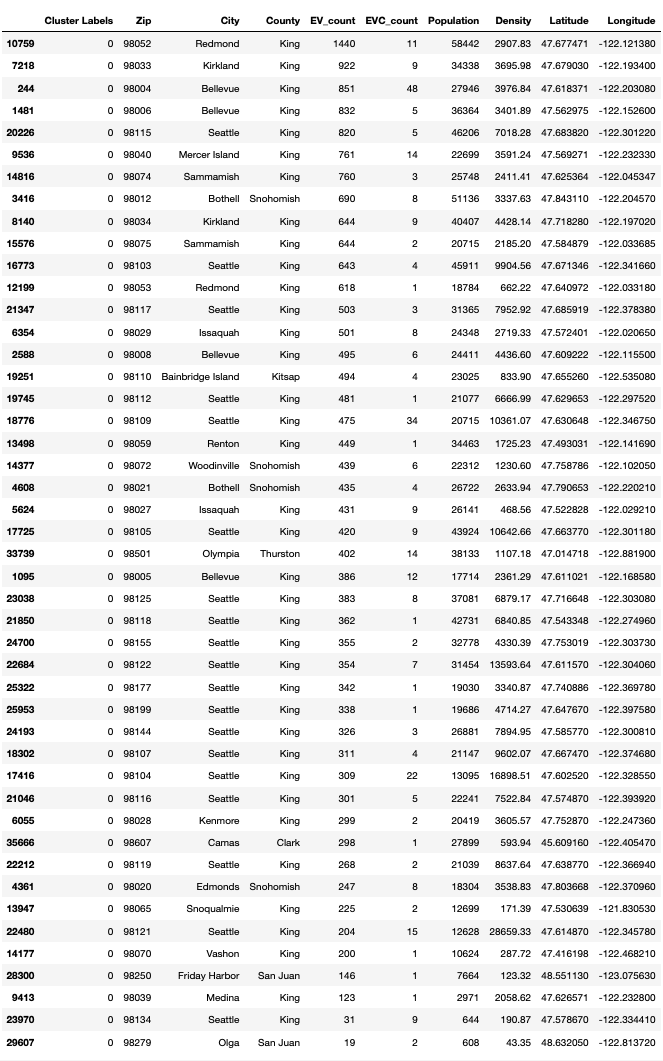
Figure 5. Washington state map with clustered zip codes.

**Results**

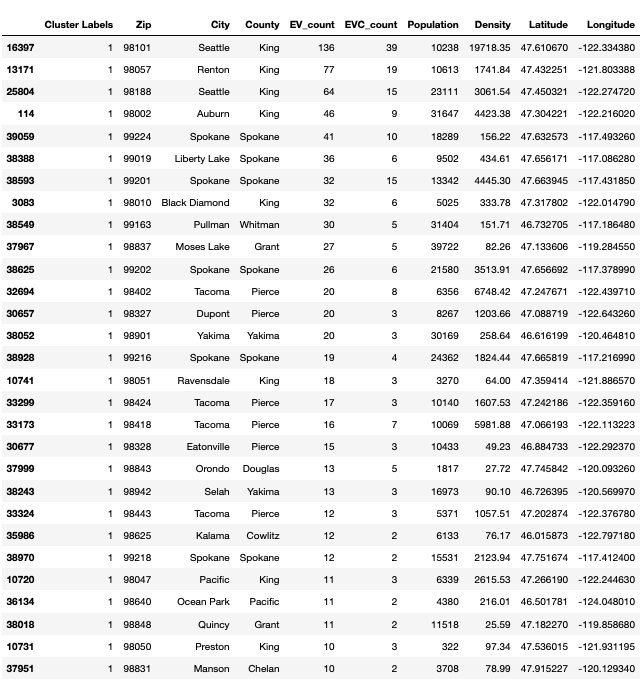
Analysis shows that quantity and location of electric vehicles in the Washington state can be defined within 4 clusters. This clusters represents the concentration of EVs with regard to availability of charging stations, population and population density. Based on the results it is clear that cluster 1 is the most efficient option for EV service deployment. This cluster among other three, has the highest density of EVs and population on the relatively small geographic area.

Nevertheless, cluster 4 is the second by EV density but the largest, in terms of geographic size, which makes it less efficient option.

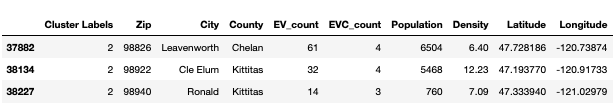
**Cluster 1**

****

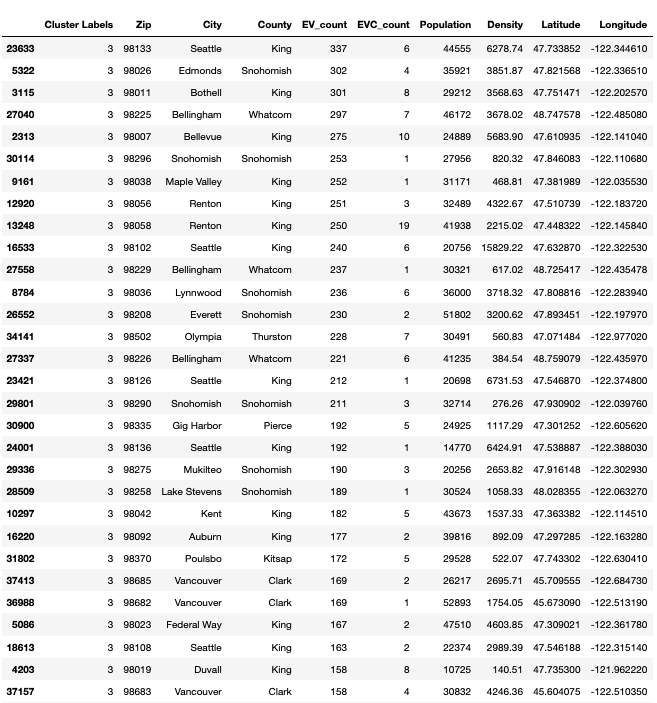
**Cluster 2**



**Cluster 3**



**Cluster 4**



207 rows × 10 columns

**Discussion**

Based on the performed analysis, we are able to identify the estimate location within the largest cluster of electric vehicles in the Washington state where EV service deployment would be the most efficient. However, an additional analysis would help to predict the potential customer growth. For instance, EV data analysis relatively to all other type of vehicles registered in Washington state for the past several years will help to reveal EV market dynamic in specific cluster. Therefore, this insight will help us to take more competent approach in the case of EV service facility deployment.

**Conclusion**

The data analysis was performed to identify the most optimal location for electric vehicle service in the Washington state. During the analysis we explored and analyzed some important statistical figures of each zip code in Washington state where any electrical vehicle registered as of May 2020. In addition, clustering analysis helped us to identify and highlight the groups of optimal zip codes. Eventually, cluster 3 which mostly situated within King County of Washington state chosen as the most optimal option for EV service that need more detailed analysis.