Distribution of EV charging stations in San Francisco Bay Area

Youcan Liang

May 2023

Abstract

This paper examines the distribution of EV charging stations in the San Francisco Bay area. It utilizes spatial analysis and clustering method to assess their availability, equity, and efficiency. The results reveal disparities in charging station distribution across different census tracts and across different driving time frames.

1 Introduction

Transitioning to electric vehicles(EVs) is an important step to mitigate the climate crisis. Low availability of charging station increases range anxiety and hesitates people from buying electric cars. A study in Netherlands shows low availability of EV charging stations in residential areas [1]. This project assesses the current charging stations in terms of availability, efficiency, and equity in the San Francisco Bay Area.

2 Methods

From Alternative Fuels Data Center, I gathered EV charging stations with exact locations in the San Francisco Bay Area. The dataset specifically includes public charging stations offering Level 2 or DC Fast charging options. Level 1 charging stations were excluded from the dataset due to their slower charging speed.

2.1 Three models for the network

I used NetworkX to build a network for these EV charging stations. Each node represents a EV charging station. A link is created between two charging stations if the distance between them is found to be less than 3 miles.

I used three methods to model the network of charging stations: Random graph model, Small world model, and Barabasi-Albert model. Calculating the average clustering coefficient allows me to see if the charging stations are highly interconnected. Comparing the average degree across models can give me an idea of which model better captures the overall connectivity of the network. Average shortest path length shows how easy it is to traverse the network and how efficiently car can get to a station from another station.

2.2 Quantify reachable stations within different driving time frame

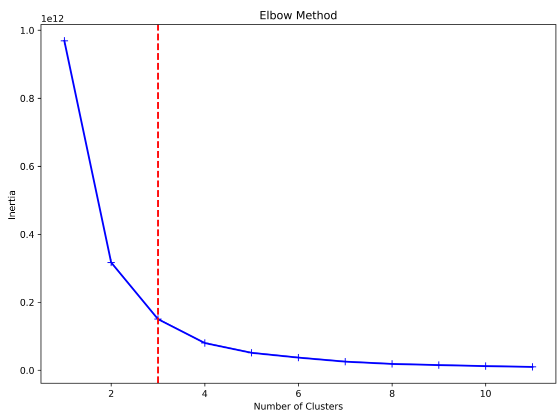
From census data, I extracted population of people who commute to work by driving cars within various timeframes(0-10 mins, 10-14 mins, . . . , 45-59 mins). I calculated the time it takes to travel from the centroid of each tract to each charging station. Driving speed was assumed to be at 40 mph constantly, considering the traffic condition and a combination of local streets and highways. Given the constraint of different time frames, I estimated the number of reachable stations on people's way to work. Note that this is an over estimation because charging stations from all directions of the centroid of a tract were considered.

2.3 Correlations and K-means

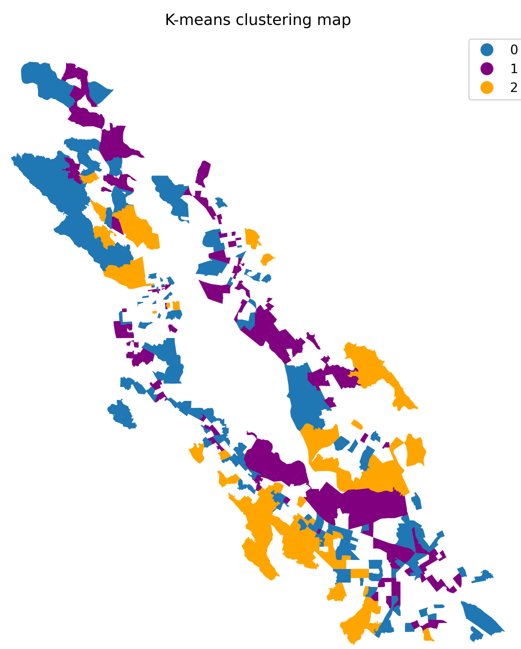
From census data, I selected the following variables for clustering analysis:

- Charging stations per person: Ratio of charging station count divide by population
- Median household income
- Median age
- Pct rented home: Percentage of households that rent for their home
- Pct bachelor: Percentage of people with a Bachelor's degree
- Pct car: Ratio of cars used in commuting divide by population

I did correlation analysis between charing stations per person and other variables. As shown in figure 1, I used the elbow method to find the appropriate number of clusters to be 3.



((a)) Elbow method



((b)) Different clusters on map

Figure 1: 3 clusters for selected variables

3 Results

None of the models well fit the empirical network, except for Barabasi-Albert model for large degree, as shown in figure 2. Table 1 demonstrates similarities between the different network models and the metrics of my network. Specifically, the average clustering coefficient and average shortest path length of my network closely resemble those of the small world model, indicating a comparable level of local clustering and overall network efficiency. Furthermore, the average degree of my network exhibits similarity to that of the Barabasi-Albert model, suggesting a comparable level of connectivity and preferential attachment in terms of node degree distribution.

Table 1: Characteristics about each model

	Avg Clustering Coefficient (C)	**Avg Degree (k)**	**Avg Shortest Path Length (l)**	**Number of Nodes**	**Number of Links**
My Network	0.784886	33.982443	11.744445	1367	46454
Random Graph Model	0.049249	67.964887	1.983302	1367	46264
Small World Model	0.738806	68.000000	10.546120	1367	46478
Barabasi-Albert Model	0.111894	34.000000	2.012644	1367	45322

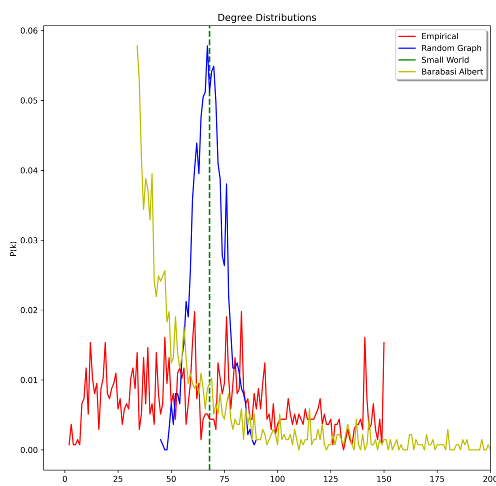


Figure 2: Degree distributions for different models

We can observe the distribution of charging stations across different census tracts in Figure 3. The figure highlights a significant disparity in the number of charging stations among the various tracts. Density of reachable stations for a various commuting time frames are shown in Appendix. From these plots, we can see the dynamical supply of charging stations. By dividing the differences of reachable stations by the differences of time frames, we find that 1 additional minute of driving leads to 4.8% increase of reachable charging stations. Further calculations could be applied to study the demand for charging stations based on people's commute time to work from census data. Then we could compare the supply and demand and figure out the underserving communities.

The correlations between charging stations per person and variables are found to be minimal, as shown in the top row of figure 4. We infer younger generations are more likely to adopt electric vehicles, as indicated by the negative correlations between median age and charging stations per person. In figure 5, the K-means method reveals minimal differences among the three clusters. Therefore, no definitive conclusions can be drawn regarding the social identities that contribute to the disparity of charging stations.

mean	8.5
std	23.9
min	1
25%	1
50%	2
75%	8
max	322

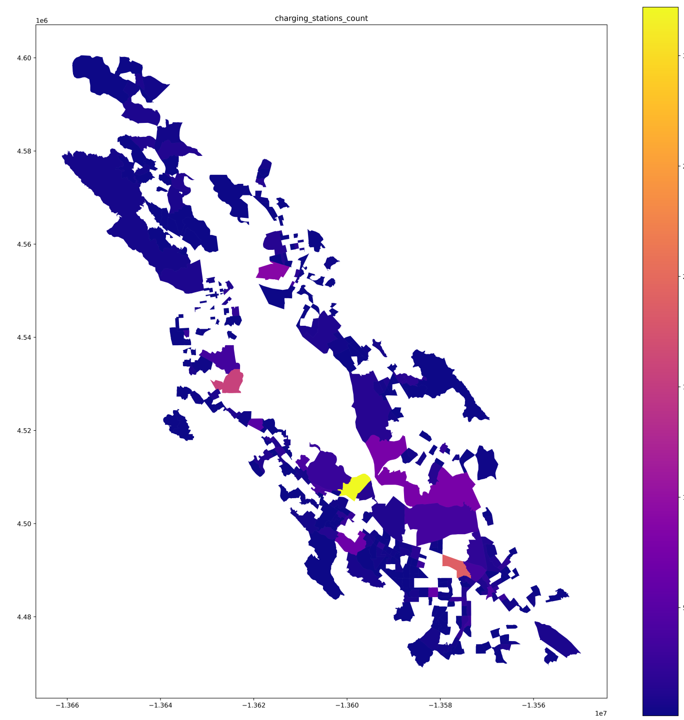


Figure 3: Charging station count by census tract

4 Conclusion

This paper provides insights about the distributions of EV charging stations in terms of availability, efficiency, and equity. Future studies should find more suitable indicators for modeling charging station network. More precise method for estimating reachable stations is needed.

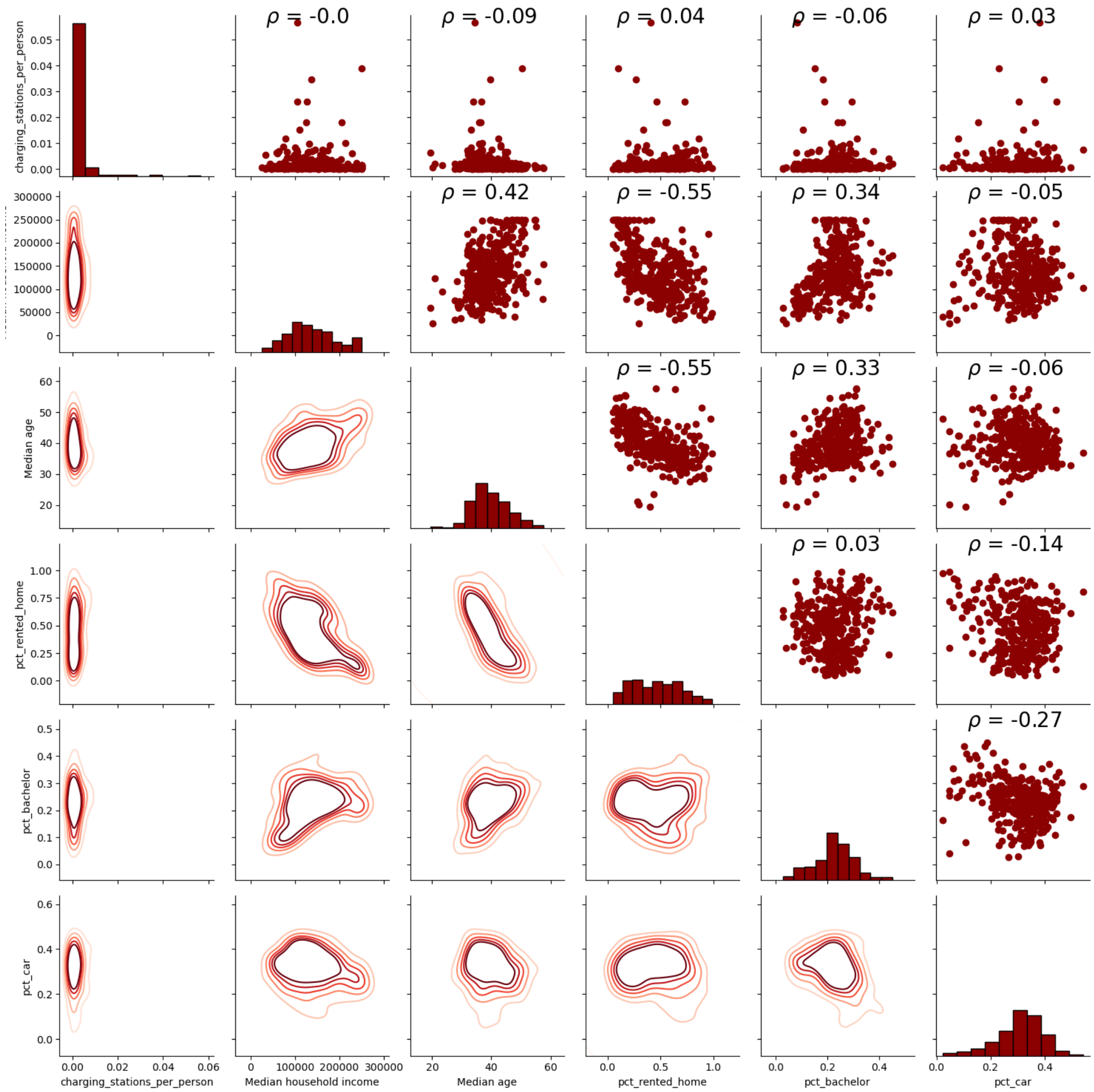


Figure 4: Correlations between charging stations and other variables

Acknowledgments

I would like to acknowledge Professor Marta Gonzalez, TA Martin Liu, and TA Giuseppe Perona for offering valuable suggestions and guidance on this project.

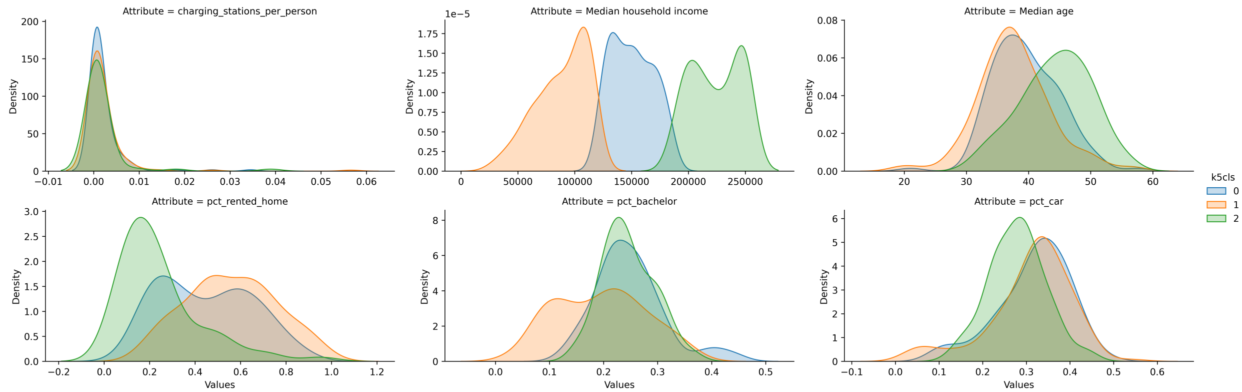
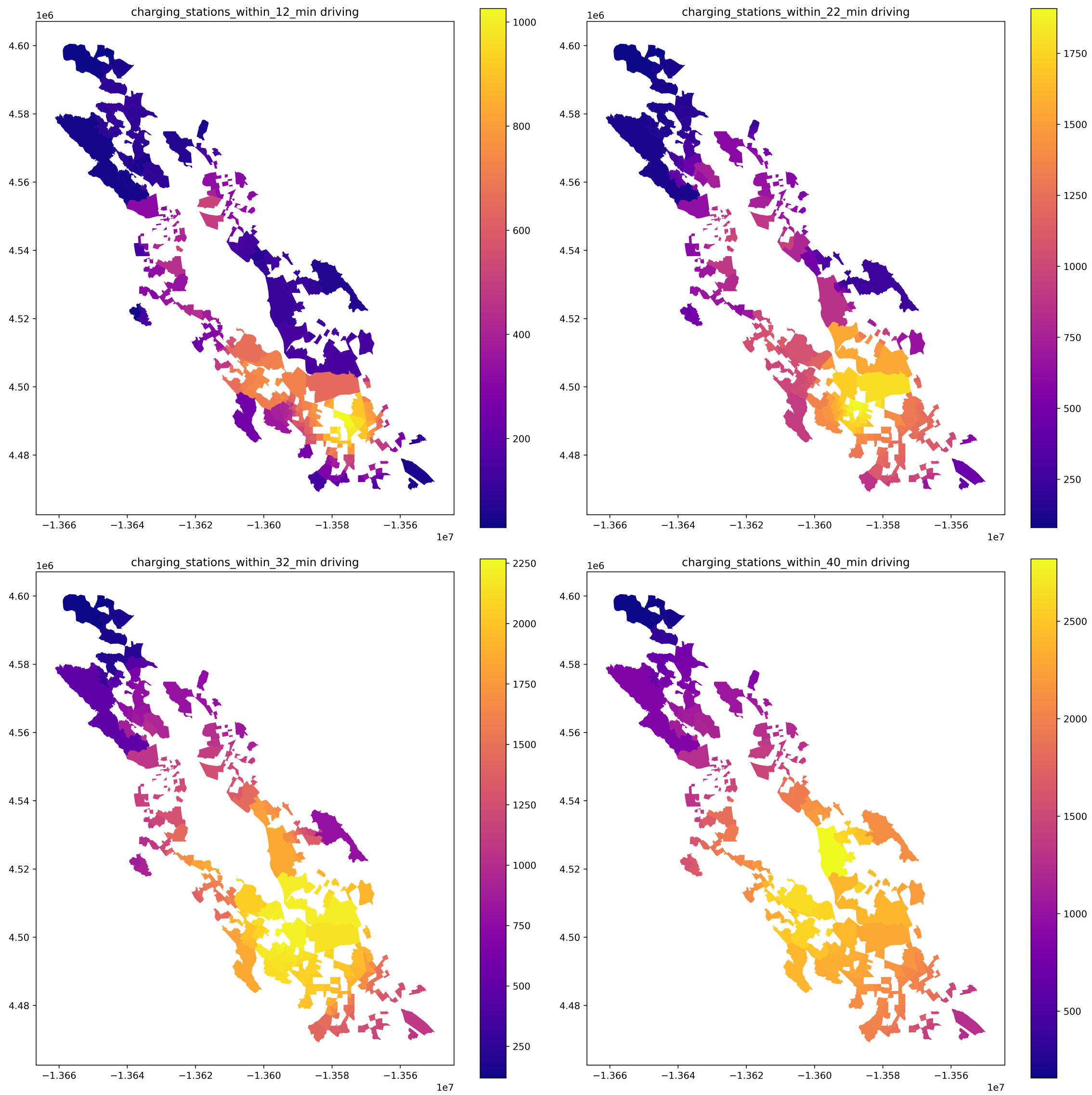


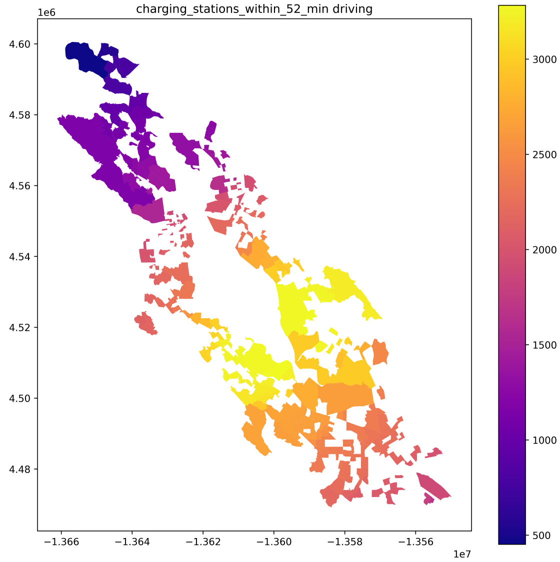
Figure 5: K-means clustering

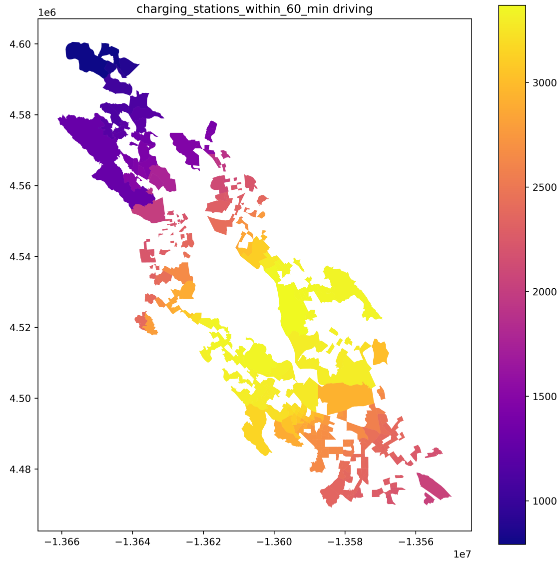
References

(1) Lucas, A.; Prettico, G.; Flammini, M. G.; Kotsakis, E.; Fulli, G.; Masera, M. Indicator-based methodology for assessing EV charging infrastructure using exploratory data analysis. *Energies* **2018**, *11*, 1869.

A Appendix: Reachable charging stations for different driving time frames




charging_stations_within_52_min driving
1e6
4.60
4.58
4.56
4.54
4.52
4.50
4.48
−1.366
−1.364
−1.362
−1.360
−1.358
−1.356
1e7
3000
2500
2000
1500
1000
500


charging_stations_within_60_min driving
1e6
4.60
4.58
4.56
4.54
4.52
4.50
4.48
−1.366
−1.364
−1.362
−1.360
−1.358
−1.356
1e7
3000
2500
2000
1500
1000