SDGE EV Infrastructure Research

Jason Gu
jig036@ucsd.edu

Phi Nguyen
pnguyen@sdge.com

Abstract

As the adoption of electric vehicles (EVs) continues to rise across the United States, the strategic placement of EV charging stations becomes a critical factor in supporting this growth. In collaboration with the Clean Transportation Team at San Diego Gas & Electric (SDG&E), this research explores EV infrastructure within SDG&E service territories. We leverage a combination of datasets, including DMV vehicle registration data, the Alternative Fuels Data Center (AFDC) API, and census data such as the American Community Survey (ACS), to analyze EV ownership patterns, charging station density, and their correlations with demographic and economic factors.

Key analyses include the growth of public EV charging stations, fitting a Poisson Distribution for EV registrations per zip code, and regression analysis to explore the relationship between EV ownership and charger availability. Furthermore, we conduct geospatial analyses to evaluate the relationship between EV charger density and demographic factors such as population size and median household income across ZIP codes. By employing Monte Carlo simulations, we model the expected EV registration per ZIP code.

The results highlight trends such as the accelerated adoption of EVs in recent years and areas where EV charger density correlates with its surrounding infrastructures. These insights inform potential guidance for equitable and efficient expansion of EV infrastructure, contributing to the smooth adoption of a cleaner future.

Code: https://github.com/JingChengGu/DSC180A-SDGE-Q1

1 Introduction

As the incentive to purchase electric vehicles (EVs) surges and the adoption process grows across the United States, the need for an efficient and equitable infrastructure of EV charging stations has become increasingly critical. The integration of EVs into urban and suburban areas presents opportunities for sustainable transportation, but it also poses challenges for energy providers and city planners. San Diego Gas & Electric (SDG&E) serves as a key stakeholder in addressing these challenges within its service areas. I have teamed up with its Clean Transportation Team to understand and optimize EV charger placements to ultimately support the broader goals of clean energy adoption and reaching a cleaner tomorrow.

With a surge in EV ownership, particularly in the years following 2019, disparities have emerged in the accessibility of EV charging stations across neighborhoods. Factors such as population density, median household income, proximity to highways or airports, and much more influence the placement and utilization of these stations. This research aims to address these disparities by combining time-series analysis, geospatial analysis, and predictive modeling to identify patterns in previous EV charger placements and forecast future EV charging station demands across SDG&E service territories.

In this paper, I utilize multiple datasets, including DMV vehicle registration records, the Alternative Fuels Data Center (AFDC) API, and American Community Survey (ACS) census data. By analyzing EV ownership patterns, charging station distributions, and socioeconomic factors, I developed a broad understanding of EV infrastructure across SDG&E service territories. Our EDA methodology includes statistical modeling, Poisson regression, Monte Carlo simulations, and geospatial visualization, allowing me to evaluate the current infrastructure and potentially project future demands.

The outcomes of this research provide insights for SDG&E and other stakeholders to enhance EV infrastructure planning. These findings support efforts to promote equitable access to EV charging stations, particularly in underserved communities, while ensuring the scalability aspect of clean transportation within SDG&E territories.

2 Exploratory Data Analysis Methods

2.1 Data Collection

In order to provide insight into the placement of EV Charging stations within SDG&E territories, I considered a lot of different factors and tested correlations that may have connections with the previous stations that were built around San Diego. Firstly, I gained access to AFDC data that contained existing EV charging station infrastructures. I used the personal key from the AFDC API to gather the dataset. Secondly, I gathered 4 years of DMV data via DMV API, which contained EV registration data per ZIP code from 2019 - 2023. Lastly, I downloaded census data from the United States Census Bureau that contained median income per ZIP Code and population density per ZIP Code.

2.2 AFDC Dataset EDA

The AFDC data I gathered contained all alternative fuel stations across the United States. However, as I was only interested in EV charging stations in SDG&E covered territories, there where a lot of filtering done to clean the AFDC data and have it ready for analysis. One positive thing about AFDC was how extensive its EV charging station data was, and I was able to make good use of this advantage within my EDA. Using Pandas, Matplotlib, Plotly, and OSMnx, I was able to perform basic time series analysis and categorical distribution analysis.

After the data was cleaned and truncated down the columns into ones that were relevant to me, I plotted a time series analysis that shows how much EV charging station count has grown over the years from the early 2000s to today using plotly.

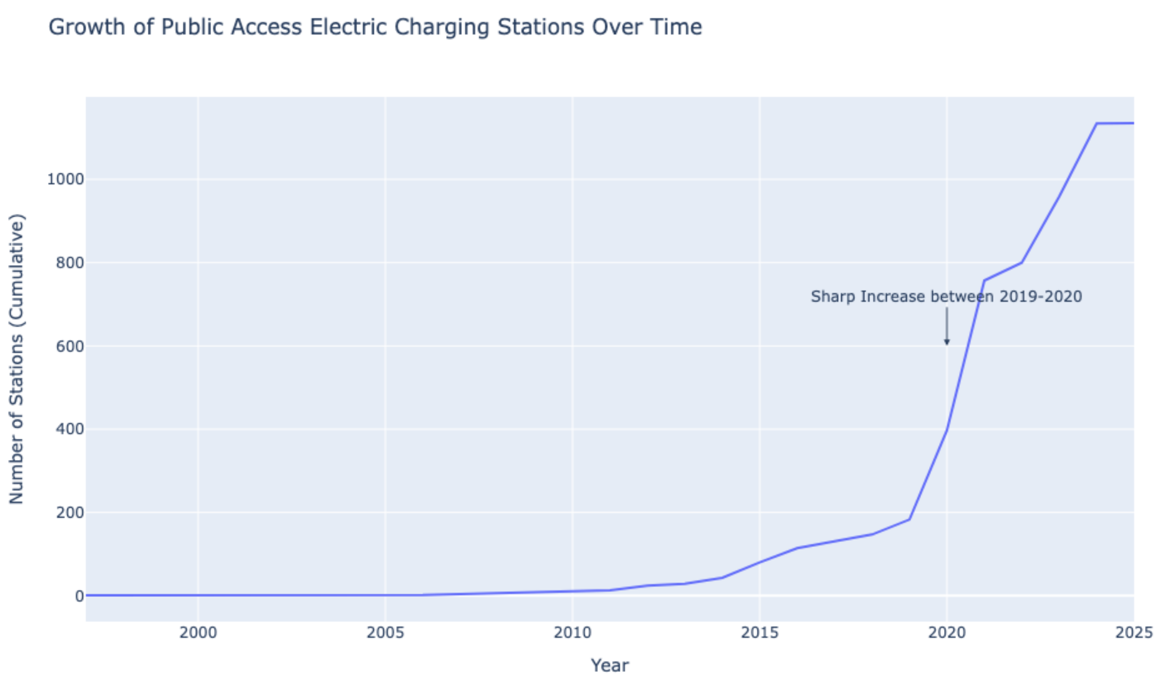

Figure 1: Growth of EV Chargers in SDG&E Territories

I then performed a categorical distribution analysis to explore the presence of different EV charger providers across the region. ChargePoint Network emerged as the dominant provider, with Tesla Destination and Non-Networked chargers following.

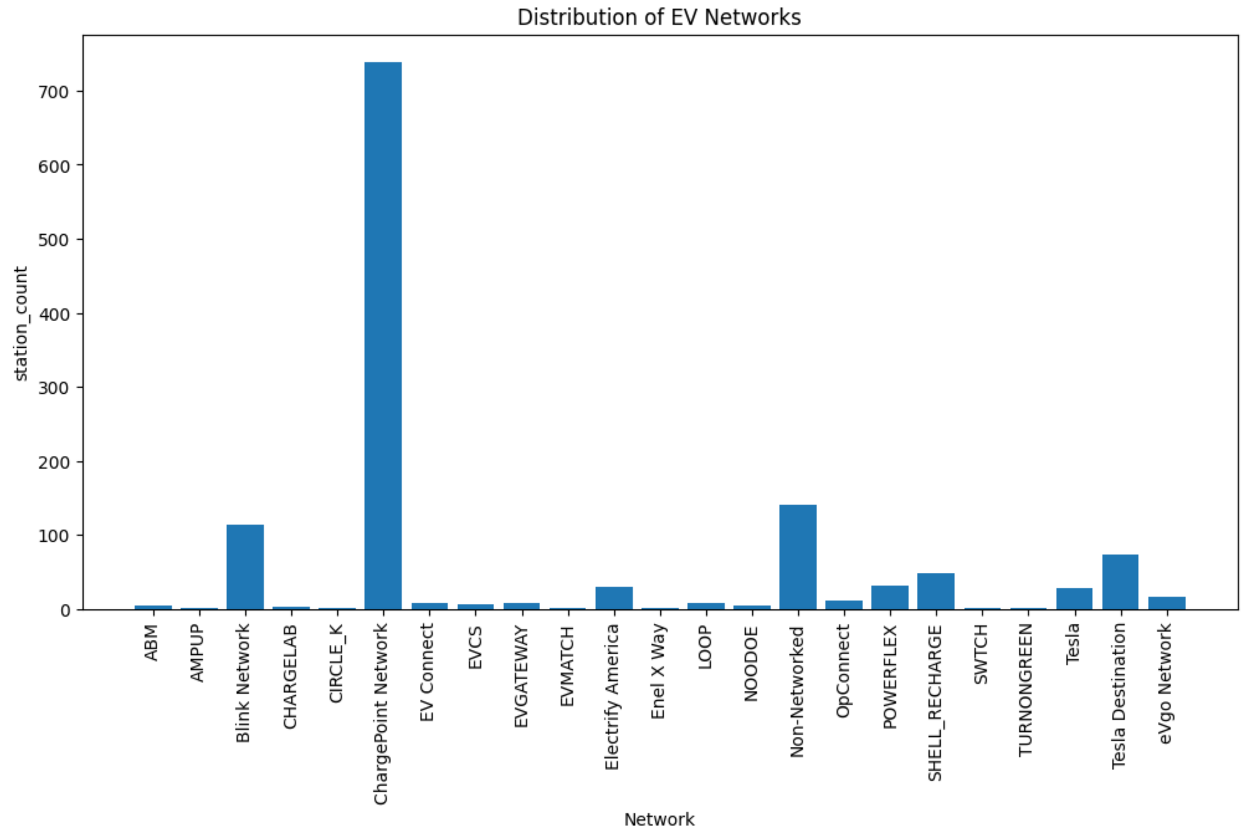

Figure 2: Distribution of EV charger providers in SDG&E territories.

The growth of charger types in SDG&E territories was analyzed over time in three main categories: Level 1, Level 2, and DC Fast chargers. The cumulative count demonstrates a significant rise in Level 2 chargers, with DC Fast chargers also showing steady growth.

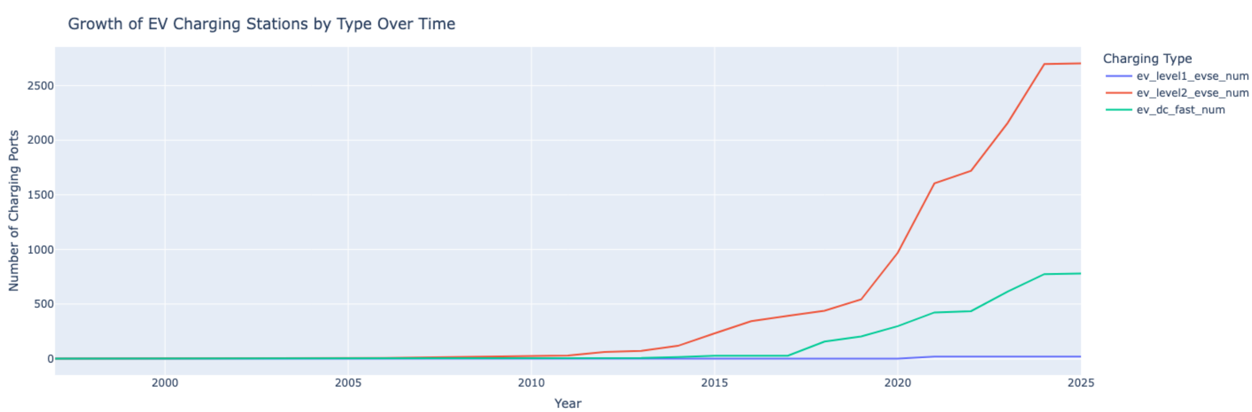

Figure 3: Growth of EV charging stations by type over time.

Using the OpenStreetMap network, I performed a driving distance analysis using OSMnx between a UCSD Hopkins charging station and the SDG&E Headquarters. This analysis highlights the calculated driving distance of approximately 13.8 miles (22.2 km).

Figure 4: Driving distance analysis between UCSD Hopkins charger and SDG&E site.

2.3 DMV Dataset EDA

The DMV dataset that I extracted via DMV API included all vehicles registered in California from 2019 - 2024. The DMV dataset, similar to the AFDC dataset, needed some initial cleaning before it was usable for analysis. After cleaning irrelevant columns and focusing on SDG&E territories, I started my analysis process which consisted of using Pandas, Matplotlib, Plotly, the statsmodels API, and Scipy for fitting Poisson distribution for predicting yearly vehicle registrations.

The following line graph illustrates the growth pattern in EV registrations from 2019 to 2023 in the SDG&E service area. This trend reflects the rapid adoption of EV technology in recent years, with a sharp rise in the number of EVs registered annually.

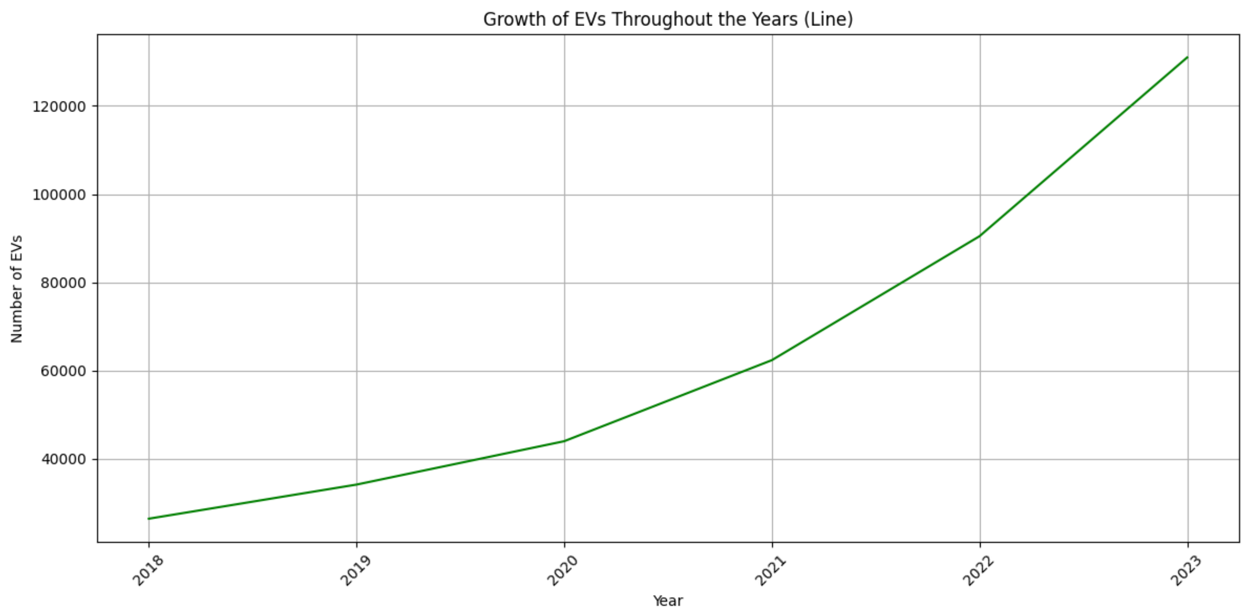

Figure 5: Growth of EV registrations in SDG&E territories over the years

In addition to cumulative growth, the yearly increase in EV registrations was analyzed. The bar graph below highlights the year-over-year growth, showing significant increases in EV adoption, especially between 2022 and 2023.

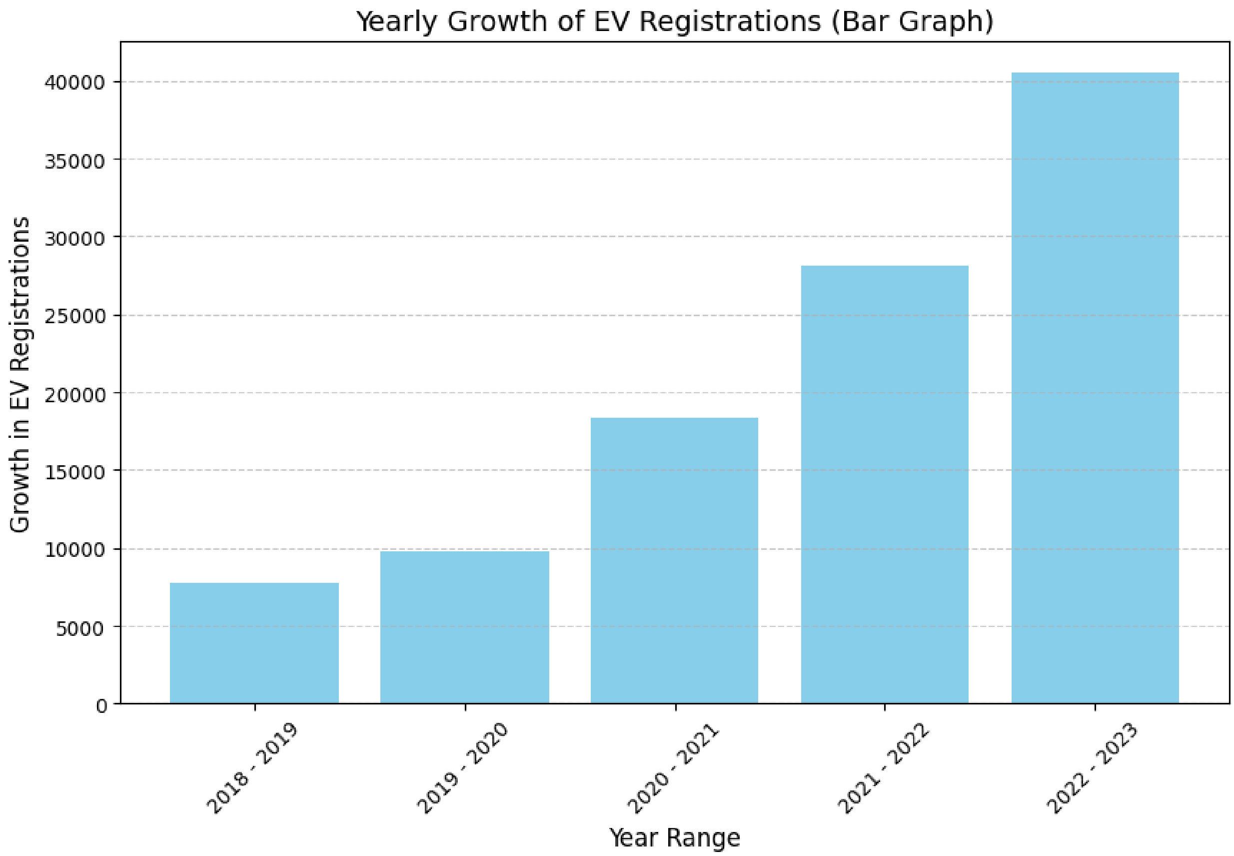

Figure 6: Yearly growth of EV registrations in SDG&E territories

To further understand the distribution of EV ownership, the top 20 ZIP codes with the highest number of EV registrations were identified. The bar graph below showcases these ZIP codes.

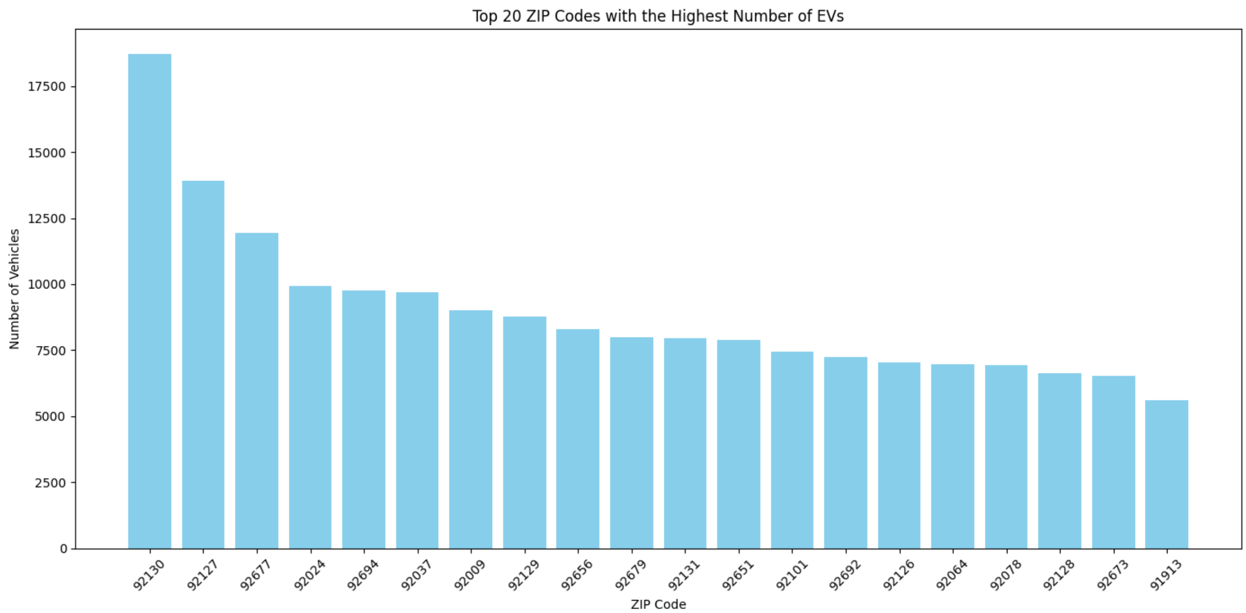

Figure 7: Top 20 ZIP codes with the highest EV registrations

A scatter plot was created to analyze the correlation between the number of EV owners and the availability of charging stations in SDG&E ZIP codes. While there is a positive trend, the scatter plot indicates significant variability, suggesting that charging station deployment may not always correlate with EV density.

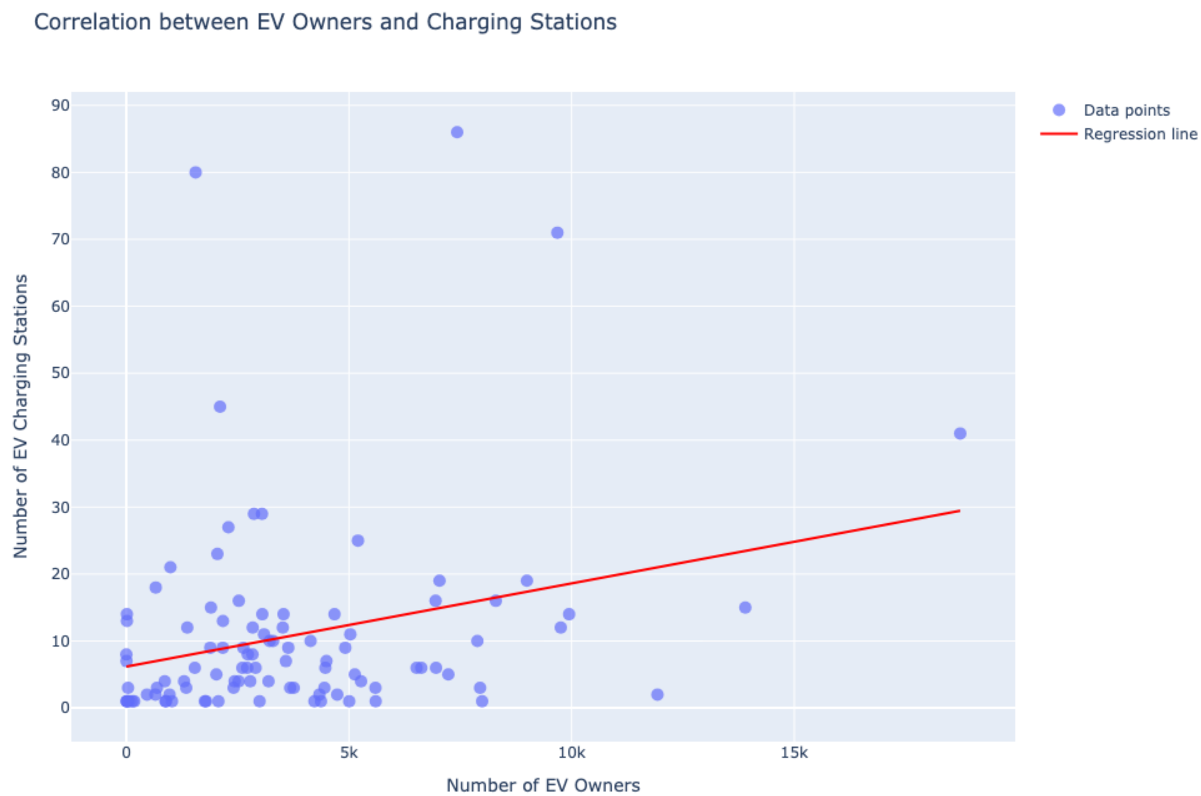

Figure 8: Correlation between EV owners and EV charging stations in SDG&E territories

Lastly, a histogram was created to visualize the Poisson samples for ZIP code 92122, derived from Monte Carlo simulations. This graph helps to model the distribution of registered EVs in this region, providing insights into the expected range of EV ownership in the future.

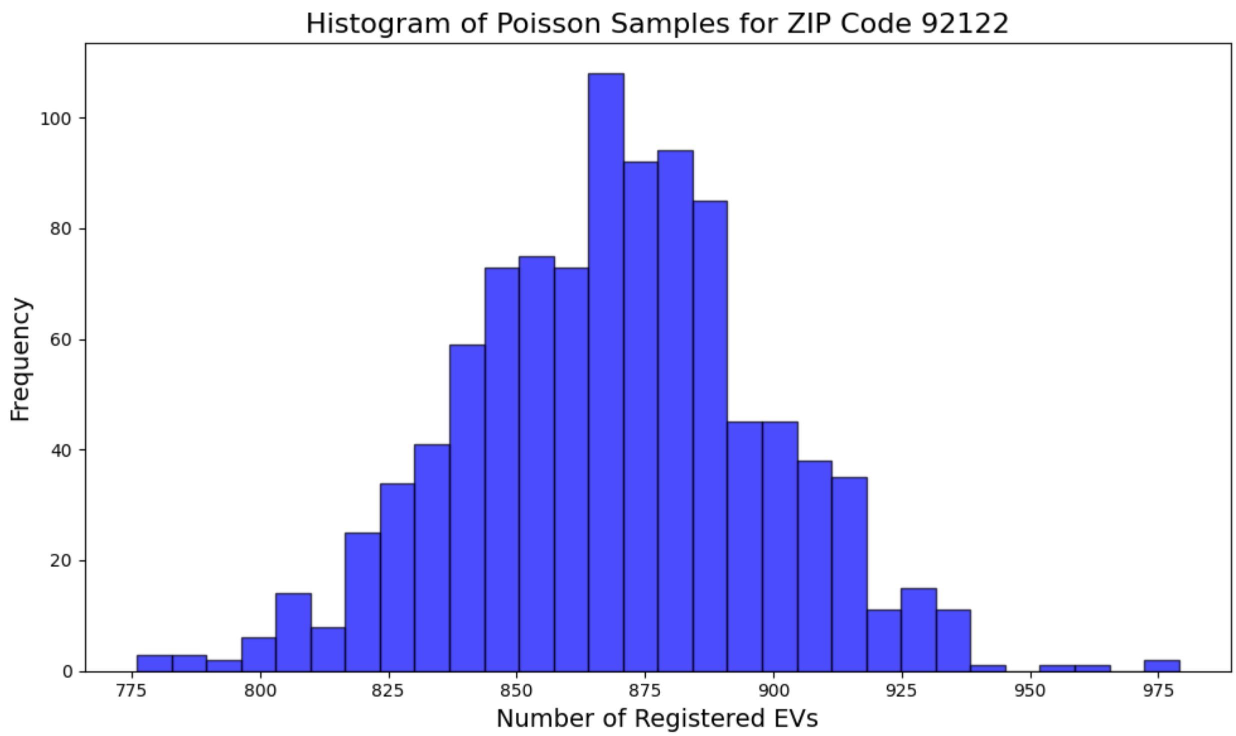

Figure 9: Histogram of Poisson samples for ZIP code 92122

2.4 Census Data EDA and Cenpy

The Cenpy library was used to retrieve census data related to median household income, and population density. The data was then used to find potential census correlation with EV charging station density within the SDG&E service territories. By merging census data with geospatial boundaries using the ZCTA shapefile, a geospatial analysis was conducted on EV charger placement using folium.

The first map visualizes the density of EV chargers across ZIP codes within SDG&E territories. This visualization helps identify areas with high and low availability of EV chargers, forming a basic understanding of EV charging station distribution at a glance.

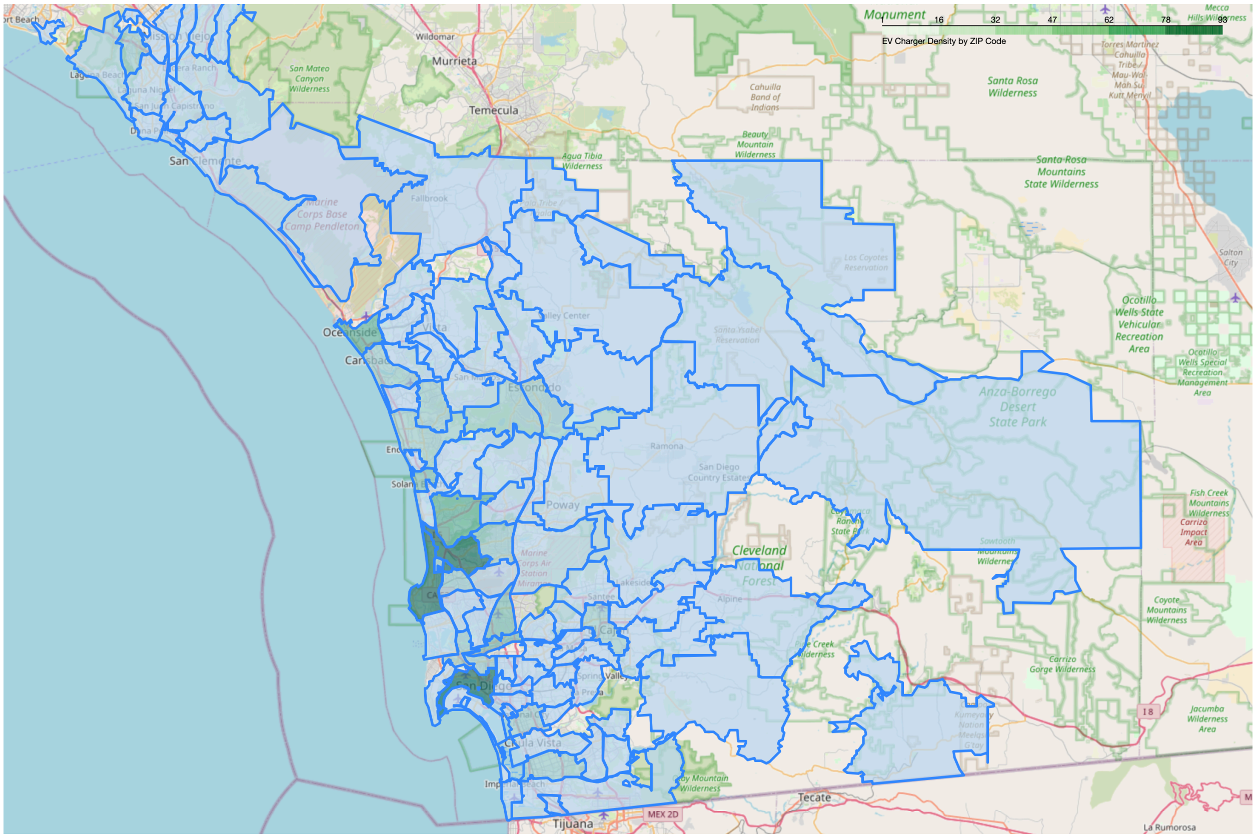

Figure 10: EV Charger Density by ZIP Code in SDG&E Territories.

The second map layers median household income with EV charger density. This analysis helps understand whether income levels influence the availability of EV charging stations in different ZIP codes.

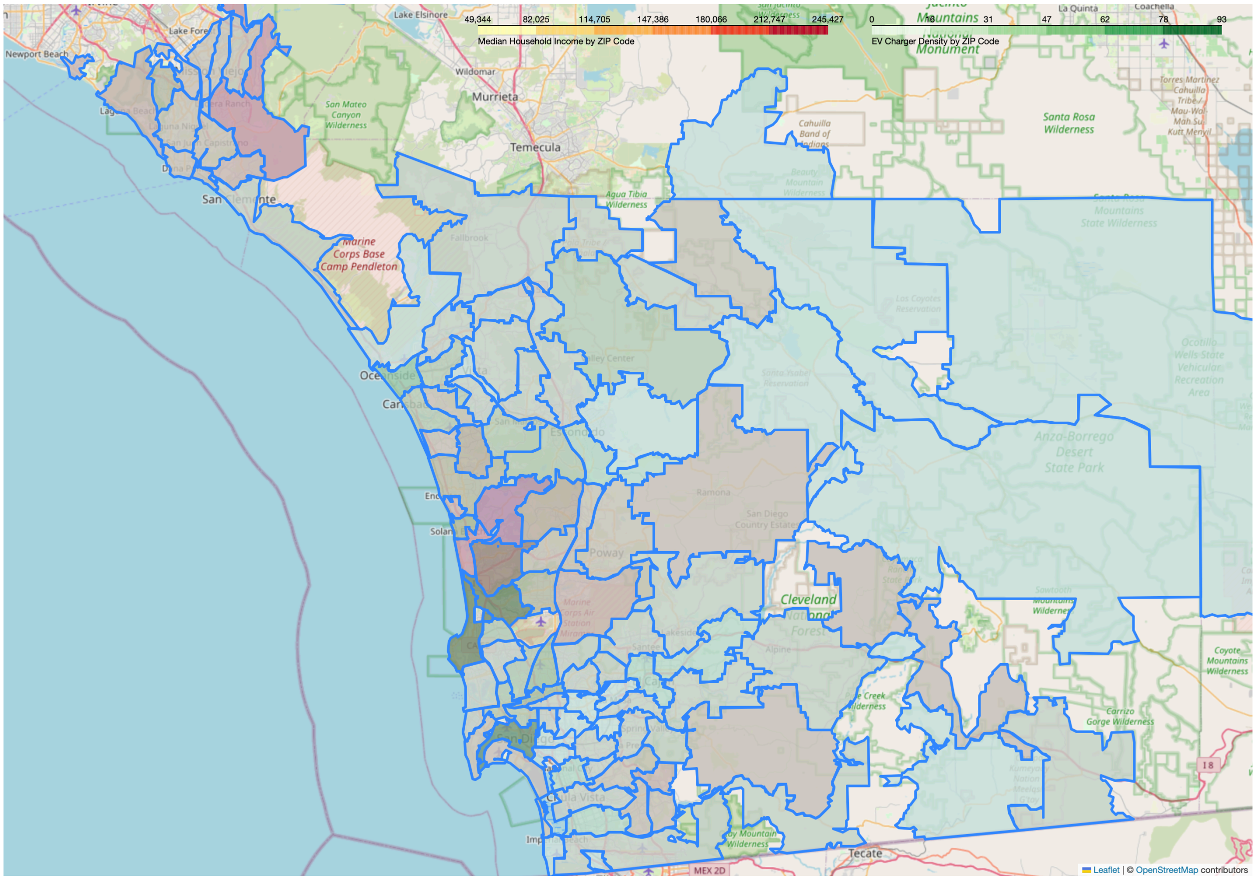

Figure 11: Median Household Income and EV Charger Density by ZIP Code.

The third map overlays population density and EV charger density. This visualization helps determine if areas with higher populations have proportionally more EV chargers available.

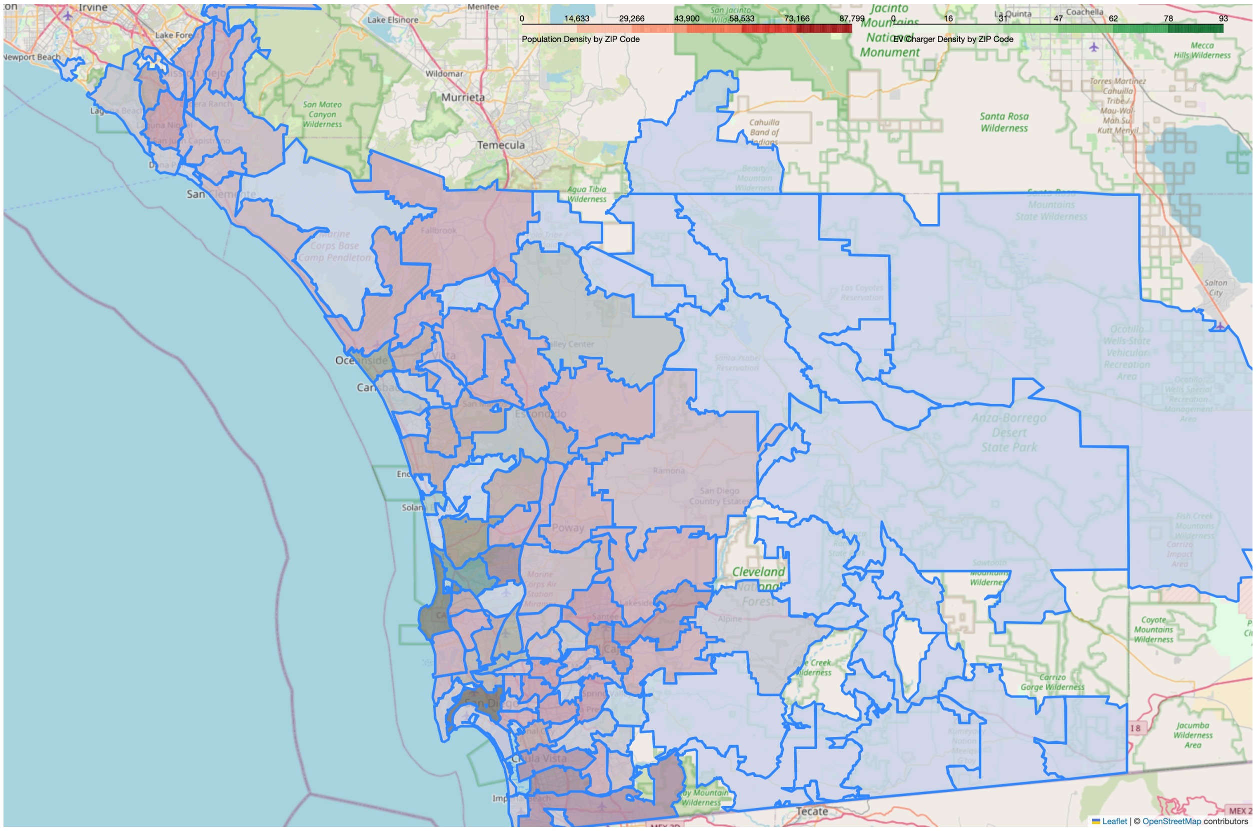

Figure 12: Population Density and EV Charger Density by ZIP Code.

The fourth map visualizes the correlation between median household income and EV charger density. The correlation score is calculated as:

$$\text{Correlation Score} = \frac{\text{Median Income}}{\text{Max Median Income}} \times \frac{\text{Charger Count}}{\text{Max Charger Count}}$$

This map identifies ZIP codes where income strongly correlates with the availability of EV chargers.

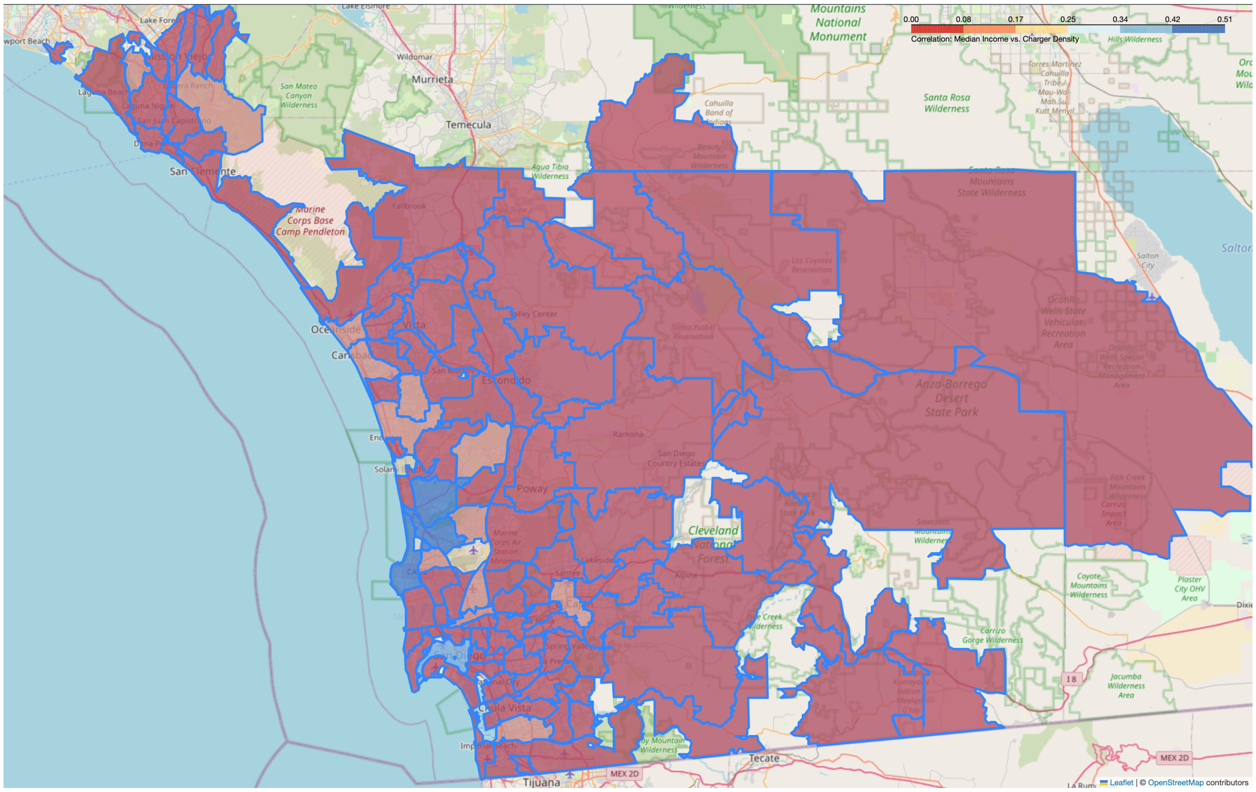

Figure 13: Correlation Score Map for Median Household Income vs. EV Charger Density.

The fifth map visualizes the correlation between population density and EV charger density. The correlation score is similarly calculated as:

$$\text{Correlation Score} = \frac{\text{Population Density}}{\text{Max Population Density}} \times \frac{\text{Charger Count}}{\text{Max Charger Count}}$$

This map highlights ZIP codes where population density influences EV charger availability.

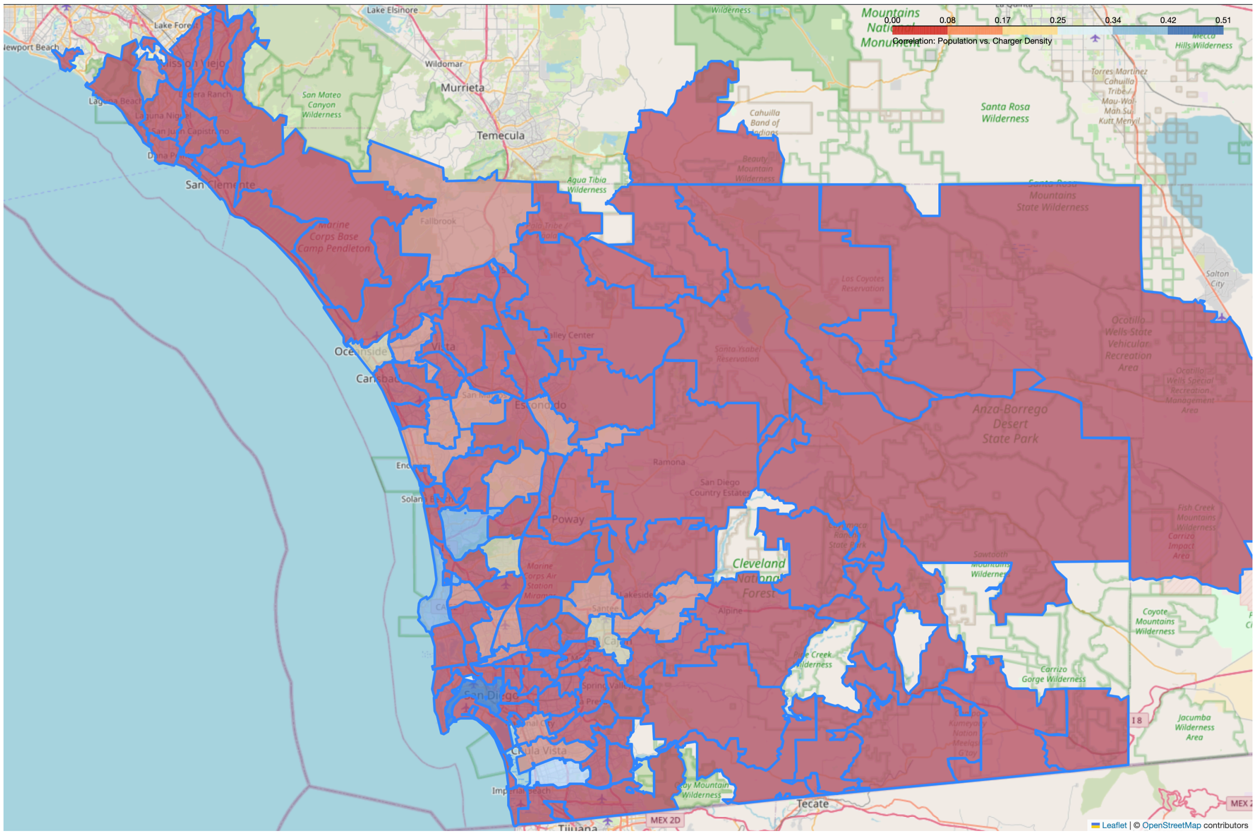

Figure 14: Correlation Score Map for Population Density vs. EV Charger Density.

3 Results

3.1 AFDC Results

Using the AFDC dataset, I developed insights upon the overview of EV charging station infrastructure in the SDG&E service territory. Figure 1 illustrates the steady growth of public EV charging stations over the past decade, with a sharp increase between 2019 and 2020. This aligns with the broader trend of EV adoption within the United States. Figure 2 highlights the dominance of certain EV network providers, such as ChargePoint, across the region. With a broad understanding of popular providers in the market, SDG&E's Lovelectric team would have a better understanding of their competitors. Figure 3 further breaks down this growth by charger type, showing that Level 2 chargers are the most widely installed, followed by DC fast chargers. This makes sense since Level 1 chargers are mainly for private home charging.

3.2 DMV Results

The DMV dataset provided insights into EV ownership patterns across ZIP codes within SDG&E territories. Figure 5 depicts the stable growth of EV registrations from 2018 to 2023. Figure 6 emphasizes the significant annual increase in EV registrations, particularly between 2022 and 2023. It is shown that every year, the increase of EV vehicles is on the rise consistently. Furthermore, Figure 8 reveals a moderate positive correlation between the number of EV owners and the number of charging stations per ZIP code. Lastly, after fitting the Poisson distribution on the EV registration per year, the Monte Carlo simulation provides a rough guideline of the potential increase in the area, as seen in Figure 9. However, there exist limitations to this method since the Poisson distribution depends on a constant increase over the years but our DMV EDA proves an acceleration in the increase yearly. Therefore, the Poisson distribution can only be used as a rough guideline for yearly increase, and for more accurate predictions, the Non-Homogeneous Poisson Process could be used alternatively to address the variability in its increase per year.

3.3 Census Results

The census data analysis revealed patterns in how socio-economic factors relate to EV charger placement. Figure 10 shows the distribution of EV charger density across ZIP codes. From the geospatial graph, it is shown that there exist a high density of EV charging stations near the San Diego Airport, around UC San Diego, UC San Diego Hospital, as well as the UTC mall. Figure 11 layers median household income with EV charger density, suggesting that there is no immediate correlation between higher-income areas with access to public chargers. Figure 12 overlays population density with EV charger density, showing a greater correction compared to the median income analysis but still has no immediate correlation. Correlation score maps in Figures 13 and 14 quantitatively assess the relationships between income, population, and charger density, showcasing correlation for all ZIP Codes relative to each other in SDG&E territories.

4 Discussion

The exploratory data analysis conducted on the AFDC, DMV, and census datasets provides valuable insights into EV adoption trends and charger placement in SDG&E territories. However, several key findings and limitations deserve further exploration.

Firstly, the AFDC analysis highlights the steady growth in public EV chargers over the past decade, with Level 2 chargers dominating the market due to their practicality for public use. Moreover, understanding market dominance by network providers, as shown in the AFDC results, offers strategic insights for SDG&E to evaluate competitors. However, this analysis does not account for the utilization rates of these charging stations, which may vary significantly based on location and time.

Furthermore, with OSMnx being a powerful tool to find road distances between locations, it could be leveraged to find distances between chargers to gauge charger density within ZIP codes and see which particular areas are under-served within ZIP codes. OSMnx could also be used to find distances between highway exits to existing chargers to see if there is a correlation between the two. OSMnx being such a versatile tool could be an incredible way for engineers to perform geospatial analysis to deep-drive AFDC data even further.

The DMV data analysis revealed a clear upward trend in EV registrations across ZIP codes, with a consistent acceleration in EV ownership in recent years. The Monte Carlo simulation based on a Poisson distribution offers a rough guideline for estimating future growth; however, it is limited in its ability to account for non-linear trends and variability in annual increases. As the DMV data suggests an accelerating adoption rate, a Non-Homogeneous Poisson Process or other time-series models may be more appropriate for accurately modeling these trends.

Census data analysis provided insights into the socio-economic and demographic factors influencing EV charger placement. The geospatial visualizations demonstrate high charger density near major hubs, such as airports, hospitals, and malls, reflecting infrastructure and urban planning priorities. However, the correlation analysis showed no significant relationship between higher-income areas and public EV chargers, as wealthier individuals tend to rely on private home charging infrastructure. Similarly, the analysis of population density revealed no immediate correlation with EV charger density.

These findings emphasize the importance of considering additional variables, such as commuting patterns, energy consumption behavior, and charger utilization rates, to improve the accuracy of charger placement models in the future. Furthermore, the limitations of using static datasets for dynamic phenomena underscore the need for real-time data integration.

5 Conclusion

This study provides a foundational analysis of EV adoption and charger placement in the SDG&E service territory using AFDC, DMV, and census datasets. The results reveal important trends in the growth of EV infrastructure, ownership patterns, and the socio-economic factors influencing public charger placement. While these insights serve as a valuable starting point, this analysis highlights the need for further investigation. Future studies should incorporate dynamic models, such as Non-Homogeneous Poisson Processes, and integrate real-time data on charger utilization to better capture the evolution of EV adoption.

Despite its limitations, my findings and insights demonstrate the potential of data-driven approaches to support clean transportation adoption. By building on these findings, SDG&E can make more informed decisions to optimize EV infrastructure and foster equitable access to public charging, ultimately accelerating the transition to sustainable clean transportation.