Evaluating Electric Vehicle and Charger Distribution in California's SDGE Territory: Insights from DMV and AFDC Data (DSC180)

Ethan Deng
ezdeng@ucsd.edu

Phi Nguyen
pnguyen@sdge.com

Abstract

The adoption of electric vehicles (EVs) is a critical step toward reducing carbon emissions and promoting sustainable transportation. However, effective EV growth requires a well-distributed network of charging stations to meet demand. This project explores the distribution of EVs and charging infrastructure in California's San Diego Gas and Electric (SDGE) territory using comprehensive Department of Motor Vehicles (DMV) registration data and Alternative Fuel Data Center (AFDC) charger data. Existing analyses often focus on either vehicle adoption or charging station availability separately, overlooking their dependence on each other. My work bridges this gap by conducting an exploratory data analysis (EDA) that visualizes EV-to-charger density per ZIP code and highlights potential disparities in infrastructure. Leveraging geospatial data, I illustrate regional patterns, identify underserved areas, and propose considerations for balancing EV adoption with infrastructure expansion. The results provide actionable insights for policymakers and stakeholders looking to optimize resource allocation to support the growth of EVs.

Code: https://github.com/ethandengg/SDGE_EV_EDA

1 Introduction

The shift toward electric vehicles (EVs) has accelerated in recent years, driven by environmental concerns, technological advancements, and supportive policies. California, a leader in clean energy initiatives, has skyrocketed in EV ownership. However, the success of this transition hinges not only on the number of EVs but also on the adequacy and distribution of charging infrastructure.

Despite increased efforts to expand the charging network, gaps remain in understanding the relationship between EV density and charger availability at the local level. Previous studies have either concentrated on mapping EV adoption or assessing the coverage of charging stations but rarely addressed both aspects together. Our project aims to fill this gap by analyzing DMV vehicle registration data to quantify EV counts and combining this with AFDC charging station data to map charger density across ZIP codes within the SDGE service area.

I conducted exploratory data analysis (EDA) with geospatial, time-series, and scatter plots using OenStreetMap NetworkX (OSMnx) to visualize the distribution of EVs and chargers, uncover patterns, and identify areas that may lack sufficient charging infrastructure relative to vehicle demand through census data using Cenpy. Through this approach, I highlight opportunities for targeted infrastructure investment and policy adjustments to support the equitable growth of EVs. The data-driven insights derived from this analysis serve as a valuable resource for local governments, utility companies, and urban planners striving to enhance the sustainability and accessibility of EVs.

2 Exploratory Data Analysis (EDA) Methods

2.1 AFDC Dataset

The AFDC dataset was gathered through the AFDC API using a personal key. And this data consisted of all the alternative fuel stations in California, where each row was either an EV charger, natural gas station, ethanol pump, hydrogen pump, etc. However, many of the columns were irrelevant for our analysis because they consisted of features that pertained to the other types of fuel sources, but I was only interested in EV charging stations. Luckily, most of the data in was EV chargers, meaning most of the data was rich in the features I was looking for. I also only gathered the columns that were useful for my EDA, which include connector types, open date, ZIP code, etc. I did this through the pandas library by truncating down the data frame to the relevant columns, and then extracted only the rows that contained ELEC in the fuel_type_code column, which means EV charger in the data.

After truncating the data down to just EV chargers and the relevant columns, I plotted a time-series analysis graph using plotly.express to showcase the growth in EV chargers over time after the year 2000. I chose to only include data after the year 2000 because the data before that year is too small to see in comparison to the recent years, and it is just not relevant data. I also added the "chargers added that year" line to showcase how much each year is growing in comparison to the previous year.

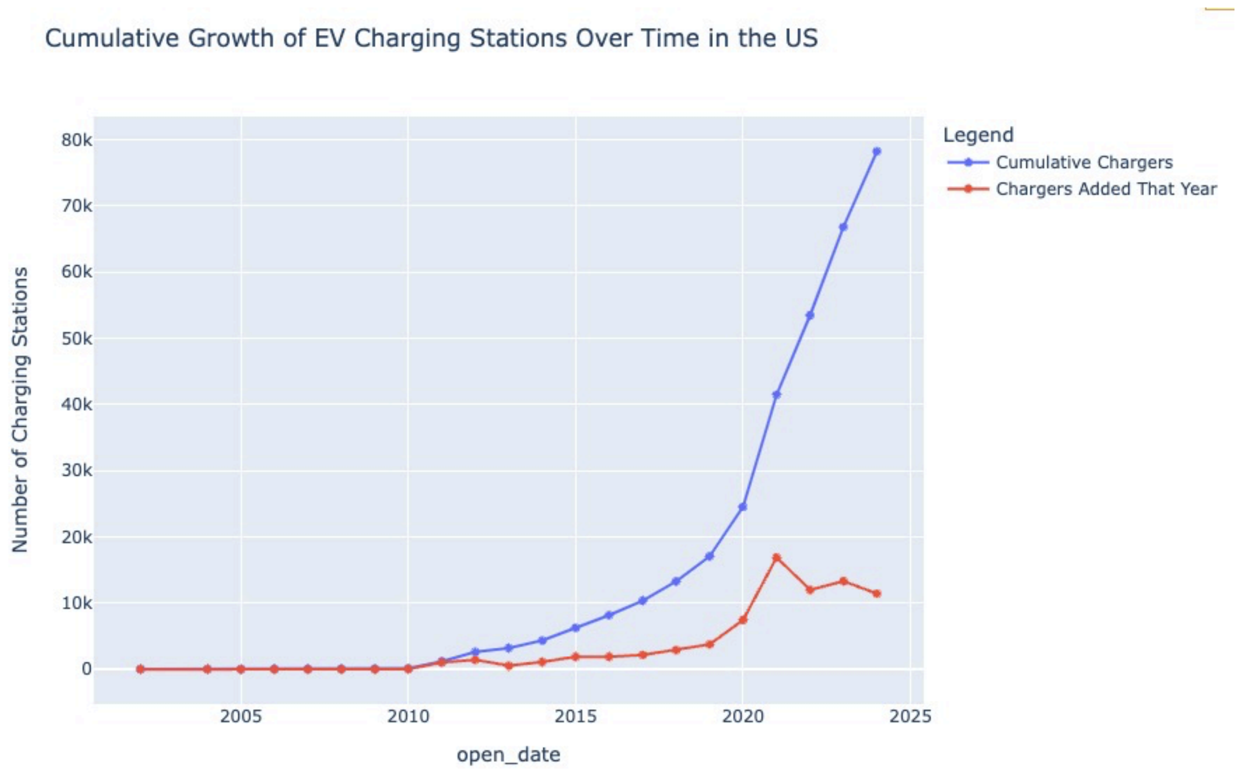

Figure 1: Growth of EV chargers in the US

I also plotted the same graph for all EV chargers in the SDGE territory and found a very similar trend, where the main spike in growth occurred in 2020. However, there are no data points before the year 2011, except for one point in 2006 which is interesting.

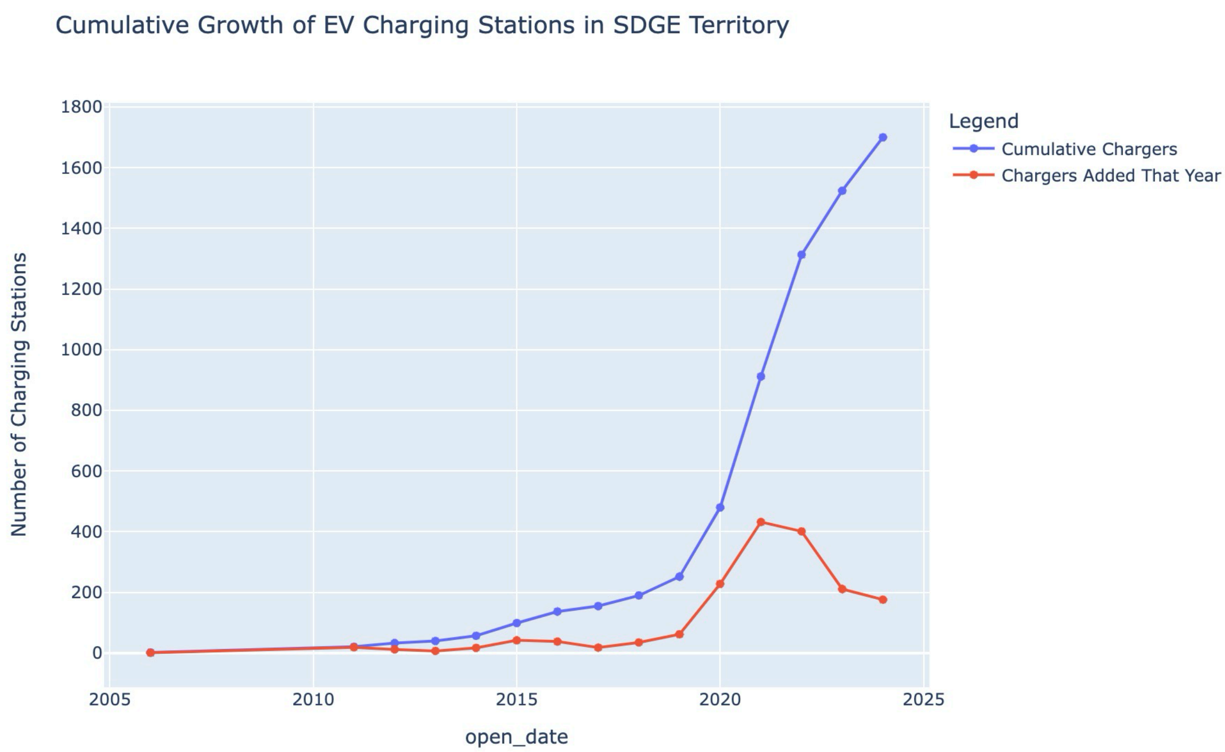

Figure 2: Growth of EV chargers in the SDGE Territory

I also plotted the growth of different EV connector types over time in the US, and we can see that the J1772 Plug is far more popular than the other plugs surprisingly, considering that the Tesla plug is North American Charging Standard (NACS). The really weird part about the growth of these chargers is that Tesla is growing the least amount after 2022, despite that being the same year it became the NACS.

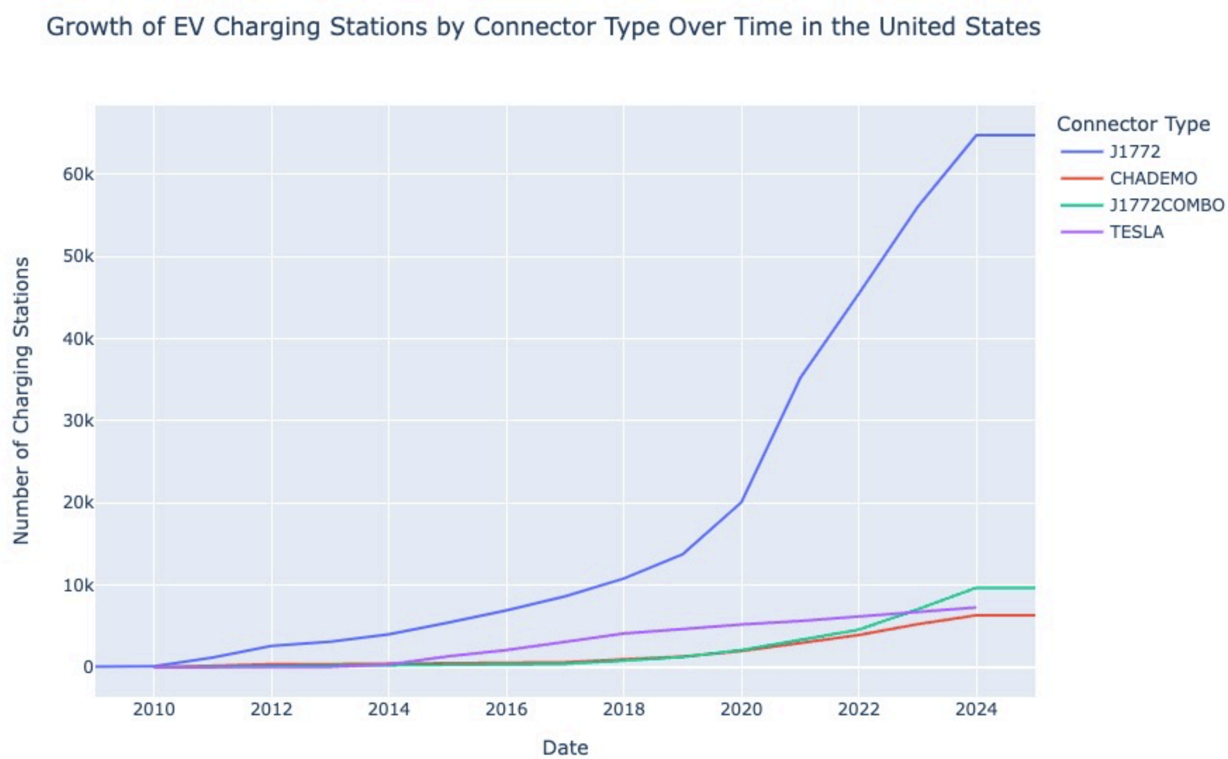

Figure 3: Growth of EV chargers in the US by connector type

Moreover, using ZIP code data, I gathered the number of charging stations located in each ZIP code in the SDGE territory to illustrate the distribution of chargers. I got that data by getting a list of unique ZIP codes in the SDGE territory and filtered the AFDC dataset down to just those rows. The geospatial choropleth plotting was done using folium and Matplotlib. This is to better visualize where the density of EV chargers is in SDGE, and the disproportial concentration in certain areas versus others. Here the more orange and red an area is, the more EV chargers it has.

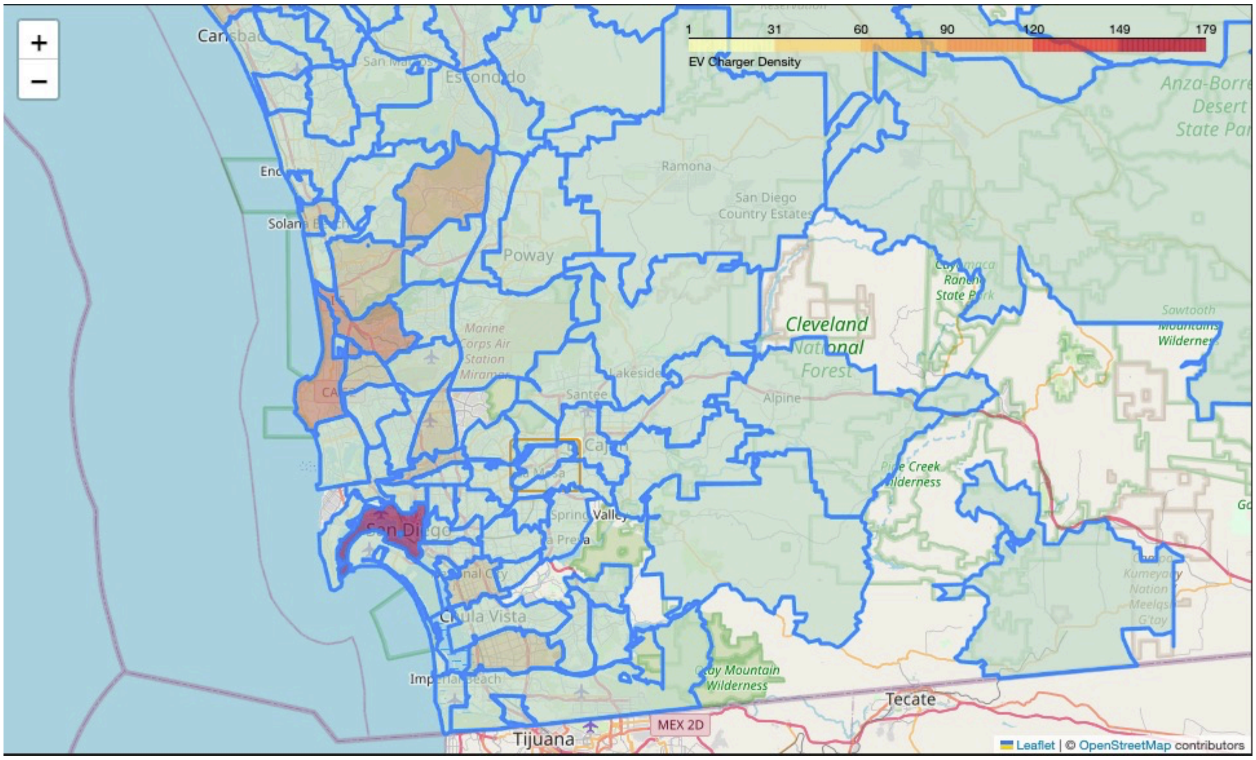

Figure 4: Number of chargers per ZIP code

Further analysis includes gathering census data from cenpy such as the median household income and total number of EV chargers there are per ZIP code in the SDGE territory to see if there was a correlation between those features. I gathered the information about the median income with cenpy.products.APIConnection('ACSDT5Y2022') and then I truncated the data down to the SDGE ZIP codes using pandas. I used a scatter plot to represent the spread and correlation between the different ZIP codes, where each point represents a ZIP code. This was done with plotly.express

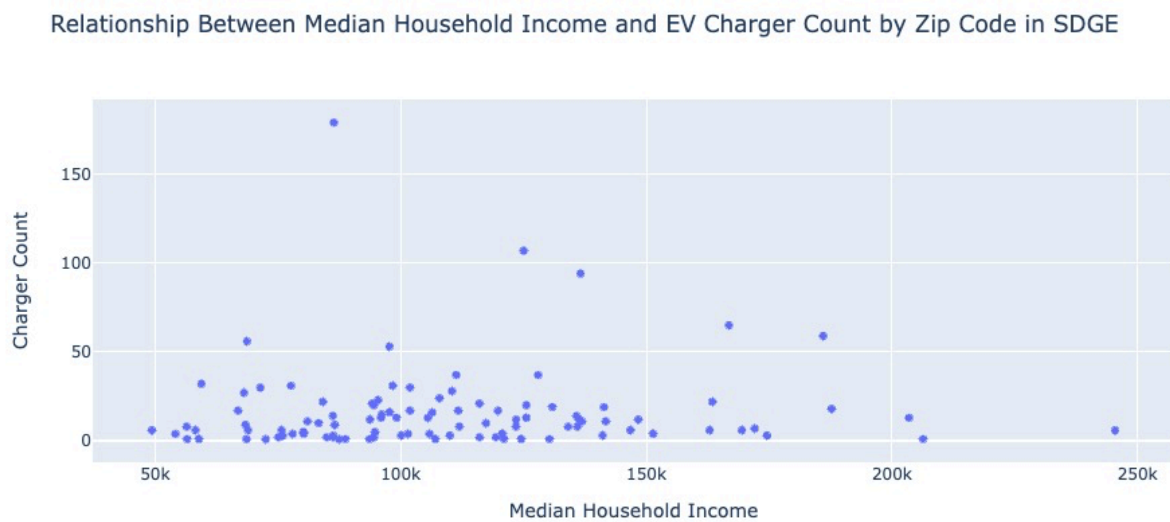

Figure 5: Correlation between Median Income and number of EV chargers per ZIP code

Here we can see the correlation between Total Vehicles per and the number of EV chargers there are per ZIP code. I again used cenpy to gather this data, and we can see that there is only a very slight correlation between the two.

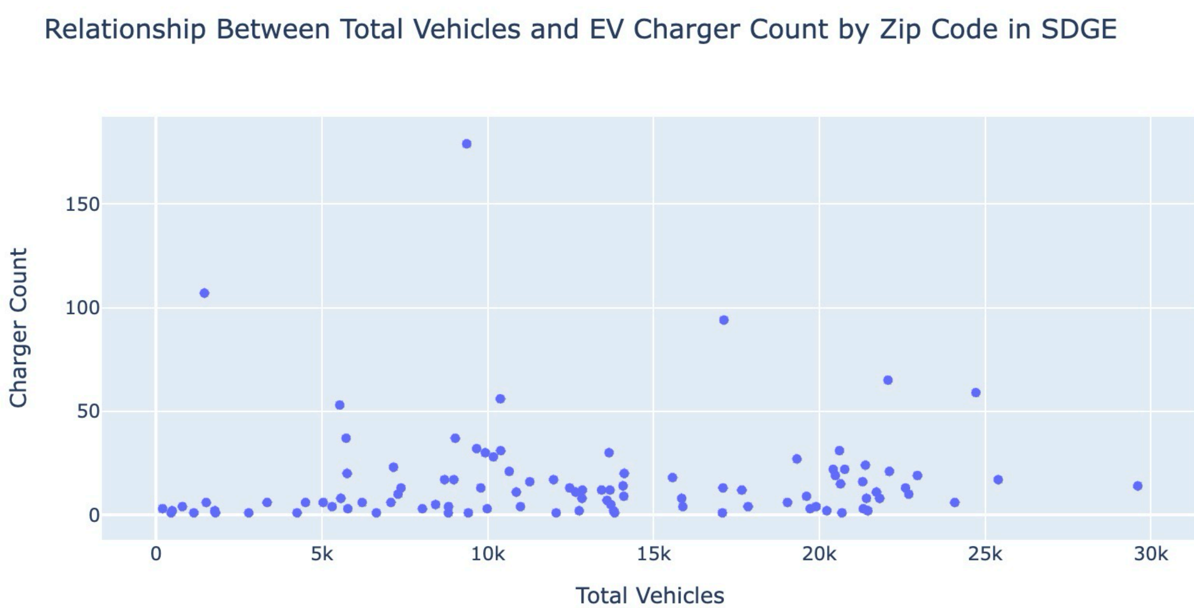

Figure 6: Correlation between number of total vehicles

I also used OSMnx to find the shortest path through the network of roads between a charger I found and SDGE headquarters. I used folium to get the geospatial plot, and the network in OSMnx to get the path between the 2 coordinates. The coordinates were found using

Google Maps.

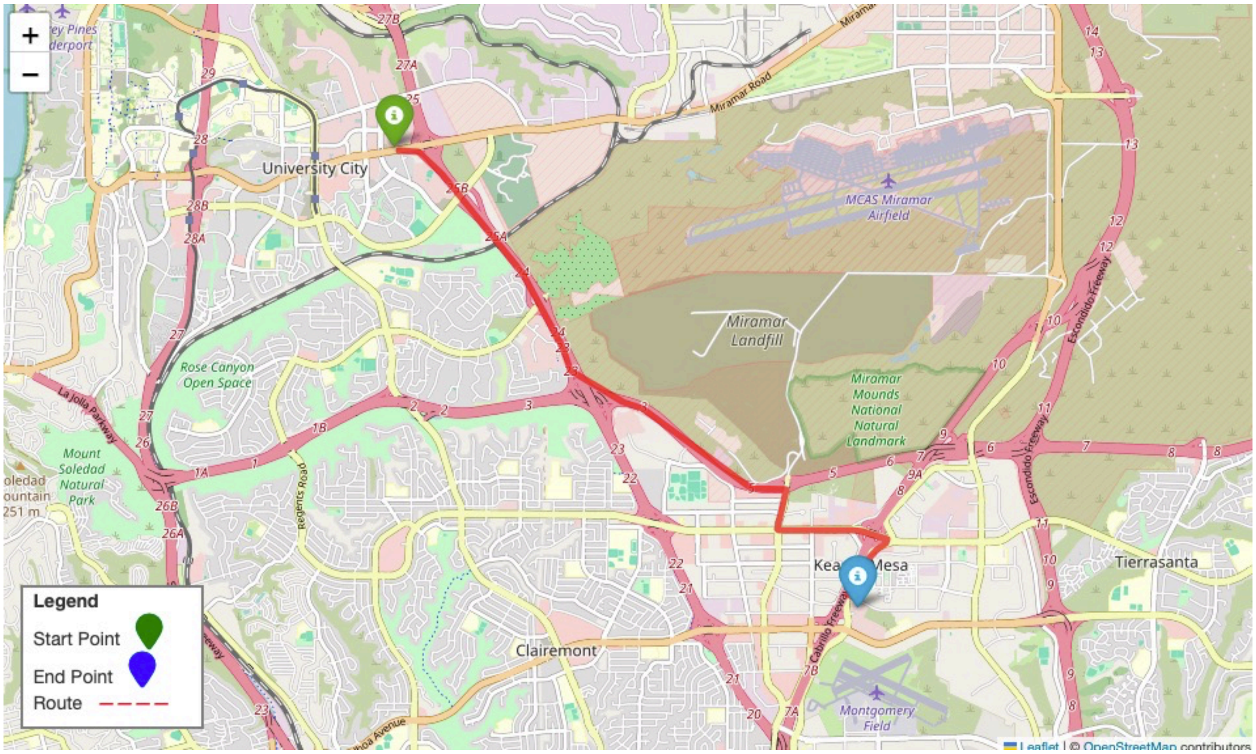

Figure 7: Distance between my charger and SDGE headquarter

2.2 DMV EDA

I gathered the DMV dataset through an API. This dataset includes all vehicles registered in California from 2019 to 2024, and they are categorized by ZIP code, fuel type, and many other features that are not relevant for our analysis. With this data I found trends in total vehicle registration as well as trends in EV registration.

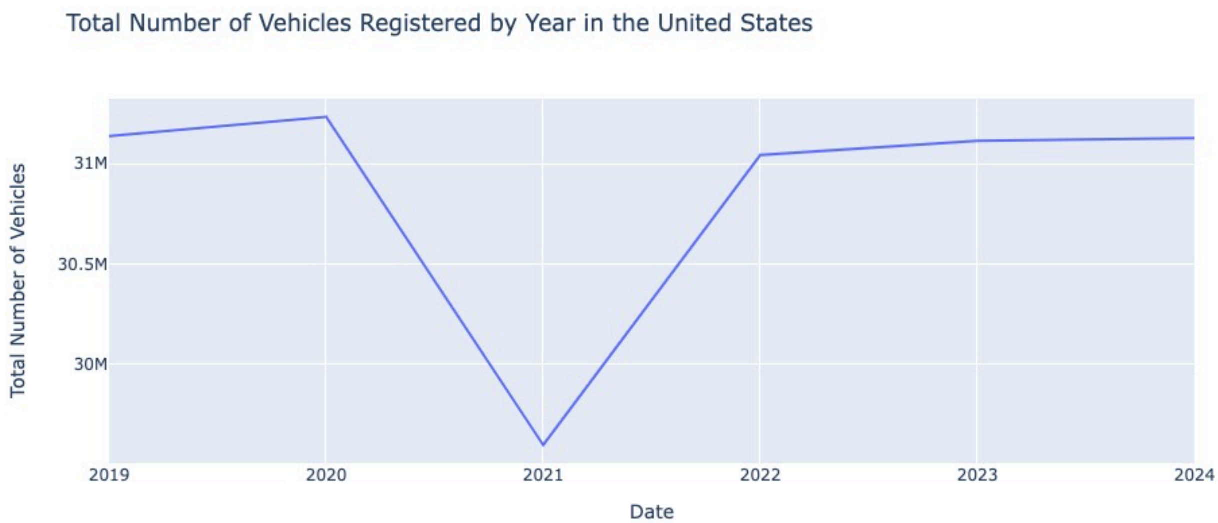

Figure 8: Total Vehicle registration from DMV in California

Here is the trend I found in the EV registration

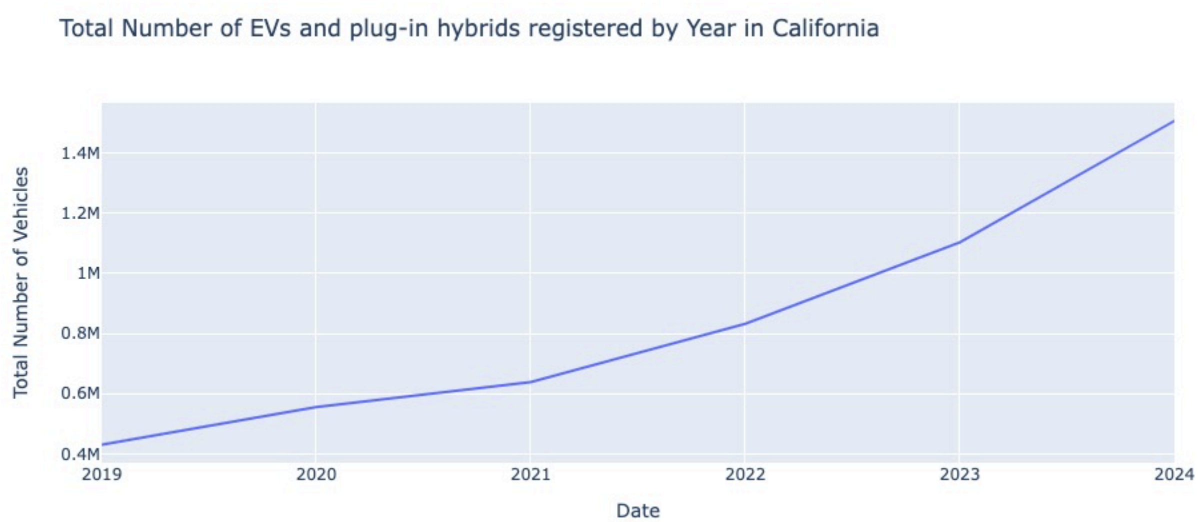

Figure 9: EV registration in California

Here I ran Monte Carlo simulation on a Poisson distribution on the number of EVs and plug-in Hybrids per ZIP code and year. I did this using stats models and random sampling.

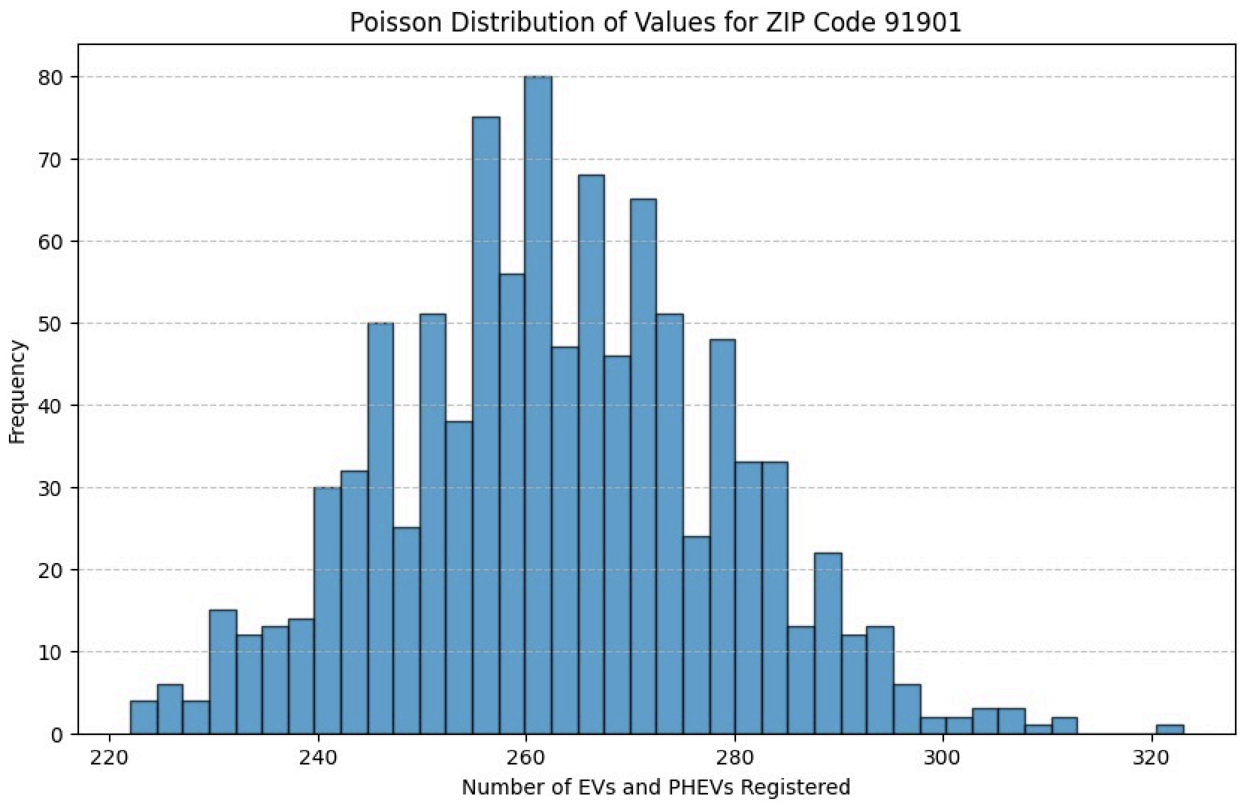

Figure 10: Monte Carlo Simulation on Poisson Distribution for ZIP code 91901

I also investigated the trend between the total EV registration and charger counts from AFDC in each ZIP code in the SDGE territory and plotted the correlation with plotly.express and added a line of best fit with statsmodel

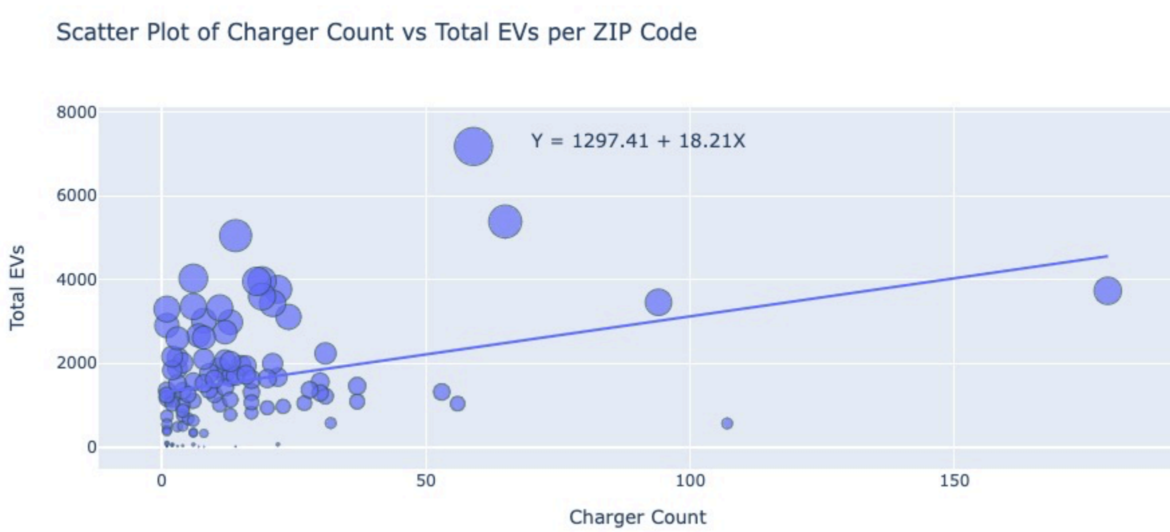

Figure 11: Correlation between EV Chargers to number of EVs per ZIP code in SDGE

Finally, I used folium to plot the ratio of EVs to EV chargers per ZIP code in SDGE after merging data frames and aggregating in pandas. What I did was create a new column in the merged data frame to evaluate the ratio of $\frac{EV}{EV Chargers}$. I did this to get a better understanding of which areas are underserved in terms of not having enough EV chargers for their community, which can be easily seen on a map.

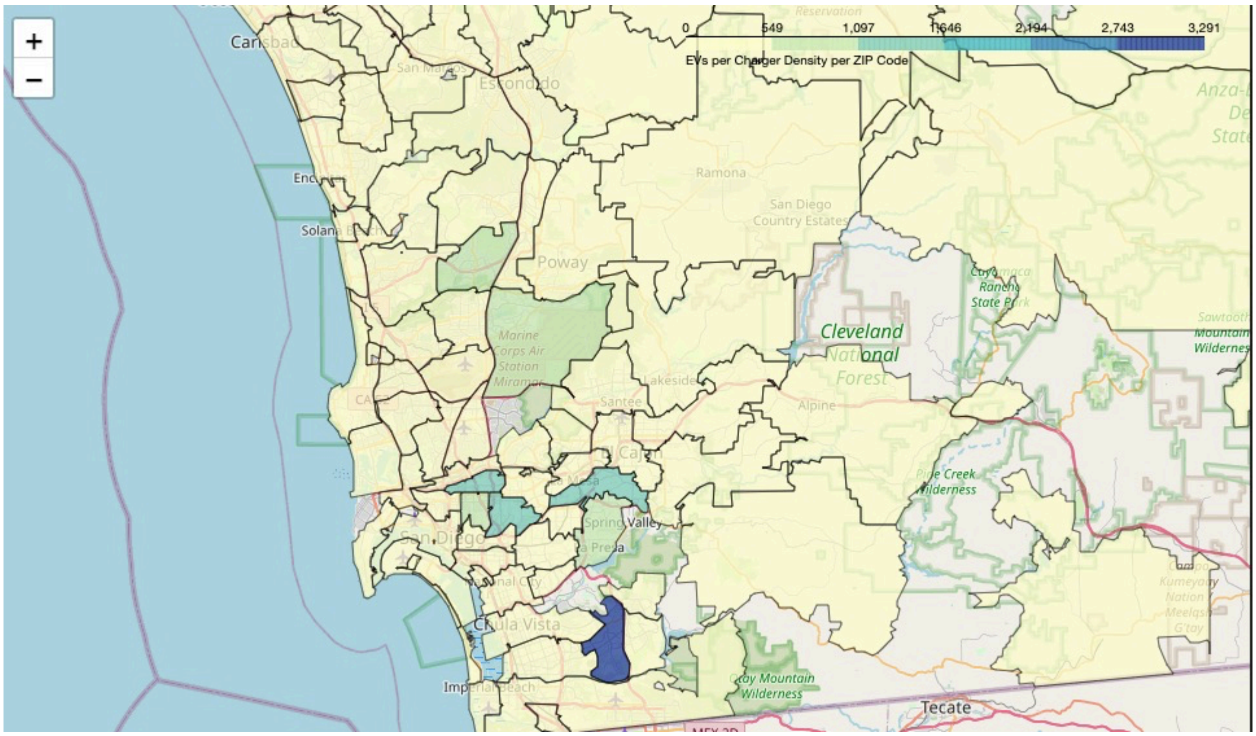

Figure 12: EV to charger ratio

3 Results

3.1 AFDC Results

In this EDA I found that there is a massive increase in EV chargers being built around the whole US that started rapidly increasing around 2020, surging from 24k chargers in 2020 to 78k in 2024 in Figure 1, which is a very good sign for moving towards cleaner transportation. The growth found in SDGE is also similar to the growth of the whole country, just scaled down. The graph showing the growth in different connector types is interesting because J1772 is much further ahead than the other plugs, with 64.75k J1772 plugs and the other 3 chargers sitting around 8k plugs each in Figure 3.

The map showing the EV charger distribution per ZIP in Figure 4 code highlights that the majority of chargers are in high traffic areas, like downtown San Diego with 178 chargers, La Jolla and Sorrento Valley have 94 and 107 chargers each respecivly, and Black Mountain Ranch and Mission Valley have 65 and 53 chargers each. Having a bunch of chargers in downtown makes a lot of sense because there are very few homes if any where people can park their car in a garage and charge at home, and also the attractions in this area are what drive high traffic and make it essential to have many EV chargers.

The census data correlation between median household income and charger count is very weak and almost non existent, it seems there is a slight positive correlation, but the majority of ZIP codes cluster around having a low number of chargers between 1-15. This can be seen in Figure 5

The census data correlation between total vehicles and charger count per ZIP code is also fairly weak, but slightly positive. There are also a huge cluster of ZIP codes that don't have many chargers. This can be seen in Figure 6.

3.2 DMV Results

In this EDA I found a steady increase in the number of EVs registered in California, despite the total number of vehicles being registered to remain fairly stable. There were only 430k EVs and plug-in hybrids registered in 2019, and that number has climbed to 1.5m in the beginning of 2024. I decided to include plug-in hybrids along with EVs because they do require charging occasionally. This can found in Figure 9

I ran a Monte Carlo simulation on a Poisson distribution in the ZIP code 91901 and found that the most likely occurrence of the number of EVs and PHEVs registered is around 265. This can be seen in Figure 10. This is just 1 simulation in 1 ZIP code, but I have a table with a poisson distribution for each ZIP code, and each one can have a Monte Carlo simulation ran on it.

There is also a small increasing trend per ZIP code in SDGE where there are more EV chargers in ZIP codes with more EVs. The general trend being that there are 18.21 more chargers per EV registered in that ZIP code. This can be found in Figure 11

In the map being displayed in Figure 12, we can pinpoint general areas that have a lot of EVs, but not very many EV chargers. The 91913 ZIP code (Chula Vista) is a great example of this, where there are 2899 EVs registered there, but only 1 public EV charger in that area. These are exactly the kinds of areas that we need to target because there is a growing EV market with a lack of infrastructure to support the growth for adopting EVs.

4 Discussion

The general increase in EV chargers is a great thing and the fact that we have been picking up the pace since 2020 means we are moving to a future with cleaner transportation. However, to hit the goals of having 500k public EV chargers in the US by 2030, we still need to step up the rate of production drastically. If we move at the pace we have been going for the past 4 years, we will only get to about 170k chargers by the year 2030, which is barely a third of our goal. The general trend can be seen in Figure 1

The growth among the different connector types is what is shocking, considering that J1772 is far ahead of the all the other connectors, despite Tesla being the North American Charging Standard (NACS) since 2022, J1772 plugs continue to be produced. I also mentioned earlier that even the CHADEMO and CCS chargers are growing faster than the Tesla chargers after the year 2022, which just shows how much more we need to focus on enforcing standardizing the plug for easier adoption. This can be seen in Figure 3

Regarding the distribution of EV chargers among the ZIP codes in this figure 4, it's important to note which areas are the most concentrated and which areas may benefit from having more access to EV chargers, especially places where EVs are growing in popularity. We also need to consider if there is enough power generation available for these areas even if we build chargers there. Ideally, this power would come from clean energy sources like wind or solar, and able to be stored in either batteries or other forms of energy such as pump electro or stored as potential energy in concrete towers. In order for this power to be transferred, we also need more high voltage transmission lines and distribution lines so that this power can be transferred to places in need. There are so many factors that go into what places need EV chargers, and how we can get that power to those places so the chargers there can be effective.

It is also interesting to see that there is a fairly weak correlation between the median household income and the charger count per ZIP code, as well as the correlation between total vehicles and charger count seen in this Figure 5. We already know that there is a clear correlation between EV adoption and household income, as income rises, so do the chances of buying an EV. However, this is not the case with EV chargers. The most reasonable hypothesis for this is that these EV owners likely own chargers at home, and we also know that nearly 90% of EV owners charge their car at home, which explains the low demand of public EV chargers. As far as the correlation with total vehicles, there is a slight correlation, but the reason why the correlation is so weak is because this is between ALL vehicles and charger count, not just EVs. This makes sense because EVs are still such a small part of the total market of cars.

The Monte Carlo simulation ran on the ZIP code 91901 in Figure 10 can tell us information on how many EVs and PHEVs there could be in that ZIP code in the future. I also a Poisson distribution of all the other ZIP codes, so there could be a monte carlo simulation generated for each one to tell us more about what to expect in the future in terms of how many EVs and PHEVs will be in the area. This will give us a better understanding as to how we need to improve the current infrastructure to accommodate for the growth in EVs

In the map in Figure 12, we can see that there are certain areas that have a high concentration of chargers, but this needs to be better distributed to accommodate for areas that have fewer chargers. There are some areas that are heavily underserved as not only do they have a low charger count, but there is a growing adoption of EVs in that area. This combination can lead to the trouble of not having the infrastructure to support a growing community in that area.

The overall growth in EVs and PHEVs found in Figure 9 is positive, but we need a lot more EVs to be sold in order to hit our target number of 7 million EVs in California by 2030. We also have an ambitious target of hitting 8 million EVs in California by 2035, and based on the trajectory of this line, we are not going to make that goal. California climbed from 500k EVs and PHEVs in 2020 to 1.5 million in 2024. That means we will only have about 3 million EVs by 2030 if we grow at this rate, so we have to exponentially increase our growth if we want to hit our targets.

5 Conclusion

In this analysis, we observed significant progress in the growth of electric vehicle (EV) adoption and charging infrastructure across California, particularly within the SDGE territory. The rapid increase in the number of EV chargers since 2020, as highlighted in Figure 1, reflects a positive shift toward achieving sustainable transportation goals. However, the current pace of charger production must be accelerated to meet the ambitious target of 500,000 public chargers in the U.S. by 2030.

While the dominance of the J1772 connector, as shown in Figure 3, underscores its widespread adoption, further investigation into the long-term implications of having multiple connector standards is warranted, especially with Tesla's NACS designation.

The geographical analysis revealed critical disparities in charger distribution, with underserved areas such as Chula Vista (ZIP code 91913) demonstrating the need for targeted investment in public EV charging infrastructure. As illustrated in Figures 4 and 12, high traffic areas like downtown San Diego are well-served, while suburban and rural areas remain under-equipped to support increasing EV ownership.

Finally, the correlation between EV registrations and charger counts across ZIP codes, seen in Figure 11, highlights the need for a balanced approach to infrastructure deployment that aligns with growing EV demand. Although California's total EV registrations, as shown in Figure 9, are increasing steadily, substantial growth is required to reach the state's target of 7 million EVs by 2030.

In conclusion, while progress has been made in the adoption of EVs and the expansion of charging infrastructure, critical gaps in equitable charger distribution and production rates remain. Policymakers, utility companies, and stakeholders must collaborate to address these challenges, ensuring that the transition to clean transportation is both sustainable and accessible to all communities.