MIMS 2021 Capstone Final Report

GEAR

Gaps Explorer for Accessibility and Readiness: California Electric Vehicle Charging Infrastructure

Alison Kuo

Wenqi Luo

Yin Qiu

Chuhan Wang

Advisor: Marti Hearst

Table of Contents

1. Project Context

Transportation accounts for 40% of the greenhouse gas (GHG) emissions in California (CARB, 2020). To fight against climate change and reduce air pollution, California must rapidly transition its vehicle fleets to zero-emission in the next decade. The latest California Air Resources Board (CARB) decarbonization scenario modeling suggested 70% of new cars sales in 2030 need to be Zero-Emission Vehicles (ZEVs) (CARB, 2020). The share must reach 100% by 2035, as mandated in 2020 by the Governor (CA State Government, 2020). It is an important marker in phasing out internal combustion engine (ICE) vehicles and continuing to reduce the reliance on fossil fuels for its economy. As a result of the new sales goal, it is then estimated that 8 million light-duty ZEVs will be on the road by 2030 (CARB, 2020) and 1.5 million public and shared private chargers are needed (CEC, 2021). These are much more ambitious than the 5 million ZEVs and 250,000 chargers goals set by the former Governor in 2018.

The current gaps are prominent. On the one hand, although the EV market share climbs from 3% in 2016 to 9% in 2021 (CEC, 2021), we need a ten-fold increase in ZEV sales to reach 100% by 2035. On the other hand, access to EV charging infrastructure is the top 3 consideration when drivers consider owning an EV (NREL, 2019). California currently has about 73,000 public chargers available (CEC, 2021), less than 5% of what is needed by 2030. All of these require strong public support to accelerate deployment, including the Governor's proposed budget of 1 billion dollars for charging infrastructure (CA State Government, 2021) and the ZEV Market Development Strategy to guide the implementation by agencies (CA State Government, 2021). These essential steps will help break the “chicken-and-egg” problem in the early days for would-be EV adopters who refrain from buying EVs due to a lack of public charging stations, and for would-be charger operators who have not entered the market due to currently low station utilization and immature business models (Union of Concerned Scientists, 2021).

The next steps for policy makers to think about are, first, how to provide the correct incentives and reduce the need for public investment to close the gaps in *scale*, and, second, how to make plans and allocate resources to ensure charger *equity* across different demographic groups and geographic regions. Our project is designed at this critical time to explore the gaps in scale and equity about access to charging infrastructure in order to monitor our readiness. How we understand these gaps are fundamental to the success of electrifying California's vehicle fleets.

2. Project Scope and Objectives

Our project was proposed to address the gaps of both accessibility and resilience of EV charging infrastructure in California, as seen in the original title of GEAR (Gaps Explorer of Accessibility and Resilience), but now narrowed down to the accessibility aspect to go more in depth in monitoring our readiness. To align with the primary goal of ZEV Market Development Strategy to accelerate large-scale, affordable and equitable ZEV market development, we have the following project objectives:

- Understand EV users' persona and their current issues with charging
- Explore gaps in charger equity in terms of demographic attributes of communities
- Explore gaps in charger scale compared to forecasted demand and understand how policies can help close the gaps faster

We utilized information visualization methods and web development frameworks to build a web-based visualization tool to present our findings. We also applied natural language processing for an aspect-based sentiment analysis on EV users' reviews of charging infrastructure, and the results are in Chapter 5.

3. Information Visualization Methods Design

3.1. Datasets

- 2019 California Vehicle Survey

Conducted by the California Energy Commission, the 2019 California Vehicle Survey collected 6,549 Californian light-duty fleet owners' vehicle preference and usage experience data. The dataset consists of 4,248 residential surveys and 2,301 commercial surveys. For the purpose of the current study, we only used residential survey data for analysis, which contains responses from 718 PEV (plug-in electric vehicles) owners on their economic and demographic characteristics, including income, race/ethnicity, and type of housing. Besides, it recorded different aspects of EV drivers' charging behaviors, and choice & preference of vehicle and charger type. We mainly used this dataset to build EV driver persona and understand their charging experiences including any pain points related to the availability of public charging infrastructure.

- Charger Data

The charger data was acquired by calling NREL (The National Renewable Energy Laboratory) Data API. The dataset contains the most up-to-date information on public chargers, including their geographic location, type, and quantity. We leveraged this dataset to analyze the status quo of public charging infrastructure at scale so as to further quantify the gap to reach the 2035 goal.

- 2017 Census Data (from American Community Survey)

Census Data were obtained through the Data API of the United States Census Bureau. We selected data on race/ethnicity, household median income, type of housing, and population density for 58 counties and 8057 census tract. We incorporated this dataset with charger data in order to (1) first map the current distribution of charger from a demographic perspective and (2) test if there is a charger access disparity among different socio-economic/demographic groups.

- Urban/Rural Code

We used urban/suburban/rural code organized by the California State Association of Counties that assigned each county in California an urbanization level.

- CalEnviroScreen 3.0 Score and SB 535 Disadvantaged Communities

Each census tract in California gets a score for its vulnerability to pollution effects by the California Office of Environmental Health Hazard Assessment. The metric takes into the "cumulative impacts" of various environmental, health and socioeconomic factors for different communities, and thus can be considered a more holistic measure than using poverty, income, or PM 2.5 alone. The dataset also provides classification labels of SB 535 Disadvantaged Community, defined as "top 25% scoring areas from CalEnviroScreen along with other areas with high amounts of pollution and low populations" (OEHHA, 2018). Disadvantaged communities are identified for better targeting funding to the most needed to reduce greenhouse emission and improve health & economic outcomes.

- Forecasted number of chargers in San Francisco and Los Angeles

We used the results of charging infrastructure needs in San Francisco and Los Angeles from the two case studies by the International Council on Clean Transportation (ICCT, 2020; ICCT, 2021). The granularity of the forecast data is at zip code level for charger types including public level 2 chargers, DC fast chargers and workplace chargers.

3.2. Exploratory Data Analysis

3.2.1. EV Driver Persona

Before exploring the gaps of California EV charging infrastructure, we wanted to first understand the persona of the users - EV drivers. We used the results of the 2019 California Vehicle Survey to investigate the EV drivers[1] (NREL, 2019). Our hypotheses include: 1) EV drivers are mostly in the mid to high income levels, 2) EV drivers' choices of EV brands, 3) EV drivers mostly live in a single-family house, 4) EV drivers mostly use Level 2 chargers.

The data were provided either in binary (Yes/No) or quantitative (scale from 1-5 to represent the number of times). The main data transformations included creating pivot tables, which were easier for building Tableau visualizations. The chart can be presented:

- All public charger type questions
- All charging frequency problems

The first hypothesis about EV drivers' income levels was supported. The survey results showed that the majority of EV drivers are in the middle-income level (36k-90k USD per year). Only 7.5% are in the low-income level. Regarding the second hypothesis, we pulled out the EV brands that the drivers own. Tesla comes to the first, about 47.5%,

[1] We focused on residential users, not commercial users.

and Chevrolet and Nissan follow behind. The rest include over 10 brands, and each only stands for 1-5%, so we decided to merge them all as others.

For testing the third hypothesis about housing type, we used House Type with Sample ID Count to see which kind of house type EV drivers most live in. The results match with our hypothesis that most EV drivers live in single-unit dwelling. Maybe when living in a single house it is easier to have a home charger, so they could prefer buying an EV than apartment residents.

Next, we explored hypothesis 4 to see which kind of charger type users use often. We use SampleID Count and Charging Frequency Pivot to show the data. It turns out that drivers often use Level 2 (240V) free chargers as their first choice. It supports our hypothesis 4 that EV drivers most often use Level 2 charger type. But we should add a note that free level 2 is the most.

3.2.2. EV driver charging behaviors & pain points

Overview

Going beyond who those EV drivers are, in this section, we want to take a closer look at their charging related issues with a focus on charging availability. Specifically, we leveraged the 2019 California Vehicle Survey to examine how this issue is impacting EV drivers' charging behaviors and potential EV drivers' concerns. For example, how long they have to wait to charge their EV, whether different populations observe different numbers of chargers nearby, and if concerns about charger availability affect their vehicle usage. We created 3 visualizations below to vividly illustrate different facets of charging availability issues.

Waiting Time for Public Charger

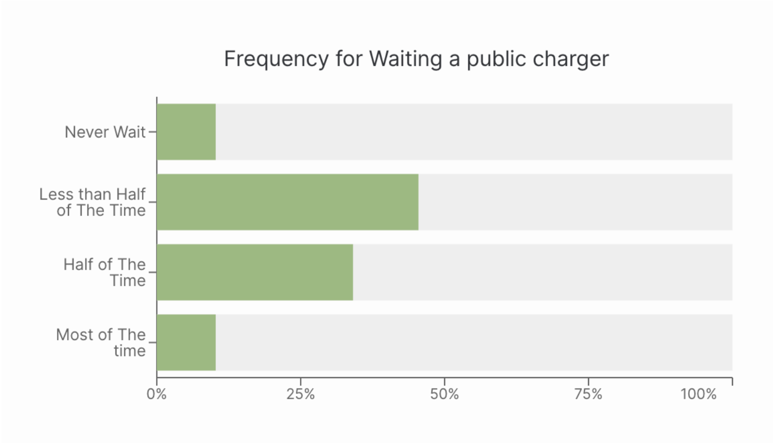


Figure 1: Frequency of Waiting for a Public Charger.

We wanted to understand whether EV drivers face a problem of always having to wait for public chargers. It’s bothersome and might cause great inconvenience for people in a rush. We looked into the survey and visualized respondents’ answers to the question “What’s your frequency of experience waiting for a public charger in the past month.” The bar chart above shows that only a few people reported to have never been in this situation. Over 50% of drivers have to wait for chargers; some even report that they have to wait most of the time.

After our hypothesis is confirmed, we try to add another variable to see if there’s any relationship between the driving distance to the public charger and waiting time. However, we found out the survey used two metrics to measure driving distance, miles and minutes, and the trend was not very significant. Besides that, it’s hard to visualize the result because most distance answers are 5,10,15. The numbers are selected, not random, so it’s hard to plot. We tried the plot chart, but it did not work. So we decided to simply show the results of the bar chart and iterated our idea. The iteration is provided in Figure18 in Appendix.

Number of observed nearby public charging stations by people of different income levels

Using a bar chart, we compared respondents' (regardless of EV ownership) reported number of EV charging spots in any of the parking facilities they frequent by their income level. Each bar is denoted by a percentage --- what is the percentage of people of X income level reported they observed Y charging stations. We put a specific income range for each income level, and a specific percentage indicated by each bar in the interactive tooltip for simplicity, considering the long legend and crowded texts made the graph look busy in previous iterations (see Figure14 and Figure15 in Appendix). Despite the different max/min percentage for each row, we consistently customize them to a range of 0%-60% for easier visual comparisons across panels. With the specific percentages hidden in the tooltip, the main goal of this chart is to show some clear trends: the higher the income level, the more charging stations people are likely to observe (in parking lots they frequent). Conversely, people who are lower in income are more likely to never observe charging stations in parking lots they frequent. This graph shows that economically disadvantaged communities may be again put into a more disadvantageous position in EV charging, who may face more inconvenience and barriers in EV adoption.

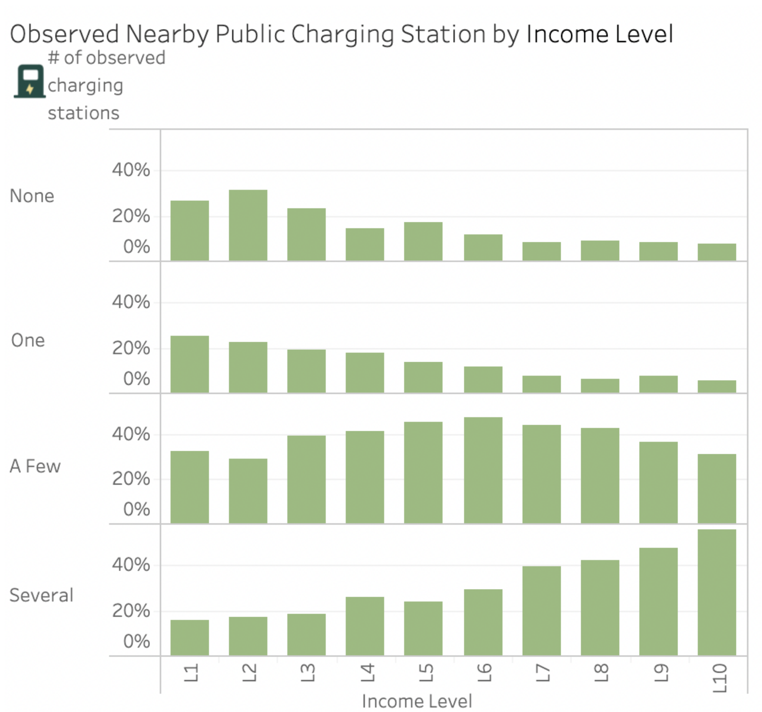


Figure 2 : Observed public charging stations by survey respondents' income level

Many EV drivers have hesitated using their EVs out of charger availability concerns

About 80% EV driver respondents expressed that they have experiences of intentionally reducing EV usage due to concerns about charging station availability. And as a result, they would take the trip by using a different vehicle or travel mode. We chose to use an isotype chart to represent such a percentage. Instead of highlighting 80 out of 100 icons which are hard to count, we filled 8 out of 10 driver icons with green color to represent EV drivers who suffer from such charging availability concerns, in order to make such proportions look more intuitive to readers. We hope to use an icon chart to make the message that “most EV drivers have experienced some degree of range anxiety and hesitate using their EV” more memorable and engaging to readers. Such concerns are not just concerns, but are real barriers in normal EV usage as well as future EV penetration.

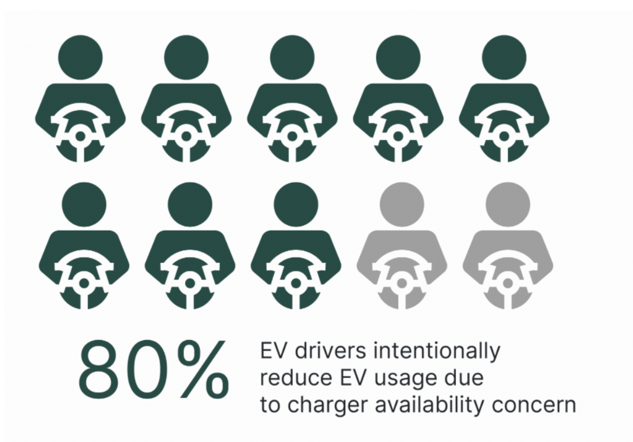


Figure 3: Charger availability concern and EV usage

3.2.3. Gaps in Charger Equity

Overview

In this section, we focus on evaluating the current charging infrastructure from an equity perspective. Including disadvantaged and minority communities is just as important for achieving widespread EV adoption. We combined California census data with public charger data to explore possible charger access disparity among different populations so as to identify populations that deserve more attention from policymakers. Several demographic attributes were examined, including housing type, population density, race and income based on past research and official reports.

Equity Analysis 1:Public charger accessibility across race

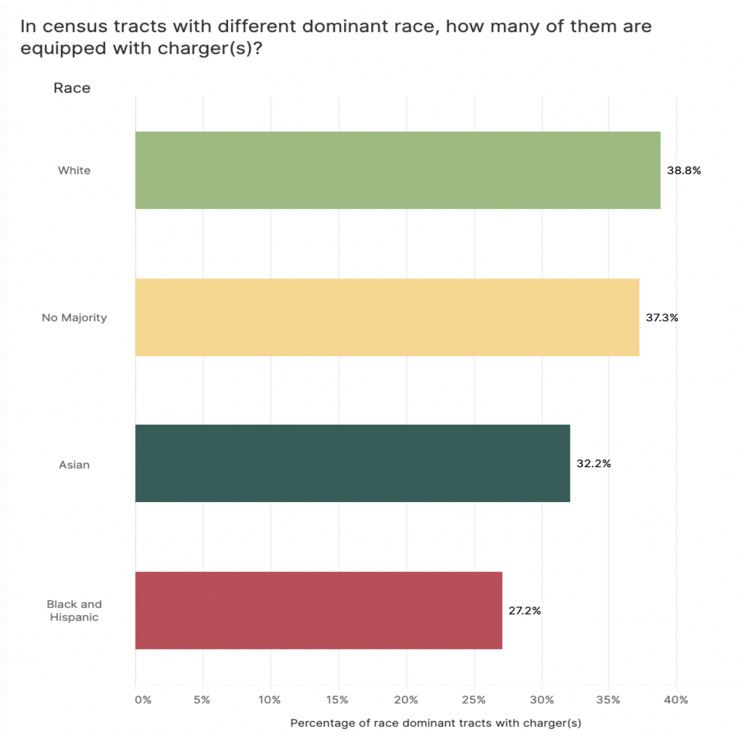


Figure 4: racial disparity in public charger access

In this graph, we used a bar chart to explore charger access disparities across races to test our hypothesis that racial minority groups have less access to public chargers than the white group. We used census tract data (smaller unit than zip code level) to have a more micro level understanding of the relationship between race & charger availability. Inspired by the work of Hsu and Fingerman (2020) who similarly used census block data in charger equity analysis, we first defined a majority race for each census tract as the race greater than 50% of the entire tract population, and assigned “no majority” to tracts without any dominant race. Since black dominants tracts are very few (see Table 1), and followed a similar trend line to hispanic majority groups in charger access suggested by past research (Hsu & Fingerman, 2021), we combined Hispanic and Black as a new group, where hispanic population and black population altogether take more than 50% of the total population. The final grouping results and racial distribution is presented in Table 2. 2723 out of all 8057 (34%) census tracts are equipped with at least one charge spot, and we calculate the percentage of each racial dominant tract that has charges. For example, as suggested in Table 3, 1183 out of 3051(38.8%) white dominant tracts are equipped with at least one public charger, which is higher than the rest racial minority groups.

Table 1: Original Grouping of Racial Dominant Groups

Racial Dominant Tract	Tract Count	Proportion out of all CA census tracts
Asian	360	4.50%
Black	66	0.80%
Hispanic	2485	30.80%
White	3051	37.90%
No majority	2161	26.80%

Table 2: Final Grouping of Racial Dominant Groups

Racial Dominant Tract	Tract Count	Proportion out of all CA census tracts
Asian	360	4.50%
Black and Hispanic	3042	37.80%
White	3051	37.90%
No majority	1604	20%

Table 3: Percentage charger-equipped tracts in each racial group

Final Grouping	Tract Count	Tracts with Charger(s)
Asian	360	116 (32.2%)
Black and Hispanic	3042	825 (27.2%)
White	3051	1183 (38.8%)
No Majority	1604	599 (37.3%)

Equity Analysis 2: CalEnviroScreen Score & SB 535 Disadvantaged Communities

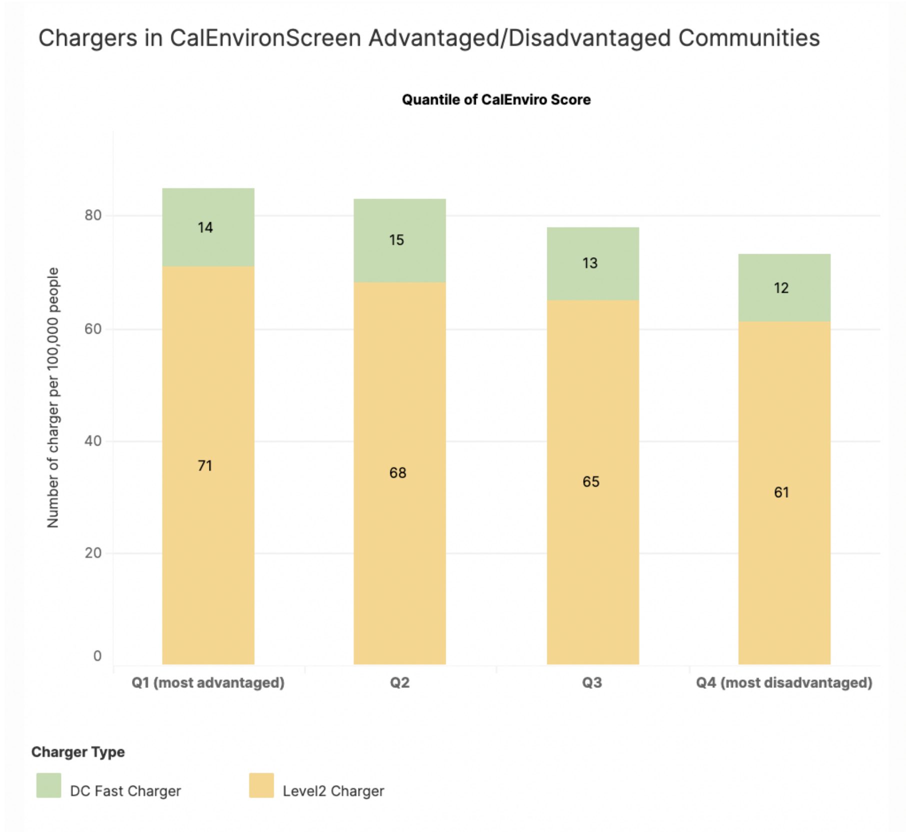


Figure 5: Charger distribution by census tract quartile of CalEnviroScreen Score

The California Office of Environmental Health Hazard Assessment used CalEnviroScreen Score to holistically measure communities' vulnerability to pollution effects, by taking into consideration environmental, health and socioeconomic factors such as poverty, employment and race. 2007 communities scoring top 25% from CalEnviroScreen Score were designated as SB 535 Disadvantaged communities (see Quartile 4 in the bar chart above). The CA Government is planning to specifically target those communities for investment to reduce pollution, while improving their health and economic outcomes (OEHHA, 2018). By grouping all communities by the quartile of CalEnviroScreen, the graph shows disadvantaged communities that are most likely to suffer from pollution, have the least amount of both Level 2 charger and DC fast charger per capita. More investments are indeed needed in building charging infrastructure for disadvantaged communities to advance an equitable as well as green transition to EV.

Equity Analysis 3: Charger disparity across different housing types

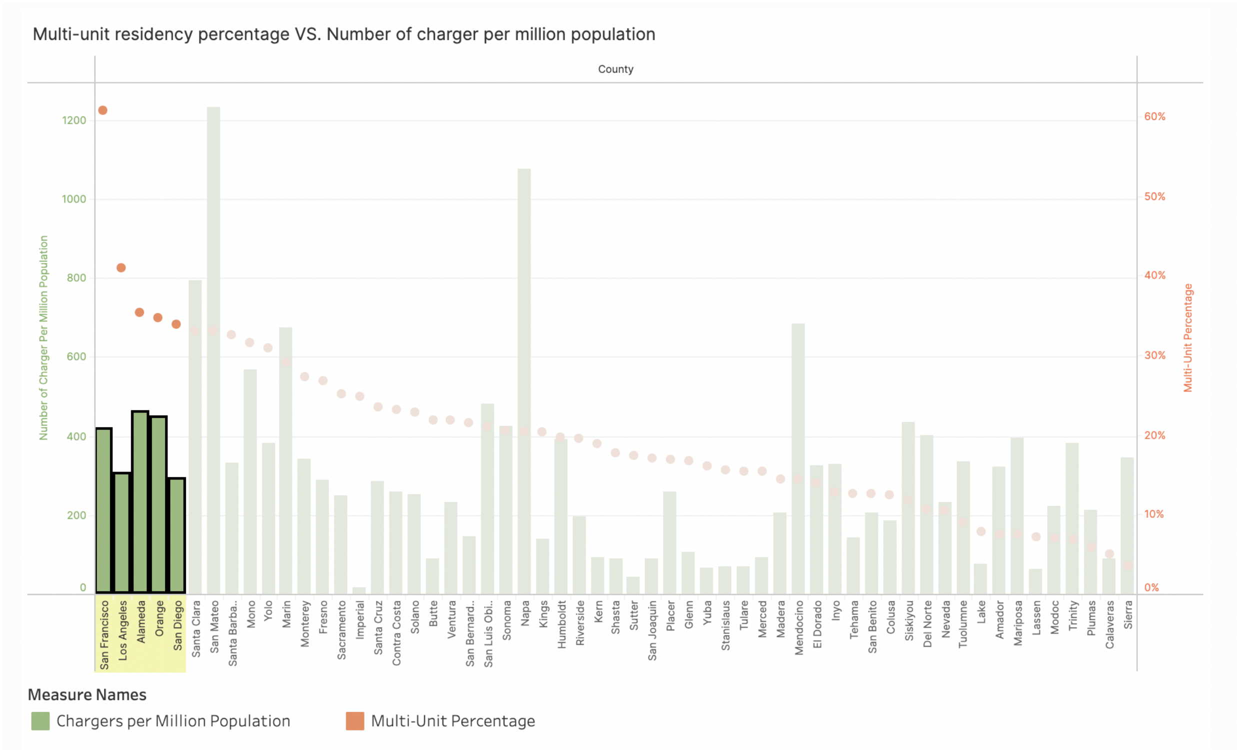


Figure 6: Multi-unit dwelling percentage VS number of charger per capita by county

EV drivers' need for public chargers can be heavily determined by their housing type. Multi-unit dwelling residents are less able to install home chargers and thus rely more on public chargers (UC Davis, 2018). Ideally, regions with more multi-unit dwelling should have greater public charging deployment; this trend is observed in some frontrunner countries of EV uptake such as the Netherlands (ICCT, 2020). To explore if such a trend also holds in California, we used a dual-axis chart to simultaneously plot the number of chargers per million population (green bars) and percentage of multi-unit residency (orange dots) at each county. Two axises are also colored in accordance with the measure's color to avoid confusion. The counties were ordered by multi-unit percentage from highest to lowest by default. Ideally, those counties on the very left with more than 30% residents living in multi-unit dwelling supposedly have most chargers per million population. However, they didn't show any advantage of charger access than other counties with a lower multi-unit percentage. We thus highlighted those counties by default to help readers quickly locate the most "problematic" regions. The old design can be found in Figure17 in Appendix, which violated some basic design principles (eg. mismatch of axis color and bars, counterintuitive measure for y axis) and was thus improved to the current version.

Equity Analysis 4: Charger Utilization Rate

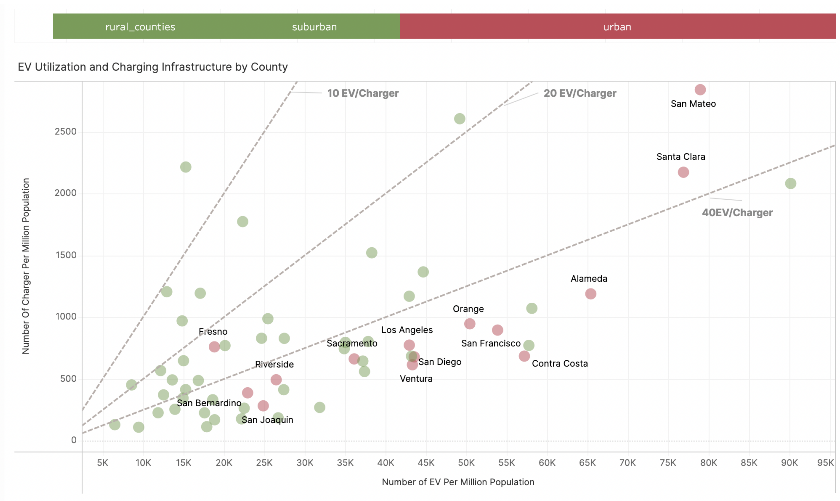


Figure 7: Default Layout

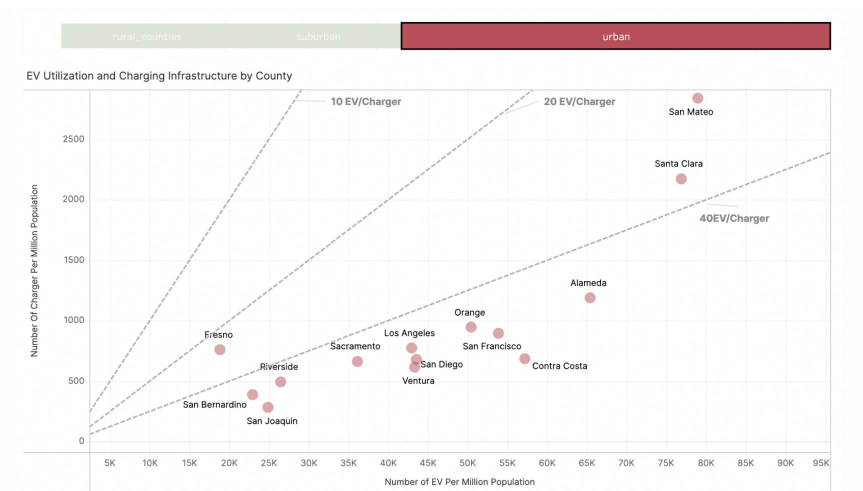


Figure 8: Use Interactive Filter Bar

In Figure7 and Figure8, we plotted each county's EV utilization rate against the number of public EV chargers per million population. Each county is represented by a dot and colored by whether it's an urban county (red) or not (green), defined by California State Association of Counties. We also drew three reference lines to suggest EV/charger ratio so readers can more quickly get a sense of the charger accessibility level. We added an interactive filter bar on the top so users can click between urban/rural regions to better explore disparity in charger accessibility. The chart suggests that most urban cities have an EV/charger ratio over 40:1. Non-urban counties are in general low in EV adoption, indicated by the X axis. It's worth noting though that the majority of those highly adopting non-urban counties still have very limited EV charger access per vehicle. Results here implied a barrier in more widespread EV adoption. There are no commensurate public chargers even if ev uptake surges as hoped.

3.2.4. Gaps in Chargers Scale

Overview

As mentioned in the project context, California currently has less than 5% of chargers that are needed by 2035. In this section, we chose two California cities as examples to demonstrate the gaps in chargers scale by comparing to what are needed in the near term (2025) and medium term (2030).

Gaps are big to have enough chargers for 2025 and 2030

From the two case study reports by ICCT (2020, 2021), we acquired the data of the numbers of chargers by 2019, and the forecasted needed chargers in the base case (without other policy interventions) in 2025 and 2030 for LA and SF. We found it can be useful to show the progress as a percentage of chargers needed in future years to highlight the gaps.

To show the percentages in an intuitive and straightforward way, we decided to use waffle charts in Figure9 below. The two charts on the left are compared to the demand in 2025 and 2030 for LA, and the two charts on the right are compared to the demand in 2025 and 2030 for SF. In early interactions, we used to have two charts on one row and have two rows for the four charts (see Figure19 in Appendix). However, we found it to be more organized to have LA on the left and SF on the right consistently. We also keep consistent color codings of red for LA and yellow for SF for the whole section. The percentage in the middle of each chart is bolded and color coded, corresponding to the information expressed in the waffle charts.

As we can see, both cities have very big gaps to reach the goals in the next 4 years and 9 years. Having LA and SF side by side, we can tell the gaps in LA are much larger than that in SF, which means LA has more pressure to grow faster in charger numbers to reach the goals.

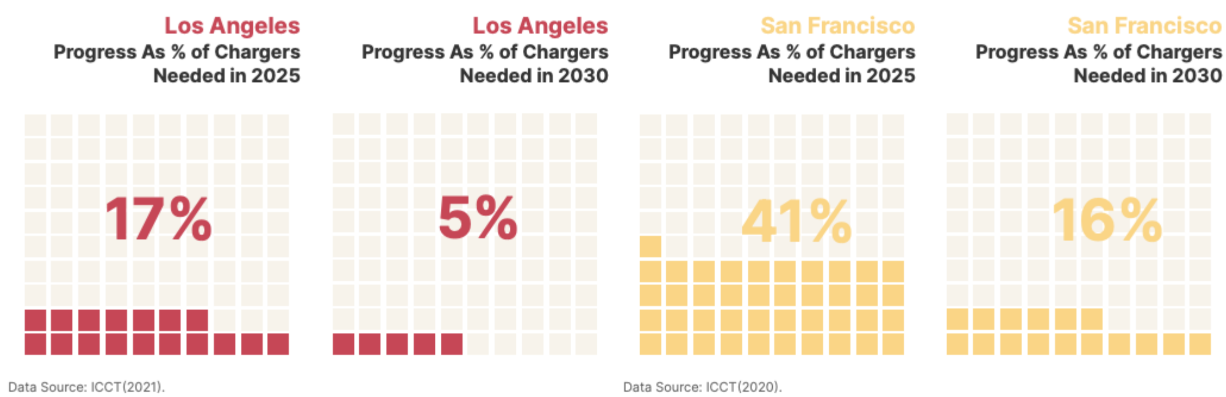


Figure 9: Progress as percentage of chargers needed in 2025 and 2030 for LA and SF.

Transport policies making the charging infrastructure goals more attainable

Since the gaps to having enough chargers for 2025 and 2030 are huge, we need to investigate ways to close the gaps quickly and realistically. One way of thinking is to accelerate the speed of building more chargers with more investment and active planning, but the speed and resources have limits. Even with all the resources committed, we may not achieve these ambitious goals in the given timeframe. What else can we do? The other way to work around the problem is to adjust the goals for 2025 and 2030 so that they are easier to achieve. The demand for 1.5 million chargers statewide in 2030 is an estimated number for the 100% ZEV sales goal. This estimation definitely has some room of adjustment if we change some parameters in the model. Keeping the efficiency and resources constant for building charging infrastructure, we can reduce the high demand of public chargers by enforcing other transport policies to reduce the demand of personal vehicle use in the city, such as encouraging sustainable trips, enforcing congestion pricing, vehicle-miles-traveled reduction and introduction of curbside chargers. There are many other policies and interventions that can help reduce the high demand of public chargers. For example, cities should make sure the building codes for new commercial and multi-family residential buildings will have EV charging infrastructure and their parking spaces are capable of accommodating future chargers. These will further reduce the number of public chargers needed in the city, and lower the expected amount for public investment.

We adopted the modeled data of forecasted number of chargers under the base case and intervention case from the two ICCT case study reports (ICCT, 2020 and 2021), categorized by charger types demand in 2025 and 2030 for both cities. The key message we would like to convey is that the impacts of policies are incredibly great, not only because they can substantially reduce the number of chargers needed, but also because their benefits are getting larger over time. These benefits are persistent and transformative to the whole transport system, making it more sustainable, integrated, clean, efficient and accessible for all. From a policy maker's standpoint, this is a really amazing future to look forward to.

We used stacked bar charts in Figure10 below to visualize the demand of chargers in 2030 for LA on the left and SF on the right. In the early iteration, we used red and yellow for the net demand of LA and SF (see Figure20 and Figure21 in Appendix), but it confused users that these variables were different. The orange part encoded the Net Demand after enforcing all the three selected policies, and the green part encoded the amount of demand reduced by different policies. If we click any policy in the legend, we will be able to see the amount of demand reduced by that policy. Looking at each charger type, we can also have a sense of the estimated demand of chargers in the base case (the length of the whole bar), a large share of demand is pushed down by the policies (green part) and the remaining Net Demand (orange part) is more realistic to achieve given the resource constraints. There is one issue with having the charts of LA and SF side by side. The scale on the y-axis is different from each other, but the website user may have an intention to compare the demand of LA with SF since they are on the same row. We also realized the problem in having the same scale for both y-axis: the bars for SF will become too tiny to view. There are also benefits in having the two cities side by side so that the flow of reading in this section can be more smooth. With these considerations, we could not come up with a perfect solution. Our current solution is to have more white space between the charts, so hopefully we can create some visual inconnection. The alternative solution can be having a line striking in between the two charts, or having a frame for each chart, so that visually they are more separate.

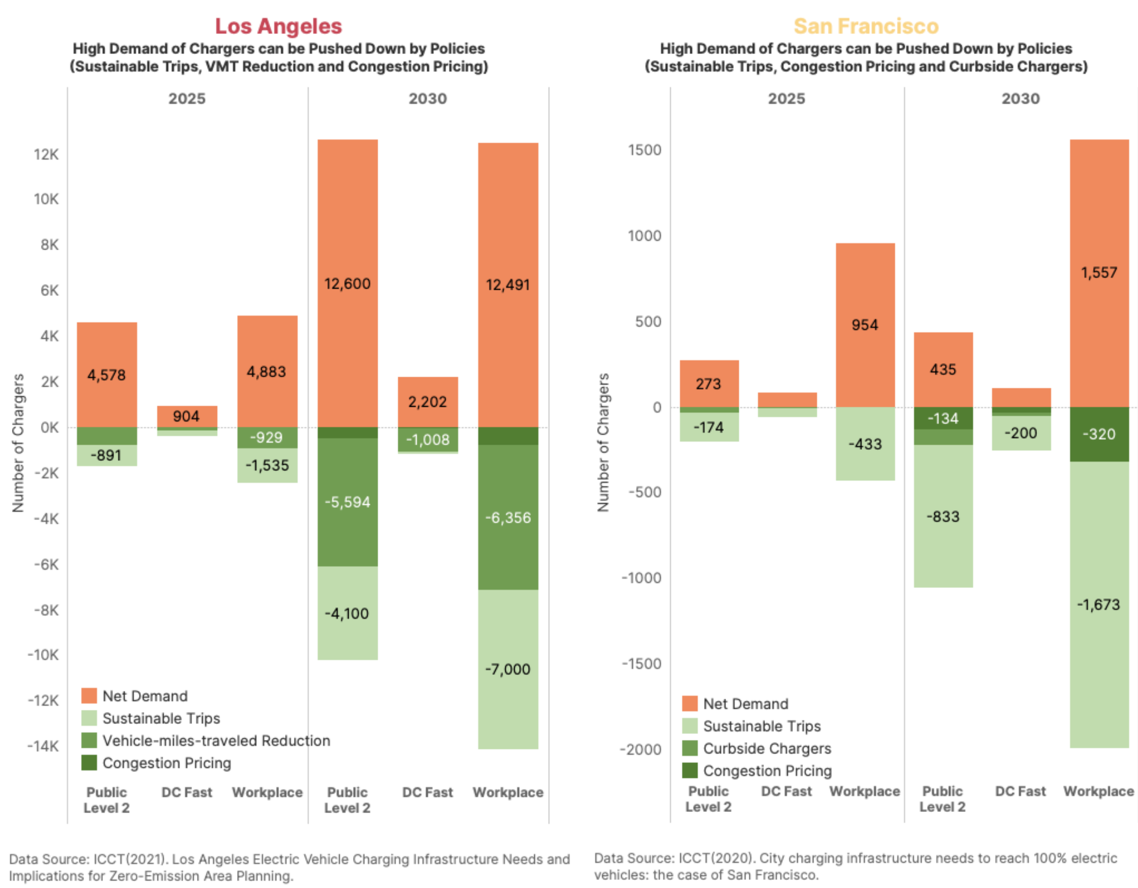


Figure 10: High demand for chargers can be pushed down by policies in both cities.

Mapping the gaps in the base case and the policy intervention case

To visualize the locations and magnitude of charger gaps in the two cities, we used two choropleth maps (see Figure11). We used inputs of forecasted demand in 2030 from the same models as the previous two charts (ICCT, 2020 and 2021). We used the current number of chargers acquired from the alternative fuels stations database (NREL, 2021). The gaps were calculated as the difference between the forecasted count and the current count of specific types of chargers aggregated by zip code level. The filter of Base Scenario and Intervention Scenario can help users visualize the substantial impacts of implementing policy interventions. The filter of Charger Type can show the count of chargers in gaps in more details. Due to the lack of data of Workplace chargers in the NREL database, we didn't include it in the Charger Type filter. It is also worth noting that the forecasting study for San Francisco was finalized in 2020. While the growth pace of new chargers is noticeably fast (by May 2021, the number of chargers almost doubled since we first requested data from NREL database in November 2020), a number of neighborhoods in San Francisco have already reached or even surpassed the forecasted demand of chargers in 2030. In these neighborhoods, we marked the gaps as 0. Lastly, we also keep the color coding for Los Angeles and San Francisco consistent with the last two charts to improve the readability of the storyline in this section.

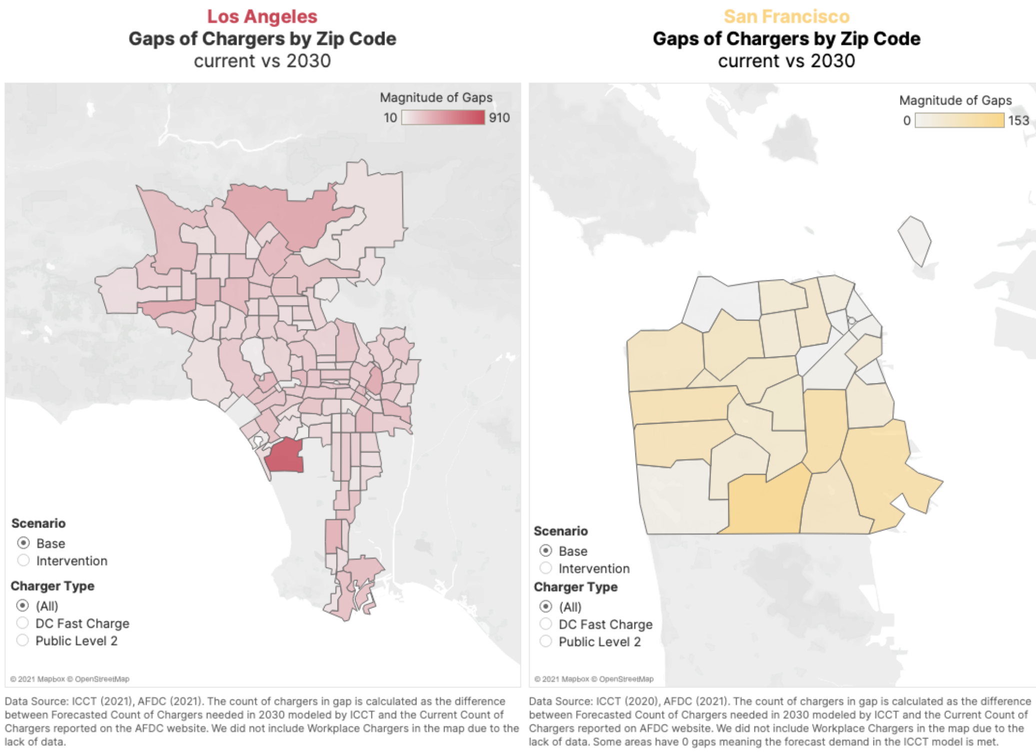


Figure 11: Gaps of Chargers by Zip Code, comparing current count and 2030 forecasted demand of chargers.

3.3. Information Visualization Heuristic Evaluation

After we had the first version of the website, we scheduled an one hour Heuristic Evaluation interview with Professor Marti Hearst, who is an expert on information visualization. We managed to integrate several important feedbacks for the infographics, charts and the overall website design. We used the visualization heuristic guideline below to improve our design.

- H1: Use correct & Meaningful visual encoding
  We chose our data type and made meaningful encoding. We are fine for H1.
- H2: Supports key visual insights
  Most of the charts support the hypothesis or insights. For example, the data value supports that higher income levels get more EV chargers. We also highlight key information, like the urban city in the capital city gap chart. But this

helps us figure out some charts do not show trends as we expected, for example, the unemployed chart has too much noise.

- H3: Uses principles and Organization and Consistency
  We label out details using a tooltip or legend. This evaluation helps us realize our first color pattern is not consistent. Using all green color patterns makes the website harmonious, but in some charts, the level 2 charger is light green while another chart is red. We modify the color pattern: all “normal data” as light green, the "continual value" from light to dark green, "discontinuous category" as a different color (red, yellow, orange), and all "text title" as dark green.
- H4: Presents Information Honestly
  In some icon charts, we omit decimals to make it more user-friendly and accessible for icons, but the trend and value are presented honestly.
- H5: Successfully Communicates
  We use narrative and tell a story for the content. Viewers can follow our story to take a journey about it.

4. Information Visualization Website

Our visualization website can be found in this link:
https://alison626k.github.io/MIMS21-GEAR/

4.1. Overview of Storyline

Our storyline of the visualization website follows this structure:

- **Landing page**: What ZEV and chargers goals do we have in California? Why is it important to understand the gaps of EV charging infrastructure?
- **EV Charger Types**: What are the three EV charger types? What are their characteristics and use cases?
- **EV Users Persona**: What kind of people own an EV in California (income level, housing type, vehicle brand decision, and charger type decision)?
- **EV Charging Issues**: What are the three prominent issues faced by EV drivers in charging (wait time, access disparity by income, charging convenience)?

- **Gaps in Equity**: What are the gaps in equity using different angles (race, income, housing type, charger utilization rate, disadvantaged communities vs others)?
- **Gaps in Scale**: What are the gaps in scale for two example cities, Los Angeles and San Francisco? How can policies help close the gaps faster/easier?
- **A Future to Look Forward to**

The first version of the storyline did not include EV charger types, but from the usability test, we found out some people would ask about the differences of charger types. Since potential EV drivers are also our target audience, we decided to add an overview of EV charger types at the beginning. Another adjustment was to bring the EV users persona forward in order to make the storyline flow more smoothly. The flow is as “Users, Problems, Gaps and Future”.

4.2. Usability Test

Our ideal target audiences of the website are those who can make policy decisions and recommendations in the e-mobility transition, or those driving the advocacy for clean air and sustainable transportation. We also appreciate the feedback from current and potential EV users in California. While their awareness and opinions about EVs and EV charging are critical in the scale-up of EV markets, our website is intended to reflect their issues at the micro-level and also the current gaps at the macro-level. It is beneficial to validate the main messages with EV users in our website iterations.

With these considerations, we invite the users from the following categories to participate in our usability test:

- Policy makers, policy advocates and researchers, who can influence decision making in clean transportation, climate change and clean energy
- Current EV owners in California, who are current users of EV charging infrastructure
- Potential EV buyers in California, including gasoline vehicle drivers who may switch to EV in the future

The usability test recruiting strategy is to start from the second and third category in the prototype phase and then invite target users from the policy maker category for their feedback when we have a more advanced version of the website. To implement this strategy, we sample from people we know that can meet the criteria but outside of MIMS to avoid convenience bias as much as we can. Therefore, we identified one EV owner and one gasoline vehicle driver who is interested in EV for our usability test.

We aimed to evaluate whether our visualization design is effective in conveying the messages in the sections of EV charging overview, charger gaps in equity and charger gaps in scale for the overall topic of infrastructure accessibility. Our test includes three specific tasks to test each section. The detailed questions and website prototype screenshots are provided in Appendix.

We planned to measure two key factors: 1) Where do users stop scrolling the webpage or spend extended time figuring out one visualization or narrative, and how much time do they spend in the confusing part. 2) What questions did they ask about our website or visualizations. The first measurement can help us mark the visualization or narrative that users get confused about, and the second measurement can help us identify what information is missing or not clear in our website.

Task 1: In the sections of EV Driver Persona and EV Driver Issue, what are the problems that EV drivers are currently facing?

> **Goal:** We are going to test if users can understand the first two sections, figuring out the EV driver issue.
>
> **Result:** The interviewees performed well for the EV Driver Persona part and did not get any confusion. However, both interviewees stopped at the EV issue section for the Waiting Time Chart, because they did not understand what the chart was talking about. They spent 1 minute on it until they figured it out through the text. They both got a little confused about the third chart of charging inconvenience and misunderstood the key message. They understood the narratives but could not relate to the chart. One of the interviewees asked about

the direction of the causation in the charging inconvenience chart. He suspected it can be interpreted as the other way around: some groups prefer gasoline cars so they avoided using EVs.

Task 2: In the EV Charger Equity section, which gap do you have most impressions on? After reading this section, which housing type and income level has the highest and lowest numbers of chargers?

Goal: We are going to test if users can recall the key messages of any specific charts, and if the chart and text can successfully show the EV charger gaps.

Result : Interviewees spent more time reading the four charts. They stopped at every chart and were not sure about meaning. After reading it carefully, the county's chart is clear for them to understand that there is a gap. One interviewee made one error about how to read the unemployment chart. He cannot tell the trend of unemployment and charger numbers. One interviewee expressed that income level and unemployment seemed a little similar because they are all related to spending power.

Task 3: After reading the EV Charger Scale section, can you tell us about the impacts of implementing policy interventions on the two cities' charging demand?

Goal: We are going to test if users can understand the information of the two bar charts.

Results: One interviewee spent almost 10 minutes reading the bar chart. He did not understand what the minus number means and the relationship with the positive number. After explanation, he can get the idea. Both interviewees got at least one error where they did not get the negative section of the bar chart.

After the interview, we learned that communicating too many concepts in one chart will confuse users. Presenting over 3 variables in one chart makes it look fancy, but hard to understand the main insight. Users have to think through and spend time on it to figure out the content. So less is more. Clear but easy to understand is better than complicated. Besides that, the text will help a lot. Some of our charts lack text when we conduct the interview, and users have to spend more time on it. The precise content can help users to understand each main insight. Using all green color patterns makes the website harmonious, but users cannot tell the difference if dark and light green represent different categories or a different amount.

We summarized the following feedbacks to be included in our website:

- EV persona part is fine. Will change the charger type chart to the icon. Make the context more related to “Persona”.
- Change the charging convenience chart(chart 3). Users are unable to understand the relationship between having gasoline and avoiding EV usage, so we plan to remove the gasoline preference category, just focus on talking about why drivers avoid using EVs after they bought them.
- Change the waiting time chart. The plot is confusing and users do not understand it. We found out if we want to deliver the message about waiting for a charger, we just have to present the percentage or waiting time. Another variable is too noisy.
- The EV gap on different metrics works fine. The unemployment is not significant, so we will remove that part. We will add a reference line on the county chart and change the color, so it will be more user-friendly. The income level and race chart are too complicated, so we will reduce the variables and focus on the main insight.
- The policy city bar is okay. But it lacks some explanation to help users get into it more quickly, so we plan to add more precise deception on that text and labels.
- Change the color pattern. Make dark green and light green represent different amounts, and find a new color for different categories.

4.3. Website Development Process

We used React and created React Apps to build the website and host it on the Github page. Because our website only contains front-end sites, we chose React as a common front-end development framework. Our visualization contents are original, and the website is building on top of an open-source template. Our website has gone through three iterations, from wireframe, usability test to finalization.

Based on the storyline, we made the scratch for the website wireframe on paper (see Figure22). We want to avoid making the website a boring collection with charts, so we plan to tell the concept as a narrative story. This is the first version of the website. The second iteration is before usability testing. We make a draft with charts on Notion as a prototype, and take it to conduct usability testing. After the test, we finally go on developing the website.

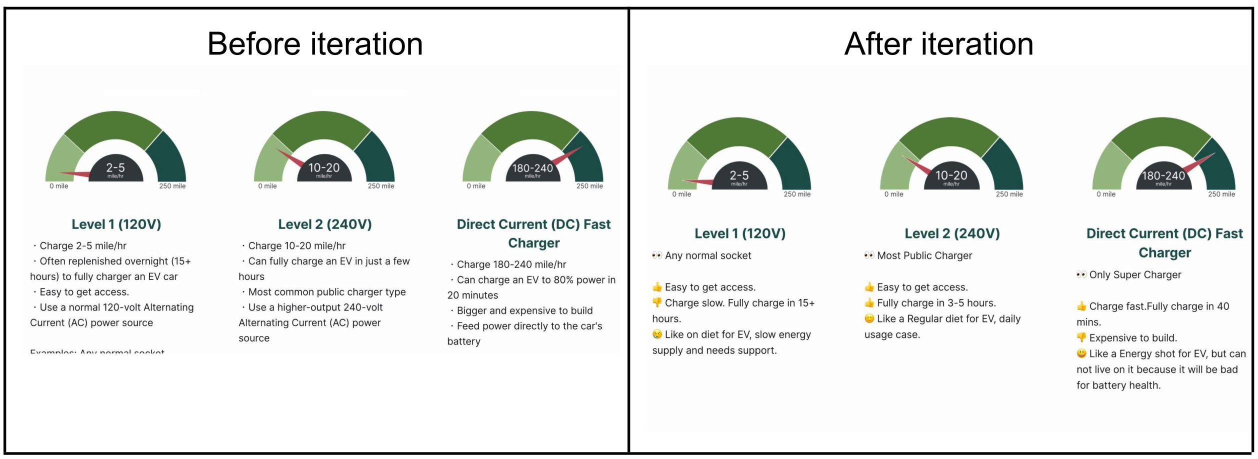


Figure 12: Example iteration for EV charter type sections: icon adding, recategorize description

Our website includes 5 main sections based on our storyline: 1) Background storytelling and animation. We try to use some animation and fade in sections to make it more accessible. 2) EV charger type, driver persona and issue. This section contains both charts and icons in the prototype version, but when we realize that the rest sections will include some complicated charts, we decided that Persona will only be icons and numbers so it's easier to read and harmonious. 3) EV charger equity gap. We plan to deliver 5 variables in the first place, but then some charts are deleted because it does not communicate any significant trends or differences. Some charts include too many axis contents, so we put it as the whole page so that users can read them. The rest of the sections, 4) City Compare and 5) NLP do not change a lot, so it speeds up our development.

When we embed the Tableau charts to the webpage, we struggle to get rid of the scrolling bars and the toolbar of Tableau. We found the solution to removing the scrolling bar by setting Tableau sheets to “entire view”, but we failed to remove the toolbar defaulted by Tableau Public. Another problem that we tried to solve is that the D3 library chart we used should display animation when loading, but we failed to conduct the intersection observer. So we can only use fade in fade out as default. We make the website responsive, but some charts are unable to do so. These are some issues due to our technical limitation and would like to solve it afterward.

The last thing to iterate for the website is our color pattern. As mentioned in *3.3 Information Visualization Heuristic Evaluation*, the first idea of our color pattern is using all green so it stays harmonious, and only red and yellow for highlight. But then we realize we need more color to separate title, icons, data, and category. The modified version is: all “normal data” as light green, the "continual value" from light to dark green, "discontinuous category" as a different color (red, yellow, orange), and all "text title" as dark green. The color pattern from first to infinalization is attached in Figure23 in the Appendix.

5. Sentiment Analysis on Charging Experiences

5.1. Data

We leveraged the datasets from Georgia Tech researchers Ha et. al.’s study (Ha et al., 2021) and their GitHub repository (Asensio Lab, 2021). It is a publicly available dataset that contains 10652 reviews on electric vehicle charging stations in the US. Each user review was labeled as one (or more) of eight main aspect categories by expert annotators and crowd-sourcing. For the scope of our study, we will only focus on aspect-based sentiment classification and will directly utilize those pre-labelled aspects as the ground truth. Also, in this study, the authors predefined eight aspect categories as Functionality, Range anxiety, Availability, Cost, User interactions, Location, Service time and Dealership (Ha, 2021).

After thoughtful consideration, we decided to narrow down our scope to only five categories, excluding user interactions, dealership and location. The reasons behind are as follows: firstly, these three aspects are not so aligned with the scope for our capstone project. Secondly, reviews containing these three categories are extremely sparse, which means very few electric vehicle owners mentioned user interactions, dealership or location in their reviews. Lastly, during our annotation process, we found that most of the sentiment expressed in these three categories are neutral (86%, 84%,73% respectively), meaning that the drivers were merely sharing information on these three categories. Looking into these three categories might not add much to our existing knowledge compared to the other aspects such as functionality and availability. Moreover, it might be just sufficient enough to use a majority vote model to predict the sentiment in these three categories.

5.2. Methodology

We use pre-trained GloVe word embeddings (Stanford University, 2014) to featurize the words expressed in each review. For each word, we take the corresponding 300 dimensional word embedding from GloVe. For words that are not in the GloVe embedding corpus, we randomly initialize the word embeddings and update them during the training process. We chose three window sizes: 3,4,5 respectively, 50 filters for each region size and 1-max pooling. We also used dropout during the training process to prevent overfitting. The structure of this CNN model is illustrated in Figure13. The diagram was inspired by the study of Zhang & Wallace (2017).

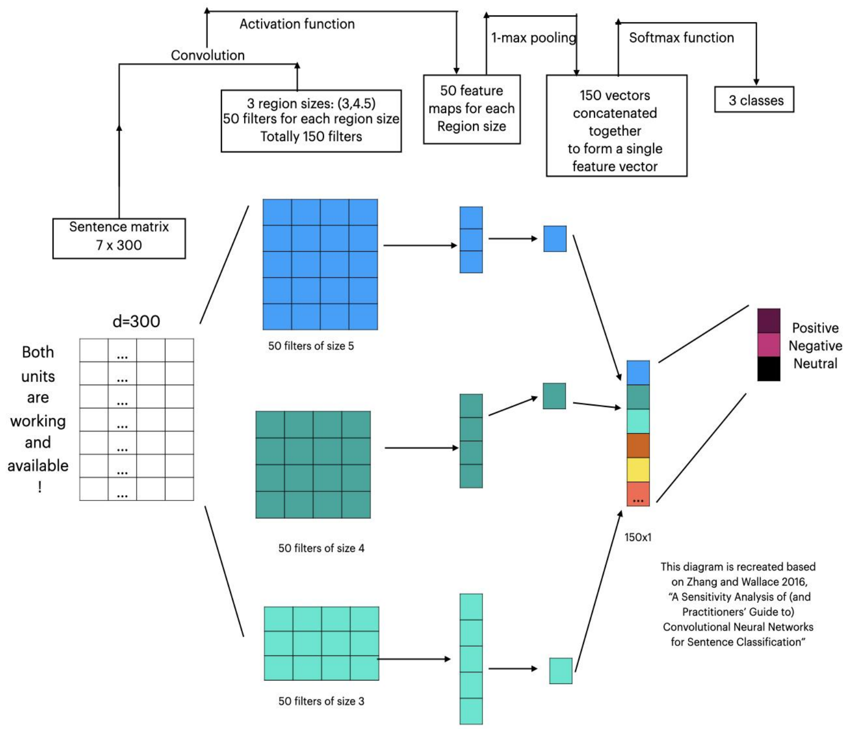


Figure 13: Illustration of CNN model.

In terms of evaluation, we calculated the precision and recall score for our CNN and LSTM model output. Since we have three labels(positive, negative and neutral) in our output space, we calculated the precision and recall score metrics for each label.
For the logistic regression model, we used a simple bag-of-words transformation as our model input and compared the model performance to the other models. We later used this logistic regression model as our baseline model for comparison.

5.3. Results and conclusion

We were able to identify some key concerns faced by EV drivers:
Here are some of the key findings from NLP section:

- Nearly half (more than 40%) of the EV drivers have faced negative experiences during charging.

- 57% of charging station reviews showing electric vehicles being ICEd.(EV spots taken by gas car)
- Functionality, availability and service time are the most frequently mentioned topics.
- 50% negative reviews regarding availability, 22% for service time, 23% for functionality.
- Here we also include our model performance here, evaluated by precision, recall and F Score.

Table 4: Performance of Sentiment Classification Models.

Functionality, LSTM

Label	Precision	Recall	Fscore
Negative	80%	65%	0.72
Positive	56%	79%	0.65
Neutral	56%	32%	0.41

Functionality, CNN

Label	Precision	Recall	Fscore
Negative	64%	43%	0.52
Positive	65%	57%	0.61
Neutral	67%	29%	0.40

Functionality, Logistic Regression

Label	Precision	Recall	Fscore
Negative	60%	44%	0.51
Positive	63%	75%	0.68
Neutral	46%	45%	0.45

Service Time, LSTM

Label	Precision	Recall	Fscore
Negative	83%	61%	0.70

Positive	82%	34%	0.48
Neutral	73%	94%	0.82

Service Time, CNN

Label	Precision	Recall	Fscore
Negative	88%	77%	0.82
Positive	87%	70%	0.77
Neutral	87%	90%	0.89

Service Time, Logistic Regression

Label	Precision	Recall	Fscore
Negative	95%	67%	0.79
Positive	80%	56%	0.67
Neutral	81%	98%	0.89

Availability, LSTM

Label	Precision	Recall	Fscore
Negative	77%	74%	0.70
Positive	61%	34%	0.44
Neutral	0%	0%	0.00

Availability, CNN

Label	Precision	Recall	Fscore
Negative	74%	52%	0.61
Positive	90%	19%	0.32
Neutral	0%	0%	0.00

Availability, Logistic Regression

Label	Precision	Recall	Fscore
Negative	60%	76%	0.67

Positive	59%	51%	0.55
Neutral	40%	15%	0.22

6. Conclusion and Future Work

Our capstone project starts with a user-centered approach to understand the status quo of EV drivers and their issues and experiences with the data of survey results and online reviews. This helps us not only assess what has been done or to be done, but actually sketch the true picture of how they are impacting EV drivers' charging experiences and how they will possibly affect would-be EV adopters' trust and interest in EVs in a positive or negative way.

Since project scoping, we have increasingly realized *equity* for the access to charging stations is so important. It is recognized by the California Government as the number one principle in every aspect of EV infrastructure planning. We are looking at more than an ambitious EV sales goal to reach or a gap to fix in the number of chargers. Planning infrastructure for equitable access is an impactful development agenda that not only brings about environmental benefits and mobility improvement but also job opportunities and social justice. We also realized that many factors that we independently studied in our equity analysis might correlate with each other, and the marginalized communities could be disadvantaged in multiple ways at the same time. For example people in less-privileged socio-economic status may also be more likely to live in rented apartments or even in cars, and thus face double barriers in charger access. The different angles we used to look at equity can hopefully help policy makers better target the resources to the people in need, so no community is left behind.

In terms of *scale* of charging stations, we are very excited to see the quickly rising number of newly built charging stations in California. We started the project in late 2020 when there were around 7000 public and private charging stations recorded in the NREL database, and it almost doubled the number when we requested the data from the API again lately. We are also confident about the political will and the momentum of the private sector in advancing the ZEV transition. The EV market share also grows quickly to 9% in the first quarter in 2021, almost unaffected during the COVID pandemics. When assessing the gaps in charger scale, we want to make one important point: it is not the more chargers the better. While there are many benefits in electrifying all vehicles and making sure we have enough charging stations to support people's mobility in the years to come, we do not expect e-mobility alone is the solution to a sustainable future. We need a multi-modal, integrated, efficient, equitable and accessible transport system. The rise of EVs should not undermine the role of transit in urban development. For this reason, we hope to suggest policy makers use very thoughtful planning for infrastructure and use a variety of policies to reduce the number of chargers to build. It is also critical to save public resources for those in most need.

Regarding limitations of the project, we find that the dynamic nature of charger data may affect the forecasting of future demand and our evaluation of gaps. With the development and implementation of new policies, the forecasted need for chargers may also change accordingly. In this light, it would be especially helpful to improve the gap explorer to automatically take and process the most up-to-date data in real time. Also, our website design can benefit from more iterations of improvements and including policy makers for our user testing in the future.

Future work is much needed to support more precise and powerful measurement of the charger gaps. For equity analysis, we might consider other transportation related factors such as access to highways and public transit that might cause or deepen charger access disparities across different communities. Moreover, future developers or researchers can further examine the interplay between different transportation, demographic or socio-economic factors maybe via more rigid statistical model building. If time allows, we are also interested in building a dynamic gap explorer with streaming analytics, so whenever there's an update in charger data or forecasted charger needs, policymakers can always safely refer to the latest results to track progress and make evidence-based decisions. Finally, since the current project only focuses on EV charging infrastructure, future work could incorporate hydrogen fuel cell fueling stations into the gap assessment to have a better picture of ZEV infrastructure in California.

7. Team Contributions

Table 5: Team Contributions.

Team Member	Contributions
Alison Kuo	Website Development and UI Design, including building design guidelines, prototyping and final implementation Data exploration and building visualizations, mainly on EV driver persona and EV charging issues Usability Testing Project Management Report Writing
Wenqi Luo	Data collection, processing and exploration for EV driver behavior issues, EV market share & registration and equity analysis (30%) Build visualizations to show on website and write section associated texts Literature review Natural language processing, mainly literature review, human annotations of review sentiment and result analysis Report writing
Yin Qiu	Literature review and project scoping Building visualizations and storyline, mainly on gaps analysis of EV charger scale and EV charger types Data collection, processing and exploration Report writing Natural Language Processing, mainly on literature review and data annotation Usability testing
Chuhan Wang	Natural Language Processing, mainly on literature review, data collection and preprocessing, model training and tuning, result

	evaluation and analysis Literature review Report writing

8. References

Asensio Lab. (2021). *transformer EV topic classification*. https://github.com/asensio-lab/transformer-EV-topic-classification

CARB. (2020). *CA GHG inventory 2000-2018*. https://ww2.arb.ca.gov/ghg-inventory-data

CARB. (2020). *Draft 2020 Mobile Source Strategy*. https://ww2.arb.ca.gov/sites/default/files/2020-11/Draft_2020_Mobile_Source_Strategy.pdf

CA State Government. (2020, 09). *CA State Government Executive Order N-79-20*. https://www.gov.ca.gov/2020/09/23/governor-newsom-announces-california-will-phase-out-gasoline-powered-cars-drastically-reduce-demand-for-fossil-fuel-in-californias-fight-against-climate-change/

CA State Government. (2021). *California ZEV Market Development Strategy*. https://static.business.ca.gov/wp-content/uploads/2021/02/ZEV_Strategy_Feb2021.pdf

CA State Government. (2021). *2021-2022 State Budget*. https://www.gov.ca.gov/2021/01/08/governor-newsom-proposes-2021-22-state-budget/

CEC. (2021). *Assembly Bill 2127 EV Charging Infrastructure Assessment*. https://efiling.energy.ca.gov/getdocument.aspx?tn=236237

CEC. (2021). *EV Market Share Statistics*. https://www.energy.ca.gov/data-reports/energy-insights/zero-emission-vehicle-and-charger-statistics

Ha, S., Marchetto, D., Dharur, S., & Asensio, O. (2021). Topic classification of electric vehicle consumer experiences with transformer-based deep learning. *Patterns*, *2*(2). https://www.sciencedirect.com/science/article/pii/S2666389920302658

Hsu, C.-W., & Fingerman, K. (2021, 1). Public electric vehicle charger access disparities across race and income in California. *Transport Policy*, *100*, 59-67. https://www.sciencedirect.com/science/article/pii/S0967070X20309021

ICCT. (2020). Charging infrastructure in cities: Metrics for evaluating future needs. https://theicct.org/publications/EV_charging_metrics_aug2020

ICCT. (2020). City charging infrastructure needs to reach 100% electric vehicles: The case of San Francisco. https://theicct.org/publications/sf-ev-charging-infra-oct2020

ICCT. (2021). Los Angeles Electric Vehicle Charging Infrastructure Needs and Implications for Zero-Emission Area Planning. https://theicct.org/publications/LA-charging-infra-feb2021

NREL. (2019). *2019 California Vehicle Survey*. https://www.nrel.gov/transportation/secure-transportation-data/tsdc-2019-california-vehicle-survey.html

NREL. (2021). Alternative Fuels Stations API. https://developer.nrel.gov/docs/transportation/alt-fuel-stations-v1/

OEHHA. (2018). *CalEnviroScreen 3.0*. https://oehha.ca.gov/calenviroscreen/report/calenviroscreen-30

Stanford University. (2014). *GloVe: Global Vectors for Word Representation*. https://nlp.stanford.edu/projects/glove/

UC Davis. (2018). Observed Charging Rates in California. https://escholarship.org/uc/item/2038613r

Union of Concerned Scientists. (2021). *Gov. Newsom's Proposal to Invest $1 Billion in ZEV Infrastructure is a Smart Proposition*. https://blog.ucsusa.org/jason-barbose/gov-newsoms-proposal-to-invest-1-billion-in-zev-infrastructure-is-a-smart-proposition

Zhang, Y., & Wallace, B. (2017, 12 1). A Sensitivity Analysis of (and Practitioners' Guide to) Convolutional Neural Networks for Sentence Classification. *Proceedings of the The 8th International Joint Conference on Natural Language Processing*, 253-263. https://www.aclweb.org/anthology/I17-1026.pdf

Appendix

1. Deliverables links

   Website: https://alison626k.github.io/MIMS21-GEAR/

   MIMS webpage:https://www.ischool.berkeley.edu/projects/2020/gear

2. Usability Testing Guide

Introduction

We are trying to get feedback on our website for a Berkeley course assignment about information visualization. We're grateful that you can take the time to help us improve the user experience. Please think aloud and tell us your thoughts as you go through this usability test. Before we start, we would kindly have your consent and signature on the informed consent form. All the information collected in this usability test interview will be used and kept only for the stated purpose. There are no right or wrong answers for the questions. Please use your best judgement to provide your answers and feel free to be as critical as you would like.
We're testing the website, NOT you, so anytime you feel something odd about it - feel free to share, be as critical as you like.
Do you have any questions before we begin?
Let's begin with some background questions.

Test: Background Questions

We asked the selected participants a set of pre-test questions which provided us further understanding of their relation to EV :

1. Do you own an electric vehicle?
   1. If so, how often do you drive it?
   2. If not, are you planning to get one or do you feel interested in getting access to it? Have you driven an electric vehicle before?
2. What is the most common transport mode you use?

Moderator Script

We are now going to see the website wireframe. It's just a draft and sketch, so please ignore the layout. Again, feel free to point out anything you feel odd about as you're going through it.

Questions

Task 1: Looking at the EV driver persona and EV driver issue, what are the problems of EV drivers currently facing?

- What do you first see in the Persona section?
- What do you first see in the EV driver issue section?
- What information do you get from the last “charger convenience “chart?
- After reading these two sections, can you tell us what problems EV drivers are currently facing?

  Note for the interviewer: We’re looking for if users can understand the first two sections, figuring out the EV driver issue. They should answer: 1) Ev drivers often wait in line for a public charger, 2) higher income level groups get more public chargers in surroundings, and 3)Charging is not convenient enough so some EV drivers regard buying EVs and still prefer gasoline vehicles.

Task 2: Now let's scroll to the next section. In the EV Charger Gap section, which gap do you have most impressions on?

- What do you first see in the EV Charger Gap section?
- Can you tell us what you learn from this section?
- After reading this section, which capital, housing type and income level has the highest and lowest numbers of chargers?

  Note for the interviewer: We’re looking for if users can memorize any specific charts, and also if the chart and text can successfully show up the EV charger gaps.

Task 3:After reading the last section, can you tell us about the impacts of implementing policy interventions on the two cities' charging demand?

- Do you find out anything different or in common about the charger situations in LA and SF?
- What are your main takeaways about the last section?

Note for the interviewer: We're looking for if users can understand the information of the two bar charts. They should be able to answer the implementation of the policy and reduce the demand and gaps of public chargers.

Thanks so much, with that we've concluded our testing process!
Do you have any last comments/closing feedback on this website? Anything is helpful!
Thank you so much for helping us to investigate the usage of the project. We really appreciate all your time and input!

3. Early Iterations of Visualizations

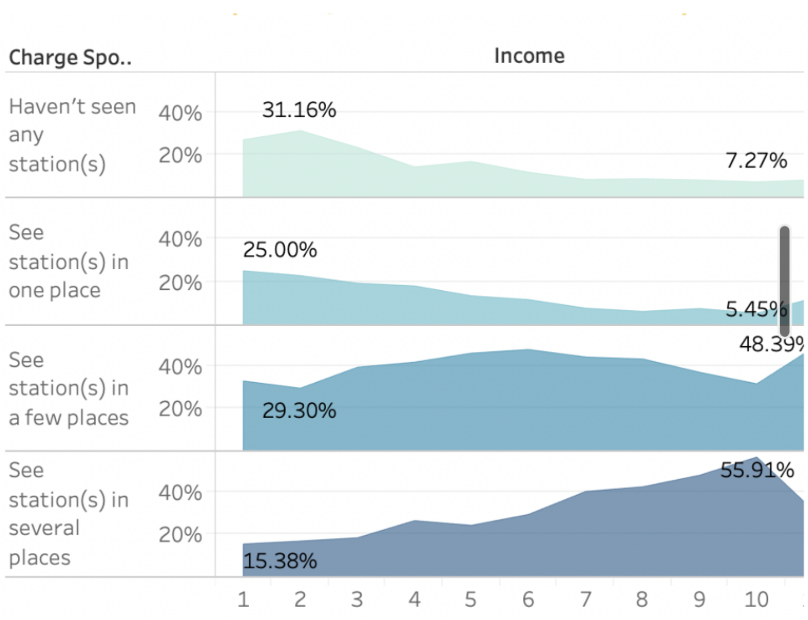


Figure 14: Early iteration 1 of income chart.

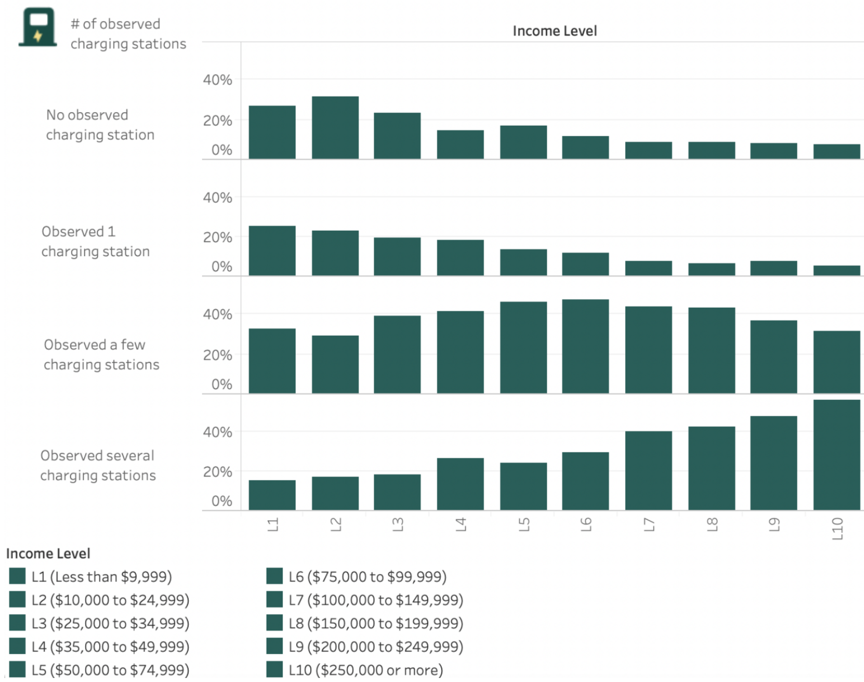


Figure 15: Early iteration 2 of income chart.

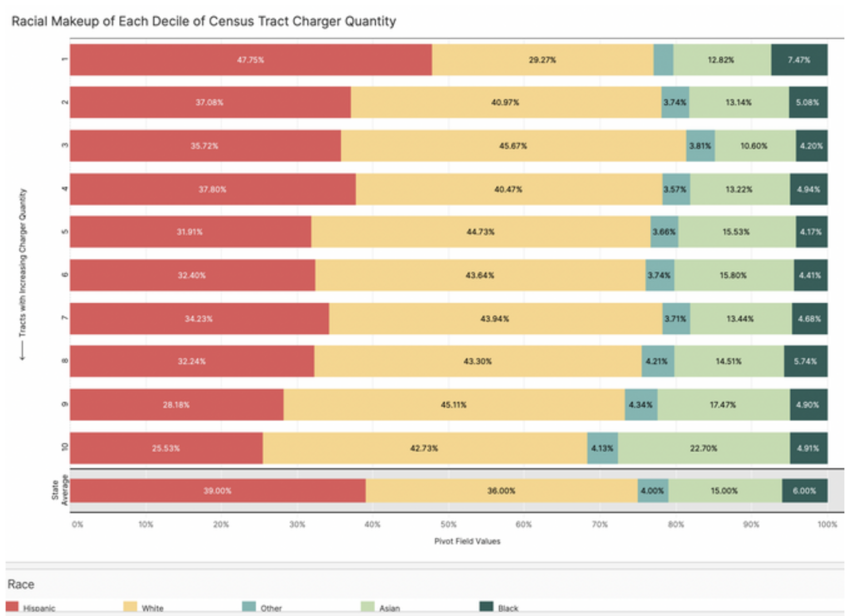


Figure 16: Early iteration of race chart.

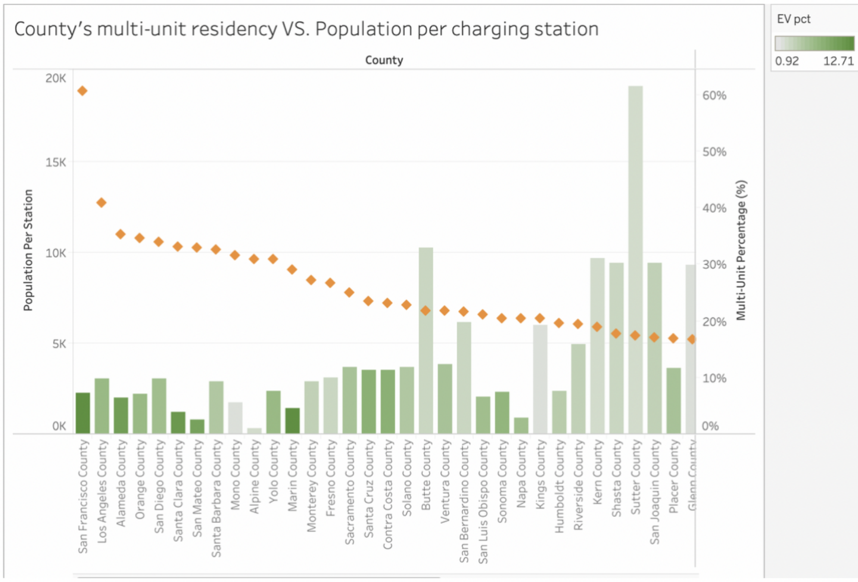


Figure 17: Early iteration of multi-unit housing chart.

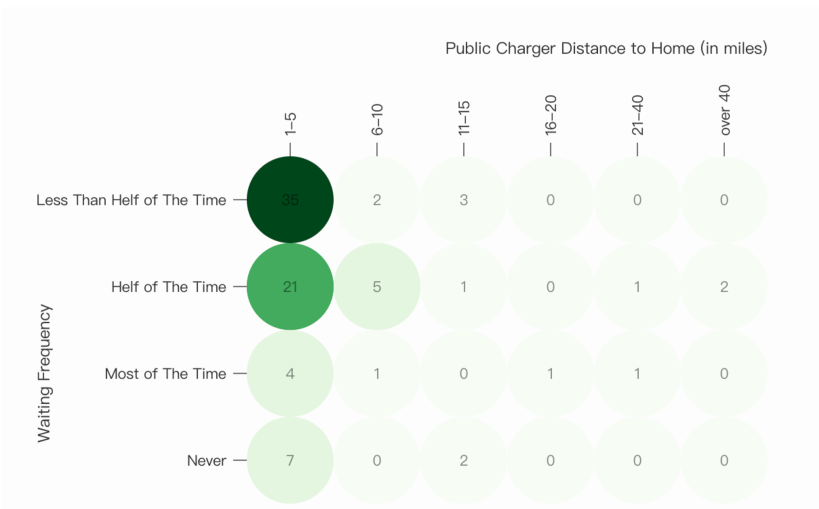


Figure 18: Early iteration of wait time chart.

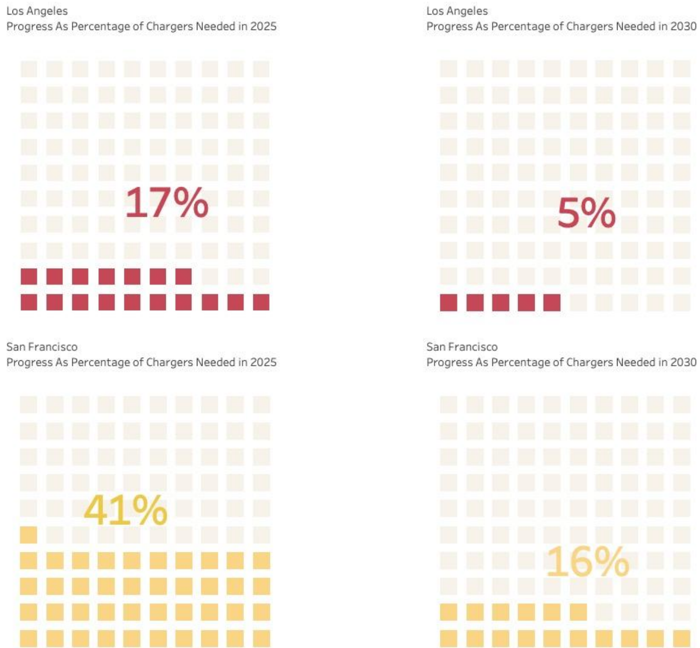


Figure 19: Early iteration of progress % waffle charts.

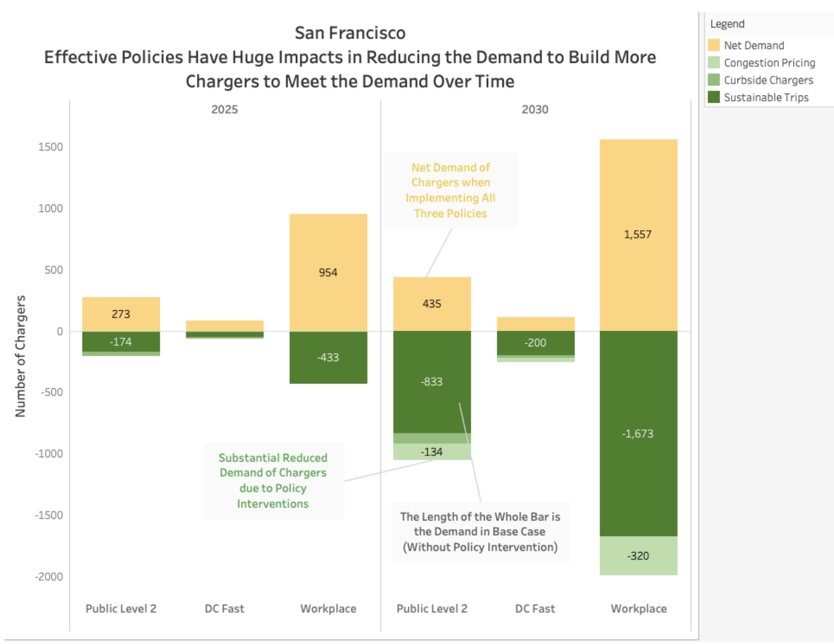


Figure 20: Early iteration of SF forecast demand chart.

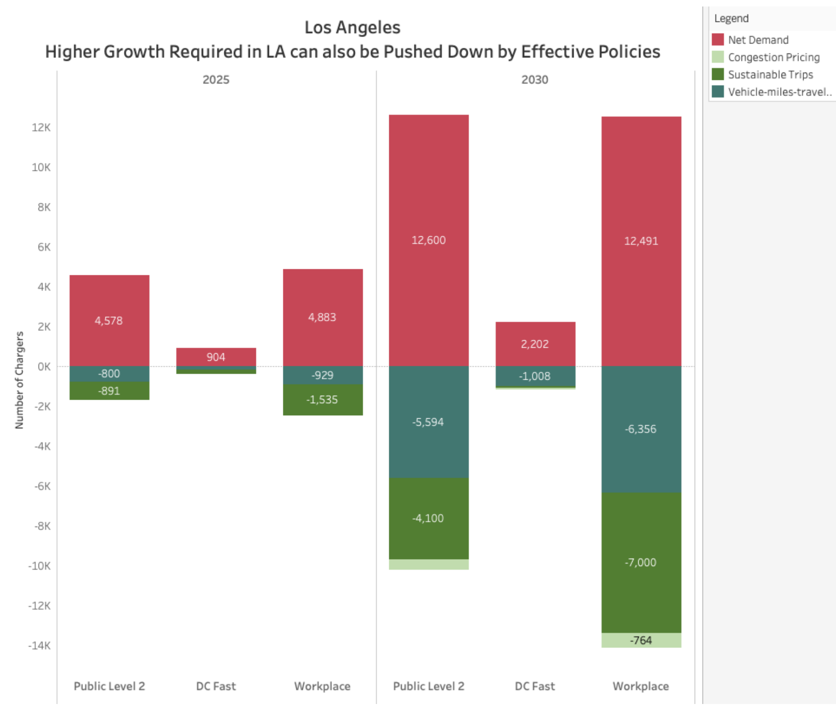


Figure 21: Early iteration of LA forecast demand chart.

9.4 Website Iterations

Website Wireframe

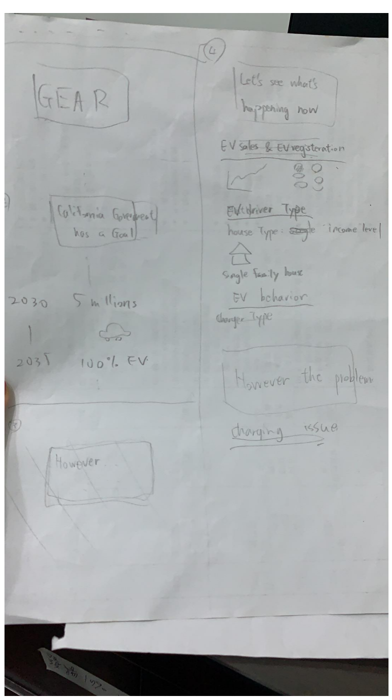
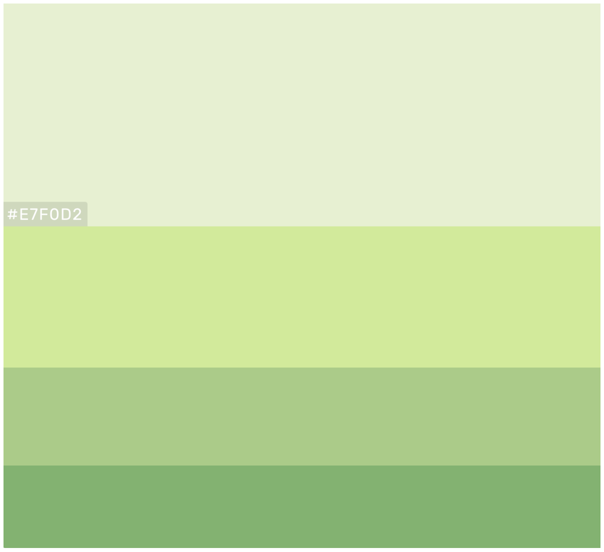


Figure 22: Early iteration of website wireframe.

Color palette iteration

Before

Main



Secondary

After

```
light: (
    1: #fdfdfd,
    2: #Cdcdcd,
    3: #A09f9f
),
dark: (
    1: #151719,
    2: #25282C,
    3: #33363A
),
primary: (
    1: #427771,
    2: #184d47,
    3: #2A5e59,
    4: #00221E
),
secondary: (
    1: #C1dcae,
    2: #96bb7c,
    3: #719d52,
    4: #527E32
),
alert: (
    error: #c64756,
    warning: #fad586,
    tiffany: #8ac4d0,
    lightblue: #a2d5f2,
    darkblue: #3282b8 ,        You,
    orange: #f08a5d
```

Figure 23: Color palette for website.