Investigating and Comparing User Experiences at EV Charging and Petroleum Fueling Stations

Abstract

Light duty electrification requires DC charging stations to take over the current role of petroleum stations. However, charging and fueling systems differ substantially in technology, user behavior, and business models. This study analyzes thousands of reviews for petroleum stations and DC charging stations. Using a mixed-methods approach combining quantitative rating, sentiment analysis, and Large Language Model (LLM)-assisted topical classification, factors that shape user experience are identified. Charging and fueling are perceived as fundamentally different activities. Fueling reviews are dominated by fuel price and staff interaction, while charging reviews emphasize equipment reliability, location convenience, and access to nearby amenities. Manufacturer-owned charging networks are perceived more positively than open networks. Open networks suffer from small station sizes and higher outage rates which reduce effective capacity and generate negative feedback. Across networks, proximity to amenities strongly influences perceived convenience and satisfaction. These findings identify actionable pathways for improving charging experiences and perceptions.

1. Introduction

In recent years, as the Battery Electric Vehicle (BEV) market has grown and matured, a charging network has grown and evolved alongside it. Although BEVs and their associated infrastructure are relatively new, they are not *de novo*. BEVs and BEV charging infrastructure are analogous to Internal Combustion Engine Vehicles (ICEVs) and ICEV fueling infrastructure, both of which are ubiquitous in the developed world. ICEVs are central to western society, particularly in the US, to the point that land use and travel patterns have come to reflect their capabilities. Further, nearly every BEV driver will have extensive experience with ICEVs. Thus, when considering the performance of BEVs and charging networks, it is crucial to compare findings to those for the ICEV analogue.

Although charging and fueling are analogous, they are not identical. Where nearly all ICEV drivers will purchase nearly all of their fuel from public-access fueling stations, BEVs drivers have the option to charge at different rates and in different contexts. Specifically, BEV charging may be separated into AC and DC charging. AC charging involves energizing the BEV's on-board charger and proceeds at a rate between 6.6 kW and 19.2 kW. DC charging involves directly charging the battery from and off-board charger and proceeds at a rate of 50 kW up to 350 kW Hybrid - EV Committee (2024). Even at the highest speeds available, BEV charging is substantially slower than ICEV fueling. However, AC charging can be accomplished in most buildings allowing for charge events to take place at home or at work. Most BEV charging occurs at home Hardman et al. (2025). If available, home and work charging allow for BEV drivers to attain or exceed convenience parity with ICEV drivers for routine driving Rabinowitz et al. (2023). For long trips/tours, BEV drivers experience the relatively long charging sessions and the relative sparsity of charging infrastructure as substantial inconveniences Dong et al. (2024).

The differences in charging and fueling behavior, in turn, inform the design of the respective networks. Private-access AC charging is, to this day, more convenient than reliance on public AC or DC charging stations. However, in the initial stages of the BEV rollout it was essential. Most BEV purchasers to date have access to home charging Hardman et al. (2025); Ramadoss et al. (2025); Robbennolt et al. (2025) and those without are substantially more likely to desist from BEV ownership than Lee et al. (2023). DC charging network operators have access to lower traffic and less revenue than petroleum station operators and have found it difficult to realize a sustainable business model Jenn et al. (2025); Gamage et al. (2023); Kim et al. (2022); Madina et al. (2016). In order to spur the development of DC charging networks, substantial private and public subsidies have been deployed. Charging-specific companies have largely responded to public subsidies and incentives Jenn (2025), and large networks such as EVgo and Electrify America are the direct result of a legal settlement. These Open networks are, and must be, usable for any vehicle with a compatible charging port. Private companies, such as Tesla and Rivian, have built and operated Proprietary networks which, until recently, were software-locked for the exclusive use of their vehicles. Tesla's DC charging network is the largest in the US by port count Alternative Fuels Data Center (AFDC) (2023) and, before 2025, operated at a loss. The resulting Open and Proprietary networks have very different structures. As seen in Figure 1, Proprietary networks consist of a smaller number of larger stations as compared to Open networks. 95% of Open DC charging stations in California contain fewer than 10 ports while 65% of Proprietary DC charging stations contain 10 or more ports.

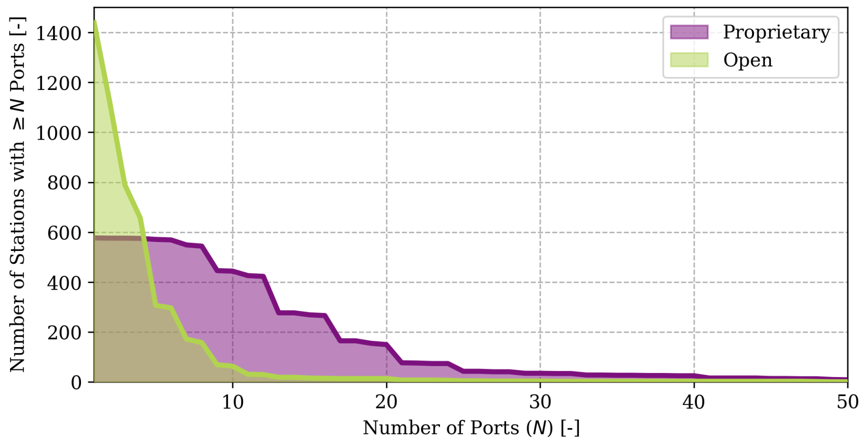

Figure 1: Survival functions of charging ports per station for Open and Proprietary networks in California

A second issue is non-charging revenue. Where ICEV fueling networks have coalesced on a business model which co-locates a collection of fuel pumps, a convenience store, and, often, a car wash, DC charging networks have coalesced around a much lower cost model of placing chargers in existing parking lots adjacent to unrelated businesses. BEV drivers do spend substantial amounts of money at adjacent businesses Zheng et al. (2024) but little or none of the revenue is returned to the charging station operator. Without ownership of the land on which the stations sit, and with high capital costs, limited revenue opportunities, a priority on network expansion, and thin operating margins, charging networks have earned a generally negative perception among existing users and potential future customers. This study answers the following research questions:

1. Are DC charging and fueling perceived as distinct or analogous experiences?
2. Are all DC charging networks perceived similarly, or as distinct populations?
3. What areas, if any, offer the greatest potential for improving perceptions of DC charging infrastructure?

This study addresses this gap and answers these questions utilizing reviews left on Google Places. Google Place reviews are part of the larger Google Reviews platform and can be added by users within the Google Maps application. Unlike charging-specific platforms that primarily attract experienced EV drivers who actively seek or troubleshoot charging stations, Google Places captures feedback from a much wider cross-section of the driving public. Its reviewers include both BEV and ICEV users who interact with charging or fueling locations as part of their broader travel routines. Google Places hosts reviews for both fueling stations and charging stations within the same interface, allowing for a direct comparison of ratings and user experiences across these two types of infrastructure. Additionally, this study makes use of modern Natural Language Processing (NLP) techniques, specifically LLMs in coding reviews.

2. Literature Review

Several studies have attempted to measure sentiment about Electric Vehicle (EV) charging networks using crowd-sourced data. Two studies have use a large dataset taken from PlugShare. This dataset contains over 127,000 reviews collected before 2020. Asensio et al Asensio et al. (2020) used Supervised NLP (CNN classifier) for sentiment and topic classification; statistical analysis. The study found that nearly half of app-using EV drivers reported negative charging experiences. Contrary to expectations, public (government-run) charging stations performed as well as private-run stations in user ratings. Stations in dense urban areas showed a 12 - 15% higher rate of negative sentiment than those in non-urban areas, pointing to service quality issues in cities (e.g. more malfunctions and congestion). Free charging sites received better reviews than fee-based sites (fee-based stations drew more poor reviews). Ha et al. Ha et al. (2021) used transformer-based deep learning (BERT/XLNet) for automated multi-label topic classification. The study identified common complaint topics (e.g. station outages, wait times, location issues) and provided evidence that less-populated areas may be underserved i.e., potentially fewer stations or lower quality service in rural/sparsely populated regions.

Liu et al. Li et al. (2024) used another large reviews dataset spanning several Southeast Asian countries and 10 years (2011 - 2021). Utilizing cross-lingual text classification (English BERT-uncased model), the study analyzed consumer charging behavior via reviews across Asian markets. Findings included government-operated charging sites having higher reported failure rates and more negative user feedback than private-run sites in Asia. This contrasted with U.S./Europe, where public vs. private station performance was more equal. The study also discovered that networked stations with communication capabilities offer better service quality (higher user satisfaction) than non-networked stations, supporting policies to promote connected chargers. Results showed marked regional differences in both overall performance and issue types. Also in Asia, Guo et al. Guo et al. (2025) analyzed a relatively small dataset of 2,265 reviews from Amap (Gaode) public charging station comments in Guangzhou. Using automated Sentiment extraction the study provided a focused look at EV charging user satisfaction in a major Chinese city. Found an overwhelming majority of feedback was negative, signaling significant user frustration. Key complaints were insufficient charging spots, slow charging speed, and high and inconsistent cost. Positive reviews were usually centered on high charging speed. These findings underscore that even in a city with >50% EV penetration, availability and reliability shortfalls are a dominant concern.

Most recently Ashby et al. Ashby et al. (2025) looked at over 100,000 Twitter/X posts using semi-automated thematic analysis (BERTopic pipeline and other Large Language Models (LLMs) to model and explore the topics with Manual "human-in-the-loop" intervention). The study examined public attitudes and perceptions regarding EV charging, aiming to surface common barriers and successes perceived by consumers. Through thematic coding of discussions, the study identified recurring concerns: insufficient charging infrastructure, charging reliability problems, high costs or unclear pricing, and convenience factors. Positive success themes included appreciation for expanding charger networks and instances of reliable, convenient charging experiences. The authors note that understanding these attitudes is crucial for policy and adoption public sentiment often highlights pain points (like charging anxiety beyond just range anxiety) that need to be addressed to build confidence in EVs.

The surveyed studies differed in dataset and methodological approach. The findings in the studies were largely consistent: (1) predominant overall negative sentiment (2) insufficiency of infrastructure, (3) unreliability, and (4) unclear pricing. In addition, several of the studies noted significant differences between groups of chargers treating them as separate populations. What these studies do not do is place user experiences at charging stations in context. It is not uncommon for prior studies to claim that their results represent the state of the DC charging network. This claim is not defensible. All of the studies rely on crowd-sourced data. There is substantial evidence that users of review platforms are much more likely to leave a review for a salient experience than a typical experience, and more likely to view negative experiences as salient Amsterdamer and Milo (2015). This is especially important for fueling/charging infrastructure because stopping to re-energize a vehicle is an inconvenience no matter how it is done. However positive the experience is, most would prefer an ideal situation where stopping is never required.

Some issues with public charging are unavoidable consequences of physical and market realities, others may be identified and corrected. The extent to which public BEV charging needs to be improved, and the extent to which it can possibly be improved can only be understood in the context of the analogous ICEV refueling infrastructure. One should be very careful in selecting a data-source. Most academic analyses Asensio et al. (2020); Ha et al. (2021); Li et al. (2024); Guo et al. (2025) rely on charging-specific review platforms that cultivate experienced and engaged user bases with distinct norms and expectations. These platforms capture perceptions of charging stations that are specific to that community and do not include reviews for ICEV fueling stations. To date, no study has examined how users experience BEV charging within the broader context of vehicle re-energizing infrastructure, including ICEV fueling.

3. Dataset

Reviews were pulled from Google Reviews for charging stations and petroleum stations in California via the Google Places API. Google Places maintains a list of place types (e.g. shopping center, chinese restaurant, ect.), and individual places can be marked as belonging to one or more types. Both charging stations and fueling stations are place types. It is not possible, via the API, to return the IDs of all places of a given type within a geographic area. Rather, the number of returns is limited to 20. Thus, it was necessary to search for stations over many API calls.

The Google Places API uses a nonlinear pricing structure. For each function and tier of usage, a given amount of calls per month will be free followed by a decreasing price per call as call volume increases. Thus, data collection was designed to minimize API calls. This was accomplished using the NearbySearch function. The NearbySearch function returns place IDs and details for places meeting a set of criteria and near to a specified set of coordinates. For this analysis, the query specified that places be returned in order of proximity to the query point.

In order to efficiently locate most of the charging stations in California, an initial set of sample points was determined from the locations of unique DC charging stations in the Alternative Fuels Data Center (AFDC) dataset Alternative Fuels Data Center (AFDC) (2023) (2,016 locations), road network nodes in the National Highway System shapefile of Transportation Statistics (2025) (59,621 locations), and the centroids of incorporated cities and towns in California US Census Bureau (2025) (1,594 locations). The NearbySearch function returns up to 20 places which can be up to 5 km from the query point. It is not likely that many DC charging stations will be more than 5 [km] from a major road, town centroid, or known location of a DC charging station. Thus, queries were conducted at known station locations, place centroids, and at 3 km intervals along the road network were not covered by the previous categories. In total, 15,360 query points were used.

The queries returned 1,609 unique DC charging stations and 10,730 unique petroleum fueling stations. While the number of petroleum stations lines up with the expected number in the state per the California

Energy Commission California Energy Commission (CEC) (2022), the number of DC charging stations returned is considerable lower than the number of unique DC charging stations found in AFDC. Because of the differences in data sourcing, there will be disagreement between platforms on the numbers and locations of charging stations. The differences in coverage between Google and AFDC vary by network as seen in Figure 2.

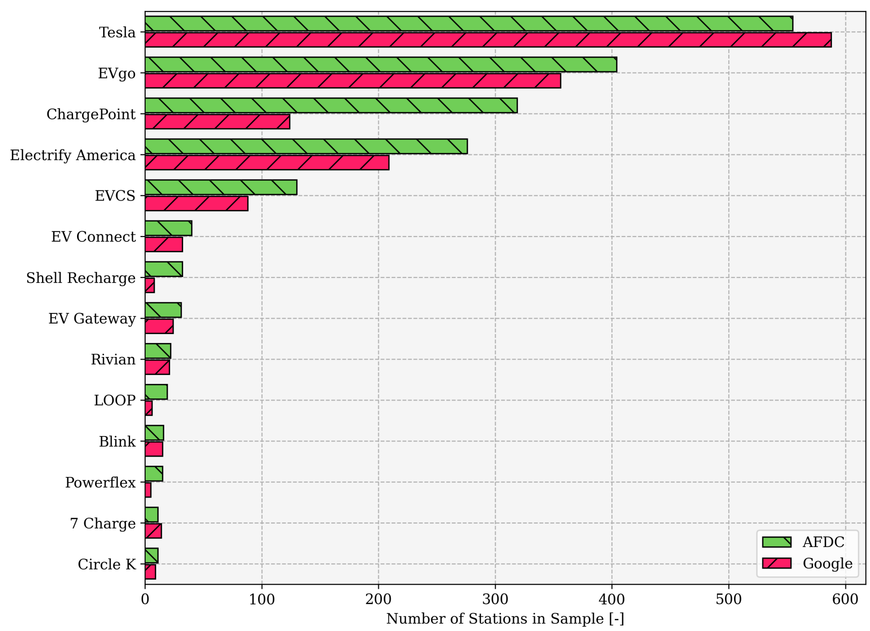

Figure 2: Differences in representation of networks between AFDC and Google

The numbers of stations returned from Google Places were lower than those found on AFDC across networks. As each unique AFDC location was a query point it is unlikely that the disparity is due to stations missed. Rather, the Google Places platform seems to list fewer stations than AFDC for most networks. As Google Places is crowd-sourced it follows that very low volume stations may not be entered into the database. The largest difference is for the ChargePoint network. The unusual scale of this difference is due to differences in attribution of stations. The AFDC database lists stations which use ChargePoint equipment as ChargePoint stations where the Google database only lists ChargePoint operated stations as ChargePoint stations.

The collected dataset contains reviews from 6,230 reviewers, the majority of which are from 2024, for DC charging stations and 45,116 reviews for petroleum stations. The Google Places API is only able to return up to 5 reviews per station. By default, the reviews returned are the 5 most "relevant" reviews based on a relevance score. The relevance score weights recency, the rank of the reviewer, the length of the review, and the feedback the review has received from other users. As a result, most stations in the dataset returned 5 reviews as shown in Figure 3.

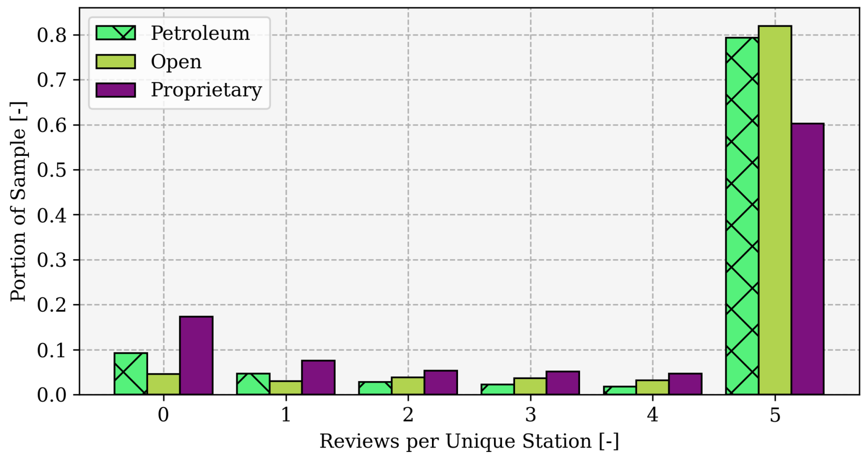

Figure 3: Distribution of number of reviews pulled by unique station

The collected dataset contains reviews from 6,230 reviews from 4,991 unique reviewers for DC charging stations and 45,116 reviews from 39,786 unique reviewers for petroleum stations. The populations are largely separate with an overlap of 219 reviewers. In both cases, the vast majority of reviewers reviewed once in the sample as shown in Figure 4.

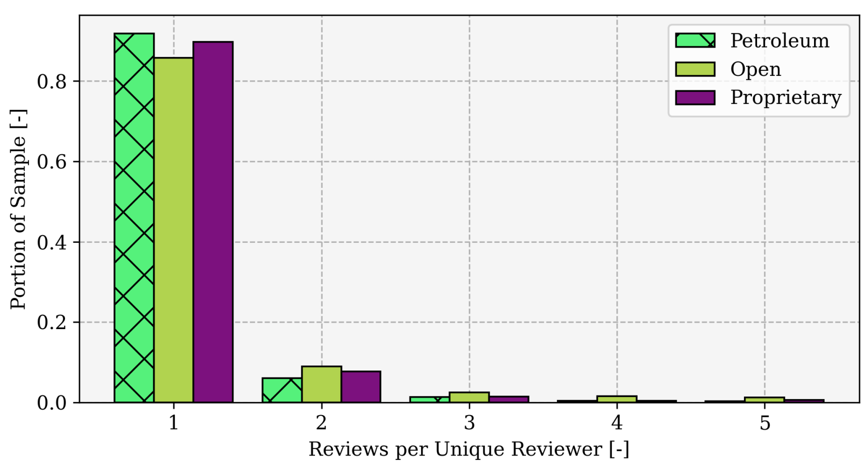

Figure 4: Distribution of number of reviews pulled by unique reviewer

Another consequence of the prioritization of "relevant" reviews is that the reviews were, on average, text-rich. The median review in the dataset is 179 characters long. The lengths of reviews are shown in Figure 5.

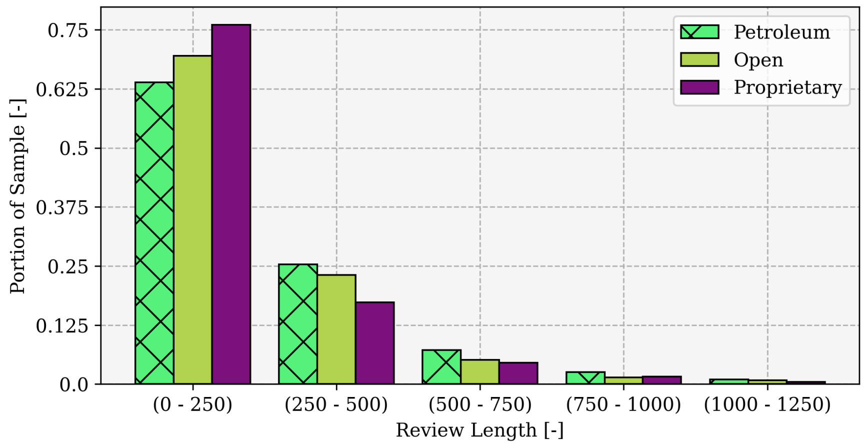

Figure 5: Distribution of character count by review

Compared to those seen in the reviewed literature, the dataset in this study presents advantages and drawbacks. The dataset substantially smaller than the largest seen in literature which contain over 100,000 reviews and are national in scale. This limitation is downstream of certain constraints: 1) funding limitations and 2) transparency and dataset dissemination. The limits on funding resulted in a limitation in geographic scope to California and the economic search described earlier. The requirement to produce a dataset which could be disseminated to other researchers prevented the team from either utilizing proprietary data provided by a platform scraping a larger dataset form a platform. The team also has access to a 90,000+ review proprietary dataset from PlugShare for California. However, despite containing many more reviews than the Google Places dataset used herein, the vast majority of the reviews are either zero-text or very short as displayed in Figure 6. Although the PlugShare dataset contains many more reviews than the Google Places dataset, the Google Places dataset contains a greater number of w of 500 characters or greater. These differences are the result of Google Places returning only the 5 most relevant reviews for a station.

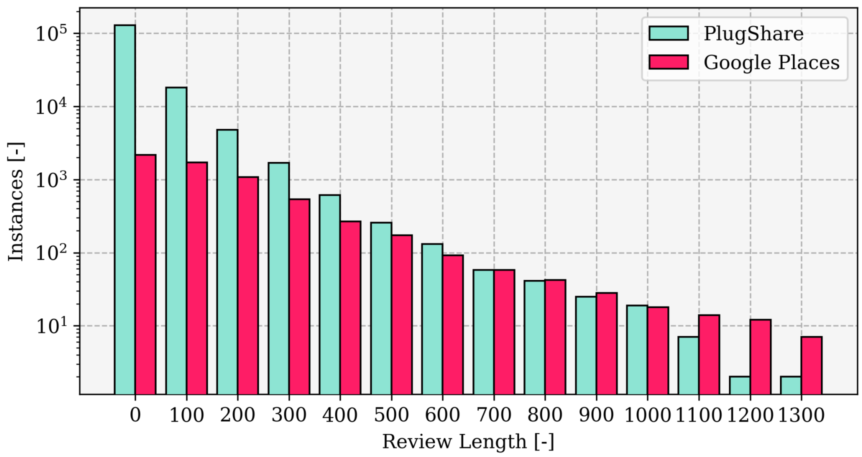

Figure 6: Comparison of review lengths between PlugShare and Google Places California DC charging stations datasets

The Google Places dataset used in this paper is insufficient to produce a rating or score for any individual station but is sufficient for drawing conclusions about the experience of users on a network basis.

Coding of Reviews

Reviews were coded to reveal topics, and sentiments contained in review text. Coding was performed using the GPT-4 LLM. The use of LLMs for review coding and thematic and sentiment analysis is recent but not unprecedented. Thematic and sentiment extraction is a task well suited to the natural language processing capabilities of LLMs and using LLMs substantially reduces the cost of the task allowing for larger datasets to be labeled. The primary hurdle is obtaining accurate results and validating these results. Recent studies have shown that LLMs can perform qualitative text coding and sentiment analysis with accuracy comparable to human coders when paired with structured prompts and manual validation. The GPT family of LLMs has been shown to perform the task of review coding in a manner which replicates human outputs accurately and consistently Gilardi et al. (2023); Dunivin (2025). LLMs have been shown to perform roughly as well as humans for both inductive and deductive coding tasks Ziems et al. (2024) but generally do not outperform humans and struggle on prompts where humans struggle Li et al. (2024). This is an artifact of the training sets provided too LLMs which are coded by humans. In general, LLM performance is sufficient but depends on prompt structure Liu et al. (2025). Interestingly, for the purposes of review coding, few-shot prompting, where examples and other information is provided, does not substantially increase the accuracy of results over zero-shot prompting Than et al. (2025).

Following this precedent, the text elements of reviews were coded using the GPT-4o LLM via the openAI

API. Topical coding was performed in two ways:

Deductive The LLM was instructed to deductively assign codes from a pre-defined list to reviews along with a confidence score. Only high confidence ($\geq 90\%$) codes were retained. The pre-defined list of codes was derived from the ChargeX Consortium Final Report [SOURCE].

Inductive The LLM was instructed to identify any an all topics related to the charging/fueling station user experience in each review text along with a confidence score. Only high confidence ($\geq 90\%$) codes were retained.

These two approaches serve two purposes. The deductive coding approach used pre-defined categories of experiences as codes in order to guarantee a small an uniform set of codes among the three populations of charging/fueling stations. The small and uniform set of deductive codes was used for direct comparative analysis. The codes for deductive labeling were:

1. Location / Convenience - Convenience or inconvenience of station location.
2. Amenities - Quantity and quality of businesses and other opportunities adjacent to the station.
3. Availability / Access - Availability and accessibility of charging/fueling equipment at the station.
4. Functionality / Reliability - functionality, reliability, and usability of charging/fueling equipment at the station.
5. Start-up / Authentication - Ability or inability to start a charging/fueling session at the station.
6. Charging Performance - Speed/performance of charging/fueling equipment at the station.
7. Information / Visibility - Quantity and quality of information and signage at the station.
8. Customer Support - Quantity and quality of customer support available for the station.
9. Comfort / Safety - Quantity and quality of lighting, overall safety, and customer behavior at the station.
10. Pricing / Billing - Quality and clarity of pricing at th station.
11. Station Employees - Quality of interactions with employees present at the station.

The inductive coding approach served to allow the LLM to identify specific topics without the need to maintain a small or uniform set between the populations. The primary purpose of the inductive codes was to provide detail and context for the analysis performed on the deductive codes.

The LLM was also prompted to assign a sentiment, where possible, to each review with text. Sentiments were chosen deductively from a set containing Positive, Neutral, and Negative.

3.1. Validation of Coding

Following coding, the coded reviews were manually validated by the research team. This process involved team members inspecting randomly selected reviews and determining whether or not they had been accurately coded as a binary outcome. A review was considered to have been coded correctly if all high-confidence themes were present and the sentiment was correct. In order to efficiently accept or reject the coding, the team made use of the Hypergeometric distribution. For a population of size N with K failures, the likelihood that there will be k failures in a sample of size n is

$$P(N,n,K,k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}} \tag{1}$$

In this case, N, n, and k are known and K is unknown. The exact total number of failures in the population K cannot be known until all reviews are checked. However, the total probability that the total number of failures is less than or equal to a given threshold K^* can be computed using the CDF of the Hypergeometric distribution

$$\begin{gathered} P(K \leq K^* \mid N,n,k) = 1 - \frac{\binom{n}{k+1}\binom{N-n}{K^*-k-1}}{\binom{N}{K^*}} \\ {}_3F_2\left[\begin{matrix} 1, k+1-K^*, k+1-n \\ k+2, N+k+2-K^*-n \end{matrix}; 1\right] \end{gathered} \tag{2}$$

where ${}_qF_p$ is the generalized hypergoemetric function. Thus, for a given level of confidence α, the greatest possible number of failures $K'(N,n,k,\alpha)$ can be determined by finding the smallest K^* such that $P(K \leq K^* \mid N,n,k) \leq \alpha$. The authors required that that coding was at least 95% accurate at a 95% level of confidence ($\alpha = 0.05$). This meant that the maximum allowable value for K' was $0.05N = 2567.3$. A program was written to randomly select a review from the sample and serve it to a team member. When the team member submitted their decision, the size of the sample n was increased by 1 and the number of failures in the sample k was updated. After each review was checked the value for $K'(N,n,k,\alpha)$ was computed. When the condition $K'(N,n,k,\alpha) \leq 2567.3$ was met, the process could be terminated and the coding accepted as valid. For the final coding, the team manually checked 1,538 reviews of 51,346 total reviews finding 1,487 to be correctly coded. At 95% confidence, the maximum possible number of failures in the population was 2,135 or 4.16% of the population, sufficiently guaranteeing a success rate of greater than 95% and validating the coding.

4. Results

The central questions in this paper are whether or not users view fueling and charging infrastructure as analogous and whether Proprietary networks are viewed similarly to or differently from open networks. These questions will be considered using ratings and coded reviews.

4.1. Ratings and Sentiments

The ratings attached to reviews are on a 1 to 5 scale. As is commonplace with crowd sourced data Amsterdamer and Milo (2015), ratings tend to be polarized. Ratings in the 2-4 range accounted for less than one third of charging station ratings and less than one quarter of petroleum station ratings. Most stations feature ratings on both extremes and, thus, the most common station average rating is around 3. This dynamic is displayed for petroleum stations, open charging stations and proprietary charging stations in Figure 7.

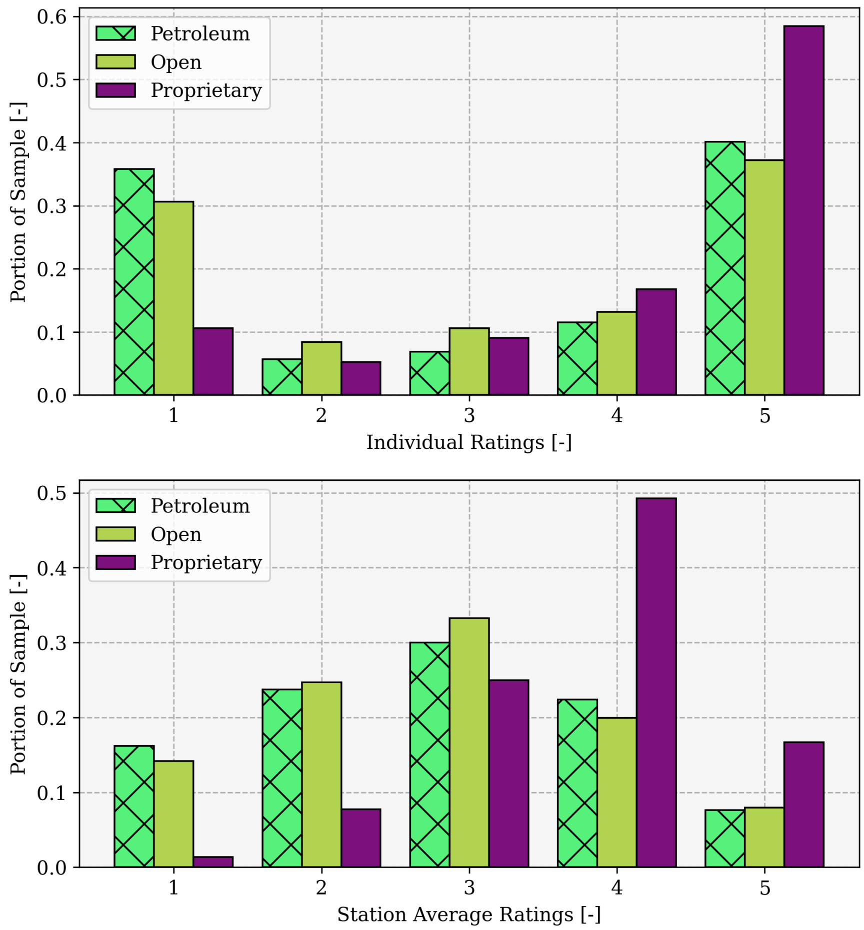

Figure 7: Probability density of individual ratings and station average ratings

While ratings follow a similar pattern for open charging stations and petroleum stations, there is a notable difference with regards to proprietary stations. Tesla stations receive higher ratings than the other populations on an individual and station average basis. Aggregating to a network level, there are statistically significant differences between populations and between networks within the open network population as seen in Figure 8.

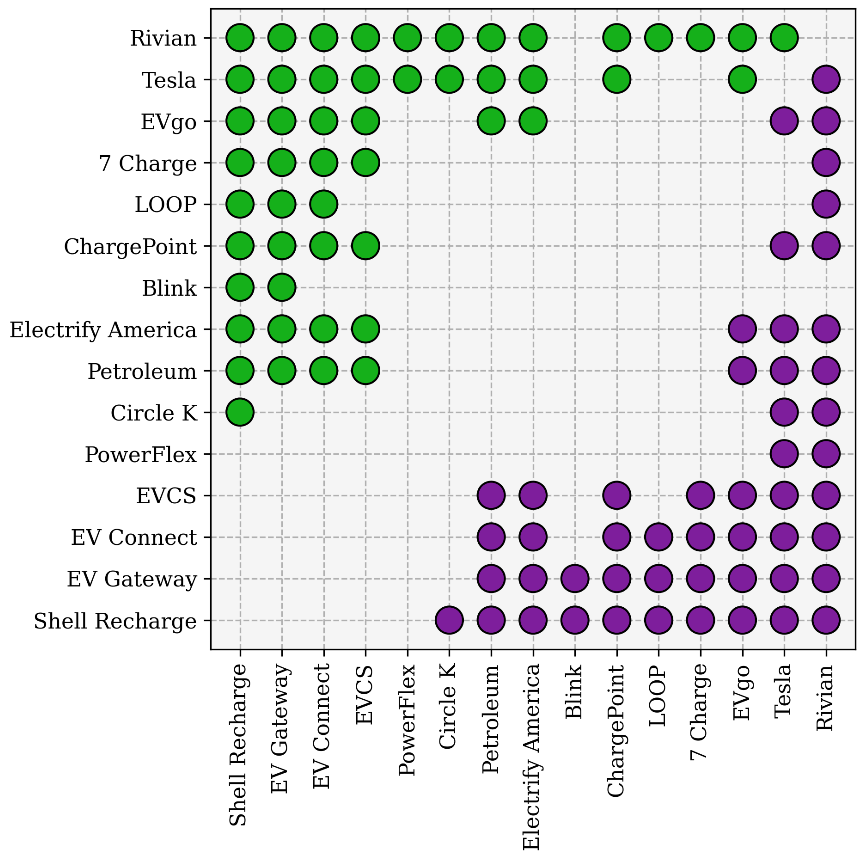

Figure 8: Significant ($\alpha = 0.05$) differences in ratings between networks. Green indicates a positive significant difference between the row and column, purple indicates the opposite.

In general, the proprietary networks, received better ratings than the open networks with Rivian outperforming Tesla. Among the open networks, a partial ordering exists with no network dominating or being dominated. Ratings, in and of themselves, should be treated with skepticism. When users provide a rating, they do so in the context of their expectations. Differences in ratings are only meaningful if user expectations for the experiences being rated are equivalent.

Rather than relying on ratings, overall sentiments were extracted from review text. Review sentiments were deductively assigned from the ordinal set, Negative, Neutral, or Positive. As seen in Figures 9, 10, and 11, the correspondence between numerical ratings to sentiments followed a consistent patters among the station populations. Most reviews with a rating of 1 through 3 expressed a negative sentiment and most reviews with ratings 4 or 5 expressed a positive sentiment. Neutral sentiments were less common than ratings in the 2 through 4 range.

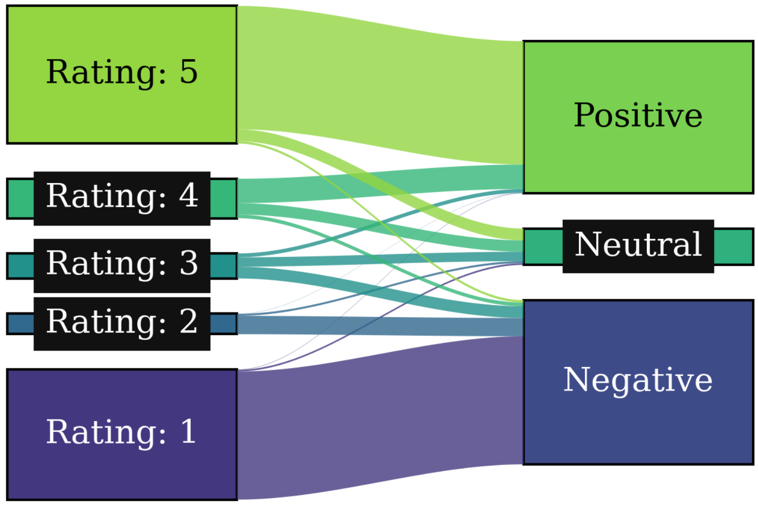

Figure 9: Correspondence between ratings and sentiments for Petroleum stations

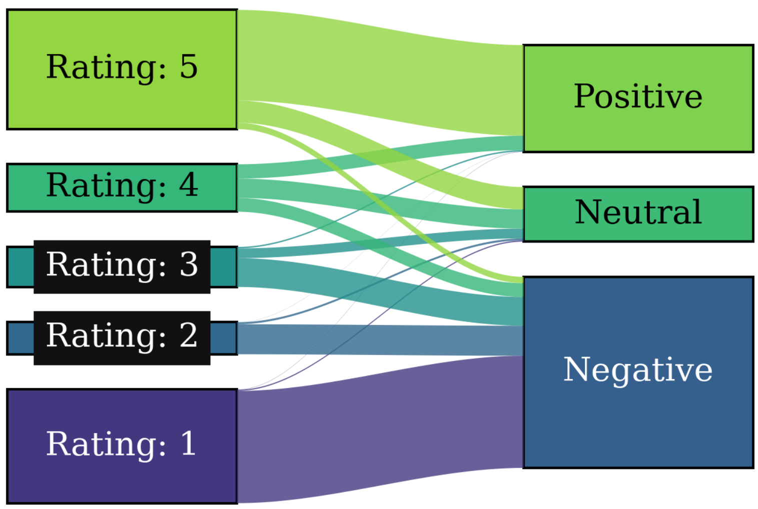

Figure 10: Correspondence between ratings and sentiments for Open charging stations

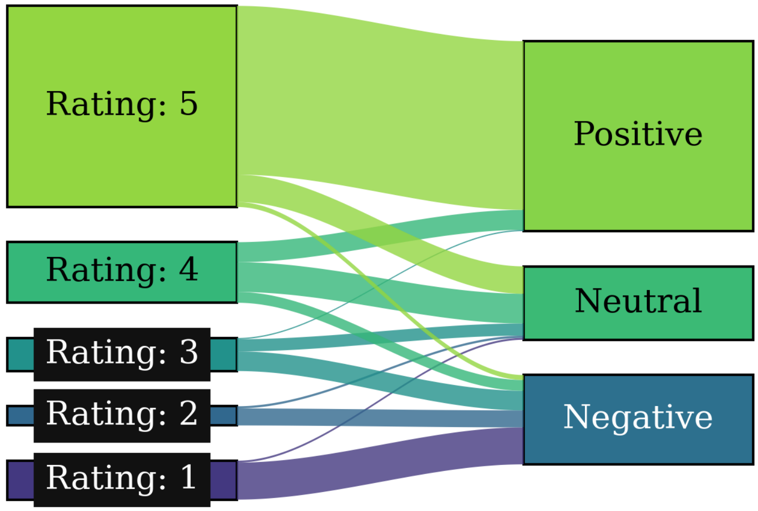

Figure 11: Correspondence between ratings and sentiments for Proprietary charging stations

There are a non-negligible number of reviews with high ratings and negative sentiment. These reviews fall into the following categories: 1) Unambiguously negative reviews assigned a good rating (perhaps by accident) 2) Reviews which state the experience at the station was positive but the usual experience at the network is negative, and 3) reviews which point out specific malfunctions and ask for them to be fixed.

4.2. Topical Analysis

The effort required to leave a text review is sufficiently high that it deters most users most of the time. When reviewers leave a text review it is because some aspect of the experience was salient. Most of the effort is in opening an app, logging in, finding the appropriate location, and starting a review. Having gone to the effort to initiate a text review, users may then go on to discuss aspects of the experience that would not, otherwise, have prompted them to leave a review. Those with knowledge of the Google Reviews platform will also know about the various incentives for longer and more informative reviews. Thus, it is worth considering not just the instances of topic codes, but also the coincidence of topic codes with one-another. Incidence and coincidence of deductive topic codes are shown in Figure 12.

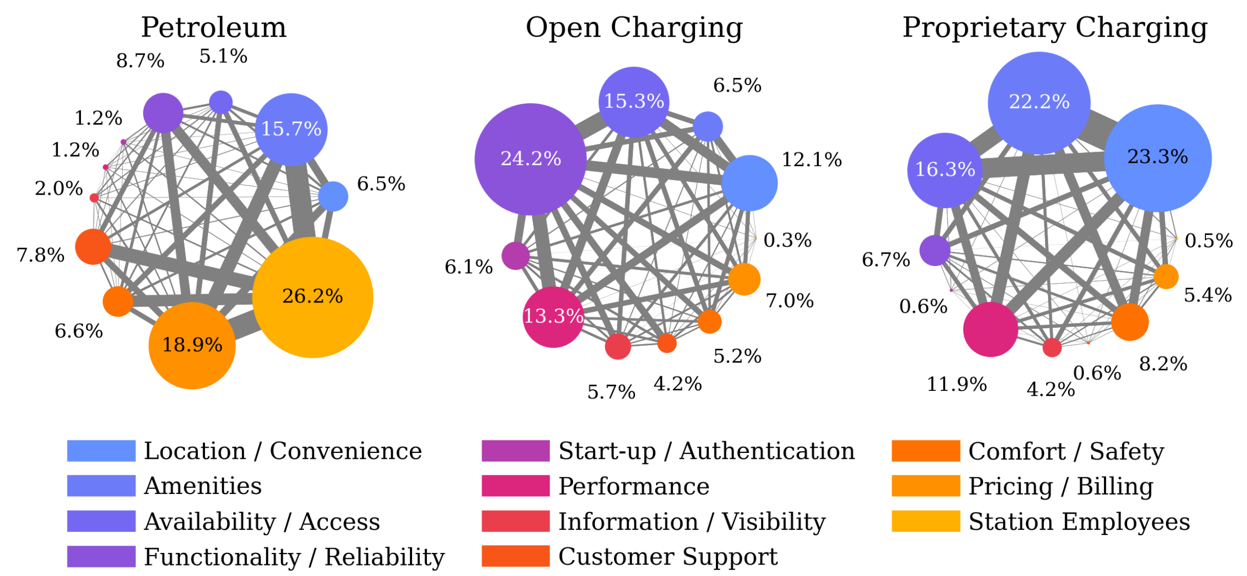

Figure 12: Graphs showing incidence and coincidence of deductive topic codes by station population. Sizes of nodes correspond to incidence and thickness of edges correspond to coincidence.

The differences in topical content between the station populations is striking. Petroleum station reviews tend to focus on aspects of the experience not directly related to fueling. The most common themes are Station Employees, Pricing / Billing, and Amenities while Performance and Start-up / Authentication occurred in less than 1% of reviews. For the open charging networks, the most common themes were Functionality / Reliability, Availability / Access, and Performance while Station Employees was extremely rare. For the proprietary networks, the most common themes were Location / Convenience, Amenities, and Availability / Access and the least common were Start-up / Authentication, Customer Support, and Pricing / Billing. Some context is necessary to interpret the raw occurrence of themes. Although California has allowed self-service fueling since the 1970s most petroleum stations feature a convenience store which will be staffed by clerks. Many customers at petroleum stations interact with these clerks to purchase fuel and/or convenience items. Most charging stations are not attached to a convenience store and have no on-site employees. Where employees are mentioned in charging station reviews, the employees in question are employees of adjacent businesses. Additionally, the equipment, and thus, fueling performance at petroleum stations is far less varied than that at charging stations leading to far fewer comments relating to performance. Finally, the widespread availability of petroleum stations means that convenience of location is less salient for them than for charging stations, where Availability/Access can strongly influence user experience.

Between the charging populations, there are also substantial differences in topical content. The most common theme for open charging stations was Functionality / Reliability, occurring in nearly one quarter of reviews. For proprietary charging stations, Functionality / Reliability occurred less than 7% of the time. Similarly, Amenities occurred far more frequently for proprietary charging stations than for open charging stations. This difference is, likely, the result of different network structures where the proprietary networks contain a smaller number of larger stations and these stations tend to be located in amenity-rich areas.

4.3. Regression Analysis

To better understand the impact of the various aspects of the charging/fueling experience in California an Ordinal Logistic Regression was performed for review sentiment. Regressors were instances of themes in the review, station and location characteristics, and station region. The assignment of station characteristics and region is discussed in the following subsections.

4.3.1. Station Characteristics

Station capacity is a metric derived from queuing theory that accounts for the number of ports at a station and the maximum power available at these ports. The station capacity metric reflects the mean number of vehicle arrivals per hour the station can handle without a durable queue forming. Queue waiting time W_q is modeled using the M/M/c queuing formula. The expected queue length for an M/M/c queue is computed as

$$W_q = L_q \lambda^{-1} \tag{3}$$

$$L_q = f_q(\lambda, \mu, c) = \pi_0 \frac{\rho (c\rho)^c}{c!\,(1-\rho)^2} \tag{4}$$

$$\pi_0 = \left[\left(\sum_{k=0}^{c-1} \frac{(c\rho)^k}{k!} \right) + \frac{(c\rho)^c}{c!\,(1-\rho)} \right]^{-1} \tag{5}$$

$$\rho = \frac{\lambda}{c\mu} \tag{6}$$

where λ is the mean arrival frequency, μ is the mean service frequency, c is the number of homogeneous servers. For any station with a known μ and c, the maximum negligible-queue arrival rate λ^* can be solved for. For this study, the negligible-queue threshold was set as $W_q \leq 60$ seconds. The quantity ρ can be conceived as station "utilization". Where ρ is low the station has excess capacity and where high the station approaches saturation. W_q approaches infinity as ρ approaches one. The impact of station size on station capacity can be seen in Figure 13.

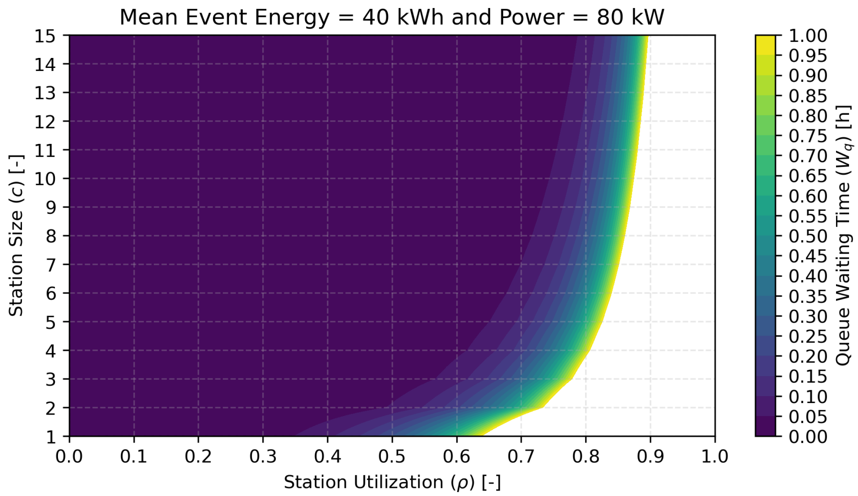

Figure 13: Queuing time with M/M/c queue model for a station with given average charging energy and average charging rate. As the size of the station c increases, queues form at higher utilization rates ρ.

Stations do not exist in isolation. Adjacent stations may serve to, indirectly, augment each-other's capacity but may, equally, serve as comparison points. In this paper, stations within a 5 minute drive of one-another are considered adjacent. Often, highly adjacent stations will be within the same shopping or business complex. Stations in different types of places serve different purposes. Thus, the median income and population density of the census tract each station is in was considered. Finally, stations which are located in rural areas and adjacent to highways were given the Corridor tag.

4.3.2. Station Region

In order to determine if there is any substantial variation in user experience by region within California, stations were also assigned one of 10 regions as in Figure 14.

Figure 14: California regions

4.3.3. Regression Results

All regressors were min-max normalized The regression analysis concerns what factors impact the sentiment of a given individual review. Separate models were fit for each population. Significant parameters of the regressed equation are displayed in Table 1.

Table 1: Significant parameters of regressed equation by population($\alpha \leq 0.05$). All regressors are min-maxed normalized. Positive coefficients imply an increased likelihood of neutral over negative and positive over neutral all else being equal. Negative coefficients imply the inverse.

Parameter	Petroleum	Open	Proprietary
Theme(s) in Review			
Location / Convenience	0.8219	0.6879	0.7704
Amenities	0.5320	0.6551	0.6059
Availability / Access		-0.2014	
Functionality / Reliability	-0.4948	-0.5186	-0.3104
Start-Up / Authentication	-0.8184		
Performance	-0.2245	0.1308	
Information / Visibility	-1.1745	-0.6336	-0.6314
Customer Support	-0.2977	-0.8006	-0.6022
Comfort / Safety	-0.5862		-0.5016
Pricing / Billing	-0.0673	-0.2860	-0.3560
Station Employees	-0.0707		
Station Characteristics			
Station Capacity		1.9714	1.4415
Number of Adjacent Stations		-0.2826	
Tract Median Income	-0.2177		
Tract Population Density			
Is Corridor Station		0.2010	
Station Region (Compared to Los Angeles)			
Northern Forest	0.3263		
Northern Valley	0.3742		
Sacramento			0.3287
Bay Area	0.1280		0.1446
Sierra	0.3574		0.3848
Central Coast	0.1748		0.4091
Central Valley	0.1670		
San Diego			
Southern Desert			0.1935

Positive coefficients in Table 1 imply an increased likelihood of neutral over negative and positive over neutral all else being equal. Negative coefficients imply the inverse. The presence of themes in reviews are strongly and directionally predictive of the sentiment of the review. Reviews which touch on the location of a station or available amenities are more likely to be positive. All other themes are more likely to be found in more negative reviews with the exception of Performance for Open network charging stations. It is worth noting where the thematic coefficients diverge between the populations. Availability / Access was significant for only Open network charging stations where it has a negative coefficient. This implies that port unavailability is a greater element of the experience at Open network charging stations but does not apply that drivers are never forced to queue at other stations. The only theme with significant coefficients of opposite signs is Performance. As mentioned earlier, differences performance at gas stations are difficult to notice unless there is some equipment or operational malfunction. Proprietary networks tend to employ standardized equipment across their stations leading to uniform performance. Open network charging stations are the only set of stations where performance is variable and high performance is a selling point.

Station characteristics, location, and region impact sentiment differently across populations. Both Petroleum and Proprietary network stations tend to receive more positive reviews when they are in more rural areas. No regional effect is seen for Open network stations, but significant effects are seen for number of adjacent stations and corridor stations. Combined with the powerful positive impact of the Location / Convenience theme, it is possible tat the more essential a station is to drivers the more willing drivers will be to excuse its flaws. Finally, the strongest positive coefficients are found for station capacity at charging stations with the effect being stronger for Open network stations.

4.4. Deductive and Inductive Topics

The deductive coding concerned a set of pre-defined topic codes. The inductive coding involved finding specific topic codes in the review text without definitions. The LLM found thousands of unique topic codes in the reviews but, after matching codes which were alternate phrasings of the same topic, it was discovered that a small subset of topics dominated the set. For petroleum station reviews, 16 inductive topics occurred in at least 1% of reviews. This number was 12 for both of the charging station populations. The purpose of the inductive coding was to provide specificity and context to the deductive coding. Figures 15, 16, and 17. The Figures show the relationship between inductive codes colored by sentiment with more positive links being green, more negative links being magenta, and more neutral links being gray. Unlike deductive codes, inductive codes are not required to have neutral phrasing and, as such, they tend to be strongly polarized in terms of sentiment.

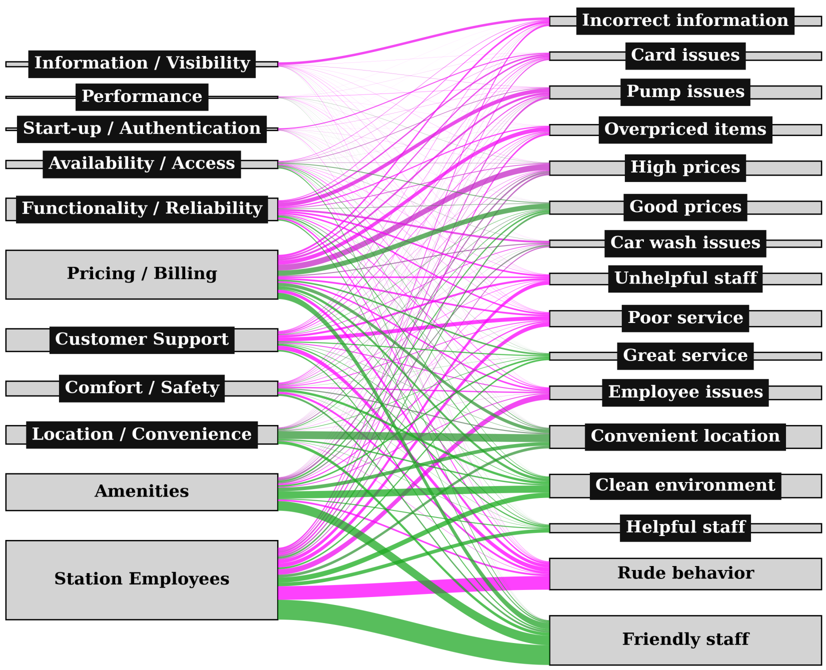

Figure 15: Deductive and inductive topic codes for Petroleum station reviews. Size of the boxes correspond to occurrence of topics. Width of connections correspond to co-occurrence of connected topics. Green colored connections indicate that reviews with both topics are majority positive, gray indicates majority neutral, and magenta indicates majority negative.

The most important aspects of the experience of ICEV fueling are those which do not relate to fueling. Of the 7 inductive codes commonly related to station employees, 3 (Friendly staff, Helpful staff, and Great service) are positive and 4 (Unhelpful staff, Poor service, Employee issues, and Rude behavior) are negative. These more or less balance in volume and explain the significant but weak regressed coefficient foe Station Employees seen in Table 1. Pricing / billing is also split between positive and negative inductive codes. This is in contrast to Location / Convenience and Amenities which mostly correspond to positive inductive topics such as Convenient location and Clean environment.

As established, the reviews for Open network charging stations are, generally, more negative than for the other populations. Functionality / Reliability is the most common deductive topic and about half of reviews which contain the Functionality / Reliability topic contain the inductive topic Charger broken. Charger broken also occurs in more than 20% of reviews for all other deductive topics other than Station Employees. There are also clear relationships between Performance and Slow charge and Information / Visibility and

App incorrect. Finally it is interesting to note that Availability / Access occurs more often with Charger broken than Site full.

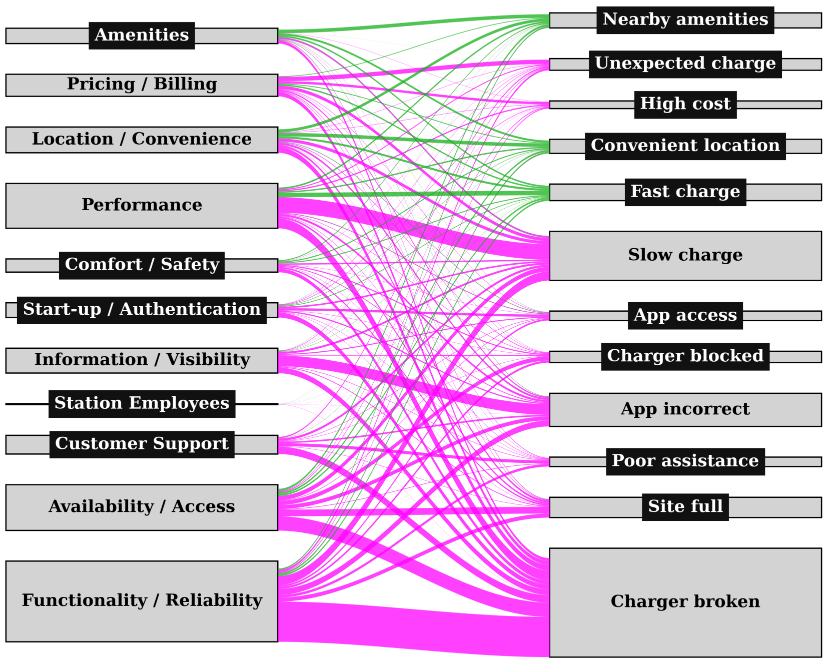

Figure 16: Deductive and inductive topic codes for Open charging station reviews. Size of the boxes correspond to occurrence of topics. Width of connections correspond to co-occurrence of connected topics. Green colored connections indicate that reviews with both topics are majority positive, gray indicates majority neutral, and magenta indicates majority negative.

Of the 12 common inductive topics for both charging station populations 9 are shared. These include 6 negative topics (Unexpected charge, App incorrect, Charger blocked, Slow charge, Site full, and Charger Broken) and 3 positive topics (Fast charge, Convenient location, and Nearby amenities). Much of the difference in user review between the charging station populations can be explained by the different rate at which these shared topics occur. In general, the positive topics occur more often for Proprietary stations and *vice versa*.

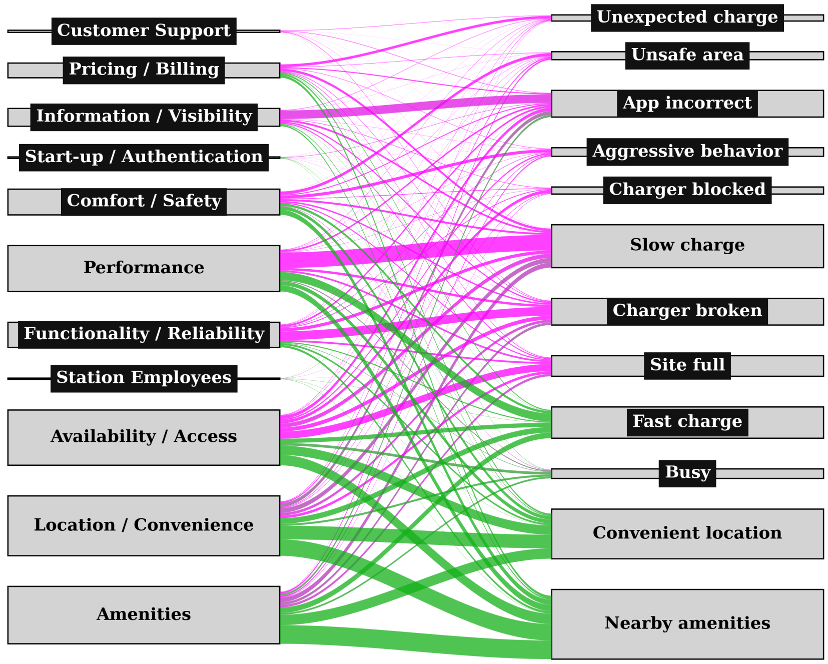

Figure 17: Deductive and inductive topic codes for Proprietary charging station reviews. Size of the boxes correspond to occurrence of topics. Width of connections correspond to co-occurrence of connected topics. Green colored connections indicate that reviews with both topics are majority positive, gray indicates majority neutral, and magenta indicates majority negative.

There is, of course, a difference between the actual occurrence of events and the frequency with which they are reported on by users of a ratings platform. What can be said is that users Proprietary and Open charging networks notice and care about many of the same things but report them occurring at different rates. It is unlikely that this stark difference in reporting does not, to some extent, reflect reality. It is probably the case that Proprietary stations are better located and are less effected by equipment unreliability (either due to higher reliability or due to larger station sizes). It is also probably the case that slow charge events, poorly functioning phone applications, and full stations, occur ar similar rates.

5. Conclusions

In the results section, thousands of reviews for the three populations of stations (petroleum, open, and proprietary) were analyzed from the points of view of ratings, sentiments, and topics. The data presented supports the following conclusions.

Charging and fueling are fundamentally different experiences. Although Google, to a certain extent, incentivizes users to leave reviews, the effort to do so is non-negligible. Users will, in general, be more inclined to comment on unusual or unexpected experiences. The nature of crowd-sourced reviews is that they will be polarized. For both fueling and charging stations, the majority of ratings were 1s and 5s and the majority of reviews were either positive or negative. These similarities are superficial. Thematic analysis shows that what users notice about the experience is substantially more similar between the charging station populations than between either population and fueling stations. As seen in Figure 12 topics related to service station employees and fuel prices dominate fueling station reviews where topics related to location, amenities, and equipment functionality, dominate for charging stations. This pattern is further borne out in the inductive topical analysis where many common inductive topics are shared between charging station populations and only one is shared with fueling stations.

Compared to Open stations, Proprietary charging stations are perceived as a better version of the same product. Comparing ratings between charging stations and petroleum stations is unfair as the expected experiences are different. However, comparing between charging station types is informative. As seen in 8, Proprietary network (Tesla, Rivian) station ratings are better, to a statistically significant degree, than most Open networks and not worse than any. As seen in Figures 10 and 11, review sentiments for open network charging stations skew negative while review sentiments for proprietary charging stations skew positive. Deductive topical analysis, as seen in Figure 12, and inductive topical analysis, as in Figures 16 and 17, indicate that BEV drivers notice negative experiences such as equipment malfunction and its consequences much more frequently at Open network charging stations, and positive experiences such as well located stations and desirable amenities more frequently at Proprietary network charging stations. However, it is important to note that the difference is in the rate of occurrence. Of the 12 most common inductive topics for each charging station population 9 are shared. The evidence presented suggests that Open and Proprietary network charging are seen as two levels of the same experience rather than as two separate experiences.

Open stations suffer from equipment unreliability and its knock-on effects. Topical analysis of open network charging station reviews shows the prominence of the Functionality/Reliability theme and the Charger Broken topic. Further, the Functionality/Reliability theme was particularly interactive for open network stations and the Broken Charger topic in a substantial portion of negative reviews touching on all themes. Open network charging stations tend to be small. Nearly 80% of open network charging stations feature 5 or fewer ports as seen in Figure 1. This contrasts to proprietary networks where nearly every station has at least 5 ports and roughly half of stations have more than 12. The knock-on effects of equipment failures are reduced station capacity and efficiency which may lead to queuing or drivers relocating to other stations. The effect of lost chargers on station capacity is nonlinear and larger stations are less effected for the same percentage loss of chargers. It's highly informative that, of the negative reviews left at open network charging stations, broken chargers were a major driver regardless of theme. It's also relevant that this dynamics was not present for proprietary chargers. This gulf may be beginning to close. As shown in Figure 18, newer Open stations are larger and this is a continuing trend.

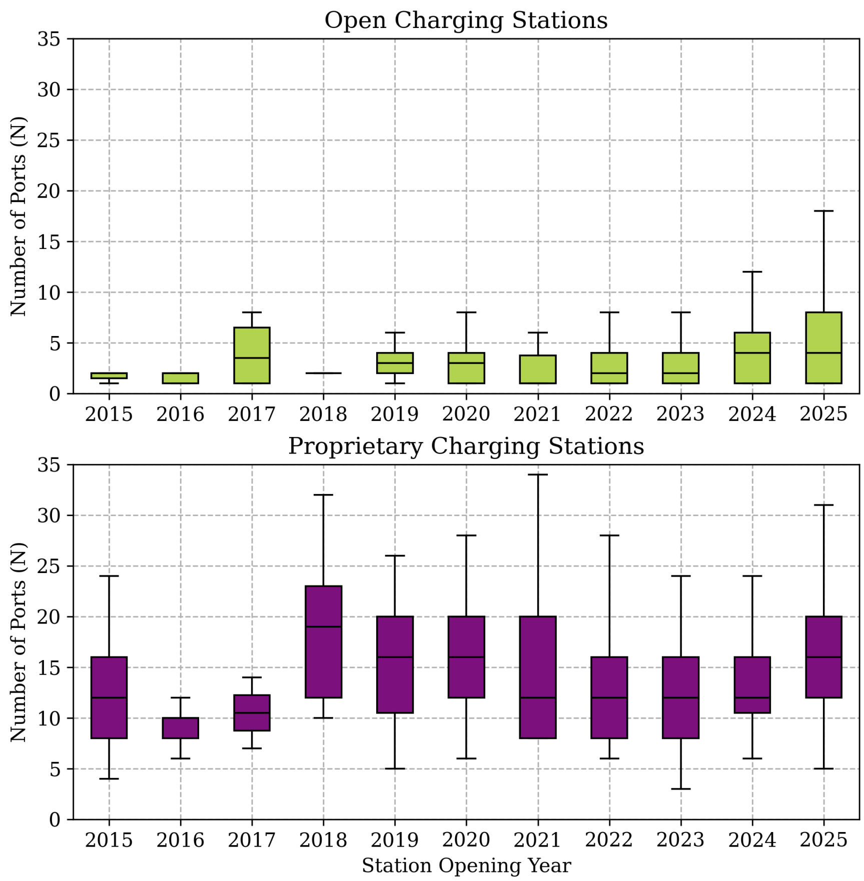

Figure 18: Boxplots of station size (number of ports) by opening year for Proprietary and Open networks

BEV drivers view nearby amenities as an important aspect of station location quality. For ICEV drivers, the fueling experience clearly boils down to the price of petroleum and the quality of the convenience store as seen in Figure 12. In California, drivers pump their own petroleum and station employees are almost entirely encountered inside the convenience store. Thus the theme Station Employees fits well with, and often co-occurs with the theme Amenities. Most charging stations are not attached to a convenience store and have no employees. In the context of BEV charging Amenities refers to surrounding businesses. As mentioned previously, open network charging station reviews are slanted negative and often concerned with equipment issues. Proprietary charging station reviews very often cover the themes of Location/Convenience and Amenities as well as the topics Convenient Location and Nearby Amenities. Further, these themes and topics often co-occur, especially in positive reviews. A possible interpretation is that better amenities improve the convenience of a location. A location becomes convenient if drivers can work a charge event into their daily routines by utilizing a nearby amenity. However, the charging station operator does not won or, in any way, control what amenities are nearby and, thus, the appeal of a station can change over time.

References

Alternative Fuels Data Center (AFDC), 2023. Alternative fuels data center home page. URL: https://afdc.energy.gov/.

Amsterdamer, Y., Milo, T., 2015. Foundations of crowd data sourcing. SIGMOD Rec. 43, 514. URL: https://doi.org/10.1145/2737817.2737819, doi:10.1145/2737817.2737819.

Asensio, O.I., Alvarez, K., Dror, A., Wenzel, E., Hollauer, C., Ha, S., 2020. Real-time data from mobile platforms to evaluate sustainable transportation infrastructure. Nature Sustainability 3. URL: https://doi.org/10.1038/s41893-020-0533-6, doi:10.1038/s41893-020-0533-6.

Ashby, C., Weir, D., Fussey, P., 2025. Understanding public views on electric vehicle charging: A thematic analysis. SSRN Preprint URL: https://dx.doi.org/10.1209/0295-5075/81/48005, doi:http://dx.doi.org/10.2139/ssrn.4945335.

California Energy Commission (CEC), 2022. California retail fuel outlet annual reporting (cec-a15) results. URL: https://www.energy.ca.gov/data-reports/energy-almanac/transportation-energy/california-retail-fuel-outlet-annual-reporting.

Dong, L., Hardman, S., Bunch, D., Mabit, S., Chakraborty, D., 2024. Cost sensitivity and charging choices of plug-in electric vehicle drivers–a stated preference study .

Dunivin, Z.O., 2025. Scaling hermeneutics: a guide to qualitative coding with llms for reflexive content analysis. EPJ Data Science 14, 28. URL: https://doi.org/10.1140/epjds/s13688-025-00548-8, doi:10.1140/epjds/s13688-025-00548-8.

Gamage, T., Tal, G., Jenn, A.T., 2023. The costs and challenges of installing corridor dc fast chargers in california. Case Studies on Transport Policy 11, 100969. doi:https://doi.org/10.1016/j.cstp.2023.100969.

Gilardi, F., Alizadeh, M., Kubli, M., 2023. Chatgpt outperforms crowd workers for text-annotation tasks. Proceedings of the National Academy of Sciences 120, e2305016120. URL: https://www.pnas.org/doi/abs/10.1073/pnas.2305016120, doi:10.1073/pnas.2305016120, arXiv:https://www.pnas.org/doi/pdf/10.1073/pnas.2305016120.

Guo, H., Ou, S., Jing, H., Qi, H., Shi, L., 2025. Decoding user experience: A study of public ev charging stations based on amap comments. doi:10.4271/2025-01-8115.

Ha, S., Marchetto, D., Dharur, S., Asensio, O., 2021. Topic classification of electric vehicle consumer experiences with transformer-based deep learning. Patterns 2, 100195. doi:10.1016/j.patter.2020.100195.

Hardman, S., Chakraborty, A., Hoogland, K., Sugihara, C., Helveston, J.P., Jensen, A.F., Jenn, A., Jochem, P., Plötz, P., Sprei, F., Williams, B., Axsen, J., Figenbaum, E., Pontes, J., Tal, G., Refa, N., 2025. Demand-side challenges and research needs on the road to 100% zero-emission vehicle sales. Progress in Energy 7, 022001. URL: https://dx.doi.org/10.1088/2516-1083/ad9f1a, doi:10.1088/2516-1083/ad9f1a.

Hybrid - EV Committee, 2024. SAE Electric Vehicle and Plug-in Hybrid Electric Vehicle Conductive Charge Coupler. doi:https://doi.org/10.4271/J1772$_2$02401.

Jenn, A., 2025. What is the business case for public electric vehicle chargers? Transportation URL: https://doi.org/10.1007/s11116-025-10651-4, doi:10.1007/s11116-025-10651-4.

Jenn, A., Chakraborty, A., Hardman, S., Hoogland, K., Sugihara, C., Tal, G., Helveston, J., Rich, J., Jochem, P., Plötz, P., Sprei, F., Williams, B., Axsen, J., Figenbaum, E., Pontes, J., Refa, N., 2025. Supply-side challenges and research needs on the road to 100% zero-emissions vehicle sales. Progress in Energy 7, 022002. URL: https://dx.doi.org/10.1088/2516-1083/ada199, doi:10.1088/2516-1083/ada199.

Kim, H., Kim, D.W., Kim, M.K., 2022. Economics of charging infrastructure for electric vehicles in korea. Energy Policy 164, 112875. URL: https://www.sciencedirect.com/science/article/pii/S0301421522001008, doi:https://doi.org/10.1016/j.enpol.2022.112875.

Lee, J.H., Cho, M., Tal, G., Hardman, S., 2023. Do plug-in hybrid adopters switch to battery electric vehicles (and vice versa)? Transportation Research Part D: Transport and Environment 119, 103752. URL: https://www.sciencedirect.com/science/article/pii/S1361920923001499, doi:https://doi.org/10.1016/j.trd.2023.103752.

Li, H., Han, Q., Bai, X., Zhang, L., Wang, W., Chen, W., Xiang, L., 2024. Electric vehicle charging station recommendations considering user charging preferences based on comment data. Energies 17. URL: https://www.mdpi.com/1996-1073/17/21/5514, doi:10.3390/en17215514.

Liu, X., Zambrano, A.F., Baker, R.S., Barany, A., Ocumpaugh, J., Zhang, J., Pankiewicz, M., Nasiar, N., Wei, Z., 2025. Qualitative coding with gpt-4: Where it works better. Journal of Learning Analytics 12, 169–185. URL: https://learning-analytics.info/index.php/JLA/article/view/8575, doi:10.18608/jla.2025.8575.

Madina, C., Zamora, I., Zabala, E., 2016. Methodology for assessing electric vehicle charging infrastructure business models. Energy Policy 89, 284–293. URL: https://www.sciencedirect.com/science/article/pii/S0301421515302202, doi:https://doi.org/10.1016/j.enpol.2015.12.007.

Rabinowitz, A.I., Smart, J.G., Coburn, T.C., Bradley, T.H., 2023. Assessment of factors in the reduction of bev operational inconvenience. IEEE Access 11, 30486–30497. doi:10.1109/ACCESS.2023.3255103.

Ramadoss, T.V., Lee, J.H., Davis, A.W., Hardman, S., Tal, G., 2025. Classifying electric vehicle adopters and forecasting progress to full adoption. npj Sustainable Mobility and Transport 2. URL: https://doi.org/10.1038/s44333-025-00049-1, doi:10.1038/s44333-025-00049-1.

Robbennolt, D., Hardman, S., Firestone, J., Bhat, C.R., 2025. A model of electric vehicle adoption and motivating reasons for adoption. Transportation Research Part D: Transport and Environment 146, 104906. URL: https://www.sciencedirect.com/science/article/pii/S1361920925003165, doi:https://doi.org/10.1016/j.trd.2025.104906.

Than, N., Fan, L., Law, T., Nelson, L.K., McCall, L., 2025. Updating the future of coding: Qualitative coding with generative large language models. Sociological Methods & Research 54, 849–888. doi:https://doi.org/10.1177/00491241251339188.

of Transportation Statistics, B., 2025. National highway system (nhs). URL: https://geodata.bts.gov/datasets/usdot::national-highway-system-nhs/about, doi:https://doi.org/10.21949/1529838.

US Census Bureau, 2025. Tiger/line shapefiles and tiger/line files technical documentation. URL: https://www.census.gov/programs-surveys/geography/technical-documentation/complete-technical-documentation/tiger-geo-line.html.

Zheng, Y., Keith, D.R., Wang, S., Diao, M., Zhao, J., 2024. Effects of electric vehicle charging stations on the economic vitality of local businesses. Nature Communications 15, 7437. URL: https://doi.org/10.1038/s41467-024-51554-9, doi:10.1038/s41467-024-51554-9.

Ziems, C., Held, W., Shaikh, O., Chen, J., Zhang, Z., Yang, D., 2024. Can large language models transform computational social science? URL: https://arxiv.org/abs/2305.03514, arXiv:2305.03514.