Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 00 (2018) 000–000

Transportation Research Procedia

www.elsevier.com/locate/procedia

World Conference on Transport Research - WCTR 2019 Mumbai 26-31 May 2019

An empirical investigation of electric vehicle user charging patterns and charging station utilization in the city of Berlin

Mitchell Klotz*, Hanno Friedrich, Benjamin Dahmen

Kühne Logistics Univeristy, Großer Grasbrook 17, 20457 Hamburg, Germany

Abstract

Research relating to Electric vehicle (EV) charging is still in its early stage, as only the last 10 years have brought significant development. Particularly, there is a need for realistic representation of charging behavior. This empirical study gathered and analyzed "real-world" public charging data from a prominent European charging network owner and operator. The researchers demonstrated a clustering method used to analyze the charging behavior of users in an urban environment, using Berlin as a reference. Furthermore, linear regression analyses regarding the utilization of charging stations were performed to verify a pair of locational factors that were correlated with the performance of certain stations. Key findings from the study included: a high number of users and an incredibly high number of charge events from users in the identified "Car-Sharing" cluster; a significant number of users and charger events from users in the "At Home Chargers" cluster; a comparable number of charge vents performed by users in the "Work Chargers" and "Private Roamers" clusters; and, within certain clusters, statistically significant results for linear regressions on station utilization, using points-of-interest within 300 meters as well as employment level in the area as regressors. Through this investigation, the we intended to support all associated parties with the development of public charging infrastructure and to contribute a foundation for further studies regarding the public charging behavior of EV users.

Peer-review under responsibility of WORLD CONFERENCE ON TRANSPORT RESEARCH SOCIETY.

Keywords: Charging behavior; Charging infrastructure; Clustering; Battery electric vehicles; Electromobility

1. Introduction

Beginning in the 1990s, emissions regulations and government incentives in countries across the globe have helped spur the most recent, and strongest, resurgence of the electric vehicle (EV) industry (Thompson, 2017). As of today, in the UK, Belgium, France, and Germany, significant purchase grants are awarded for EVs (Pressman, 2017). In the United States, tax credits are offered to help subsidize the purchase price of a new EV (Electric Vehicles, n.d.). In

* Corresponding author.
E-mail address: Mitch.Klotz@the-klu.org

Peer-review under responsibility of WORLD CONFERENCE ON TRANSPORT RESEARCH SOCIETY

currently the world's largest market for EVs, the Chinese government also offers substantial subsidies for new purchases (Zhang & Chen, 2018). As federal (and state) governments throw their weight behind the EV industry, these types of incentives are proving to be quite effective for widespread adoption (Harrison & Thiel, 2015).

Additionally, stricter emissions regulations mandated by these governments are encouraging automobile manufactures to invest heavily in EV research and development. As a result, relentless innovation is not only driving down the purchase price of these vehicles, but also resulting in EVs that are far superior to their internal-combustion-engine (ICE) counter-parts in terms of performance specs such as acceleration, safety, and operational maintenance required. Regardless which of these two – the encouraging government incentives, or attractive performance features – are primarily responsible, EV sales are taking off worldwide. According to McKinsey, 2016 experienced a 57% increase in EV sales over the previous year. Still, these 1.3 million EVs sold represented only about 1% of all new passenger-vehicles sold globally. The firm estimates this share to increase exponentially, up to 20-25% of car sales globally by the year 2030 (Chatelain et al., 2018). The International Energy Agency (IEA) expects 125 million EVs on the road by that time. According to Bloomberg, even major oil companies are anticipating significant penetration of EVs in the relatively near future (DiChristopher, 2018). Total SA, one of the world's largest producers of oil and gas, claims that up to one third of new car sales in 2030 will be accounted for by EVs (Randall, 2017).

The anticipation of such rampant growth in sales of these vehicles is leading a variety of companies to enter the market for providing the charging infrastructure necessary to support this development. Automobile manufacturers themselves, electric utility companies, and third-party charging network operators alike are establishing themselves as pioneers in this potentially lucrative business (Hoium, 2017). Regardless of their nature, in order to be successful, these companies need to understand their customers and effectively introduce and operate charging stations at strategic locations that make sense.

This can, and will be, a challenge, as recent research has demonstrated that EV users have a different mindset toward "recharging" their vehicle, as opposed to "refueling" it (Daubitz & Kawgan-Kagan, 2014). On one hand, Philipsen et al. (2016), among others, claim that "limited charging options… contribute to the so-called range anxiety." Research by Weldon et al. (2016) and Xu et al. (2017) at least partially confirm this psychological fear, demonstrating risk-aversion of EV users that lead to charging of the vehicle far more often than necessary. On the other hand, as Joseph Nagle of EverCharge, an EV infrastructure company based in California, stated just this month, "adding more chargers is essential to the continued growth of electric vehicles […; however,] continually installing more and more public chargers is simply not a sustainable or sensible solution." In the same article, he emphasizes that what EV drivers really want is access to at-home or at-work charging (Nagle, 2018). Further support of these conclusions comes from empirical work done by Jabeen et al. (2013), who emphasized that "drivers preferred to charge [their] EV at home or work rather than at a public charging station."

The goal of this study was to support those companies who are developing charging infrastructure by providing a clearer understanding of customer behavior. Current research is rather limited to stated preference surveys and analyses of EV trials and other experimental projects. Our purpose was to provide a practical analysis of natural charging behavior in an urban environment. As, to our knowledge, the first to use real data on public charging in a city, we contributed an analysis of this charging behavior and identification of different types of EV charging station users based on a clustering method. Furthermore, via regression analyses regarding the utilization of charging stations, we verified a pair of locational factors that are correlated with the performance of certain stations. This study should provide preliminary benefit to industry practitioners, as well as serve as a foundation for consequent research in the field.

2. Method

In a systematic review of the literature on EV use modeling, Daina at al. (2017) emphasize a shortage of work on "the detailed spatial and temporal patterns of EV use and charging behavior." The authors of that study quite perfectly underscore the purpose of the research here, identifying that "limitations of existing data are particularly significant and although some of these limitations may be partially overcome by making use [of] stated preference data from hypothetical choice experiments, there remains a need for significantly improved datasets on real world charging behavior both in the contexts of EV trials and demonstrations and, critically, in periods of normal operation." By obtaining and analyzing one year of actual, up-to-date charge event data from EV users of a major European charging infrastructure company's charging stations in Berlin, the research and results described below provide at least one instance of satisfying this necessity.

2.1. Data gathering and cleaning

The company from which the data was obtained is a leading Charge Point Operator (CPO) in Europe, with a network of several hundred charging stations that span across Germany, Belgium, and the Netherlands. For the purpose of this research, the focus was on charging behavior in a city. Therefore, a dataset was obtained from the company's most developed location – Berlin. The dataset included all charge events on record at the 234 charge points located within the city limits. In total, 77,370 individual charge events were recorded in the raw data set. The relevant attributes associated with each event were the following: Session ID, Start Date, Start Time, End Date, End Time, Charging Duration, Consumed Energy (kWh), Average Power (kW), Charge Pole ID, Latitude, Longitude, and EVCO ID (an individual identifier of the RFID tag used to initiate the charge event).

Similar to Xu et al. (2017) and Morrissey et al. (2016), who performed an extensive analysis of public charge event data in Japan and Ireland, respectively, a cleaning process had to be undertaken before beginning the analysis. To begin with, out of the company's 234 charging stations in the city, 10 of these stations are "fast" Direct Current (DC) chargers. These stations have higher electrical power, and therefore different charging speed capabilities when compared to standard AC chargers. So, for the sake of this study, all charge events recorded at these stations were eliminated from the dataset, allowing an accurate analysis of charging at all comparable Alternating Current (AC) charging stations. Furthermore, several charge events resulted in immaterial charge durations. Any event with a charge duration of less than 3 minutes was eliminated. These events are most likely associated with a simple user error involving mishandling of the charging equipment or other confusion at the charging point. Additionally, a few charge events had to be discarded because of impossible recorded values for Average Power (kW), such as 500 and 880. Relating to the attribute EVCO ID, a small number of charge events were recorded without a value for this field. Also, there were events that were recorded by an EVCO ID associated with an account used for testing. These events were also eliminated from the dataset. Finally, the authors of this study decided to analyze one full year of the charging data. Therefore, the charge events from the most recent 12 months (April 2017 through March 2018) were included in the final dataset, as this time period encompassed the most significant number of charge events. At the conclusion of "cleaning", the final dataset used for the analyses included 38,824 charge events performed by 2,202 unique users occurring at 228 different charging locations throughout the city.

2.2. Geospatial analysis using QGIS

After cleaning the data, the next step in preparing the data for the study's investigation involved geospatial analysis relating to the individual charging stations. For this step, the open-source geographic information system software QGIS was utilized. The charge event data acquired from the company already included longitude and latitude attributes, therefore mapping the individual stations in space was straightforward. Fig. 1(a) on the next page shows a visualization of this step.

In order to perform some of the regression analyses later on, determining points-of-interest (POI) nearby each charging station was essential. The idea was to define the number of these POIs within a reasonable walking distance from the stations' locations, allowing for regressions with charge point utilization (number of charge events) as the dependent variable, and number of these POIs as the independent variable. While travel purpose and pedestrian demographics could have an impact, 400 meters is commonly used as an acceptable walking distance in transportation and public health studies (Yang & Diez-Roux 2012). Keeping this 400-meter distance in mind, buffer zones – in the form of circles with a 300-meter radius and the station as the center point – were created for each charging station. Fig. 1(b) shows each station with this buffer. The reason for the 300-meter radius, as opposed to 400 meters, was that pedestrians in the city simply don't necessarily walk in a straight line. Travelling on curved streets and turning onto perpendicular ones, a walk covering 400 meters of distance is likely to result in the pedestrian reaching their destination that is actually only around 300 meters from the point of origin.

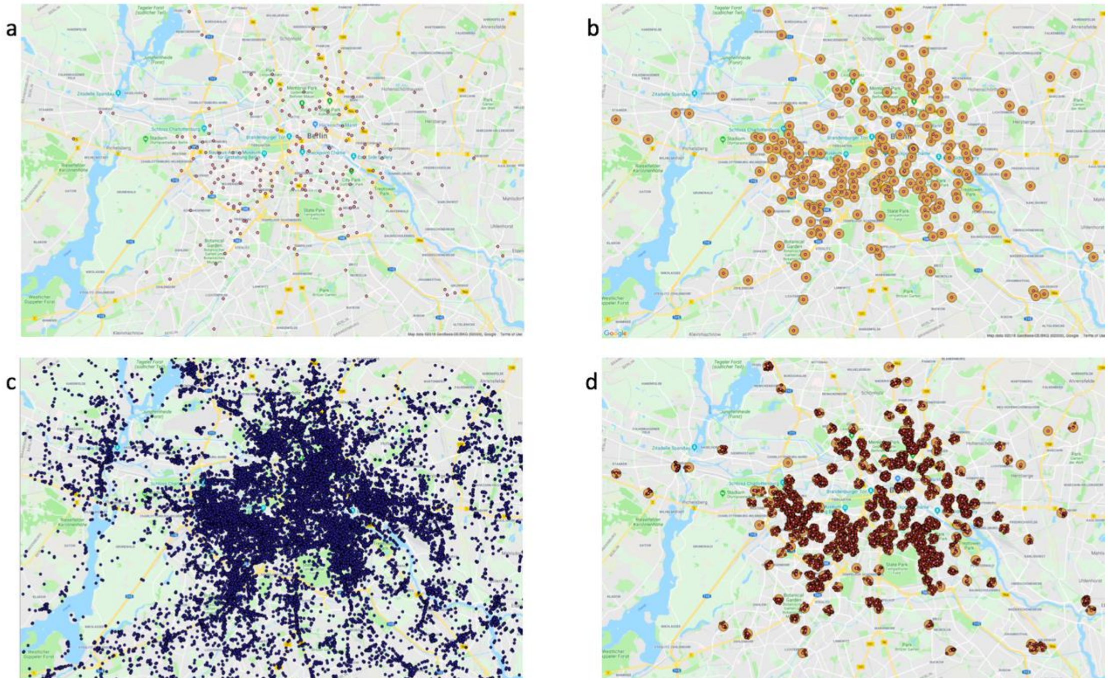


Figure 1. (a) charging stations; (b) with 300m buffer; (c) all points-of-interest; (d) relevant points-of-interest within buffer zones.

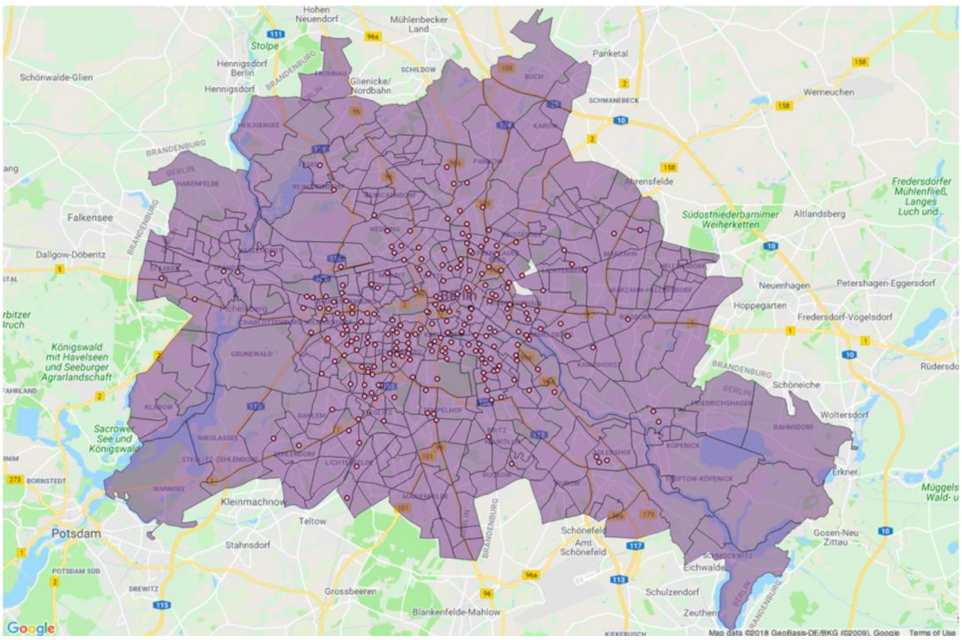


Figure 2. Map of Allego stations and Berlin Planungsräume (PORs).

The proceeding step was then to establish how many POIs were located within the buffer zone of each station. To do so, a POI shape file for the city of Berlin was obtained from Geofabrik GmbH’s free download server, which makes extracts from the OpenStreetMap project that are normally updated on a daily basis (“OpenStreetMap Data Extracts”, n.d.). This original file included a total of 53,435 POIs throughout the city. However, for the purpose of this study, some subjective filtering was performed. Items under categories such as “bench” or “toilet” were purposely excluded. A map of all POIs that were considered can be seen in Fig. 1(c). These included categories such as shops, restaurants, tourist info centers, stadiums, and parks. Finally, using a “count points in polygon” algorithm under the processing toolbox of QGIS, the number of POIs in each buffer zone (within 300 meters of each charging station) was determined. A map of each station, its 300-meter buffer zone, and the POIs located within this buffer zone can be seen in Fig 1(d).

Lastly, the authors of this paper were interested in examining a correlation between the utilization of a charging station and the level of employment in the area. To allow for this, we contacted the statistics office of Berlin-Brandenburg (Amt für Statistik Berlin-Brandenburg) and acquired from them a dataset that listed the 447 Planungsräume (singular: Planungsraum; PLR – the lowest-level division of planning space in each municipality) in the city along with the number of businesses, as well as the total number of employees registered in each of those PLRs. Each station was then tagged with the employment data associated with the PLR in which it was located. On the previous page, Fig. 2 shows a map of each station along with the 447 PLRs in the city. As depicted, only a portion of these PLRs were relevant to the analysis due to the fact that there were no charging stations located in several of them.

2.3. Divisive hierarchical clustering with manual splits

For the first part of the analysis, in order to analyze EV user charging behavior in the city, a divisive (“top-down”) approach to hierarchical clustering was used as the method. We began with the entire dataset of charge events and had the intention of sorting users into distinguishable clusters based on similar charging behaviors. To do so, rather than using a predefined algorithm, we performed manual splits of users at each level of clustering. These splits were made based on the evaluation of certain elements of their charging behavior. As the individual “points” in the data set corresponded to unique charge events, and our goal was to rather sort users, it was necessary to perform calculations and analysis from this data with respect to each individual user in order to determine these elements. Important to note is that throughout this paper, the term “user” does not necessarily refer to an individual person. Rather, it refers to a unique EVCO ID, or user ID, as given in the original dataset, associated with the charging account used to perform the charge events.

The reason for manual splitting was that this method had never been used before in this context. Without examples of effective heuristics and metrics for a sorting algorithm, it was necessary to allow a degree of subjectivity in determining how to split the clusters. For splitting, histograms of certain charging behavior elements, such as number of distinct charging stations for each user, were used to support the determination of criteria. Manual splits were made based on observations of clusters or obvious thresholds in these figures. Our intuition also played a role, and in one instance of sorting, no histogram was generated. Rather, we considered what previous sorting had been made to reach the current level and intuitively assigned ad-hoc criteria to perform sorting at that level. Each level of sorting is explained in further detail starting in the next paragraph. The process concluded when it was deemed that the identification and labelling of discernable end clusters provided relevant, valuable insight and no further splitting was necessary.

Using the diagram in Fig. 3 to help visualize this process, we started with the complete data set at node 1. Here, users were sorted based on the number of charging events they performed over the one-year period. In order to generalize behavior and make conjectures about charging habits, the authors considered that the user had to have performed at least 10 charge events. Therefore, users with less than 10 charge events were sorted to node 2, while users with 10 or more charge events were sorted to node 3. Fig. 4 visualizes the histogram of users’ sum of charge events and sorting based on the criterium of 10 charge events. In total, 64.21% of users were sorted to node 2 and 35.79% of users were sorted to node 3.

Having sorted a cluster of users with enough charge events to begin analyzing, the first element of charging behavior that we were interested in was how many different charging stations each user utilized relative to the total number of charge events they performed. It was expected that this insight would give an indication as to which users had particular patters and preferences regarding certain stations, and which users were simply charging everywhere. We were interested in digging deeper on those users who did have patterns of using particular stations. To establish a criterium, a simple ratio of charge-events-to-distinct-stations-used was applied for each user. For example, if a user

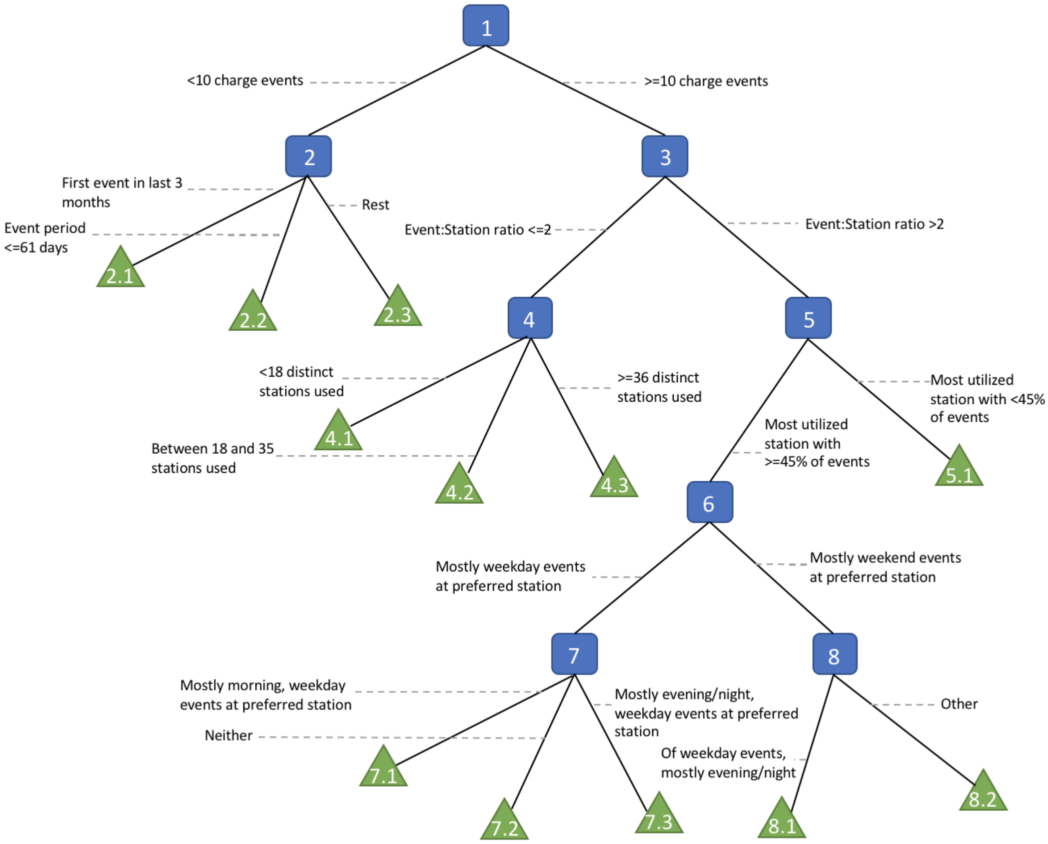


Figure 3. Clustering Tree.

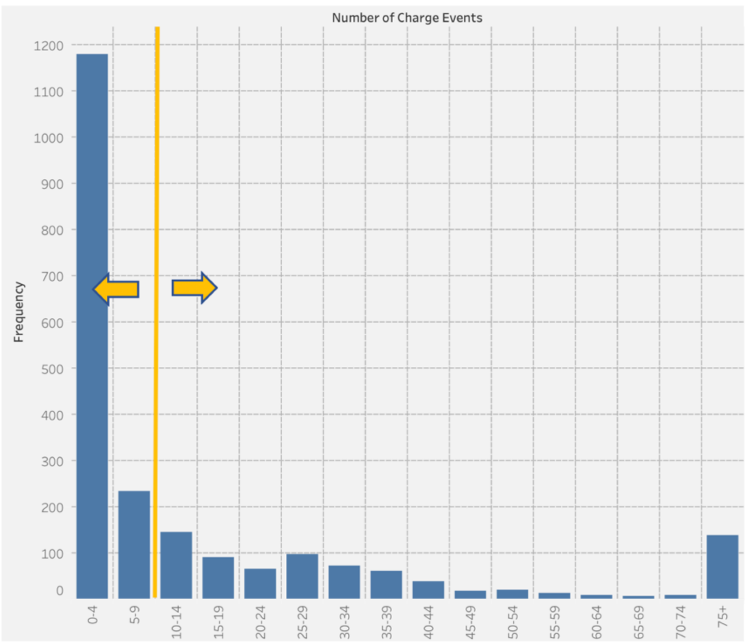


Figure 4. Histogram of users' number of charge events

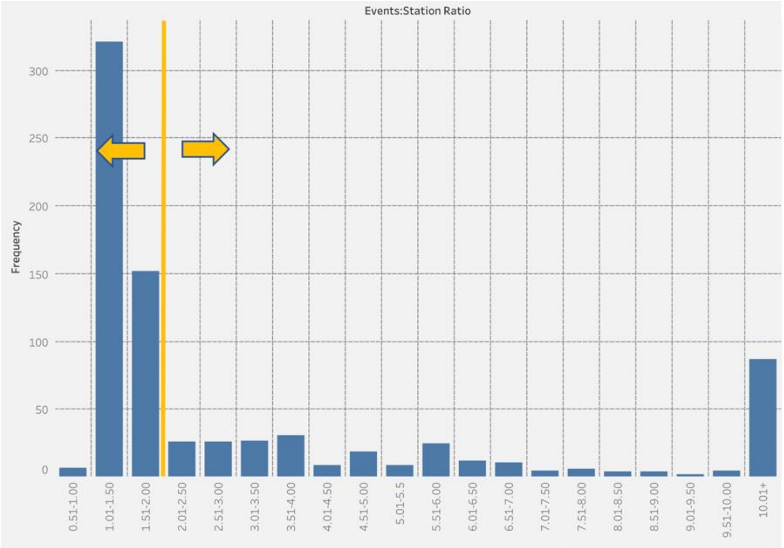


Figure 5. Histogram of node 3 users' charge events-to-station ratio.

performed a total of 70 charge events during the period using 10 different stations, the user had a ratio of 7. If a user performed 120 charge events at 80 different locations, the user had a ratio of 1.5. Looking at the histogram in Fig. 5, there was an obvious cluster of users with a ratio less than or equal to 2.00. Therefore, the manual split was made based on this notion. Out of the 788 users before the split, the 480 users with a ratio less than or equal to 2.00 were sorted to node 4 and the 308 users with a ratio greater than 2.00 were sorted to node 5.

Recapping the path to node 4, these were users with at least 10 charge events over the analyzed period, and a charge event-to-station ratio of 2.00 or less. Again, this low ratio implied that the user was utilizing a high number of different charging stations, but rarely concentrating a number of events at one particular station. Using Tableau to help visualize this behavior, Fig. 6 shows a map of events for one of these users as an example.

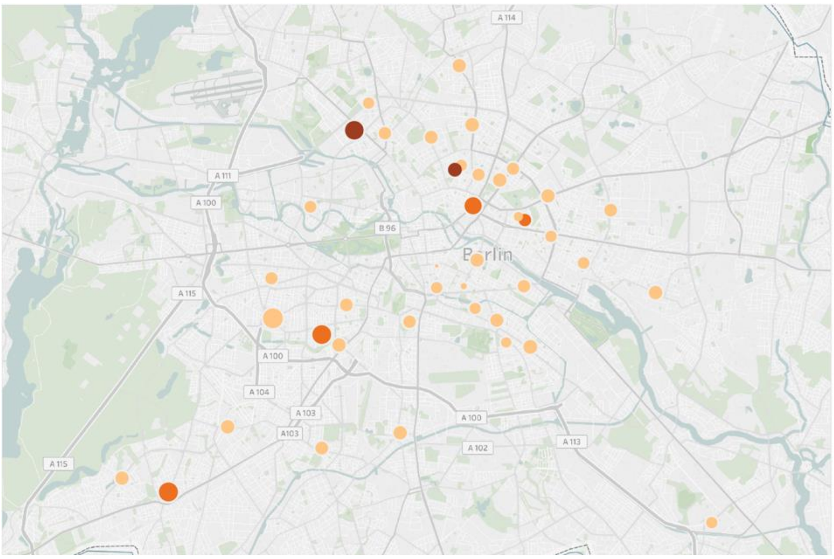


Figure 6. Example of one node 4 users' station utilization. Each circle represents a charging station, the color tint indicates the relative number of events performed at the station, and the size indicates the relative consumed energy (kWh) at the station.

We were aware that, as of last year, three-quarters of electric vehicles in Berlin were owned and operated by businesses ("Berlin: Capital of electric mobility," 2017). The article from which this information came from noted the city's local water company as an example, with nearly 10% of its entire fleet composed of electric vehicles. Furthermore, it described a handful of car-sharing companies operating in the city who use electric vehicles as part of their fleet. Given the general nature of car-sharing, as well as the shared usage of company cars, we assumed the charging events associated with these vehicles to resemble the visual in Fig. 6. However, we were also conscious of the idea that individual EV owners might also behave in such a manner, performing charge events at stations located throughout the city, although presumably to a milder extent because of the lower number of different individual drivers operating the vehicle.

As a potential distinguishing element, we analyzed the total number of charging stations utilized by each user in this group. Fig. 7 displays the histogram. As seen, there was an evident cluster near the left side of the distribution. Specifically, the cluster was observable between 18 and 35 on the x-axis. Boundaries were therefore established at these points and users were sorted into end clusters, depicted by the green triangles in Fig. 3, based on the following criteria. Users with more than 35 distinct charging stations were sorted to 4.3 and labelled as the "Car-Sharing" cluster. Users with 18 to 35 distinct charging stations were sorted to 4.2 and labelled as the "Business Roamers" cluster. And users with less than 18 distinct charging stations were sorted to 4.1 and labelled as the "Private Roamers I" cluster. We justify the labelling of these clusters by presuming that such a high number of different charging stations would not reasonably be utilized by one or two individuals, as would be the case with the private ownership and charging of an EV associated with one user ID. It must be that the partition of users with over 35 distinct stations was the result of car-sharing, where presumably numerous different individuals performed charging events using the same user ID. Similarly, the case of business-owned "company cars" would explain the middle cluster of users in the histogram, where it was assumed that multiple individual operators of the same vehicle, however a lower number, used the same user ID to perform charge events. And finally, users with less than 18 distinct stations were assumed to simply be one or two, in the case of a shared household vehicle, individuals performing charge events at a rather high number of different stations throughout the city.

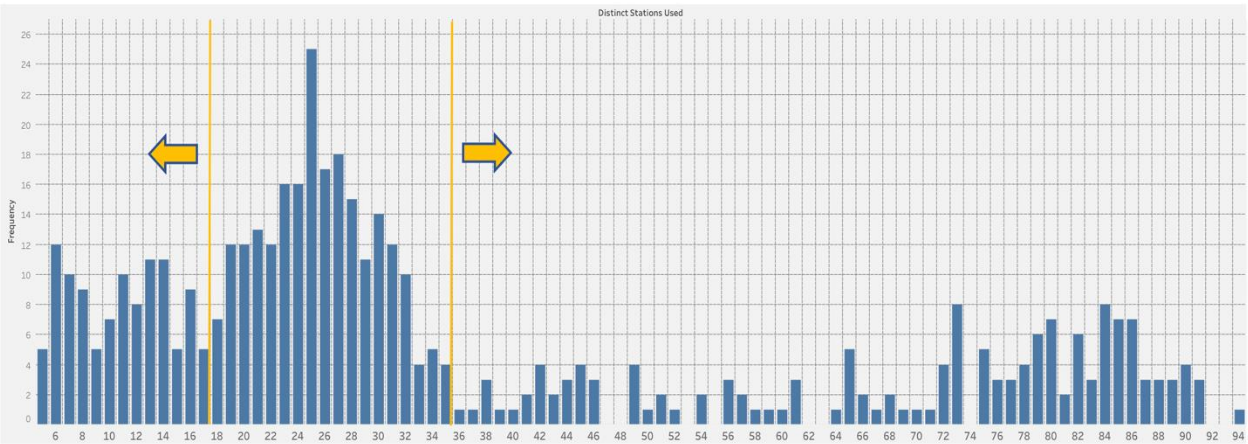


Figure 7. Histogram of node 4 users' number of distinct stations used.

Referring now to node 5, these were users with at least 10 charge events and a charge event-to-station ratio of more than 2.00. This relatively higher ratio was an indication that these users had some degree of concentration of charging events at a limited number of different locations, potentially revealing some sort of pattern. At this point, we were concerned with digging deeper and determining whether users had an evident preference for charging at any one station in particular. Here, the criterium for the next level of sorting considered the highest utilized station of each user. What is meant by highest utilized station is the station at which the user performed the highest number of charge events relative to the other stations used. For each user, a calculation was made to determine the percentage of that user's events which were executed at its highest utilized station. Again, a histogram of this element is displayed in

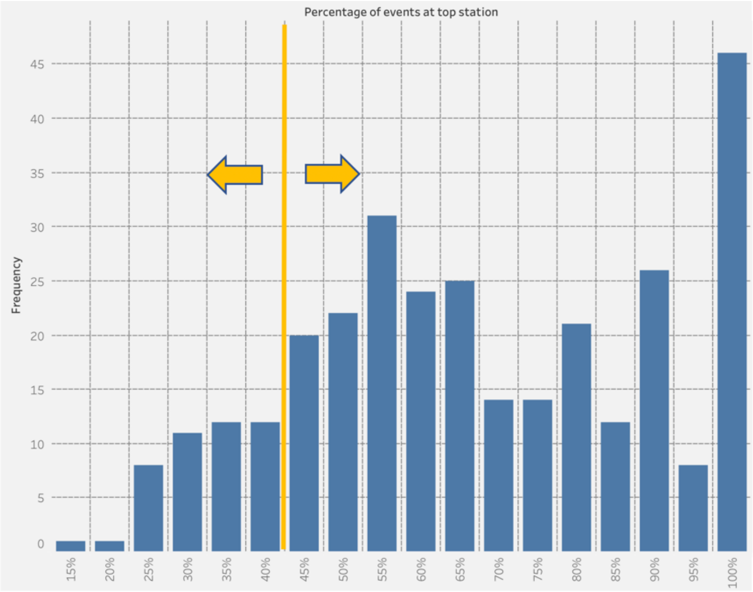


Figure 8. Histogram of node 5 users' percentage of events performed at their most utilized station.

Fig. 8. Using the graphic as reference, a manual split was made at the 45% mark. It was considered that if the highest utilized station of a user encompassed 45% or more of the user's charge events, this station was an obvious preference for the user, and further investigation should be done regarding these users. Users for which this was the case were then sorted to node 6, while users who did not have an obvious preferred location (no station that was the location of more than 45% of events) were sorted to 5.1 and labelled as the "Private Roamers II" cluster, with similar justification as described for the labelling of the "Private Roamers I" cluster. All of the users tagged in the "Private Roamers II" cluster utilized 19 or less different charge stations. Nevertheless, for each of these users, none of the stations were the location of a suggestively high proportion of the user's charge events.

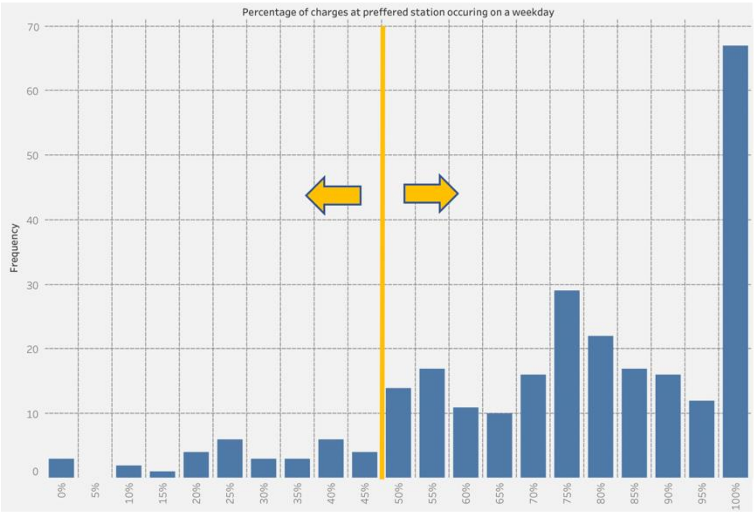


Figure 9. Histogram of node 6 users' percentage of charges at preferred station that occurred during the week.

At node 6, having recognized users with a preferred station, the criterium for sorting at this level was whether the majority of the user's charge events performed at their preferred stations occurred during the week or the weekend. The reasoning behind this was that, through this and one further level of sorting, we anticipated easily distinguishing and identifying end clusters of users as "At Home Chargers" and "Work Chargers." These clusters were of particular interest because, as mentioned in the introduction of this paper, claims have been made that EV users prefer to either charge at home or at work. Although the data in this study pertains to charging performed at public stations (no access to private work or home charging data), we still imagine that users in an urban city environment have limited access to either of these private options. Therefore, it was supposed that at work and at home preferences could still be revealed through behavior at these public charging stations.

Returning back to the criterium used for sorting at this level, considering only each user's events that occurred at their preferred station, we calculated the percentage of these events that occurred during the week. Fig. 9 shows the distribution of these percentages and where the cutoff was made for sorting. Users with a percentage lower than 50% were sorted to node 7, while users with a percentage of at least 50% were sorted to node 8.

At node 7, having distinguished users who performed at least 10 charge events, had a charge event-to-station ratio greater than 2, a preferred charging location, and the majority of events at this preferred location performed during the week, the final criterium here was related to the time of day at which the weekday charge events at the user's preferred station were performed. As previously mentioned, we intended to distinguish and identify clusters of users who were demonstrating a preference for charging while at work or at home, using public stations to do so. The we believed this final sorting criterium would allow for a clear distinction and insightful understanding of the number of users in each of these clusters.

So, for each user, now considering only the events at their preferred location that occurred during the week (excluding events at that station that occurred on a Saturday or Sunday), two calculations were made. First, for each user, the percentage of these events that started in the morning between 5:00 and 10:00 was calculated. We assumed this to be a logical time frame for when a user would plug in their vehicle before heading into work. If a majority of the user's weekday events at their preferred station began during this time frame, it was deemed apparent that they were utilizing the particular station for charging while at work. For each user, the second calculation was made on the percentage of weekday charge events at the preferred station that started in the evening after 16:00. Again, a fairly subjective assumption, but if a majority of the user's respective events occurred during part of the day, it was logical to believe the user was performing these events while at their place of residence. Fig. 10 displays the distributions of these two elements, with moderately obvious visual evidence supporting the manual sorting criterium here.

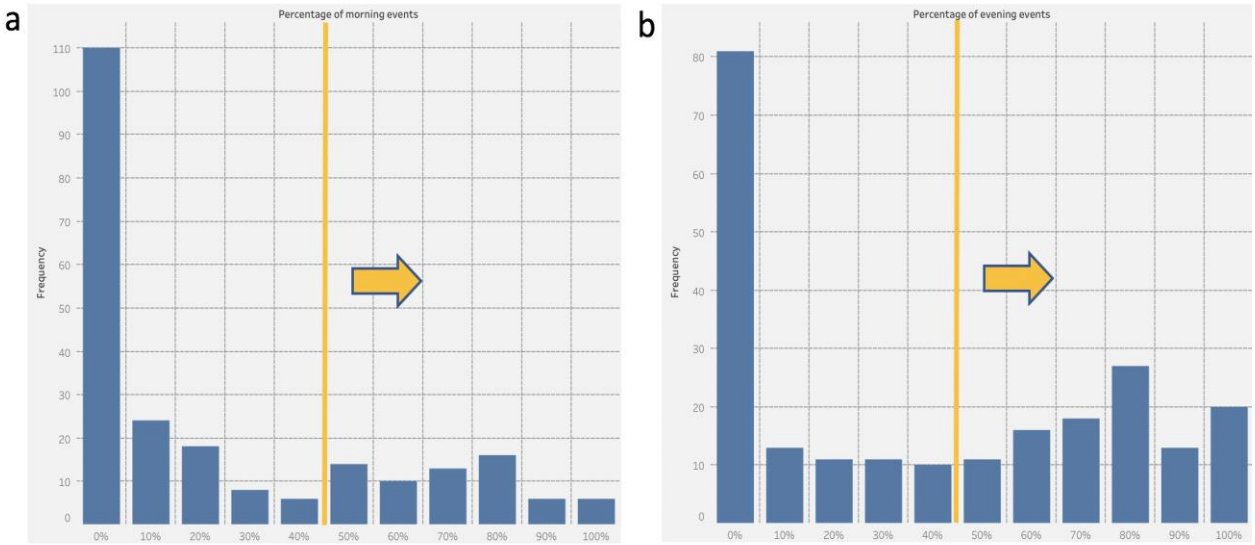


Figure 10. Histograms of node 7 users' (a) percentage of weekday events at preferred location that occurred in the morning; (b) percentage of weekday events at preferred location that occurred in the evening.

Regarding the first calculation, users with a percentage greater than 50% were sorted to 7.1 and labelled as the “Work Chargers” cluster. Regarding the second calculation, users with a percentage greater than 50% were sorted to 7.3 and labelled as the “At Home Chargers I” cluster. It was impossible for a user to have satisfied both of these criteria, but it was possible that neither were satisfied. Users with neither an obvious pattern of charging at their preferred station in the morning nor the evening were sorted to 7.2 and labelled as the “Undefined I” cluster. Perhaps it was the case that these users had performed the majority of charging events while at work or at home, simply at atypical times during the day. However, it could have also been the case that their preferred station was unrelated to either of these behaviors. We left this end cluster of users as is and reflected the ambiguity in the label.

At node 8, these were users with the same characteristics as those at node 7, but the charge events at their preferred station occurred mostly on the weekend rather than during the week. This too produced a degree of ambiguity. Nevertheless, the possibility that the charging behavior of some of these users had demonstrated a preference for charging while at home was still considered. It was deemed not completely necessary that the user’s majority of charging at the preferred station had to have occurred during the week for the user to have exemplified this preference.

Therefore, one final criterium was used to sort users at this node and potentially distinguish another end cluster of users with a preference for charging while at home. Although at least 50% of charge events at their preferred station occurred during the weekend, for each user, a calculation was made regarding those events at these stations that occurred on a weekday. Of these events, the percentage that occurred in the evening after 16:00 was calculated. To clarify via an example, let us say a user performed 75 charge events at their preferred location, with 50 of these events occurring on the weekend. The calculation was made based on those 25 events that occurred during the week, determining the percentage that occurred in the evening after 16:00. If this percentage was greater than 50%, the user was sorted to 8.1, where this cluster of users was labelled “At Home Chargers II.” The others were sorted to 8.2 and labelled as the “Undefined II” cluster, with the reasoning following the same logic as described for the “Undefined I” cluster described previously. Fig. 11 once again displays a histogram of this element. With only a low number of users to begin with at this node, the distribution displayed should be considered conservatively. Nevertheless, the figure provides support for the sorting criterium.

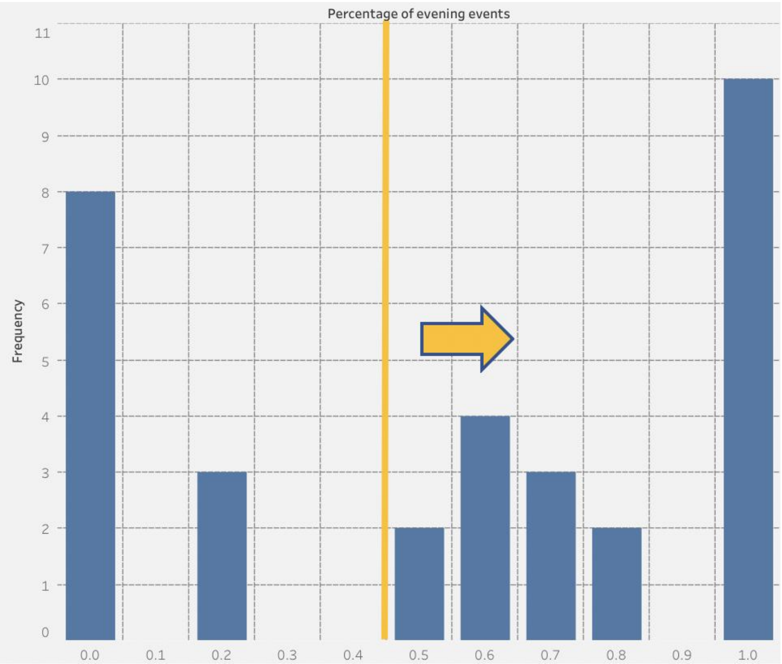


Figure 11. Histogram of node 8 users’ percentage of events at preferred station that occurred during the evening.

Finally, we refer all the way back to node 2. Although it was deemed there was not enough data to derive generalizations about these users (less than 10 charge events during the one-year period from April 2017 to March 2018), we were motivated to identify them. To begin, it was imagined that some of these users had performed so few charging events simply because they were new users and just started using public charging stations. Therefore, one criterium was subjectively determined to sort these users. If the date of the user's first charge event occurred in the most recent three months of the period (after 1st of January 2018), the user was sorted to 2.1 and included in the cluster labelled "New Users." Digging deeper, we were was aware that some users were exposed to promotional incentives and could have had access to trial periods of one or two months of free charging. Therefore, for establishing another sorting criterium to identify this circumstance, a calculation was made for each user on the number of days between the dates of their first and last charge events. Considering that element, if the number of days was 61 or less (maximum equivalent of two months), the user was sorted to 2.2 and included in the cluster labelled "Trial/Change of Mind." All other users were sorted to 2.3, the cluster labelled "Other Infrequent Users."

2.4. Regression

For the second part of the analysis, the focus was on identifying possible locational factors that were correlated to the utilization of individual charging stations. As mentioned earlier in the second sub-section, number of POIs within 300 meters, as well as employment level in the area were considered for this analysis. To reiterate, in order to consider employment level for these regressions, each station was tagged with the number of employees employed in the PLR in which the station was located. Furthermore, rather than using this raw number as the independent variable, the number of employees were divided by the square-kilometer surface area of the PLR to provide a more relevant employees-per-square-kilometer value for each PLR (and station).

In particular, regressions focused on three specific user clusters: the "Car-Sharing" cluster, the "Private Roamer I & II" cluster ("Private Roamer I" and "Private Roamer II" combined), and the "Work Chargers" cluster. For the "Car-Sharing" cluster, regressions were performed on both the charge event data of only these users, as well as the charge event data of all users that occurred only at the stations used by this cluster. Total charge events represented the dependent variable of focus, while number POIs within 300 meters of the station was the regressor in both cases.

For the "Private Roamer I & II" cluster, regressions were similarly performed on both the charge event data of only these users, as well as the charge event data of all users that occurred only at the stations used by this cluster. Again, total charge events represented the dependent variable of focus, while number POIs within 300 meters of the station was the regressor in both cases.

For the "Work Chargers" cluster, regressions were also performed on the charge event data of users within this cluster. The dependent variable of total charge events was again the same. However, for this cluster, the above-mentioned employees-per-square-kilometer value was used as the regressor. The results of these regressions are explained later in the next section of this paper.

3. Results

3.1. Clusters

As described in the previous section, individual users were sorted into clusters based on certain criteria. For example, a user with 10 or more charge events, a charge event-to-station ratio of more than 2, a preferred station where more than one-third of the user's charge events were performed, the majority of charge events at the preferred station occurring during the week, and the majority of these mentioned events occurring during the morning between 5:00 and 10:00 was sorted into the "Work Chargers" cluster. A user with more 10 or more charge events, a charge event-to-station ratio of 2 or less, and a utilization of 36 or more stations was sorted into the "Car-Sharing" cluster. Fig. 12 (a) shows a table of all criteria used for clustering.

In total, 12 individual clusters were identified. However, for the sake of further analysis, we proceeded with grouping a few clusters in particular. "At Home Chargers I" and "At Homers Chargers II" were further considered as one cluster. The same applied for "Private Roamers I" and "Private Roamers II," as well as the two "Undefined" a

	Clustering Criteria														
Cluster	<10 Charge events	First event in last 3 months	Event period over <= 61 days	>=10 Charge events	Events:Station ratio <=2	<18 distinct stations used	Between 18 and 35 distinct stations used	>=36 distinct stations used	Events:Station ratio >2	User's most utilized station with >=45% of events	Mostly weekday events at most utilized station	Mostly morning, weekday events at most utilized	Mostly evening/night, weekday events	Mostly weekend events at most utilized station	Of weekday events at most utilized station, mostly evening/night
At Home Chargers I	-	-	-	✓	-	-	-	-	✓	✓	✓	-	✓	-	-
At Home Chargers II	-	-	-	✓	-	-	-	-	✓	✓	-	-	-	✓	✓
Business Roamers	-	-	-	✓	✓	-	✓	-	-	-	-	-	-	-	-
Car-Sharing	-	-	-	✓	✓	-	-	✓	-	-	-	-	-	-	-
New Users	✓	✓	-	-	-	-	-	-	✓	-	-	-	-	-	-
Other Infrequent Users	✓	-	-	-	-	-	-	-	✓	-	-	-	-	-	-
Private Roamers I	-	-	-	✓	✓	✓	-	-	-	-	-	-	-	-	-
Private Roamers II	-	-	-	✓	-	-	-	-	✓	-	-	-	-	-	-
Trial/Change of Mind	✓	-	✓	-	-	-	-	-	✓	-	-	-	-	-	-
Undefined I	-	-	-	✓	-	-	-	-	✓	✓	✓	-	-	-	-
Undefined II	-	-	-	✓	-	-	-	-	✓	✓	-	-	-	✓	-
Work Chargers	-	-	-	✓	-	-	-	-	✓	✓	✓	✓	-	-	-

b

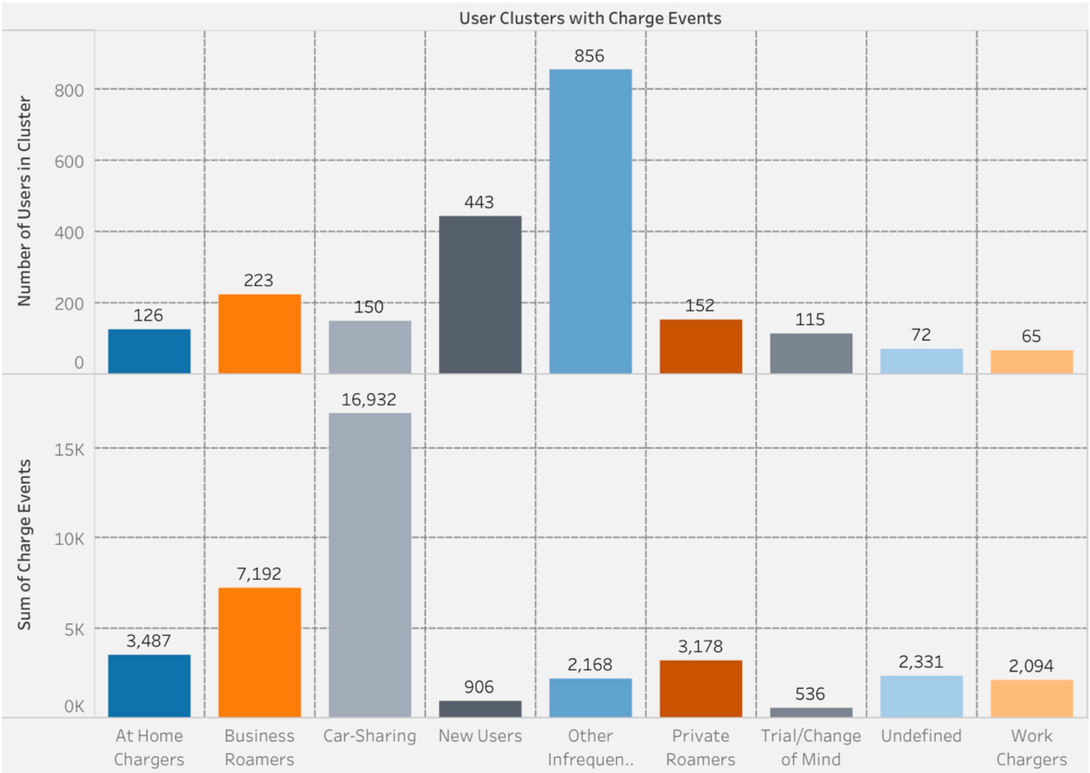


Figure 12. (a) table of criteria used for clustering; (b) clusters with number of users and sum of charge events.

clusters. Fig. 12 (b) gives a representation of these 9 clusters, identifying the number of users in each cluster in addition to the sum of charging events performed by the users in each cluster.

As seen in the bottom graph, one of the most obvious takeaways is the overwhelming number of charge events performed by the "Car-Sharing" cluster. Of the 38,824 in total, this cluster performed well over 40% of events. Furthermore, these events were performed by a relatively few number of users. On average, each user performed 112 charge events. Fig. 13 (a) shows a map of the charge events for one such user in this cluster. The display was a pattern among users in this cluster. Again, it should be noted that our treatment of the word "user" is this paper refers to the EVCO ID associated with each charge event, not necessarily a particular induvial person. Looking at this map, it is very probable that these events were performed by a high number of different individuals driving the same BMW i3 during their DriveNow journeys, for example.

The "Business Roamer" cluster also executed a large number of charge events, but in comparison, with a lower average number of events per user. This is because, of the cluster groups whose users performed at least 10 charge events, the "Business Roamer" cluster had the highest number of users. "Berlin: Capital of electric mobility" (2017) mentions that the local water company in the city, for example, has a number of electric vehicles in its fleet. We imagine this as an example of some of the "Business Roamers" cluster users, where the company's employees are

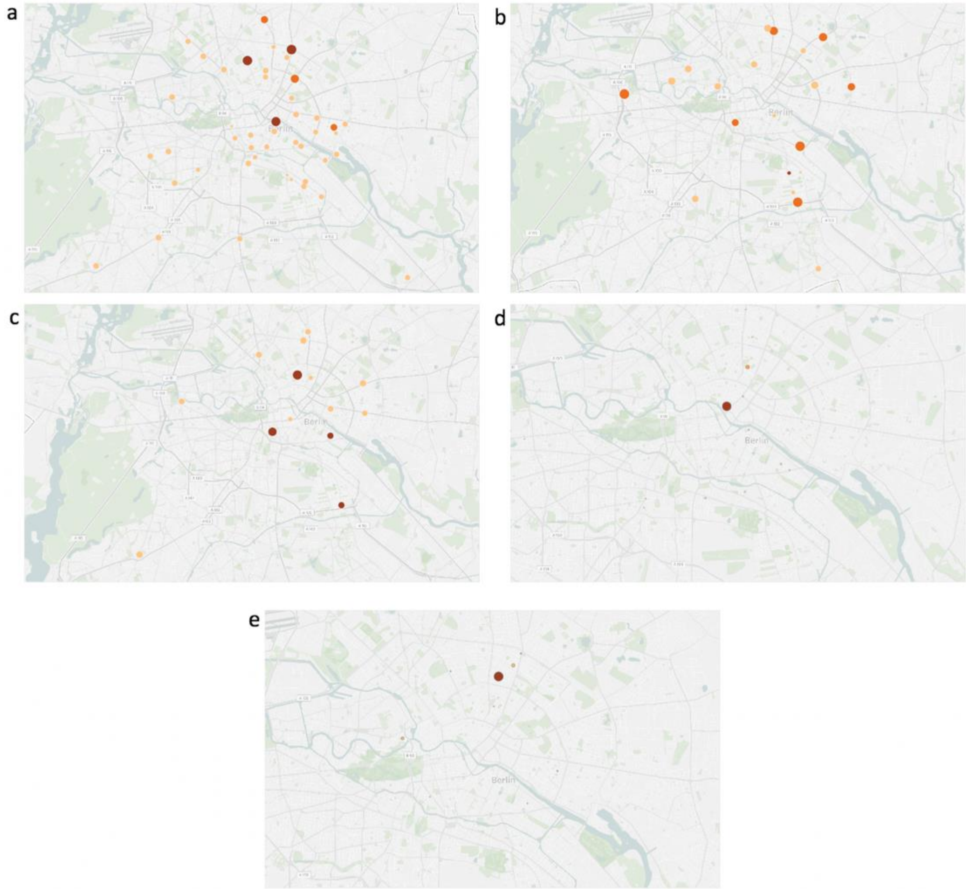


Figure 13. Maps of charging behavior for one (a) “Car-Sharing” user; (b) “Business Roamers” user; (c) “Private Roamers” user; (d) “Work Chargers” user; (e) “At Home Chargers” user.

travelling to designated locations throughout the city as part of their daily work and utilizing public charging on occasion.

Also of particular interest is the “Work Chargers” cluster. As displayed, the number of users in the cluster were the lowest. However, the number of charge events performed by these users was fairly significant. On average, these users performed over 32 charge events during the period, nearly double the average of the users in the “Private Roamers” cluster and slightly more than the average for users in the “At Homers” combined cluster. Furthermore, it could very well be that many of the users who were sorted to the “Undefined” combined cluster are in fact more similar to the “Work Charger” cluster than any other. To recap, these were users with a high number of charge events performed at one particular preferred location. Because the charge events for these users at those preferred stations did not follow a particular pattern in terms of time of day, they were not sorted to either the “At Homers” or the “Work Chargers” clusters.

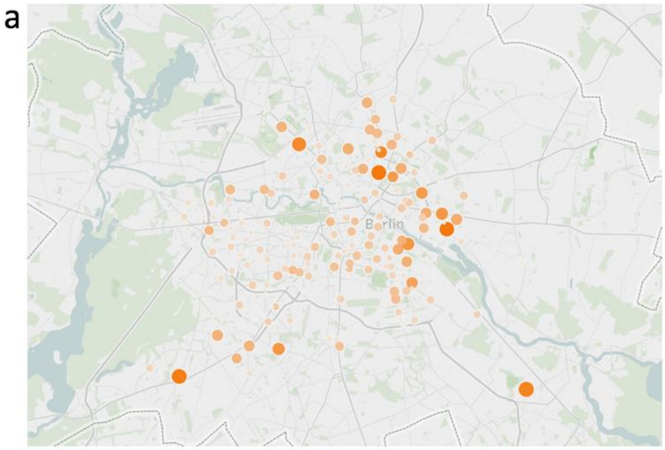
Not surprisingly, apart from the “Car-Sharing” and “Business Roamers” clusters, users in the “At Home Chargers” cluster were also significant, both in terms of number of users as well as total charge events. A total of 126 users in this combined cluster performed an average of almost 28 charge events over the one-year period. Fig. 13 (c) and (d) show examples of the charging events for users in the “At Home Chargers” and “At Work Chargers” clusters. The stark contrast between the charging behavior of these users, who have obvious preferred stations, and the behavior of users belonging to the other clusters is easily recognizable.

Finally, as seen in Fig. 12 (b), the cluster “Other Infrequent Users” includes the highest total number of users, however, a very low relative total number of charge invents. This is explicit, as these users, along with the “New User” cluster, performed less than 10 charge events during the analyzed period. Nevertheless, such a high number of relative users in the cluster lead us to at least propose further explanations. As the stations analyzed in this study were exclusively owned and operated by one company, it could be that these are users who only infrequently charge at the company's stations, while performing the bulk of other charge events at stations owned and operated by other companies in the city. Another explanation is that some of these users are private home-owners, with access to charging hardware inside their place of residence. Over the course of the year, they could have performed the bulk of their charging events privately in their parking garage, for example, while only rarely using an openly available charging station from this company. Finally, some of these users could also be just visitors. Either as tourists taking holiday in the city, professionals staying in the city during a business trip, or other journey-goes simply stopping by en route to a further destination, these users are only rarely utilizing these particular charging stations.

3.2. Regressions

Building on the clustering analysis, we intended to explore correlations relating to the utilization of charging stations within these clusters. While the clustering analysis took the perspective of the user and attempted to generalize charging behavior in that respect, this part of the analysis took the perspective of the charging stations. Simple linear regressions were performed within three particular clusters: the "Car-Sharing" cluster, the "Work Chargers" cluster, and the combined "Private Roamers" cluster.

For the "Car-Sharing" cluster, the linear regression encompassed the charge event data from only this cluster and used number of POIs within 300 meters as the regressor and total charge events as the independent variable. Fig. 14 shows the results of the regression, as well as a map of all charge events from this cluster group. The regression resulted in a statistically significant, positive coefficient of 0.3365 with a p-value of 0.0057. This result supports the intuition that among these users, stations that are located in more appealing, urbanized areas (a higher number of POIs within reasonable walking distance) are utilized more than stations in less commercialized areas.



b

	Coefficient	Std. Error	t-ratio	p-value	
Constant	64.7718	8.94129	7.244	<0.0001	***
POIswithin300m	0.336503	0.12035	2.796	0.0057	***

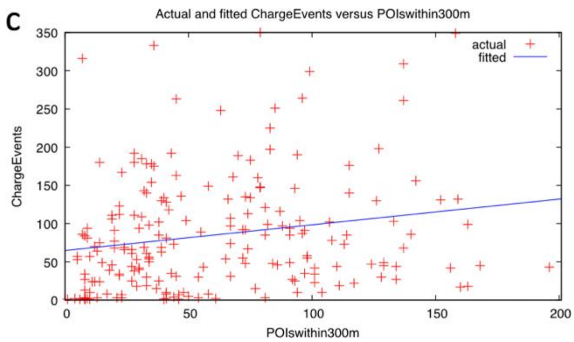


Figure 14. "Car-Sharing" cluster (a) map of charge events; (b) regression results for POIs within 300m; and (c) scatter plot of stations with regression line.

For the "Work Chargers" cluster, the dependent variable was the same; however, employment level was used as the regressor. As mentioned before, this variable was derived by first determining the number of registered employees in each of the 447 PLRs in Berlin. This number was then divided by the square-kilometer surface area of the PLR to get a respective employees-per-square-kilometer value for the PLR. Each station was then tagged with this value of the PLR in which it was located. Fig. 15 displays the results of this regression. With a statistically significant p-value close to zero, the model predicts an increase of about 4 charge events at the station for an increase of 10,000 employees per-square-kilometer of area. This result supports our intuition that for this cluster, charging stations located in heavily industrialized areas are utilized more than those in comparably less industrial districts. Looking at Fig. 15 (c), however, the left portion of the chart shows some ambiguity. There are several highly utilized stations in areas with a relatively low number of registered employees-per-square-kilometer. This could simply mean that a number of these users are charging while at work, but that their workplace is in a more urbanized location. The users in this cluster are not necessarily exclusive to industrial areas. Another explanation could be that the total land space of the PLR in which the station is located is mostly undeveloped – for example, a large park or body of water occupies most of the space. The station could be located in the densely industrialized portion; however, the employees-per-square-kilometer value is low because of this overall land use.

Figure 15. "Work Chargers" cluster (a) map of charge events; (b) regression results for employees-per-square-kilometer; and (c) scatter plot of stations with regression line.

The final regression was performed within the combined "Private Roamers" cluster. As explained in the method section describing the clustering criteria, these were users without an obvious preference for one particular station. They had a high number of charge events spread out across a number of different stations in the city. Therefore, we were curious to see if the number of POIs within 300 meters was also correlated to the utilization of the stations used by these clusters. The same regression was performed as with the "Car-Sharing." Fig. 16 displays the results. Here, the regressor was statistically significant with a p-value close to zero. From the results, it is quite clear that, among users in this cluster, stations with a relatively higher number of POIs nearby are utilized more than those with less.

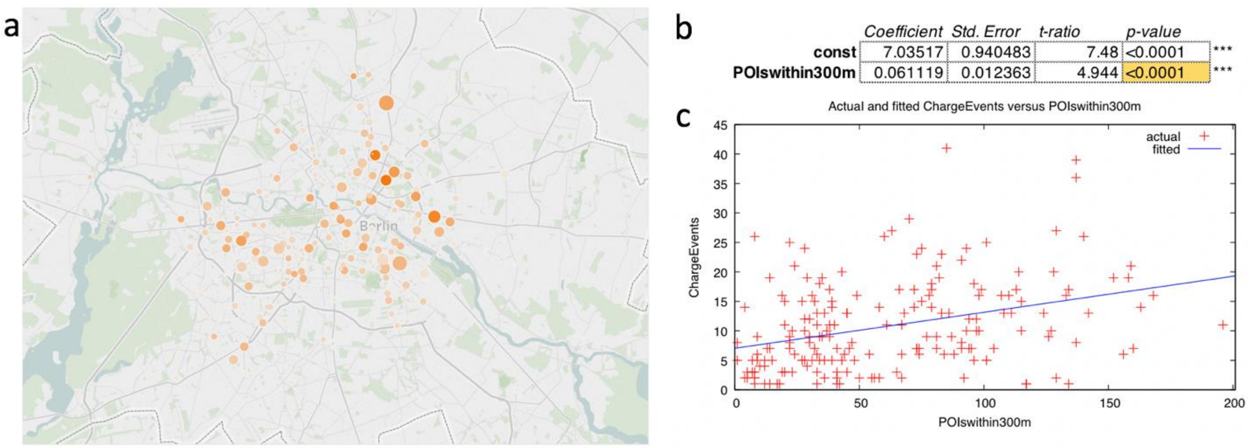


Figure 16. "Private Roamers" combined cluster (a) map of charge events; (b) regression results for POIs within 300m; and (c) scatter plot of stations with regression line.

4. Discussion

The purpose of this study was to analyze authentic charging data for EV users of public charging stations in an urban setting and group users into clusters based on their charging behaviors. Furthermore, some regression analysis was performed within these clusters to test correlations between locational factors of stations and their utilization.

One of the most interesting findings was the high number of users and overwhelmingly high number of charge events performed by users in the "Car-Sharing" cluster. Of the clusters whose users performed 10 or more charge events over the one-year period, this cluster included the second highest number of users. More importantly, these 150 users performed over 50% of all charge events made by users with 10 or more events. This finding is crucial because most of the research to date on charging behavior and charging preferences seems to ignore the concept of mobility services such as car-sharing. Rightfully so, as it has only recently become a sweeping trend in the automotive industry. Nevertheless, it is essential to recognize the idea that charging events associated with one vehicle can be the result of up to hundreds of different users. Both for further research as well and infrastructure development and planning, the treatment of stated and revealed preferences of users and their charging behavior should therefore consider this phenomenon.

This also applies to the finding regarding the "Business Roamers" cluster. This cluster included the single highest number of users and the second highest number of charge events performed, referring to all clusters whose users performed at least 10 charge events. To recap, these users were tagged as "Business Roamers" because of their moderately high number of distinct charging stations used – a cluster evidenced in Fig. 5. (a). Similar to what was mentioned in the previous paragraph, we expect the events of these "users" to have been performed by multiple different individual operators of the vehicle. As companies around the world continue striving for positive brand image by boasting "green" technologies and pursue other benefits of electromobility, it may well be that businesses own and operate significantly large fleets of EVs. The charging behavior and preferences of these users should also be considered in further research and infrastructure development.

Another important finding relates to the relative number of users and charge events by users in the "Work Chargers," combined "At Home Chargers," and combined "Private Roamers" clusters. As described in the introduction, researchers and industry practitioners suggest that EV drivers have a preference to charge either at home or at work. This study was focused solely on charging behavior at public stations, so it was impossible to verify these claims through an analysis of charging behavior inside personal residences or privately on site at workplaces. However, there are two reasons that the findings here are still insightful. First, in many areas of the city, especially near the dense city center, we imagine very few residents (who are most likely living in apartment buildings, as opposed to private homes with a garage) to have any capability of charging inside their personal residence. Second, based on personal knowledge of the subject, we are aware that workplaces with private charging hardware on campus are, to this date, rare. Therefore, we supposed that users in the context of this study would still exhibit such preferences via charging behavior at these public charging stations.

That being said, the "At Home Chargers" cluster included 126 users who performed 3,487 charge events, while the "Work Chargers" cluster represented 65 users who performed 2,094 events. In total, these clusters equaled 191 users who performed 5,581charge events. This compared to the "Work Chargers" cluster, which represented 152 users who performed a total of 3,178 charge events. Considering these results, even regarding the usage of public charging stations, there is an evident preference for users to charge while at home or at work. These findings should provide valuable insight to all companies contributing to the development of public charging infrastructure.

Concerning the regression analysis on utilization of stations, we found significant results regarding both number of POIs within 300 meters and employees-per-square-kilometer. Within the "Car-Sharing" cluster, we saw a significant positive correlation between number of charge events and number of POIs within 300 meters. Although this was more of a general surface-level analysis, rather than specific to any one category of POI, this finding suggests that these users are charging at popular, urban areas in the city. Furthermore, the same regression within the combined "Private Roamers" cluster produced similar results, although with a lower regression coefficient. Again, compared to the combined "At Home Chargers" cluster and the "Work Chargers" cluster, who had obvious preferred charging locations, these users utilized a number of different stations throughout the city. Further studies could certainly work towards identifying particular categories of POIs that impact station utilization more or less than others.

Within the “Work Chargers” cluster, the significant correlation between employees-per-square-kilometer and number of charge events encourages the provision of charging stations in industrialized districts of the city. While some companies have the ability to provide private charging access on premise to their employees, it will still be necessary for charging station operators to cater to users in this cluster who prefer to charge while at work but are unable to do so directly on site.

Lastly, it is important to touch on the high number of users who performed less than 10 events over the one-year period in which charging data was obtained. In total, the “New Users” and “Other Infrequent Users” cluster represented nearly 60% of the 2,202 users in the data set. This finding has minimal significance both because of the low number of charge events performed by these users, and again, because the data set analyzed was associated with charging stations owned and operated by strictly one firm. Electric utilities such as RWE as well as other mobility service providers operate a significant number of stations in the city. Therefore, it was simply that these users did not perform enough charge events at the stations from which data was gathered, so their charging behavior was unidentifiable. Suggestions were made earlier in this paper regarding these users – that perhaps they have access to private charging at home and rarely seek public chargers, or that they are simply rarely using the stations operated by the company from which the data was obtained and instead performing charging at stations operated by other firms. At the moment, this is impossible to confirm. Nevertheless, we imagine that an analysis of aggregated charge events at all public stations in the city (not just the ones owned and operated by the company from which the data was obtained) would produce key findings of similar nature.

5. Conclusion

The results of this study should contribute a foundation for further research regarding the charging behavior of EV users in an urban environment. To date, much of the research in this field draws conclusions from stated preference studies. In their systematic review of the literature on EV use modeling, Daina at al. (2017) describe that “amongst the most critical [limitations,] there is the lack of realism how charging behavior is represented.” The authors would agree, as studies on EV use and charging behavior such as those from Jabeen et al. (2013) and Xu et al. (2017) feel rather synthetic. Observations and analysis were made regarding behaviors during the Western Australia Electric Vehicle *trial* and from *probe Battery Electric Vehicles* in Japan, respectively. This study provides at least one instance of a more natural, realistic representation of charging behavior.

We see this study as a basis for further research. The key findings should be compared with similar analyses of charging behavior in other cities around the world. We imagine the relative numbers of users in each of the cluster groups to be comparable in some cities, while possibly drastically different in others. A realistic understanding and representation of EV charging behavior will provide valuable support to the further development of charging infrastructure as these vehicles take center stage in the automotive industry.

References

Berlin: capital of electromobility. (2017, February 28). Retrieved from https://www.diamona-harnisch.com

Chatelain, A., Erriquez, M., Mouliere, P., Schäfer, P. (2018, March). *What a teardown of the latest electric vehicles reveals about the future of mass-market EVs*. Retrieved from https://www.mckinsey.com

Daina, N., Sivakumar, A., Polak, J.W. (2017). Modelleing electric vehicles use: a survey on the methods. *Renewable and Sustainable Energy Reviews, 68(2017),* 447-460.

Daubitz, S., Kawgan-Kagan, I. (2014, September). *Integrated charging infrastructure: cognitive interviews to identify preferences in charging options*. Paper presented at the 2014 European Transport Conference, Frankfurt, Germany. Retrieved from https://aetransport.org

Electric Vehicles: Tax Credits and Other Incentives. (n.d.). In *Office of Energy Efficiency & Renewable Energy*. Retrieved June 29, 2018, from https://www.energy.gov/eere

DiChristopher, T. (2018, May 30). *Electric vehicles will grow from 3 million to 125 million by 2030, International Energy Agency forecasts*. Retrieved from https://www.cnbc.com

Harrison, G., Thiel, C. (2015). Comparing European member state electric vehicle uptake incentives and charging infrastructure provision, European Transport Conference, 2015. Association for European Transport.

Jabeen, F., Olaru, D., Smith, B., Braunl, T., Speidel, S. (2013, October). *Electric vehicle battery charging behavior: findings from a driver survey*. Paper presented at the Australasian Transport Research Forum 2013, Brisbane, Australia. Retrieved from http://www.patrec.org/atrf.aspx

Morrissey, P., Weldon, P., O'Mahony, M. (2016). Future standard and fast charging infrastructure planning: An analysis of electric vehicle charging behavior. *Energy Policy, 89(2016)*, 257-270.

Nagle, J. (2018, June 21). *Electric vehicles have a charge problem, just not the one you think.* Retrieved from https://www.linkedin.com/pulse

OpenStreetMap data extracts. (n.d.). Retrieved from https://download.geofabrik.de

Philipsen, R., Schmidt, T., van Heek, J., Ziefle, M. (2016). Fast-charging station here, please! User criteria for electric vehicle fast-charging locations. *Transportation Research Part F, 40(2016)*, 119-129.

Pressman, M. (2017, September 2). *Electric car incentives in Norway, UK, France, Germany, Netherlands, & Belgium*. Retrieved from https://cleantechnica.com

Randall, T. (2017, April 25). *The electric-car boom is so real even oil companies say it's coming.* Retrieved from https://www.bloomberg.com

Thompson, C. (2017, February 15). *The fascinating evolution of the electric car*. Retrieved from https://www.businessinsider.de

Weldon, P., Morrissey, P., Brady, J., O'Mahony, M. (2016). An investigation into usage patterns of electric vehicles. *Transportation research Part D, 43(2016)*, 207-225.

Xu, M., Meng, Q., Liu, K., Yamamoto, T. (2017). Joint charging mode and location choice model for battery electric vehicle users. *Transportation Research Part B, 103(2017)*, 68-86.

Yang, Y., Diez-Roux, A.V. (2012). Walking distance by trip purpose and population subgroups. *American Journal of Preventive Medicine, 43(1),* 11-19.

Zhang, Y., Chen, S. (2018, February 13). *China raises subsidies to reward longer range electric cars*. Retrieved from https://www.bloomberg.com