Electric Vehicle Charging Locations for Uber in New York City

Victor Salazar Ramos, Andreea Mateescu, Elisabeth Fokker, Jacky P.K. Li, and Sandjai Bhulai

Faculty of Science
Vrije Universiteit Amsterdam
Amsterdam, The Netherlands
Email: vss920@vu.nl, mmu680@vu.nl, efr330@vu.nl, jacky.li@vu.nl, s.bhulai@vu.nl

***Abstract*—Electrification is widely considered an attractive solution for reducing the oil dependency and environmental impact of road transportation. This paper examines the viability of the current charging stations with the assistance of taxi service strategy optimization. We study charging station data along with ride pickup data. Our methodology is to model the pickup location data and the charging station data by K-means clustering and to determine the optimized distance between the pickup locations and the charging stations. Consequently, we shed light on the number of charging stations that can improve the efficiency of using electric taxis.**

***Index Terms*—Electrification; K-means clustering; electric taxis; charging stations.**

I. INTRODUCTION

Transportation accounts for nearly 70 percent of U.S. oil consumption and 28 percent of the country's greenhouse gas emissions, according to U.S. Energy Information Administration [1]. This makes it a prime target for technological improvements that can reduce emissions and combat climate change. Electrification of transportation represents one of the highest impact strategies to help achieve that goal. In particular, public vehicles (e.g., taxis) provide a crucial opportunity for electrification. Despite the benefits of eco-friendliness and energy efficiency, the adoption of electric taxis faces several obstacles. This includes constrained driving range, long recharging duration, limited charging stations, and low gas prices, all of which impede taxi drivers' decisions to switch to electric taxis. To contribute to a sustainable future, Uber Technologies Inc. has recently released the 'Uber Green' option, which connects potential customer(s) with electric or hybrid-electric vehicles [2]. However, the 'Uber Green' option can be an alternative to fuel-based driving only in the presence of an optimally distributed web of charging stations.

Previous studies have focused on developing electric taxi strategy optimization compared to conventional taxis with an internal combustion engine. Tseng et al. [3] studied empirically under various battery capacities and charging conditions to conquer the long charging period and optimize the number of taxis needed to satisfy the demands.

Raboaca et al. [4] propose a new operational model for the mobile charging station through temporally stationing it at different places for certain amounts of time. Their model is built by a queuing process. The goal is to place a minimum number of temporary service centers to minimize operating costs. Jung et al. [5] compare the revenue differences between electric and non-electric taxis. The research concentrates on individual versus whole-fleet policies. This concludes the advantage of non-electric taxis over electric taxis due to the long charging periods.

Wang et al. [6] investigate the effectiveness of the charging stations over the taxi demand. Their research contains five years of taxi transaction data, charging station data, and the distance traveled after each charge. This research provides the evolving patterns of electric taxi networks and also charging stations. Wang et al. [7] also research Large-Scale Electric Bus Fleets. Their research invents a real-time charging scheduling system to optimize the charging time and usage for the electric bus fleet.

Scorrano et al. [8] investigate the taxi data of Florence, Italy. The city of Florence and many European cities have lower vehicle tax in order to encourage taxis to switch from a combustion engine to fully electric. Their research evaluates the impact of the annual distance traveled, purchase subsidy, and new revenue loss.

In terms of using K-means clustering research in electric taxi fleeting, Zhang et al. [9] create two stages of charging stations for electric taxis and electric vehicles to solve the long charging period. Dong et al. [10] concentrate on the optimal location model in the service region to calculate the optimal sites of the charging stations to maximize the operational efficiency and charging convenience. Jia et al. [11] investigate the large-scale Cellular Signaling Data and illustrate the method to generate the 24-hour travel demand for each electric vehicle.

This paper examines the viability of the current charging stations with the assistance of taxi service strategy optimization. We investigate the distribution of the charging station data along with the Uber pickup trip data. Our methodology is to model the pickup location data and the charging station data by K-means clustering and to determine the optimized distance between the data. Consequently, we shed light on the number of charging stations that can improve the efficiency of using electric taxis.

The paper is structured as follows. In Section II, we analyze

 ISBN: 978-1-61208-891-4

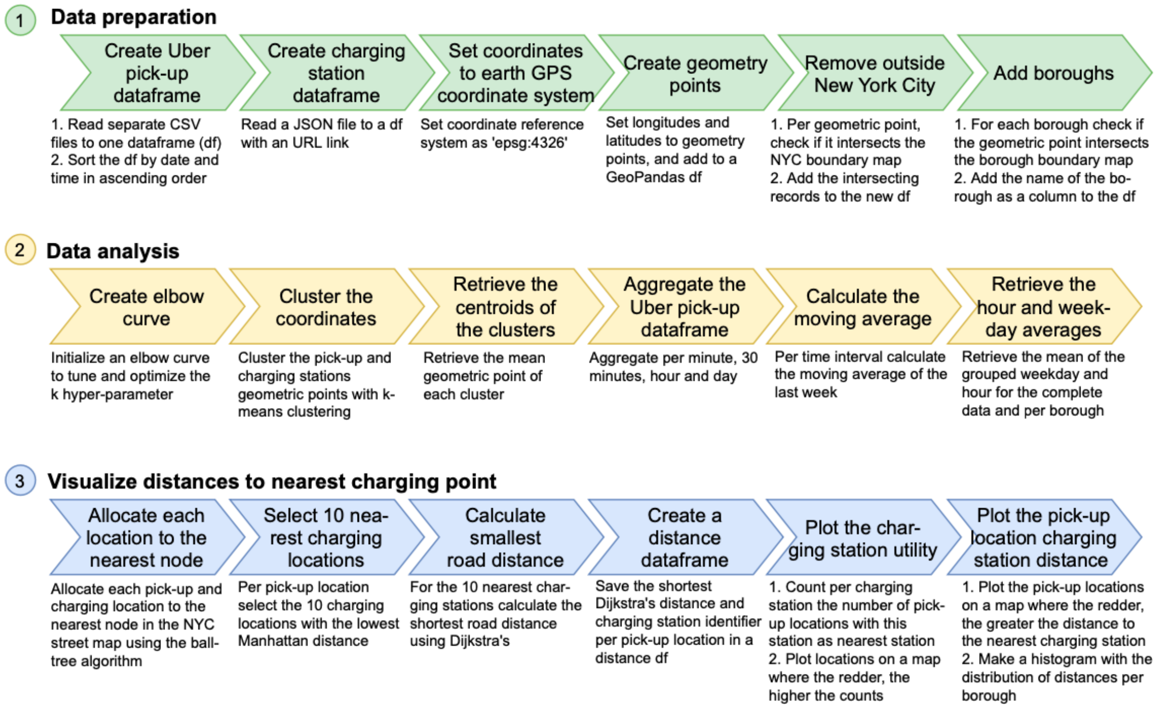

Figure. 1. Data wrangling methods.

the Uber dataset from April 2014 to May 2015 and the charging station data of New York City of 2021. This provides the input for our analysis and visualizations, which are explained in Section III. Finally, the paper is concluded in Section IV.

II. DATASET

In our research, we are investigating the ride-hailing service and the charging stations in New York State. The first data source consists of Uber trip data from April of 2014 to September of 2014. The data are separated by month for New York State. The data consist of 4.5 million ride demand records which contain the date, time, and detailed location with the latitude and longitude of each pickup. This dataset is retrieved from Kaggle [12]. The second data source consists of the electric vehicle charging station records in New York State. The data contains 2,347 registered stations as such and various location information, including the zip code, the borough, latitude, and longitude. It can be found on the government website of the New York State [13] dated 29th of January 2021. In order to answer the research question, we assume all Uber vehicles were operated by electric vehicles in 2014, and multiple data wrangling techniques were engaged. Figure 1 summarizes these techniques.

A. Data preparation

To prepare the data, the Uber pickup data was initially separated by month. Hence, we concatenated the monthly data into one data frame and ranked the date by the index, the date, and the time of the pickup location. The charging station data was read as a JSON file.

In order to work with spatial data, we converted all the dataframes in GeoDataFrames. This was done by creating a GeoSeries column that is referred to as the GeoDataFrame's geometry and indicates the latitude and longitude of a GPS coordinate using the geopandas library [14]. Next, we removed all the coordinates outside New York City. We only added the five boroughs of New York City: Bronx, Brooklyn, Manhattan, Queens, and Staten Island. To visualize the map of New York City with both the pickup points and the charging stations, the OSmnx package [15] for spatial visualizations was used. The background for this visualization was the street map of New York City, which can be interpreted as follows: the nodes represent intersections, and the edges represent street segments.

B. Data analysis

To have a better understanding of the hot spots for the time-dependent pickup locations and charging stations, the K-means clustering algorithm [16] is employed. The objective is to build a clustering model of the pickup data and the charging stations, by which we generated elbow curves to estimate and optimize the hyperparameter K. Based on the results, we assign five different clusters for the pickup data, which we display in Figures 2 and 3. There are four different clusters for the charging station data, which we display in Figures 4 and 5. The similarity of the centroids of the K-means clustering analysis suggests a high concentration of

 ISBN: 978-1-61208-891-4

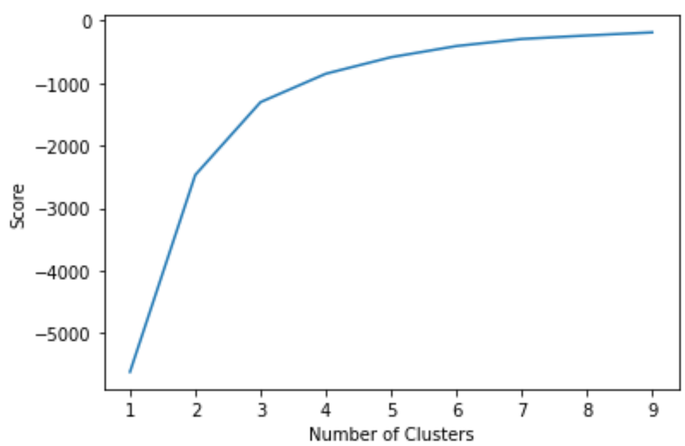

Figure. 2. Elbow curve for the Uber pickup location data.

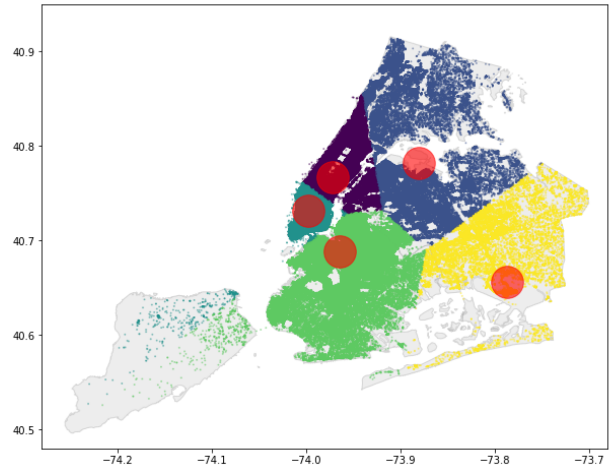

Figure. 3. Clusters for the Uber pickup location data.

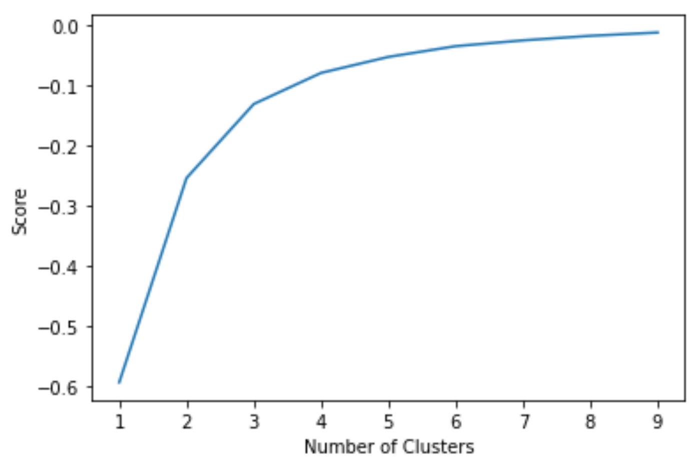

Figure. 4. Elbow curve for the charging station location data.

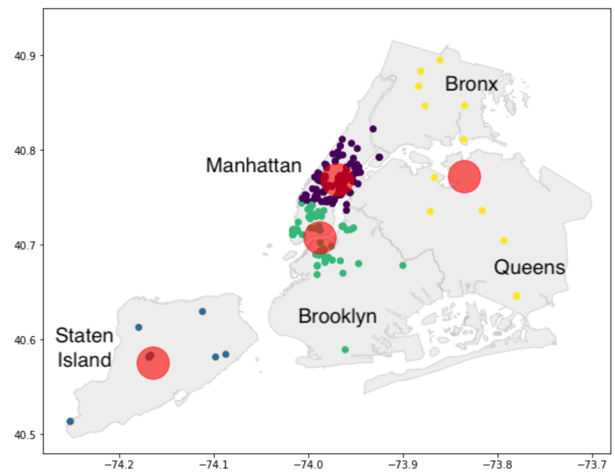

Figure. 5. Clusters for the charging station location data.

both pickups and charging stations in Manhattan and the west of Brooklyn.

When analyzing the distribution of the minutely arrivals of Uber, the data present substantial noise and could not depict a solid trend of the moving average [17]. However, the trend becomes strongly positive when moving to hourly and daily distributions of arrivals, suggesting the increasing popularity of Uber services in Figures 6, 7, and 8.

Next, we analyze the average number of Uber pickups per weekday in Figure 9. We notice the presence of two peaks: the morning rush hour and the leaving-work in the afternoon rush hour. We observe a clear distinction for Fridays and Saturdays. There is a distinct peak during the night, probably caused by the happy hour on Friday and stay out later in the evening. When comparing averages for the five boroughs, there is a distinguishable indication that the majority of trips take place in Manhattan, followed by Brooklyn, Queens, and ultimately the Bronx and Staten Island.

III. Visualize Distances to Nearest Charging Point

In order to assess the improvements that can be made at the charging station level, we must investigate the distances between pickups and the charging stations. To calculate the nearest charging point from each pickup location, we allocate each coordinate to the nearest node in the graph. We then minimize the travel times between the pickup and the charging locations using Dijkstra's algorithm [18]. However, due to the large number of calculations required, this approach would take approximately 13 years of run time, which is not feasible for a research as such. Therefore, we used a good enough approximation method with the help of the ball tree algorithm [19].

We approximated the ten nearest charging stations with the Manhattan distance [20]. We managed to reduce the running time from 13 years to 3.5 days when running Dijkstra's algorithm on the ten nearest stations. The data consists of 4.4 million pickup locations and 266 charging stations. The results are somewhat intuitive, as suggested by the hot spot analysis. As expected, Manhattan has by far the shortest travel distances,

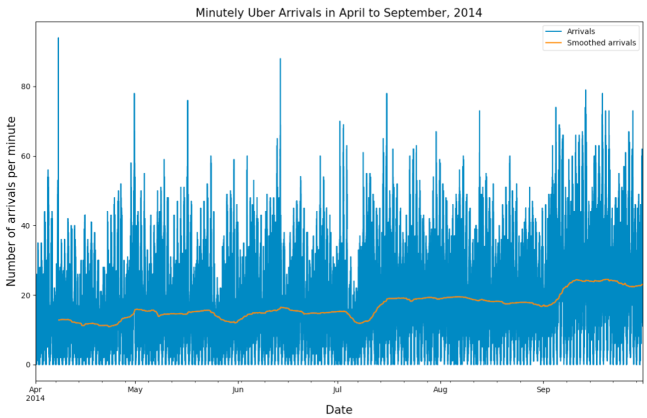

Figure. 6. Minutely Uber arrivals from April to September, 2014.

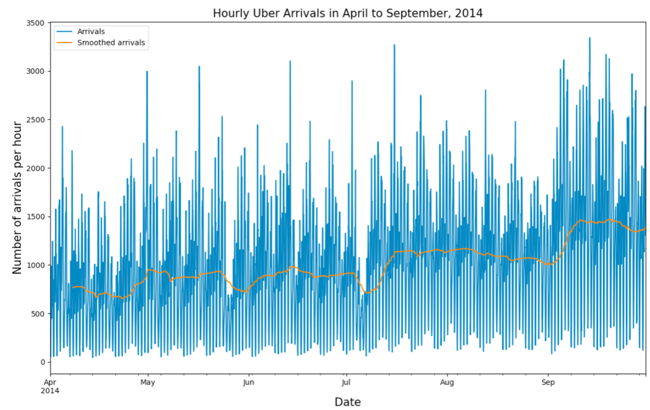

Figure. 7. Hourly Uber arrivals from April to September, 2014.

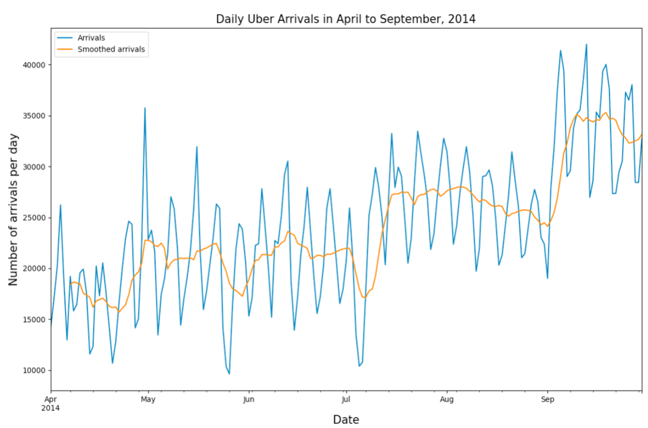

Figure. 8. Daily Uber arrivals from April to September, 2014.

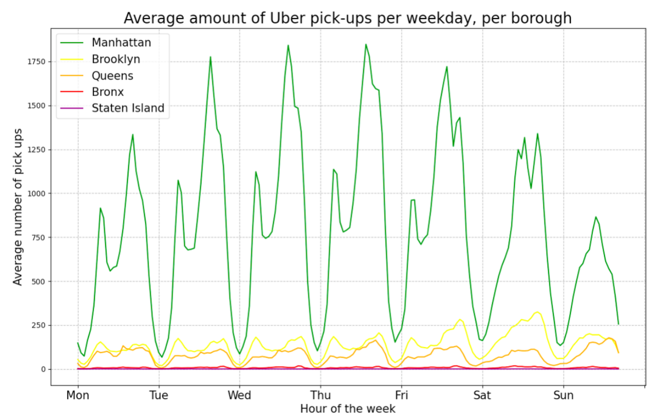

Figure. 9. Average number of Uber pickups per weekday per borough.

hinted by the fact that it had the most pickup locations and a high concentration of charging stations. Table I presents the number of times the rank of the stations with the smallest Manhattan distance have the smallest Dijkstra's distance. The charging station with the 10th smallest distance according to the Manhattan distance covers only 0.07% of the smallest Dijkstra's distance, indicating that ten nearest charging stations is a good cut-off point.

Next, we investigate which charging stations are the most utilized and in which places there is a need for more charging stations. In Figure 10, a visual utility analysis plots four categories of charging stations, ranging from highly utilized indicated by dark red toward a light cream color to indicate slightly used charging stations. A conclusion drawn from the utility analysis suggests that more charging locations are needed at the Southside of Manhattan, the Westside of Brooklyn, and the center of Queens.

In Figure 11, a complementary visual analysis plots the distances from the pickups to the nearest charging stations. The dark green color of the figure represents short distances between pickup points and charging locations. Furthermore, the red color represents longer distances between the pickup points and the charging locations. From this analysis, we depict that overall, Manhattan has small distances from pickup points and charging stations which makes common sense as Manhattan has the highest density in North American Cities. The predominantly yellow coloring in the Southeast of Queens and the North of Staten Island, culminating with intense red coloring in the Southeast of Queens, clearly suggests that more charging stations are needed. To provide a better understanding of the pickup locations versus the charging stations, we separated each Borough and created 5 histograms that represent the cumulative amount of the distance between pickup location and charging station at each Borough.

Figures 12, 13, and 14 provide the shortest distance to the nearest charging stations in Staten Island, Bronx, and Queens, respectively. These three figures provide insufficient evidence about the distance due to the insufficient number of pickup locations for one year in those areas. On the other hand, Figures 15 and 16 indicate that the majority of the pickup locations of Uber in 2014 are in Brooklyn and Manhattan. The majority distance between the pickup location and the charging stations is less than 2,000m. In order to investigate further, Figure 17 displays a complementary visual analysis from the pickups to the nearest charging stations in Manhattan. The figure indicates the further distance in upper Manhattan

 ISBN: 978-1-61208-891-4

TABLE I
PERCENTAGE OF OCCURRENCES WHERE THE TEN NEAREST MANHATTAN DISTANCES EQUAL THE SHORTEST PATH DISTANCE BASED ON DIJKSTRA'S ALGORITHM.

Rank station smallest Manhattan distance	1	2	3	4	5	6	7	8	9	10
% station smallest Dijkstra's distance	61.64%	19.28%	8.50%	6.67%	1.98%	0.99%	0.48%	0.22%	0.16%	0.07%

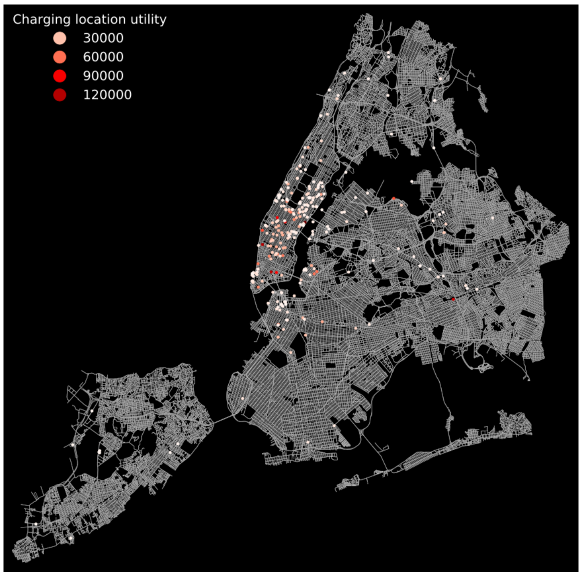

Figure. 10. Charging station utility map.

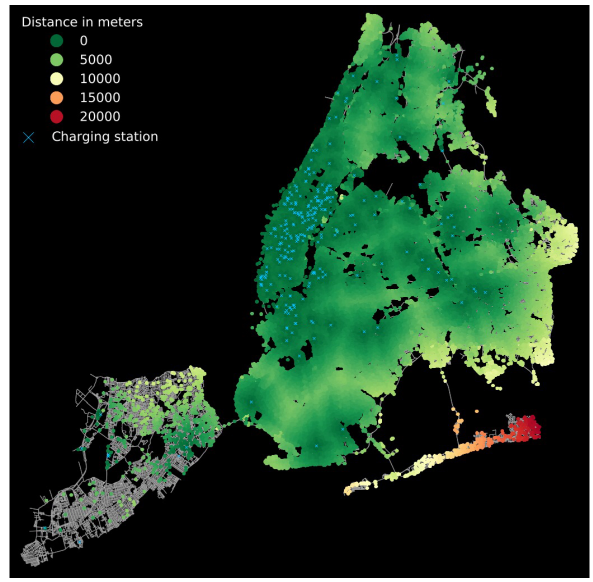

Figure. 11. Distance to nearest charging station map.

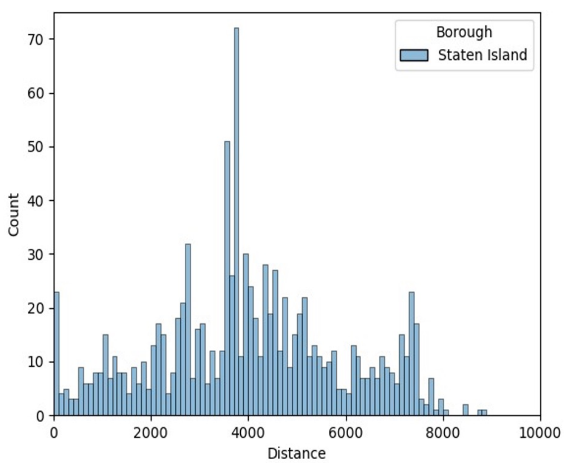

Figure. 12. Shortest distance to nearest charging station in Staten Island.

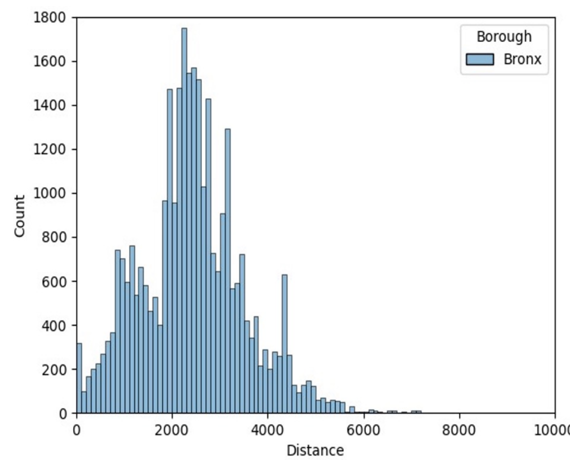

Figure. 13. Shortest distance to nearest charging station in the Bronx.

and east of Manhattan (East Harlem area).

IV. CONCLUSION

This research examines the viability of the current charging stations with the ride-hailing service of Uber by addressing (1) the distributions of charging and Uber pickup locations and (2) the distances between the pickup locations and their nearest charging station. According to cluster centroids, the pickup locations and charging locations are mostly distributed in Manhattan and the West of Brooklyn. Manhattan has the smallest charging station distances, while the distances from pickup locations in the East of Queens, the South of Queens, and the North of Staten Island to the nearest charging station is further away.

In conclusion, we performed multiple visual and hot spot

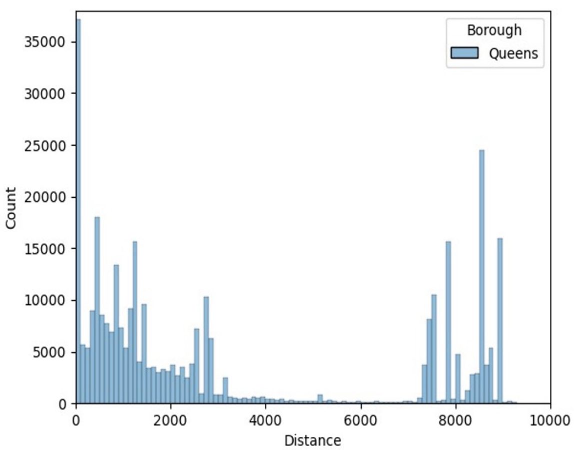

Figure. 14. Shortest distance to nearest charging station in Queens.

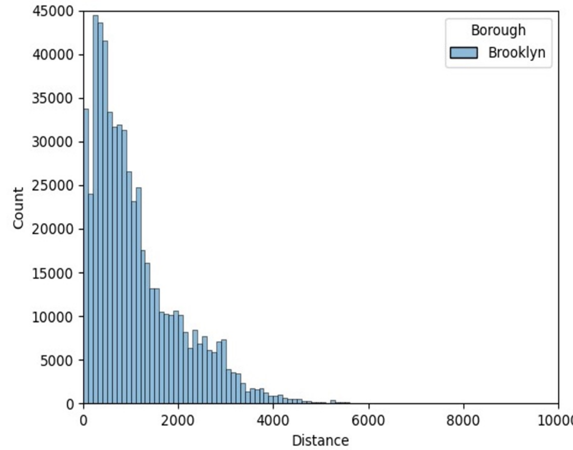

Figure. 15. Shortest distance to nearest charging station in Brooklyn.

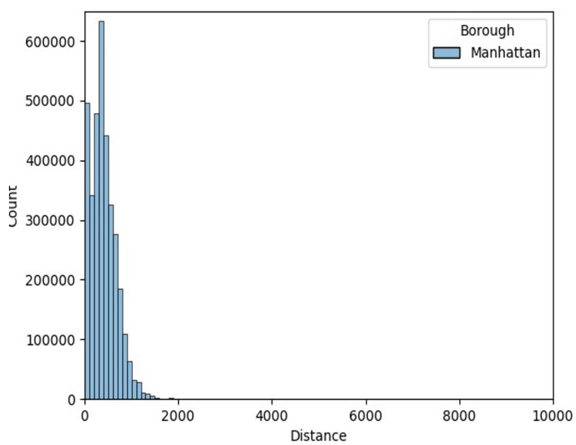

Figure. 16. Shortest distance to nearest charging station in Manhattan.

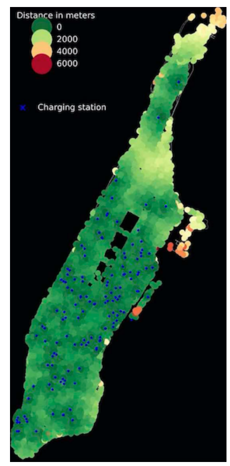

Figure. 17. A visual analysis of the shortest distance to the nearest charging station in Manhattan.

analyses, which conclude that more charging stations are needed in the Northwest of Brooklyn, upper Manhattan, middle of the Bronx, the Southwest of Queens, and the North of Staten Island. The favorable behavior of middle Manhattan shows that there are enough charging stations for Uber Inc. to be fully electric based on the 2014 data.

As for future discussion, Uber pickup data consists of 4.4 million data points measuring the closest distance with 10 nearest charging stations out of 266 charging stations using Dijkstra's algorithm. This creates the first glance of electrification with ride-hailing in New York City. With 4.4 million ride-hailing data points which only counters 15% of total New York City pickups in 2014 [21]. To fully understand the significance of the charging stations and the movement of the New Yorkers, a future study can combine the New York taxi data with Uber data to analyze and optimize the driving distance and charging period of each vehicle. This study attempts to present an image of the electric driving scene for Uber, and the granularity of data serves the purpose well.

References

[1] EIA, "Oil and petroleum products explained: Use of oil," Jan 2021, Last access date: 30 September 2021. [Online]. Available: http://www.eia.gov/energyexplained/oil-and-petroleum-products/use-of-oil.php

[2] K. Korosec, "Uber expands green rides option to 1,400 cities – Techcrunch," Jan 2021, Last access date: 30 September 2021. [Online]. Available: https://techcrunch.com/2021/01/12/uber-expands -green-rides-option-to-1400-cities/

[3] C.-M. Tseng, S. C.-K. Chau, and X. Liu, "Improving viability of electric taxis by taxi service strategy optimization: A big data study of New York city," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 3, pp. 817–829, 2018.

[4] M.-S. Răboacă, I. Băncescu, V. Preda, and N. Bizon, "An optimization model for the temporary locations of mobile charging stations," *Mathematics*, vol. 8, no. 3, p. 453, 2020.

[5] J. Jung and J. Y. Chow, "Effects of charging infrastructure and non-electric taxi competition on electric taxi adoption incentives in New York city," *Transportation Research Record*, vol. 2673, no. 4, pp. 262–274, 2019.

[6] G. Wang, X. Chen, F. Zhang, Y. Wang, and D. Zhang, "Experience: Understanding long-term evolving patterns of shared electric vehicle networks," in *The 25th Annual International Conference on Mobile Computing and Networking*, 2019, pp. 1–12.

 ISBN: 978-1-61208-891-4

[7] G. Wang, X. Xie, F. Zhang, Y. Liu, and D. Zhang, “bcharge: Data-driven real-time charging scheduling for large-scale electric bus fleets,” in *2018 IEEE Real-Time Systems Symposium (RTSS)*. IEEE, 2018, pp. 45–55.

[8] M. Scorrano, R. Danielis, and M. Giansoldati, “Mandating the use of the electric taxis: The case of florence,” *Transportation Research Part A: Policy and Practice*, vol. 132, pp. 402–414, 2020.

[9] S. Zhang, H. Wang, Y.-f. Zhang, and Y.-Z. Li, “A novel two-stage location model of charging station considering dynamic distribution of electric taxis,” *Sustainable Cities and Society*, vol. 51, p. 101752, 2019.

[10] Y. Dong, S. Qian, J. Liu, L. Zhang, and K. Zhang, “Optimal placement of charging stations for electric taxis in urban area with profit maximization,” in *2016 17th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*. IEEE, 2016, pp. 177–182.

[11] J. Jia, C. Liu, and T. Wan, “Planning of the charging station for electric vehicles utilizing cellular signaling data,” *Sustainability*, vol. 11, no. 3, p. 643, 2019.

[12] Kaggle, “Uber pickups in New York city – Kaggle,” Feb 2018, Last access date: 30 September 2021. [Online]. Available: https://www.kaggle.com/fivethirtyeight/uber -pickups-in-new-york-city?select=uber-raw-data- apr14.csv

[13] N. Y. State, “Electric vehicle charging stations in New York — state of New York,” Feb 2018, Last access date: 30 September 2021. [Online]. Available: https://data.ny.gov/Energy-Environment /Electric-Vehicle-Charging-Stations-in-New-York /7rrd-248n

[14] T. Russell and M. Fleis, “Geopandas:,” Jan 2021, Last access date: 30 September 2021. [Online]. Available: https://github.com/geopandas/geopandas

[15] G. Boeing, “OSMnx: A Python package to work with graph-theoretic OpenStreetMap street networks,” *Journal of Open Source Software*, vol. 2, no. 12, 2017.

[16] P. S. Bradley and U. M. Fayyad, “Refining initial points for k-means clustering.” in *ICML*, vol. 98. Citeseer, 1998, pp. 91–99.

[17] J. Durbin, “Efficient estimation of parameters in moving-average models,” *Biometrika*, vol. 46, no. 3/4, pp. 306–316, 1959.

[18] E. W. Dijkstra, “A note on two problems in connexion with graphs,” *Numerische Mathematik*, vol. 1, pp. 269–271, 1959.

[19] S. M. Omohundro, *Five balltree construction algorithms*. International Computer Science Institute Berkeley, 1989.

[20] E. F. Krause, “Taxicab geometry,” *The Mathematics Teacher*, vol. 66, no. 8, pp. 695–706, 1973.

[21] T. W. Schneider, “Analyzing 1.1 Billion NYC Taxi and Uber Trips, with a Vengeance,” Jan 2015, Last access date: 30 September 2021. [Online]. Available: http://toddwschneider.com/posts/analyzing-1-1-billion-nyc-taxi-and-uber-trips-with-a- vengeance/

 ISBN: 978-1-61208-891-4