Preparing for the EV Transition in Baltimore City

IBM Data Science Capstone Project

By: Christopher Peoples – Managing Partner at PP&A

5/5/2020

I. Introduction

Today sustainability and climate-conscious behavior have become significant areas of focus for sectors and businesses around the world. One of the sectors which are looking at a substantial shake-up due to this new way of thinking is the automotive sector. This can be seen with the increased push for the electrification of vehicles. Also, a wave of innovation is also sweeping in promising to fundamentally alter the way people use and share personal vehicles by way of self-driving technology.

With these changes in the mind, the importance of electric vehicle charging stations is set to increase to meet the demand of this emerging electric fleet of vehicles. For this reason, a prudent business owner today should understand a neighborhood's level of preparedness for this changing transportation landscape to be sure their business remains at the front lines this coming change.

Baltimore is no exception to this change. As a city that is known for its big differences on a neighborhood-basis, the selection of the right location for a business is a critical success factor. For this reason, this study will explore the different neighborhoods in Baltimore to highlight the current state of the EV charging station infrastructure build-out effort as well as highlight attractive neighborhoods which are set up well for the EV transition.

II. Data

This analysis requires a variety of data sources ranging from neighborhood data, geolocational data, vehicle registration data, and charging infrastructure data. Using these data sources the study will analyze the current state of venues and EV charging infrastructure in Baltimore city to determine which neighborhoods on the front-line in terms of readiness for increased EV traffic.

A detailed description of the data elements which will be incorporated can be found below:

- NREL Charging Station Data - This is a data set that is available on a national level that includes a variety of unit-level details about charging stations around the country. These details include charging station type, location, access types, and pricing. This data set will serve as the basis for the descriptive analytics and locational placement of EV infrastructure in Baltimore.
- Baltimore City GeoJSON - To map the results of the analysis on a neighborhood basis, I will be using a GeoJSON file containing the polygon shape of neighborhoods

in Baltimore. Also, the polygon shapes will be used to derive neighborhood locations for charging stations. Further, the subsequent venue queries using Foursquare's data set will rely on the coordinates of the centroids for the respective neighborhood polygons. Also, based on the polygon shapes, the area of the respective areas will be derived on a square mile basis, which will be used in EV charger density calculations.

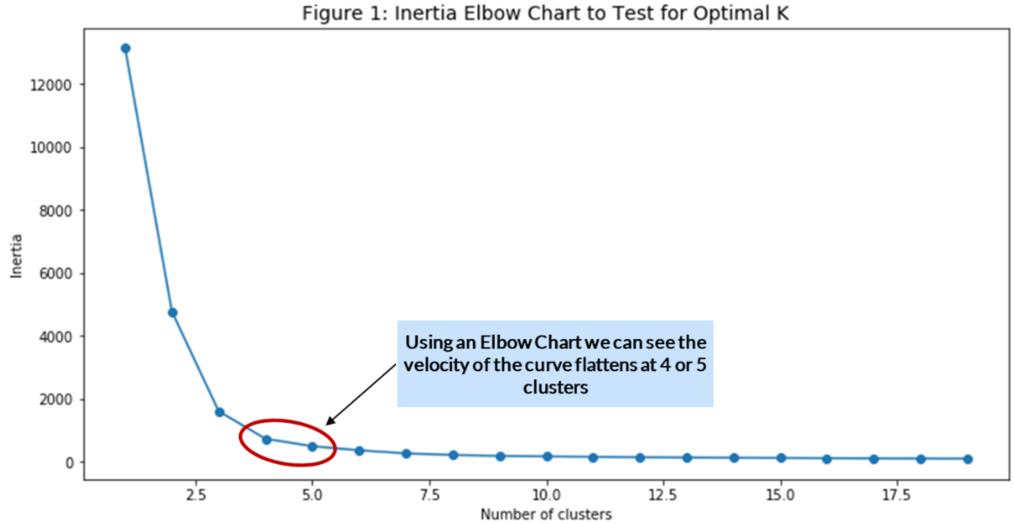
- Maryland Vehicle Registration Data - To get an understanding of the current state of EV penetration as well as to home in on the location of EVs in the Baltimore area, MVA vehicle registration is used. While not used in further elements of the analysis, it is important to validate the justification for conducting the study by examining the current number of EV registered in and around the Baltimore City area.
- Foursquare Venue Data - This data set will be used to find venue types near the center of the respective Baltimore venues to support a K-Means Clustering exercise. As one of the primary elements for the clustering analysis, this will begin to identify patterns within the respective clusters and characterize elements that may lead to EV charging station build-outs.

III. Methodology

The analysis is based on the unsupervised K-Means machine learning algorithm used to build neighborhood clusters in Baltimore city. The reason for using this methodology is that the neighborhoods in the context constitute an unlabeled data set. By employing the K-means method, the study hopes to label and classify neighborhood clusters for the purpose of finding areas ready to support a growing number of EVs.

As the K means process identified iteratively and organically to establish a pre-defined number of clusters, it is important to derive an optimal K, or number of output clusters. To determine the optimal K, the K-Means machine learning algorithm was applied to the combined dataset over a range of K from 1 to 20 to derive neighborhood clusters. Based on the Elbow chart analysis seen below in figure 1, an optimum K of 5 was selected for the analysis

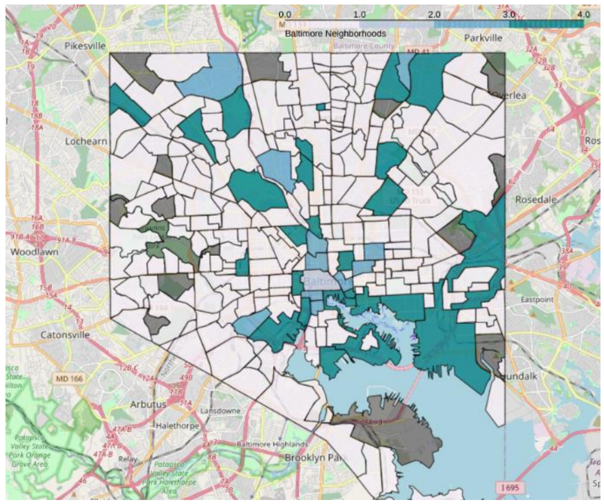
Figure 1: Inertia Elbow Chart to Test for Optimal K



IV. Results

a. Baltimore City Results

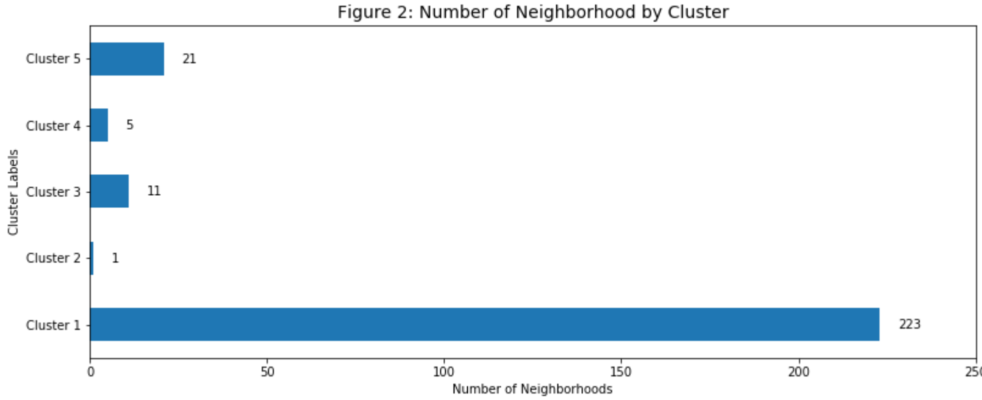
Plot 1 Baltimore City Neighborhood clusters



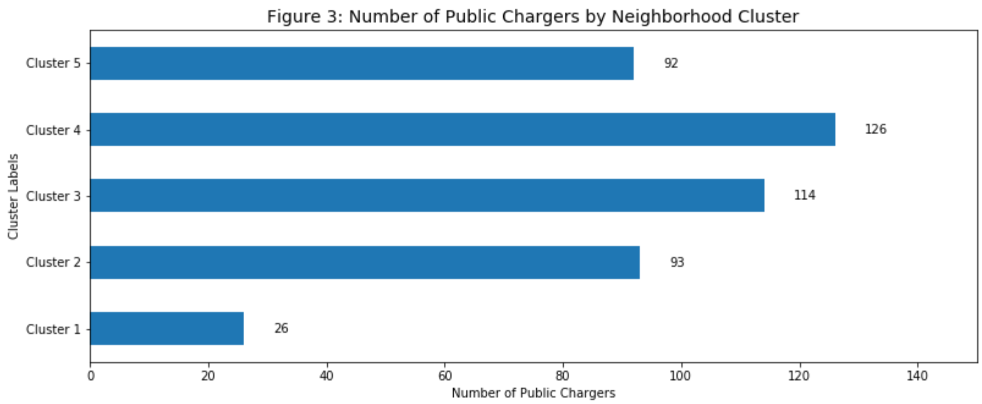
Upon running the K-Means algorithm with a K value of 5 a total of 261 neighborhoods were labeled into distinct clusters. As seen in plot 1 above, based on the heterogeneous distribution of the clusters above, the initial hypothesis around the importance of neighborhood selection would be reinforced by the differences seen in the map.

As per figure 2 below, most neighborhoods fall within the first cluster with a total of 223 neighborhoods, with the remaining 38 neighborhoods spread across the other four clusters. Cluster 2 includes only 1 neighborhood, and the other clusters 3, 4, and 5 include 11, 5, and 21 neighborhoods, respectively.

Figure 2: Number of Neighborhood by Cluster

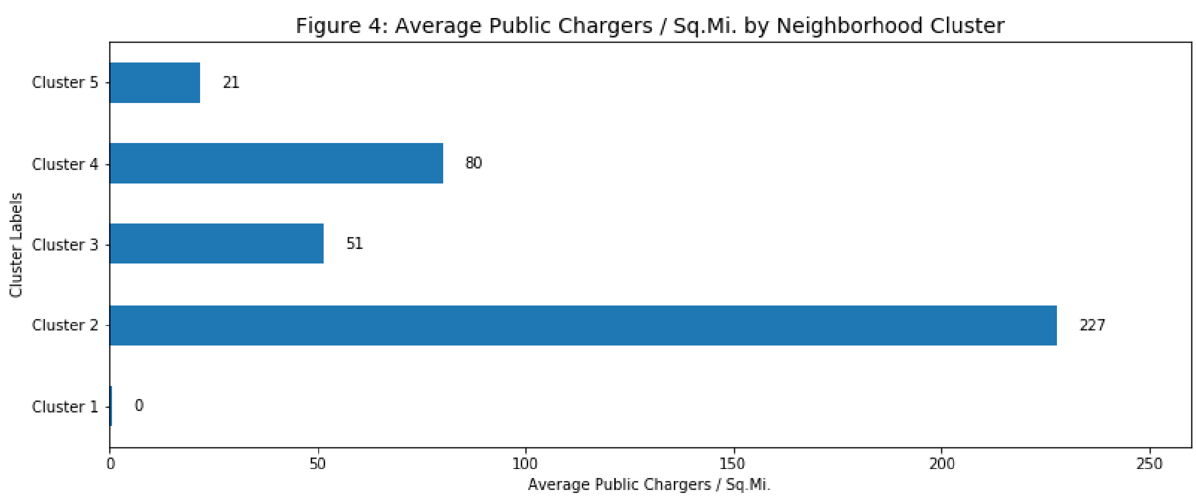


The number of EV charging stations seen in figure 3 shows a total of 451 public charging stations distributed throughout the 5 clusters. Apart from the first cluster, which contains only 26 charges, it appears that the distribution of chargers is relatively uniform among the other 4 clusters.



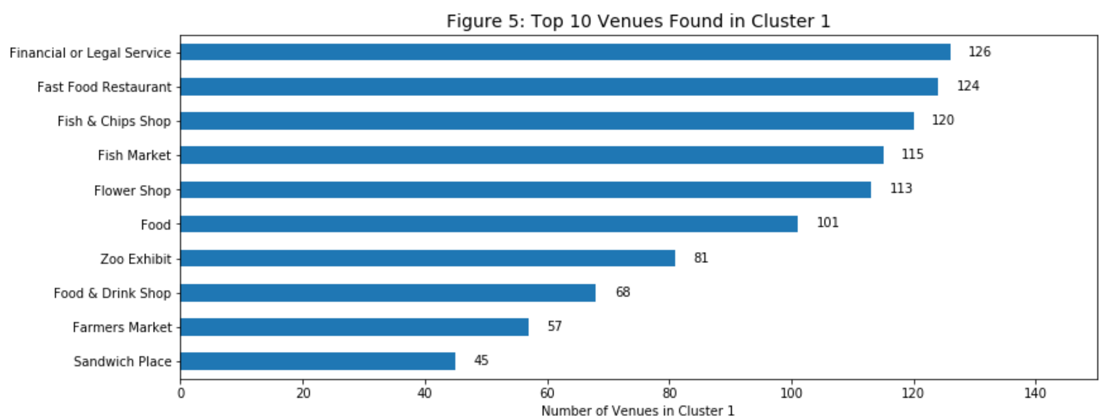
However, this uniformity is only a finding at first glance. To get a better perspective regarding the state of charging infrastructure in a cluster it is necessary to consider neighborhood area as well. The presence of EV chargers in a respective neighborhood can be supplemented with a measure of EV charger density, which must be incorporated into the analysis. A neighborhood EV charger density is calculated by dividing the number of chargers in a neighborhood by the respective total area of the neighborhood. To get a cluster average, the study then takes the average EV charger density of all neighborhoods in a cluster.

With this EV charger density, it is much easier to observe the differences in EV readiness between the respective clusters. Figure 4 shows the big differences between the cluster in terms of EV charger density with cluster 2 showing the highest level of EV charger density per square mile. Cluster 1, also due to its high neighborhood count, has relatively few chargers in comparison to its total area, and differences can also be observed amount cluster 3, 4, and 5.



b. Cluster-Specific Results

i. Cluster 1 Results



Cluster 1 is the largest cluster with the largest volume of total neighborhoods. In total, cluster 1 contains 223 neighborhoods out of the total 261 neighborhoods. However, when it comes to charging stations, this cluster contains comparably few charging stations, especially considering the number of neighborhoods included. Cluster 1 contains only 26 public EV chargers that are spread out over a large area of neighborhoods. The resulting EV charger density is also the lowest among the other clusters rounding out to less than 1 per square mile.

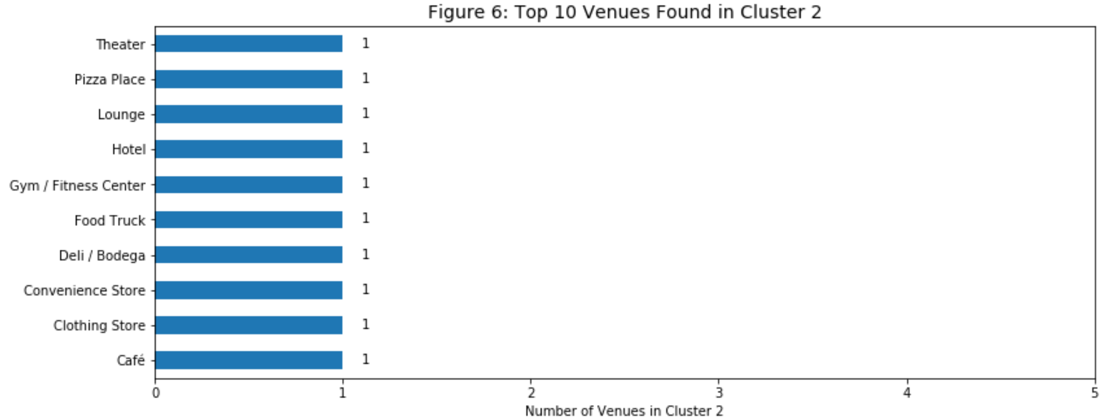
When looking at the venues typically found within the cluster, it is curious to see the top 3 occurrences being financial or legal services and fast-food style restaurants. These may be indicative of a lower likelihood for visitors to stay within the area for extended periods of time.

Table 1 10 Sample Neighborhoods found in Cluster 1

Neighborhood	Area (Sq. Mi.)	Total Public Chargers	Public Chargers / Sq. Mi.
Wyman Park	0.11	0	0.00
Frankford	2.13	0	0.00
Franklintown	0.60	1	1.62
Garwyn Oaks	0.13	0	0.00
Heritage Crossing	0.09	0	0.00
Medfield	0.26	0	0.00
Lauraville	0.59	0	0.00
Carrollton Ridge	0.22	0	0.00
Concerned Citizens Of Forest Park	0.08	0	0.00
Lakeland	0.55	0	0.00

ii. Cluster 2 Results

Figure 6: Top 10 Venues Found in Cluster 2



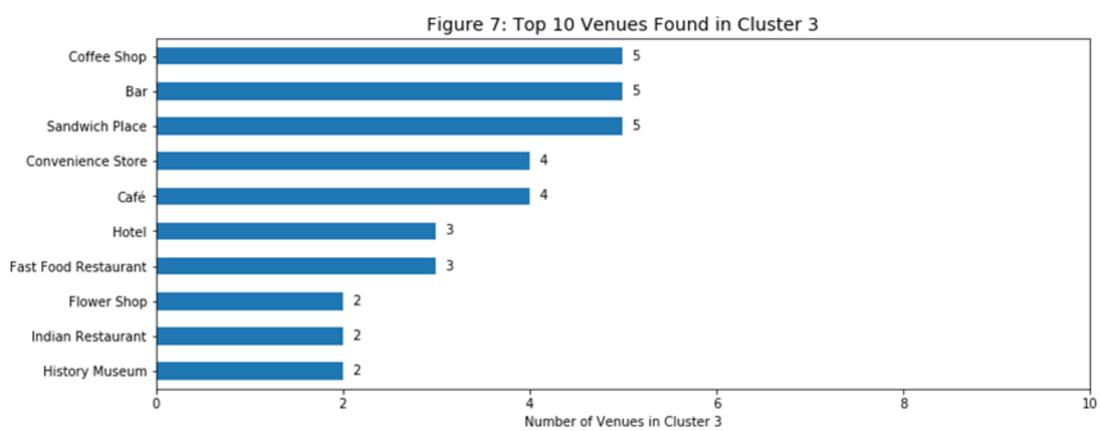
The Downtown neighborhood has been uniquely characterized as cluster 2 and has the highest EV charging station density. The downtown neighborhood itself covers an area of 0.4 square miles, but within this area, 93 of public EV chargers can be found. In addition, there are an additional 3 private stations which fall within this area. Due to the high volume of chargers with this small area, a high-level EV charger density at ~228 EV chargers per square mile can be found in the downtown area.

While the venues of this neighbor were likely less of a driver when categorizing this cluster, the downtown area is known for a high level of office activity. Unfortunately, this is one element that was not capturable via the data sources utilized in this analysis. However, despite this lacking data, the algorithm was able to capture its unique nature nevertheless.

Table 2 Cluster 2 Neighborhood

Neighborhood	Area (Sq. Mi.)	Total Public Chargers	Public Chargers / Sq. Mi.
Downtown	0.41	93	227.99

iii. Cluster 3 Results



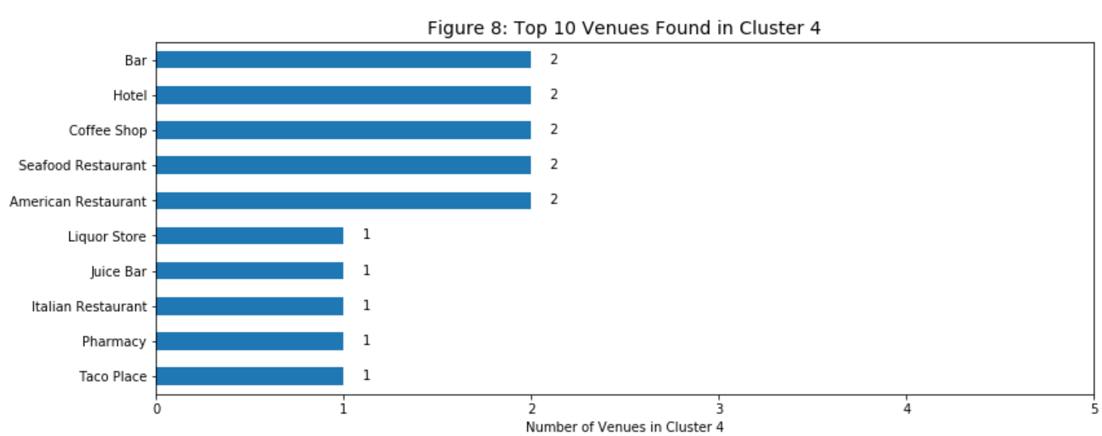
Cluster 3 appears to cater to a more affluent population in smaller areas that are comparably equipped with EV charging stations. A total of 11 neighbors have been allocated to this cluster. Within the cluster, 114 EV charging stations can be found with 1 additional private station registered.

This cluster is characterized by coffee shops, bars, and sandwich places, indicating a bit of recreational feel to the area. In terms of EV charger density, it ranks 3rd among the other clusters showing around 51 EV charging stations per square mile.

Table 3 10 Sample Neighborhoods in Cluster 3

Neighborhood	Area (Sq. Mi.)	Total Public Chargers	Public Chargers / Sq. Mi.
Carroll Park	0.34	10	29.03
Mt Pleasant Park	0.25	8	31.58
Hampden	0.62	14	22.75
Dunbar-Broadway	0.17	14	81.49
Jonestown	0.14	8	59.42
Mid-Town Belvedere	0.22	12	54.25
Middle East	0.20	10	50.43
Mount Vernon	0.22	10	44.06
Mount Washington	1.24	8	6.44
Otterbein	0.10	8	77.93

iv. Cluster 4 Results



Cluster 4 only includes 5 neighborhoods, which are typically seen as areas with a high recreational amenity. Within these 5 neighborhoods, a total of 126 public EV chargers can be found in addition to an additional 4 private stations. In terms of EV charger density, this cluster ranks 2nd among the other clusters with a density of around 80 EV chargers per square mile.

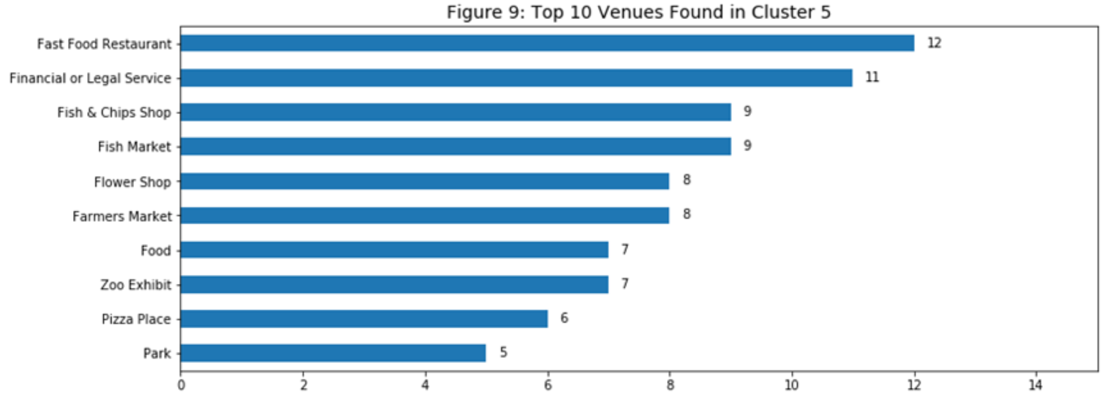
As per the neighborhoods in table 4 below, these neighborhoods include a high level of recreational and gastro entertainment venues. Based on the venues, these seem to be venues that attract crowds of people for extended periods at varying times of the day.

Table 4 Cluster 4 Neighborhoods

Neighborhood	Area (Sq. Mi.)	Total Public Chargers	Public Chargers / Sq. Mi.
Canton Industrial Area	2.49	21	8.43
Fells Point	0.29	23	79.73
Locust Point Industrial Area	0.93	32	34.27
Inner Harbor	0.24	21	86.87
University Of Maryland	0.15	29	191.78

v. Cluster 5 Results

Figure 9: Top 10 Venues Found in Cluster 5



Cluster 5 consists of a total of 21 neighborhoods. In total, 92 public EV chargers can be found within this cluster in addition to an additional 3 private charging stations. The overall EV charging station density for this neighborhood is the second-lowest among the clusters with a value of around 25 EV charging stations per square mile.

Looking at the venues, one can see that the top 3 venues are similar to the venue types, as seen in cluster 1, which are fast food types of establishments and financial and legal services. There is a difference, however, in some of the later venues which show a higher number of parks and farmer's markets.

Table 5 10 Sample Neighborhoods in Cluster 5

Neighborhood	Area (Sq. Mi.)	Total Public Chargers	Public Chargers / Sq. Mi.
Canton	0.57	3	5.25
Carroll - Camden Industrial Area	0.70	6	8.52
Johns Hopkins Homewood	0.26	5	19.47
Loch Raven	0.57	4	7.00
Port Covington	0.33	4	12.03
Pulaski Industrial Area	1.97	4	2.03
Bolton Hill	0.23	4	17.54
Central Park Heights	0.44	4	9.02
Charles North	0.14	4	28.78
Clifton Park	0.56	6	10.69

V. Discussion

As we can see from the results above, there are some significant differences when looking at the readiness of neighborhoods for the EV transition. As EVs become ever more prevalent, I would anticipate that the overall number of charging stations will continue to increase for the foreseeable future. The overall number of stations see in Baltimore is just north of 450 public stations, which currently serves the ~1,800 EV and plugin hybrids registered as of 2020 in both Baltimore City and Baltimore County. This estimate, of course, does not include EVs registered outside of these zones. With the current total number of registered vehicles in Baltimore City and Baltimore Count at close to 1 million vehicles, there is still plenty room to grow as the conversion to EVs continues. The effective distribution of EV charging stations will also evolve into a key consideration when planning for the expansion of future infrastructure build-outs.

VI. Conclusions

Baltimore City is making strides to build out EV charging infrastructure; however, the road ahead is still long. This analysis should be viewed as a starting point. It should be refined to reflect additional differences of neighborhoods better, such as demographics, crime rates, location of workplaces, and a more targeted venue selection. With these further refinements applied to the machine learning methodologies employed, I would expect to derive an improved clustering result in highlighting priority areas for quicker EV infrastructure build-out. As a business owner, looking to come to Baltimore, parking, and transportation convenience translates into ease of access for the customer. As pockets of EV charging activity are already emerging, a business owner can and should begin considering the proximity to charging infrastructure or a neighborhood's EV charging density when deciding where to place a stake in the ground.