INVESTIGATING ELECTRIC VEHICLE CHARGER LOCATIONS IN LONDON

IBM DATA SCIENCE PROJECT

INTRODUCTION

Electric Vehicles, (EVs) are seen as an environmentally friendly alternative to petrol vehicles. Electric vehicles (plug in) do not require fossil fuels, leading to reduced total fuel consumption, and do not emit greenhouse gas emissions. As the United Kingdom strives to meet its net zero carbon target and reduce its dependence on fossil fuels the electrification of transport is becoming increasingly important. There has been, in recent years, a rapid growth in the amount of EVs in both residential and public transport. EVs require supportive infrastructure in the form of charging stations. Residential charging stations refer to chargers located on an owner's private property that is used by the owner alone. Public chargers are available in public areas like sidewalks or public parking garages. These public chargers are the equivalent of petrol fuelling stations and will be required to recharge EVs during normal daily activities especially for commuters or public transport who may require re-charging as the day progresses. There are two broad types of chargers; rapid chargers and regular or slow chargers. As the name implies rapid chargers may charge vehicles in minutes to hours while regular chargers may require a few hours to possibly overnight charging. Ensuring the optimal type, location and number of chargers installed in extremely important especially in a densely populated city like London. As its workday population is significantly high during the day rapid chargers may need to be located in areas with more public parking. While regular chargers may be more suited to more residential areas. This project will therefore seek to investigate the locations of these chargers in London boroughs to determine how they are currently distributed and if there are areas for improvement. The problem statement can be phrased as:

Are the locations of different types of chargers suitable for the characteristics of those locations in London boroughs?

In order to investigate this question location data from foursquare and publicly available data that describe London boroughs will be combined. Clustering will be used to identify similar areas and the defining features will be used to attempt to describe if these groups have good charger concentrations or not. This analysis will be helpful to government planners in identifying areas that may need more of a certain type of chargers. It can also help potential EV owners (residents or commuters) identify how the borough they live or work in may be suited to their charging habits.

DATA

Data required to investigate the problem statement will be sourced from publicly available resources and from Foursquare. Data required to investigate the problem statement include:

1. Location Data will be used to determine the characteristics of each borough. Of particular interest will be the amount of parking facilities and residential buildings. This data will be obtained from Foursquare Venues data. API calls will be used to query the data for parking and residential buildings withing a few kilometres from each borough's GPS location.

Borough	Venue
Barnet	Corsini Residence
Barnet	Residence One
Barnet	Livingston Road Residence
Barnet	NRV Residents' Gym
Bexley	Commonwealth Way Flowers Estate Residence

	Borough	Venue
0	Barking and Dagenham	Car Parking
1	Barking and Dagenham	Chadwell Heath Station Parking
2	Barking and Dagenham	Liberty Parking
3	Barking and Dagenham	Tesco DC Parking
4	Barking and Dagenham	Car Parking In Clements Road

2. A list of London Boroughs along with its GPS co-ordinates, population estimates and area. This data are available at https://en.wikipedia.org/wiki/List_of_London_boroughs and will be read into a python dataframe.

Borough	Inner	Status	Local authority	Political control	Headquarters	Area (sq mi)	Population (2013 est)[1]	Co-ordinates	Nr. in map
Barking and Dagenham [note 1]	NaN	NaN	Barking and Dagenham London Borough Council	Labour	Town Hall, 1 Town Square	13.93	194352	51°33′39″N 0°09′21″E / 51.5607°N 0.1557°E	25
Barnet	NaN	NaN	Barnet London Borough Council	Conservative	Barnet House, 2 Bristol Avenue, Colindale	33.49	369088	51°37′31″N 0°09′06″W / 51.6252°N 0.1517°W	31
Bexley	NaN	NaN	Bexley London Borough Council	Conservative	Civic Offices, 2 Watling Street	23.38	236687	51°27′18″N 0°09′02″E / 51.4549°N 0.1505°E	23
Brent	NaN	NaN	Brent London Borough Council	Labour	Brent Civic Centre, Engineers Way	16.70	317264	51°33′32″N 0°16′54″W / 51.5588°N 0.2817°W	12
Bromley	NaN	NaN	Bromley London Borough Council	Conservative	Civic Centre, Stockwell Close	57.97	317899	51°24′14″N 0°01′11″E / 51.4039°N 0.0198°E	20

3. A table of number / types of EV chargers per London borough which can be found from the UK government publicly available data at http://maps.dft.gov.uk/ev-charging-map/ev-charging.html . Note this table lists number of chargers per 100000 people in each borough in the UK. This data will be read into a dataframe and the London Data will be used.

	ONS code	Local Authority	Total public charging devices	Total public rapid charging devices	Charging devices per 100,000 population
0	K02000001	UNITED KINGDOM	17947	3107	27
1	E06000047	County Durham	102	14	19
2	E06000005	Darlington	28	2	26
3	E06000001	Hartlepool	6	1	6
4	E06000002	Middlesbrough	29	6	21

4. The work day population of London per borough will give an indication of the number of commuters or the actual daily population present in London each day. It is therefore more representative of the potential demand for public chargers. This can be found Greater London Authority's London Datastore located at https://data.london.gov.uk/dataset/daytime-population-borough?resource=7c9b10fb-f8c9-45bb-8844-d5e5cd7f6dca. This data will be read into a dataframe and the appropriate columns (Work day population without tourists) will be used.

	Code	Boroughs	nan	Workday Population (excludes tourists)	In work (employee)	In work (self-employed)	Not in work
	NaN	NaN	NaN	NaN	NaN	NaN	NaN
:	E09000001	City of London	553103	431384	330622	95692	2427
:	E09000002	Barking and Dagenham	178326	164584	41039	8446	61136
.	E09000003	Barnet	356003	331094	101609	41075	108067
:	E09000004	Bexley	211551	194807	56038	12394	71273

5. Another important metric will be the job density per borough which may give some insights into which areas may be more commercial or which areas contain more workers that require rapid charging. This can be found Greater London Authority's London Datastore located at https://data.london.gov.uk/dataset/jobs-and-job-density-borough?resource=116a2961-6c12-4960-ab3a-945c7448a989.

Borough	Job Density
City of London	110.11
Barking and Dagenham	0.49
Barnet	0.67
Bexley	0.56
Brent	0.71

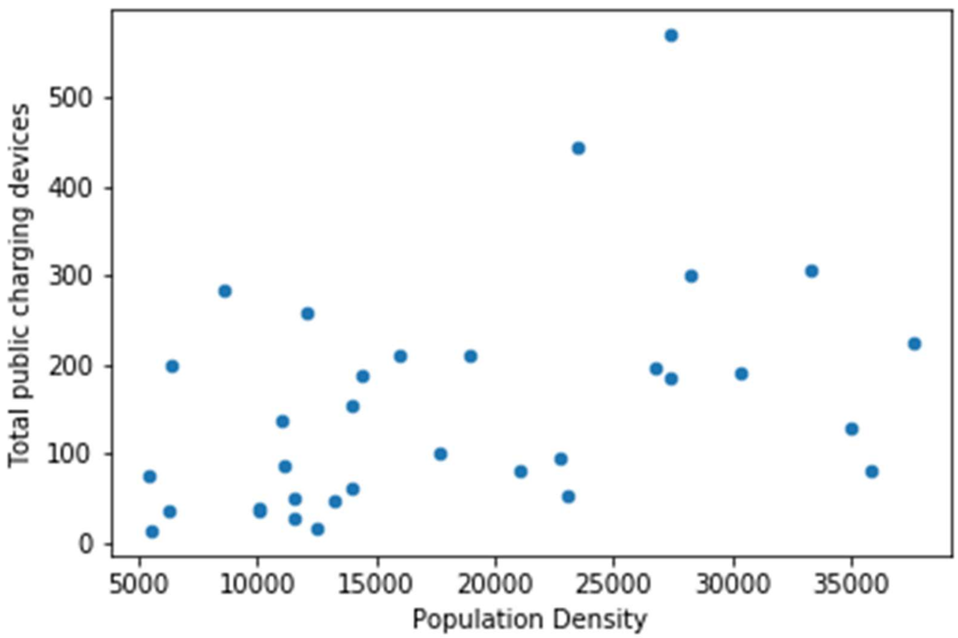
METHODOLOGY

Literature Review:

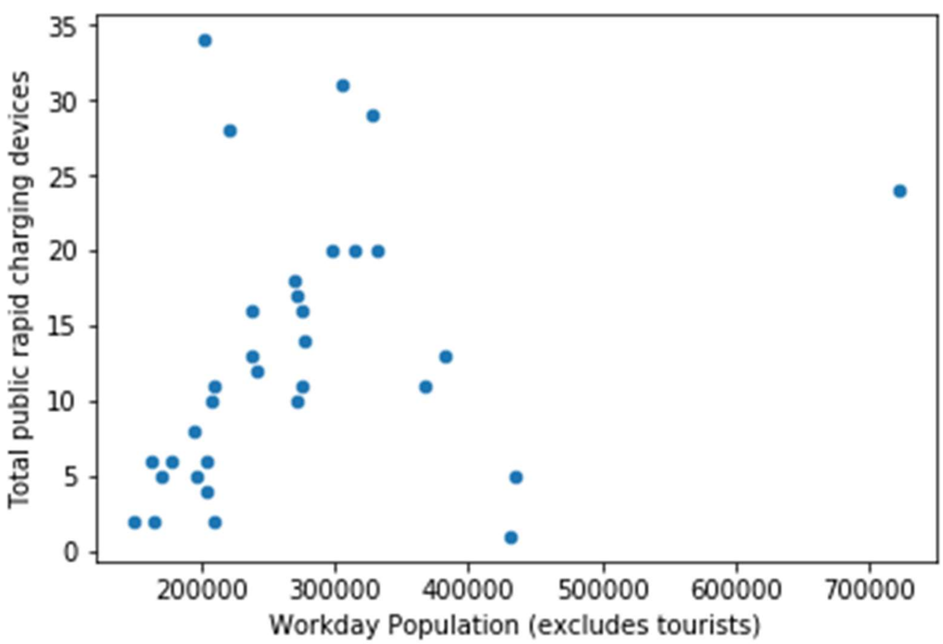
The first step of the project was to carry out a literature review to determine what are some the key feature that should influence Electric Vehicle Data. According to (Li, et al) [1] high population density can be an indicator of high EV adoption rates and therefore reflect the need for increased amounts of EV chargers. The authors also note that socioeconomic factors like job density, or amount or workers versus live-in populations are all important factors. Finally, from a logistics point of view public parking and residential buildings will require charging infrastructure.

Data Cleaning/ Wrangling and Exploration

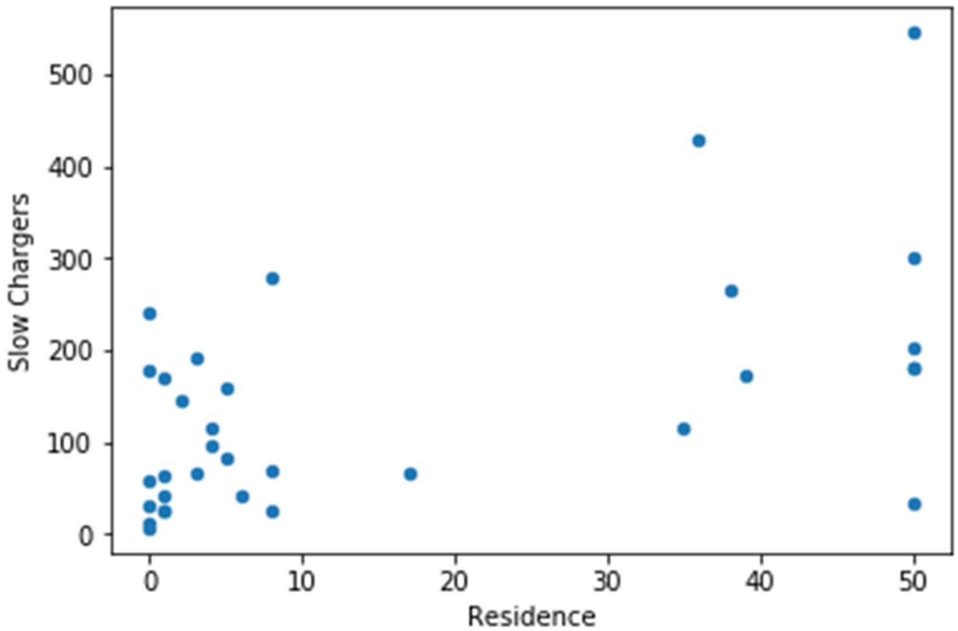
The data described above was obtained from publicly available sources and Foursquare (location data). The data where then clean/ wrangle it into a format to be used for clustering. This involved removing unwanted data, changing the formats to be appropriate (convert strings to integer, represent categorical variables with binary labels). Some initial data exploration was done to confirm some of the relationships identified in the literature review:



There appears to be a positive linear relationship between Population density and total amount of chargers.



There appears to be a positive linear relationship between workday and total amount of rapid chargers.



There appears to be a positive linear relationship between residence and total amount of slow chargers.

The data then had to be combined into a features table.

	Borough	Inner	Area	Population	latitude	longitude	Total public charging devices	Total public rapid charging devices	Charging devices per 100,000 population	Workday Population (excludes tourists)	Job Density	Parking	Residence	Population Density	Slow Chargers
0	Barking and Dagenham	0	13.93	194352	51.560699	0.1557	60	2	28	164584	0.49	5	0.0	13952.045944	58
1	Barnet	0	33.49	369088	51.625198	-0.1517	137	20	35	331094	0.67	3	4.0	11020.842042	117
2	Bexley	0	23.38	236687	51.454899	0.1505	34	8	14	194807	0.56	1	1.0	10123.481608	26
3	Brent	0	16.70	317264	51.558800	-0.2817	209	16	63	274896	0.71	13	3.0	18997.844311	193
4	Bromley	0	57.97	317899	51.403900	0.0198	75	10	23	271896	0.62	1	1.0	5483.853717	65
5	Camden	1	8.40	229719	51.528999	-0.1255	186	5	71	434279	2.17	50	50.0	27347.500000	181
6	Croydon	0	33.41	372752	51.371399	-0.0977	86	20	22	314819	0.60	2	3.0	11156.899132	66
7	Ealing	0	21.44	342494	51.513000	-0.3089	210	31	61	305316	0.66	15	0.0	15974.533582	179

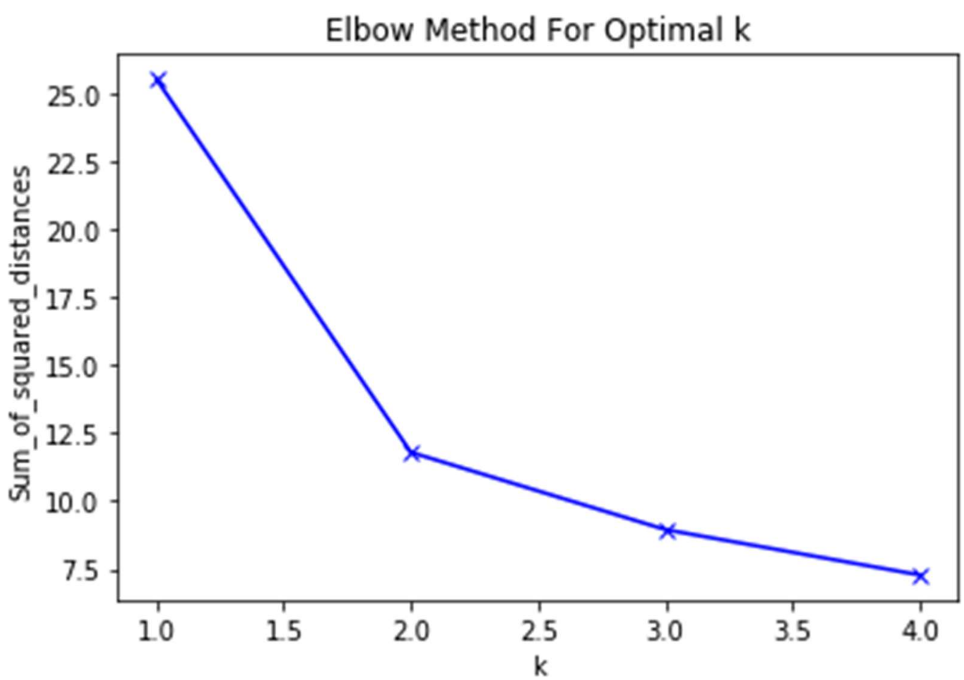
After the relevant features have been gathered there is an additional step required to normalise the data. The reason for this is to allow a proper comparison of the features. Some numbers like the population and population density are very large compared to job density so by normalising the data the relevance of each feature is equally represented. There are several normalisation techniques that can be used. For this project Min – Max normalisation was used. This is achieved by applying:

$$x^{-} = \frac{x - x_{minium}}{x_{maximum} - x_{minimum}}$$

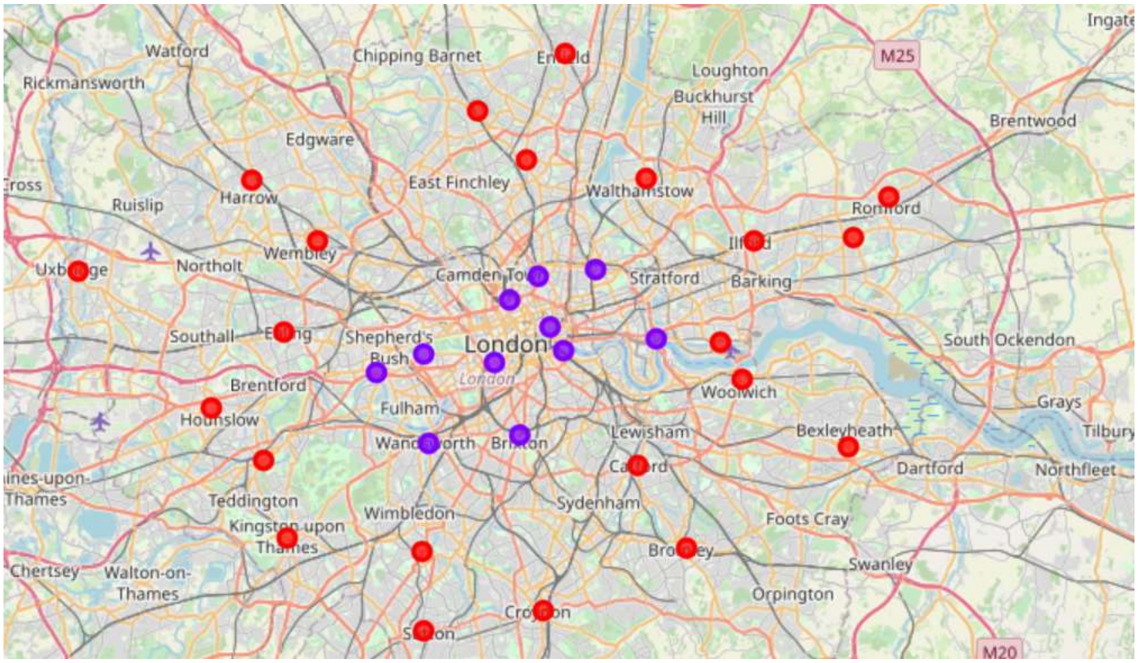
This scales each data point relative to its range resulting in numbers between 1-0.

	Area	Population	Total public charging devices	Total public rapid charging devices	Charging devices per 100,000 population	Workday Population (excludes tourists)	Job Density	Parking	Residence	Population Density	Slow Chargers
0	0.225330	0.512238	0.084381	0.030303	0.056235	0.025740	0.000820	0.081633	0.00	0.263899	0.094620
1	0.569393	0.989982	0.222621	0.575758	0.073350	0.317108	0.002461	0.040816	0.08	0.172552	0.204082
2	0.391557	0.627986	0.037702	0.212121	0.022005	0.078626	0.001458	0.000000	0.02	0.144587	0.035250
3	0.274055	0.848291	0.351885	0.454545	0.141809	0.218770	0.002826	0.244898	0.06	0.421144	0.345083
4	1.000000	0.850027	0.111311	0.272727	0.044010	0.213520	0.002005	0.000000	0.02	0.000000	0.107607
5	0.128056	0.608934	0.310592	0.121212	0.161369	0.497667	0.016133	1.000000	1.00	0.681348	0.322820
6	0.567986	1.000000	0.131059	0.575758	0.041565	0.288629	0.001823	0.020408	0.06	0.176792	0.109462
7	0.357432	0.917272	0.353680	0.909091	0.136919	0.272000	0.002370	0.285714	0.00	0.326927	0.319109
8	0.538610	0.857204	0.046679	0.393939	0.017115	0.223566	0.002005	0.000000	0.02	0.143806	0.033395
9	0.301847	0.702684	0.314183	0.818182	0.149144	0.123333	0.001367	0.142857	0.10	0.279181	0.283859
10	0.109763	0.684559	0.206463	0.333333	0.100244	0.159093	0.003281	1.000000	0.70	0.918891	0.202226

Now K-mean clustering can be used to cluster like boroughs together in order to investigate the properties of the boroughs and the EV charging numbers/ types. K-means clustering is an unsupervised technique that places centroids in the first instance randomly and uses distance from these centroids to assign instances of data into clusters. In this instance the distance metric used is Euclidian distance. After assigning clusters, new centroid positions are calculated based on the centre of the clusters and instance are reassigned. This process is repeated until the centroid positions do not move significantly. At this stage the data are clustered into unique labelled groups (clusters). In order to determine the correct number of clusters for this dataset the elbow method can be used. Basically the model is trained with different number of clusters and the sum of squared distances (an indication of how similar points in a cluster are) is plotted. The point at which segments the graph into a steep reduction on one side and a gradual reduction on the other is taken as the optimum point. This is because beyond this point there is not much gained from adding more clusters.



Two clusters were chosen and K-means used to assign the data into clusters. These clusters were then visualised using the folium library.

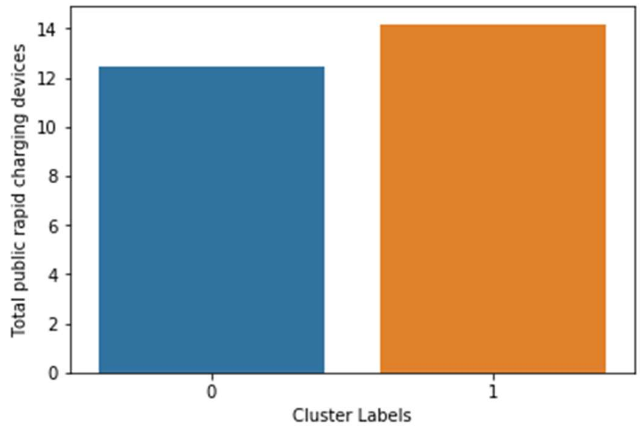


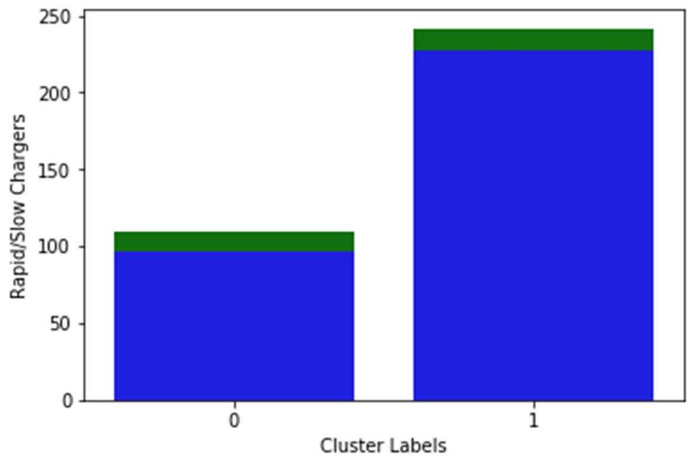
RESULTS

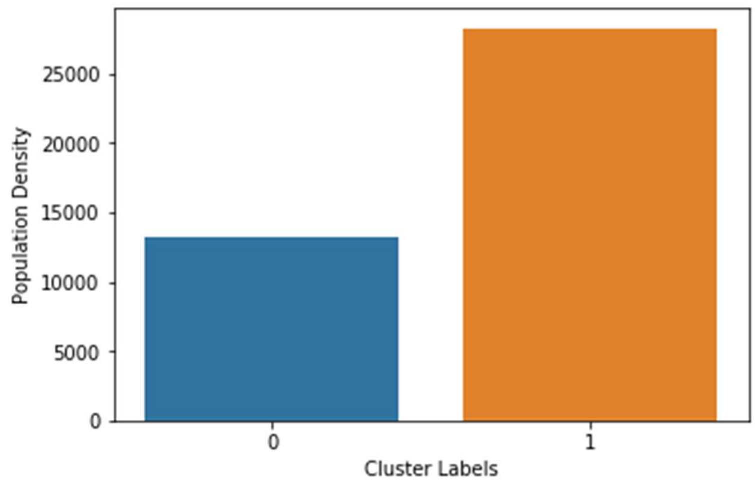
Now that we have data assigned to clusters, we can examine some plots of the data per cluster to gain some insights.

	Cluster Labels	Borough	Inner	Area	Population	latitude	longitude	Total public charging devices	Total public rapid charging devices	Charging devices per 100,000 population	Workday Population (excludes tourists)	Job Density	Parking	Residence	Population Density	Slo Charger
0	0	Barking and Dagenham	0	13.93	194352	51.560699	0.1557	60	2	28	164584	0.49	5	0.0	13952.045944	5
1	0	Barnet	0	33.49	369088	51.625198	-0.1517	137	20	35	331094	0.67	3	4.0	11020.842042	11
2	0	Bexley	0	23.38	236687	51.454899	0.1505	34	8	14	194807	0.56	1	1.0	10123.481608	2
3	0	Brent	0	16.70	317264	51.558800	-0.2817	209	16	63	274896	0.71	13	3.0	18997.844311	19
4	0	Bromley	0	57.97	317899	51.403900	0.0198	75	10	23	271896	0.62	1	1.0	5483.853717	6
5	1	Camden	1	8.40	229719	51.528999	-0.1255	186	5	71	434279	2.17	50	50.0	27347.500000	18
6	0	Croydon	0	33.41	372752	51.371399	-0.0977	86	20	22	314819	0.60	2	3.0	11156.899132	6
7	0	Ealing	0	21.44	342494	51.513000	-0.3089	210	31	61	305316	0.66	15	0.0	15974.533582	17

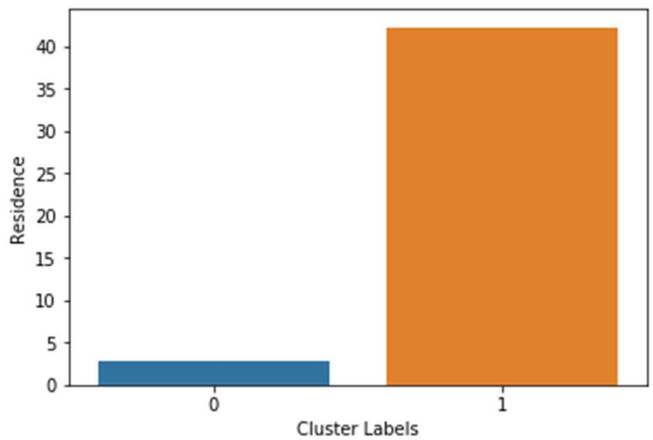
Plots of different Borough Properties for each Cluster:


Total public rapid charging devices
14
12
10
8
6
4
2
0
0
1
Cluster Labels

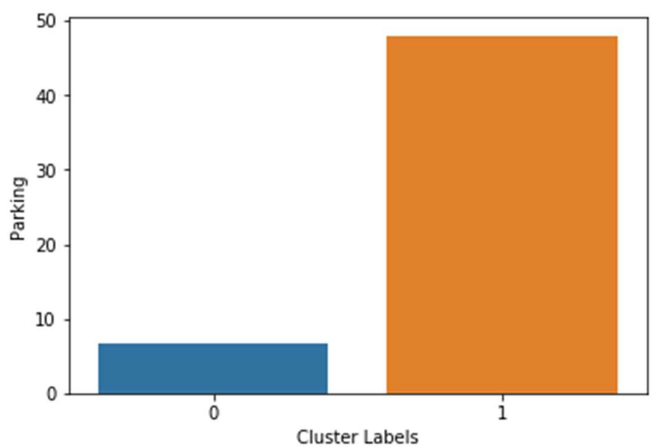

Rapid/Slow Chargers
250
200
150
100
50
0
0
1
Cluster Labels



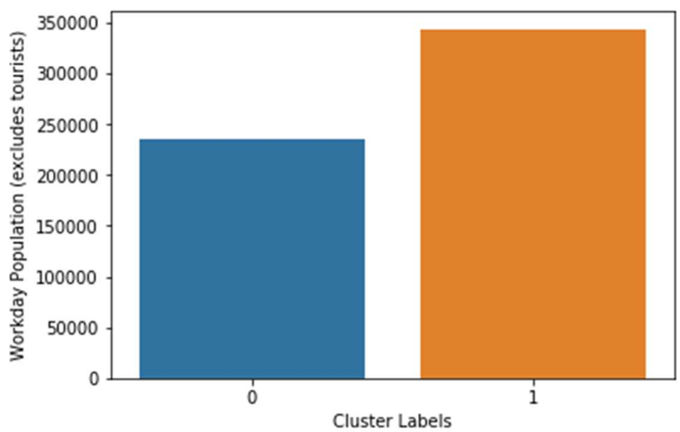
25000
20000
15000
10000
5000
0
Population Density
0
1
Cluster Labels



40
35
30
25
20
15
10
5
0
Residence
0
1
Cluster Labels



50
40
30
20
10
0
Parking
0
1
Cluster Labels



DISCUSSIONS

We see that cluster 1 is more densely populated with higher amounts of both workday and resident populations and also significantly more parking. The data shows that proportionally the number of chargers available matches this disparity, that is, there are more chargers available in cluster 1. However, despite the clear differences in each cluster the number of rapid chargers is almost the same. This is far from optimal.

For Cluster 1, it is clear that much more of both types of chargers may need to be installed in cluster 1 since it has large number of residential and work populations.

For cluster 0 while it is less densely populated its work day population relative to its proportion of residences may mean that more rapid chargers are needed to charge electric vehicles during working hours.

CONCLUSION

It was found that the distribution of EV chargers is far from optimal. It was found that more densely populated areas like city centres are more likely to require more of both types of EV chargers. While areas that have higher working populations like commercial sectors may need more rapid chargers to charge workers cars during daily working hours. A future recommendation could be to repeat this process for many different cities to compare different approaches and identify different relevant features.

REFERENCES

[1] Li, P. Chen and X. Wang, "Impacts of renewables and socioeconomic factors on electric vehicle demands – Panel data studies across 14 countries", *Energy Policy*, vol. 109, pp. 473-478, 2017. Available: 10.1016/j.enpol.2017.07.021 [Accessed 18 June 2020].