Tale of Two Cities A Data Science Take

Introduction

- A Tale of Two cities, a novel written by Charles Dickens was set in London and Paris which takes place during the French Revolution. These cities were both happening then and now. A lot has changed over the years and we now take a look at how the cities have grown.
- London and Paris are quite the popular tourist and vacation destinations for people all around the world. They are diverse and multicultural and offer a wide variety of experiences that is widely sought after. We try to group the neighbourhoods of London and Paris respectively and draw insights to what they look like now.
- The aim is to help tourists choose their destinations depending on the experiences that the neighbourhoods have to offer and what they would want to have. This also helps people make decisions if they are thinking about migrating to London or Paris or even if they want to relocate neighbourhoods within the city. Our findings will help stakeholders make informed decisions and address any concerns they have including the different kinds of cuisines, provision stores and what the city has to offer.

Data Acquisition

We require geographical location data for both London and Paris. Postal codes in each city serve as a starting point. Using Postal codes we use can find out the neighborhoods, boroughs, venues and their most popular venue categories.

In the data collection stage, we begin with collecting the required data for the cities of London and Paris. We need data that has the postal codes, neighbourhoods and boroughs specific to each of the cities.

The data looks like this:

	Location	London borough	Post town	Postcode district	Dial code	OS grid ref	
0	Abbey Wood	Bexley, Greenwich [7]	LONDON	SE2	020	TQ465785	
1	Acton	Ealing, Hammersmith and Fulham[8]	LONDON	W3, W4	020	TQ205805	
2	Addington	Croydon[8]	CROYDON	CR0	020	TQ375645	
3	Addiscombe	Croydon[8]	CROYDON	CR0	020	TQ345665	
4	Albany Park	Bexley	BEXLEY, SIDCUP	DA5, DA14	020	TQ478728	
			2.2		200	20.0	
528	Woolwich	Greenwich	LONDON	SE18	020	TQ435795	
529	Worcester Park	Sutton, Kingston upon Thames	WORCESTER PARK	KT4	020	TQ225655	
530	Wormwood Scrubs	Hammersmith and Fulham	LONDON	W12	020	TQ225815	
531	Yeading	Hillingdon	HAYES	UB4	020	TQ115825	
532	Yiewsley	Hillingdon	WEST DRAYTON	UB7	020	TQ063804	

Data Cleaning

For London, We replace the spaces with underscores in the title. The borough column has numbers within square brackets that we remove using:

```
wiki_london_data.rename(columns=lambda x: x.strip().replace(" ", "_"), inplace=True)
wiki_london_data['borough'] = wiki_london_data['borough'].map(lambda x: x.rstrip(']').rstrip('0123456789').rstrip('['))
```

For Paris, we break down each of the nested fields and create the dataframe that we need:

```
paris_field_data = pd.DataFrame()
for f in paris_raw.fields:
    dict_new = f
    paris_field_data = paris_field_data.append(dict_new, ignore_index=True)

paris_field_data.head()
```

Visualizing the Neighbourhoods of London and Paris

Now that our datasets are ready, using the Folium package, we can visualize the maps of London and Paris with the neighbourhoods that we collected.

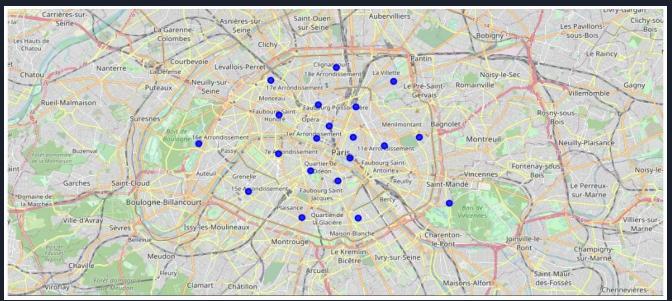
Neighbourhood map of London:



Visualizing the Neighbourhoods of London and Paris

Now that our datasets are ready, using the Folium package, we can visualize the maps of London and Paris with the neighbourhoods that we collected.

Neighbourhood map of Paris:



Model Building - KMeans

Moving on to the most exicitng part - Model Building! We will be using KMeans Clustering Machine learning algorithm to cluster similar neighbourhoods together. We will be going with the number of clusters as 5.

The data looks like this:

borough	town	post_code	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th M Comr Venue
Bexley, Greenwich	LONDON	SE2	51.49245	0.12127	4	Supermarket	Platform	Convenience Store	Historic Site	Train Station	Coffee Shop	Zoo Exhibit	Film Studio
Ealing, Hammersmith and Fulham	LONDON	W3, W4	51.51324	-0.26746	1	Grocery Store	Train Station	Breakfast Spot	Park	Indian Restaurant	Deli / Bodega	Fish Market	Exhib
City	LONDON	EC3	51.51200	-0.08058	2	Coffee Shop	Italian Restaurant	Hotel	Pub	Gym / Fitness Center	Food Truck	Sandwich Place	Beer I
Westminster	LONDON	WC2	51.51651	-0.11968	2	Hotel	Coffee Shop	Pub	Sandwich Place	Café	Italian Restaurant	Restaurant	Theat
Bromley	LONDON	SE20	51.41009	-0.05683	2	Supermarket	Grocery Store	Convenience Store	Hotel	Fast Food Restaurant	Park	Italian Restaurant	Gym / Fitnes Cente
	Bexley, Greenwich Ealing, Hammersmith and Fulham City	Bexley, Greenwich Ealing, Hammersmith and Fulham City LONDON Westminster LONDON	Bexley, Greenwich Ealing, Hammersmith and Fulham City LONDON W3, W4 LONDON EC3 Westminster LONDON WC2	Bexley, Greenwich Ealing, Hammersmith and Fulham City LONDON EC3 51.49245 51.51324 51.51324 51.51200 Westminster LONDON WC2 51.51651	Bexley, Greenwich LONDON SE2 51.49245 0.12127 Ealing, Hammersmith and Fulham LONDON W3, W4 51.51324 -0.26746 City LONDON EC3 51.51200 -0.08058 Westminster LONDON WC2 51.51651 -0.11968	borough town post_code latitude longitude Labels Bexley, Greenwich LONDON SE2 51.49245 0.12127 4 Ealing, Hammersmith and Fulham LONDON W3, W4 51.51324 -0.26746 1 City LONDON EC3 51.51200 -0.08058 2 Westminster LONDON WC2 51.51651 -0.11968 2	borough town post_code latitude longitude Cluster Labels Common Venue Bexley, Greenwich LONDON SE2 51.49245 0.12127 4 Supermarket Ealing, Hammersmith and Fulham LONDON W3, W4 51.51324 -0.26746 1 Grocery Store City LONDON EC3 51.51200 -0.08058 2 Coffee Shop Westminster LONDON WC2 51.51651 -0.11968 2 Hotel	borough town post_code latitude longitude Common Common Venue Bexley, Greenwich LONDON SE2 51.49245 0.12127 4 Supermarket Platform Ealing, Hammersmith and Fulham LONDON W3, W4 51.51324 -0.26746 1 Grocery Store Train Station City LONDON EC3 51.51200 -0.08058 2 Coffee Shop Italian Restaurant Westminster LONDON WC2 51.51651 -0.11968 2 Hotel Coffee Shop Bromley LONDON SE20 51.41009 -0.05683 2 Supermarket Grocery	borough town post_code latitude longitude Cluster Labels Common Venue Common Venue Common Venue Bexley, Greenwich LONDON SE2 51.49245 0.12127 4 Supermarket Platform Convenience Store Ealing, Hammersmith and Fulham LONDON W3, W4 51.51324 -0.26746 1 Grocery Store Train Station Breakfast Spot City LONDON EC3 51.51200 -0.08058 2 Coffee Shop Hotel Hotel Westminster LONDON SE20 51.51651 -0.11968 2 Hotel Coffee Shop Pub	borough town post_code latitude longitude Cluster Labels Common Venue Plate Ealing, Hammersmith and Fulham LONDON EC3 51.51324 -0.26746 1 Grocery Station Breakfast Spot Park City LONDON EC3 51.51200 -0.08058 2 Coffee Shop Italian Restaurant Hotel Pub Sandwich Place Westminster LONDON S	boroughtownpost_codelatitudelongitudeCluster LabelsCommon VenueCommon VenueCommon VenueCommon VenueCommon VenueCommon VenueCommon VenueBexley, GreenwichLONDONSE251.492450.121274SupermarketPlatformConvenience StoreHistoric StationTrain StationEaling, Hammersmith and FulhamLONDONW3, W451.51324-0.267461Grocery StoreTrain StationBreakfast SpotParkIndian RestaurantCityLONDONEC351.51200-0.080582Coffee ShopItalian RestaurantHotelPubGym / Fitness CenterWestminsterLONDONWC251.51651-0.119682HotelCoffee ShopPubSandwich PlaceCaféBromleyLONDONSE2051.41009-0.056832SupermarketGroceryConvenienceHotelFast Food	borough town post_code latitude longitude Cluster Labels Common Venue Comm	borough town post_code latitude longitude Cluster Labels Common Venue Venue Venue Common Venue Ven

Visualizing the clustered Neighbourhoods

Our data is processed, missing data is collected and compiled. The Model is built. All that's remaining is to see the clustered neighbourhoods on the map. Again, we use Folium package to do so.

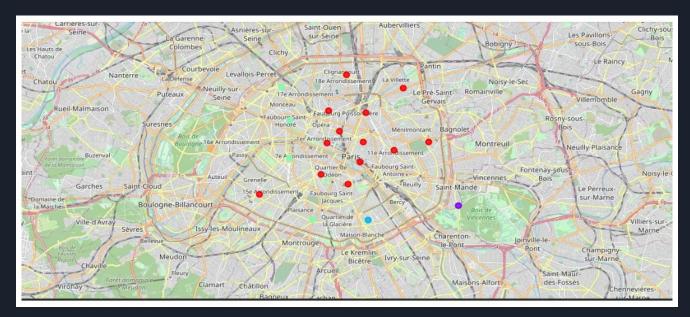
Map of clustered neighbourhoods of London:



Visualizing the clustered Neighbourhoods

Our data is processed, missing data is collected and compiled. The Model is built. All that's remaining is to see the clustered neighbourhoods on the map. Again, we use Folium package to do so.

Map of clustered neighbourhoods of Paris:



Results and Discussion

- The neighbourhoods of London are very multicultural. There are a lot of different cusines including Indian, Italian, Turkish and Chinese. London seems to take a step further in this direction by having a lot of Restaurants, bars, juice bars, coffee shops, Fish and Chips shop and Breakfast spots. It has a lot of shopping options too with that of the Flea markets, flower shops, fish markets, Fishing stores, clothing stores. The main modes of transport seem to be Buses and trains. For leisure, the neighbourhoods are set up to have lots of parks, golf courses, zoo, gyms and Historic sites. Overall, the city of London offers a multicultural, diverse and certainly an entertaining experience.
- Paris is relatively small in size geographically. It has a wide variety of cuisines and eateries including French, Thai,
 Cambodian, Asian, Chinese etc. There are a lot of hangout spots including many Restaurants and Bars. Paris has a lot
 of Bistro's. Different means of public transport in Paris which includes buses, bikes, boats or ferries. For leisure and
 sightseeing, there are a lot of Plazas, Trails, Parks, Historic sites, clothing shops, Art galleries and Museums. Overall,
 Paris seems like the relaxing vacation spot with a mix of lakes, historic spots and a wide variety of cuisines to try out.

Conclusion

• The purpose of this project was to explore the cities of London and Paris and see how attractive it is to potential tourists and migrants. We explored both the cities based on their postal codes and then extrapolated the common venues present in each of the neighbourhoods finally concluding with clustering similar neighbourhoods together.

- We could see that each of the neighbourhoods in both the cities have a wide variety of experiences to offer which is unique in it's own way. The cultural diversity is quite evident which also gives the feeling of a sense of inclusion.
- Both Paris and London seem to offer a vacation stay or a romantic getaway with a lot of places to explore, beautiful landscapes, amazing food and a wide variety of culture. Overall, it's upto the stakeholders to decide which experience they would prefer more and which would more to their liking.