A Tale of Two cities
Clustering the Neighbourhoods of London and Paris

November
2020

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.

Business Problem

- 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 model also helps people make decisions if they are thinking about migrating to London or Paris or even if they wish 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 Description

A. London

To derive our solution, We scrape our data from https://en.wikipedia.org/wiki/List_of_areas_of_London

This Wikipedia page has information about all the neighbourhoods, we limit it to London

- *borough*: Name of Neighborhood
- *town*: Name of the borough
- *post_code*: Postal codes for London.

This Wikipedia page lacks information about geographical locations. To solve this problem, we use *ArcGIS API.*

Data Description

B. ArcGIS API

ArcGIS Online enables you to connect people, locations, and data using interactive maps. Work with smart, data-driven styles and intuitive analysis tools that deliver location intelligence. Share your insights with the world or specific groups.

More specifically, we use ArcGIS to get the geographical locations of the neighbourhoods of London. Adding the following columns to our initial data set prepares our data.

- ✓ latitude: Latitude for Neighborhood
- ✓ longitude: Longitude for Neighborhood

Data Description

C. Paris

To derive our solution, We leverage JSON data available at https://www.data.gouv.fr/fr/datasets/r/e88c6fda-1d09-42a0-a069-606d3259114e

The JSON file has data about all the neighbourhoods in France; we limit it to Paris.

- postal_code: Postal codes for France
- nom_comm: Name of Neighborhoods in France
- nom_dept: Name of the boroughs, equivalent to towns in France
- geo_point_2d: Tuple containing the latitude and longitude of the Neighborhoods

Foursquare API Data

- We will need data about different venues in different neighbourhoods of that specific borough. To gain that information, we will use “Foursquare” location information. Foursquare is a location data provider with information about all manner of venues and events within an area of interest. Such information includes venue names, locations, menus and even photos. As such, the foursquare location platform will be used as the sole data source since all the stated required information can be obtained through the API.

The information obtained per venue as follows:

- Neighbourhood: Name of the Neighborhood
- Neighbourhood Latitude: Latitude of the Neighborhood
- Neighbourhood Longitude: Longitude of the Neighborhood
- Venue: Name of the Venue
- Venue Latitude: Latitude of Venue
- Venue Longitude: Longitude of Venue
- Venue Category: Category of Venue

Methodology

We will be creating our model with the help of Python, so we start by importing all the required packages. The code is available on GitHub to follow along.

Package breakdown:

- **Pandas: To collect and manipulate data in JSON and HTML and then data analysis.**
- **requests: Handle HTTP requests.**
- **matplotlib: Detailing the generated maps.**
- **folium: Generating maps of London and Paris.**
- **sklearn: To import K Means machine learning model.**

DATA COLLECTION FEATURE SELECTION

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.

DATA Preprocessing

For London, We replace the spaces with underscores in the title.

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

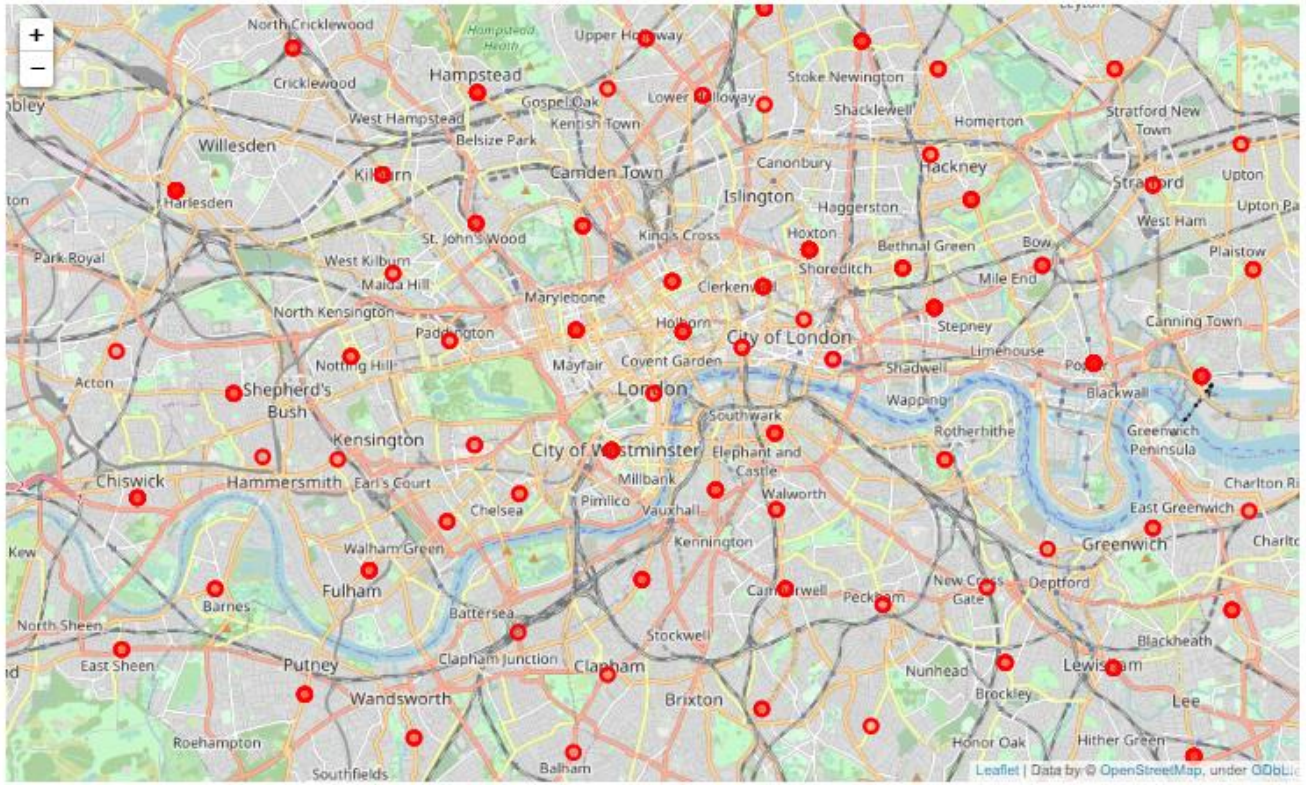
FEATURE SELECTION

For both of our datasets, we need only the borough, neighbourhood, postal codes and geolocations (latitude and longitude).

FEATURE ENGINEERING

Both of our Datasets contain information related to all the cities in the country. We can narrow down and further process the data by selecting only the neighbourhoods of ‘London’ and ‘Paris’.

Visualizing the Neighborhoods of London and Paris



RESULT AND DISCUSSION

The neighbourhoods of London are very multicultural. There are a lot of different cuisines 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 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 Bistros. 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 its 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 up-to-the stakeholders to decide which experience they would prefer more and which would more to their liking.

REFERENCES

1. The Battle of Neighbourhood — My London's Perspective by Dayo John
2. Foursquare API
3. ArcGIS API