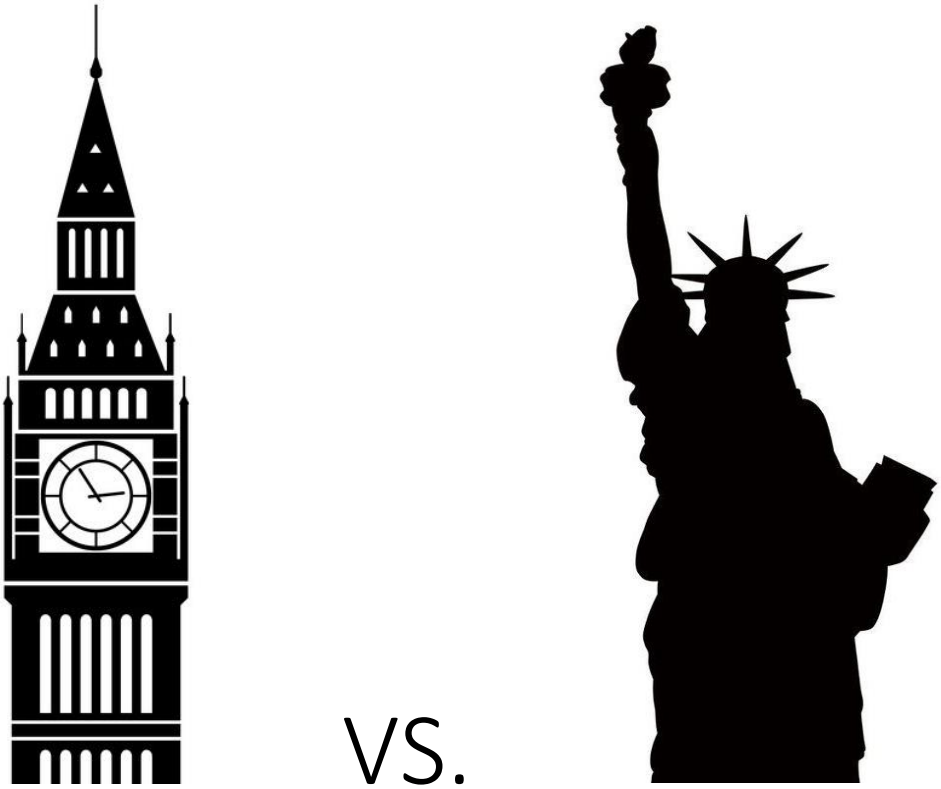
The Battle of Neighbourhoods

Michal Polski



Final Project: London vs. New York

1. Introduction / Background

The current world allows us to travel and move between the cities, countries and even continents. The world became more global than ever and the distance does not matter that much anymore. Having in mind the freedom of movement I have decided to look into two cities, which are as similar as different. Global, almost equally large, important and iconic cities. In the same time these two cities represent very different culture and history. The cities which I have chosen are:

- New York
- London

My aim is to help understand the venues landscape in each of the cities.

2. 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 also helps people make decisions if they are thinking about migrating to London or New York or even if they want to relocate neighbourhoods within the city. The project findings will help stakeholders make informed decisions and address any concerns they have including the different kinds of cuisines, activity spots, grocery stores and what the city has to offer.

3. Data Description

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.

3.1. 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 only.

a) *borough*: Name of Neighbourhood
b) *town*: Name of borough
c) *post_code*: Postal codes for London

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

3.2. 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 geo locations of the neighbourhoods of London. The following columns are added to our initial dataset which prepares our data.

a) *latitude*: Latitude for Neighbourhood
b) *longitude*: Longitude for Neighbourhood

3.3. New York Data

Data Source I (json file):

https://cocl.us/new_york_dataset

Geo-spacial data of the New York to get a better understanding of the neighbourhoods in it and their corresponding locations in the Folium map would make certain things clear for the Project. This will be achieved using the acquired data and visualize the same using Choropleth maps.

Data Source II (json file):

https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701EN-SkillsNetwork/labs/newyork_data.json

3.4. Foursquare API Data

The data is going to be collected/acquired from the Foursquare API about the various venues in each neighborhood of New York city. 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.

After finding the list of neighbourhoods, we then connect to the Foursquare API to gather information about venues inside each and every neighbourhood. For each neighbourhood, we have chosen the radius to be 500 meters.

The data retrieved from Foursquare contained information of venues within a specified distance of the longitude and latitude of the postcodes. The information obtained per venue as follows:

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

Based on all the information collected for both London and New York city, we have sufficient data to build our model. We cluster the neighbourhoods together based on similar venue categories. We then present our observations and findings. Using this data to our stakeholders, so they can take the necessary decision.