The Battle of the Neighbourhoods

Table of Contents

Introduction 1
Business problem and Target Audience 1
Data Overview 2
Reference 3

Introduction

This capstone project is a peer-graded assignment for Applied Data Science Capstone, part of the IBM Data Science Professional Certificate course in Coursera. The aim is to leverage the skills learnt during the course and apply them to a real-world exercise by defining a problem with a target audience and solving it with location data. What is location data? According to Aklson, 2021, Location data describes places and venues with their geographical location and their characteristic using geographical coordinates (latitude and longitude values) to extract information about the kind of places or venues near a given location. There are many location data providers, but we will be using Foursquare for this project. In data science, data is not necessarily available all the time for projects, so we will have to be innovative about how to source for the dataset needed.

Business problem and Target Audience

London, situated on the banks of the River Thames, is the capital and largest city of England and the United Kingdom. London was initially called Londinium and founded by the Romans; it has been a significant settlement for two millennia (London - Wikipedia, 2021). London is a multicultural city with a mixed culture from all around the world, no matter what part of the city you are in, the streets are inundated with different cultures since ancient time, and according to Chynoweth, 2013, during the Olympics, London was discovered to have communities from almost everywhere in the world. Taking advantage of London's diversity, J'aime, a big restaurant chain in France, is thinking of opening a French Caribbean Restaurant in London and has employed my service as a data scientist to look for a perfect location for their new restaurant. In data science, data is not necessarily available all the time for projects, so we will have to scrap for it to get the necessary dataset needed. The aim is to scrap data from the internet about borough, code postal, longitude and latitude of London city and leverage Foursquare location data by comparing the geographical location in

London using the geographical coordination to find a suitable place for a French Caribbean Restaurant.

Data Overview

As the dataset is not readily available, we used web scraping to data collected from Wikipedia https://en.wikipedia.org/wiki/List_of_areas_of_London containing London's borough and postcode and put it into a data frame then did some the data wrangling and cleaning. We then used the Geocoder package in Python to get the geographical coordinates: the longitude and latitude of each borough. Once the dataset was ready, we then used the Foursquare API to get the available information about the venues in the respective neighbourhoods of the dataset.

Out[3]:

	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

Fig 1: The image of the dataFrame collected from Wikipedia.

28]:

	PostalCode	Borough	Neighbourhoods	Latitude	Longitude
0	E1	Tower Hamlets	Mile End, Ratcliff, Shadwell, Spitalfields, St...	51.057920	-2.413281
1	E10	Hackney	Lea Bridge	53.768073	-2.759532
2	E10	Waltham Forest	Leyton	51.571815	0.010935
3	E11	Redbridge	Wanstead	51.577964	0.023570
4	E11	Redbridge, Waltham Forest	Snaresbrook	51.580970	0.021340

Fig 3: The image of the dataFrame after getting the geographical coordinates with the Geocoder package in Python.

2]:

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Mile End, Ratcliff, Shadwell, Spitalfields, St...	51.057920	-2.413281	The Nog Inn	51.055782	-2.410609	Pub
1	Mile End, Ratcliff, Shadwell, Spitalfields, St...	51.057920	-2.413281	Co-op Food	51.057239	-2.408917	Grocery Store
2	Mile End, Ratcliff, Shadwell, Spitalfields, St...	51.057920	-2.413281	Hooga Coffee	51.056794	-2.409374	Coffee Shop
3	Mile End, Ratcliff, Shadwell, Spitalfields, St...	51.057920	-2.413281	The Dolphin	51.056716	-2.407890	Pub
4	Lea Bridge	53.768073	-2.759532	Pig & Whistle	53.766564	-2.765502	Pub

Fig 3: The image of the dataFrame got from the Foursquare API

Reference

Aklson, A., 2021. [online] Coursera. Available at: https://www.coursera.org/learn/applied-data-science-capstone/lecture/kozeg/location-data-providers [Accessed 9 April 2021].

Chynoweth, C., 2013. Getting to know multicultural London. [online] Thetimes.co.uk. Available at: https://www.thetimes.co.uk/article/getting-to-know-multicultural-london-33d602bwxxs [Accessed 25 April 2021].

En.wikipedia.org. 2021. London - Wikipedia. [online] Available at: https://en.wikipedia.org/wiki/London [Accessed 26 April 2021].