IBM APPLIED DATA SCIENCE CAPSTONE

EXPLORING NEIGHBOURHOODS IN LONDON

by

IBRAHEEM KEKERE-EKUN

30th January, 2021

Table of Contents

1. Introduction

1.1. Business problem

London city, capital of the United Kingdom is one of the world's most important global cities. It is among the oldest of the world's great cities—its history spanning nearly two millennia—and one of the most cosmopolitan. By far Britain's largest metropolis, it is also the country's economic, transportation, and cultural centre. London is a very diverse city with a population of over 8 million people with a land area of 1,572 square kilometres.

Finding a destination, or a place to move to can be challenging. This project aims to help tourists, business owners and people that intend to move to London, find the best destinations and neighbourhoods based on the most common venues, crime rates and housing prices. This project analyzes the neighbourhoods in London and clusters them based on venues. The clustering algorithm that will be used is K Means. K Means was chosen because it is one of the simplest algorithms for clustering.

1.2. Target audience

The results of this project can be used by various stakeholders including business owners and people who want to move to London. Visual and interactive maps for anyone to view the results of this project.

2. Data

2.2. Data requirements

Data containing the boroughs and neighbourhoods is required, as well as data containing neighbourhood locations, crime rates, house prices and venues.

2.3. Data description

The data containing neighbourhoods in London was scraped from Wikipedia. This webpage contains information about neighbourhoods, postcode districts and boroughs in London.

Location	London borough	Post town	Postcode district	Dial code	OS grid ref
Abbey Wood	Bexley, Greenwich [7]	LONDON	SE2	020	TQ465785
Acton	Ealing, Hammersmith and Fulham[8]	LONDON	W3, W4	020	TQ205805
Addington	Croydon[8]	CROYDON	CR0	020	TQ375645
Addiscombe	Croydon[8]	CROYDON	CR0	020	TQ345665
Albany Park	Bexley	BEXLEY, SIDCUP	DA5, DA14	020	TQ478728
Aldborough Hatch	Redbridge[9]	ILFORD	IG2	020	TQ455895
Aldgate	City[10]	LONDON	EC3	020	TQ334813
Aldwych	Westminster[10]	LONDON	WC2	020	TQ307810
Alperton	Brent[11]	WEMBLEY	HA0	020	TQ185835
Anerley	Bromley[11]	LONDON	SE20	020	TQ345695
Angel	Islington[8]	LONDON	EC1, N1	020	TQ345665
Aperfield	Bromley[11]	WESTERHAM	TN16	01959	TQ425585
Archway	Islington[12]	LONDON	N19	020	TQ285875

Figure 1: Wikipedia page showing list of areas in London

The Wikipedia page lacked information about the coordinates of the neighbourhoods. Therefore, Geocoder library was used to fetch the coordinates for each neighbourhood.

	Neighbourhood	Borough	Area	Latitude	Longitude
0	Abbey Wood	Bexley, Greenwich	SE2	51.482490	0.119194
1	Acton	Ealing, Hammersmith and Fulham	W3, W4	51.495053	-0.220839
2	Addington	Croydon	CR0	51.344980	-0.012265
3	Addiscombe	Croydon	CR0	51.344980	-0.012265
4	Albany Park	Bexley	DA5, DA14	51.457075	0.184164

Figure 2: Geographical coordinates of neighbourhoods in London

The data containing crime rates in London boroughs was gotten from finder.com. The webpage contained information of the crime rates per 1000 people and the total number of crimes reported for each borough in 2020.

Borough	Crime count	Crime rate
Barking and Dagenham	21,189	99.5 per 1,000
Barnet	31,911	80.6 per 1,000
Bexley	18,323	73.8 per 1,000
Brent	31,725	96.2 per 1,000

Figure 3: webpage containing crime rates in London

The average housing price, as well as price per square feet in each London neighbourhood was gotten from propertydata.co.uk.

Property data for London

Prices, rents, yields, growth & more in London

<< Back to all cities

London is a major global city and world cultural capital, with strengths in the arts, commerce, education, entertainment, fashion, finance, healthcare, media, professional services, research and development, tourism, and transportation. A major settlement for 2,000 years, London is the world's largest financial centre and the 5th or 6th richest city in the world.

The property investment capital of the UK, London has a diverse array of housing and communities across a giant urban area. You can find different types of long-let investment property in various areas to suit most budgets, and being the world's most-visited city as measured by international arrivals means there is no shortage of short-let opportunities too. Having the largest concentration of higher education institutes in Europe guarantees that London has an active student property rental market as well.

There are 217 postcode districts with significant property markets in and around London, and here we break down key property data for each.

Make informed property decisions - sign up today

Area	Avg yield	Avg price	£/sqft	5yr +/-	Explore data
BR1	3.4%	£445,609	£465	+13%	Explore data
BR2	3.3%	£484,598	£475	+14%	Explore data
BR3	3.4%	£446,496	£499	+14%	Explore data
BR5	3.6%	£417,726	£426	+16%	Explore data
BR6	3.0%	£538,239	£477	+14%	Explore data
BR7	2.8%	£605,254	£501	+13%	Explore data

Figure 4: webpage showing average house prices per square feet [£] in each London borough

Foursquare api was used to get data of all the venues and their categories for each London neighbourhood. Foursquare is an American location data provider.

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Abbey Wood	51.482490	0.119194	Clam Field Recreation Area	51.479830	0.121023	Athletics & Sports
1	Abbey Wood	51.482490	0.119194	Abbey Wood Caravan Club	51.485502	0.120014	Campground
2	Abbey Wood	51.482490	0.119194	Bostall Woods	51.480026	0.115305	Forest
3	Abbey Wood	51.482490	0.119194	99 Bus to WW	51.481398	0.125636	Bus Station
4	Acton	51.495053	-0.220839	The Queen's Head	51.495975	-0.220711	Pub

Figure 5: dataframe containing venues, venue categories and venue coordinates in London