Predicting Gentrification in London

By Tim de Zwart, April 2020

Link of notebook:

1. Introduction: Business Problem

Gentrification is a common and controversial topic in politics and urban planning. Although it often increases the value of a neighbourhood, it usually results in a great demographic change, shifting the neighbourhoods composition and average household income, and with this displacing the current population of the neighbourhood [1].

Measuring gentrification is a difficult task, however income change and raise in property value over time in a neighbourhood have shown to be good indicators. Furthermore, we all know the stereotypical images of gentrified neighbourhoods: trendy coffee places, boutique shops, sustainable stores, etc. The goal of this report is to analyse the type of venues that can generally be found in gentrified areas, and then show which areas with a low real estate price are similar to these neighbourhoods, in order to predict which parts of the cities are prone to gentrification soon.

The city chosen for this analysis was London. This choice was made because it's a vibrant city with a diverse population (and therefore many different venues) that has been known to suffer from gentrification in the past. The data for London is readily available which makes it a perfect study subject for this project.

The stakeholders for this project are manifold: current inhabitants of areas that want to know what will happen with the value of their property, urban planners that want to predict what will happen to the city, real estate companies wanting to know where to make their next investment, etc.

Sources

[1] https://en.wikipedia.org/wiki/Gentrification

2. Data

Based on the problem definition, the factors that will help us determine which neighbourhood is prone to gentrification, are:

- The type of venues in a certain area. Areas with many venues that are similar as the venues in areas that are already gentrified are probably prone to gentrify soon.
- The real estate prices in London per area. Areas that currently have a low real estate price are more prone to gentrification than areas that already have a high real estate price.

The data sources that are needed to perform this analysis are:

- A list of areas and boroughs of London, which is found here: https://en.wikipedia.org/wiki/List_of_areas_of_London

- Geographical coordinates (latitude and longitude) of the listed areas in order to plot on maps and find nearby venues. These coordinates are retrieved using Python's geopy library.
- A list of the most gentrified areas over the last years in London in order to compare and cluster the areas, which can be found here: https://londonist.com/london/news/these-are-london-s-most-gentrified-boroughs-but-are-they-really. There boroughs were plotted on a map using a GEOJSON file from https://skgrange.github.io/www/data/london_boroughs.json.
- The number of venues as well as the types of these venues for each area are retrieved using the Foursquare API
- The real estate prices per borough in the UK, which can be found here: http://webarchive.nationalarchives.gov.uk/20170726163612///visual.ons.gov.uk/wp-content/uploads/2017/10/map.csv

Area data

The list of areas and boroughs is scraped of the Wikipedia page, cleaned up (for more information on how this was done, check the notebook), and filtered. For this analysis, only areas within the "Post town" of London itself were used, in order to reduce the area of analysis a bit. Python's geopy library was then used in order to find the coordinates of the corresponding areas. This leaves us with the following dataframe.

	Area	Borough	Latitude	Longitude
0	Abbey Wood	Bexley	51.491060	0.121216
1	Acton	Ealing	51.508140	-0.273261
2	Aldgate	City	51.514248	-0.075719
3	Aldwych	Westminster	51.513103	-0.114920
4	Anerley	Bromley	51.407599	-0.061939
5	Angel	Islington	51.531946	-0.106106
6	Archway	Islington	51.565437	-0.134998
7	Arnos Grove	Enfield	51.616402	-0.133287
8	Balham	Wandsworth	51.445645	-0.150364
9	Bankside	Southwark	51.508302	-0.096297

Figure 1: Snippet of used dataframe with areas, boroughs and geographical coordinates of London

Python's folium library was then used to create a map of London, with the locations of all the areas that will be analysed on top. See

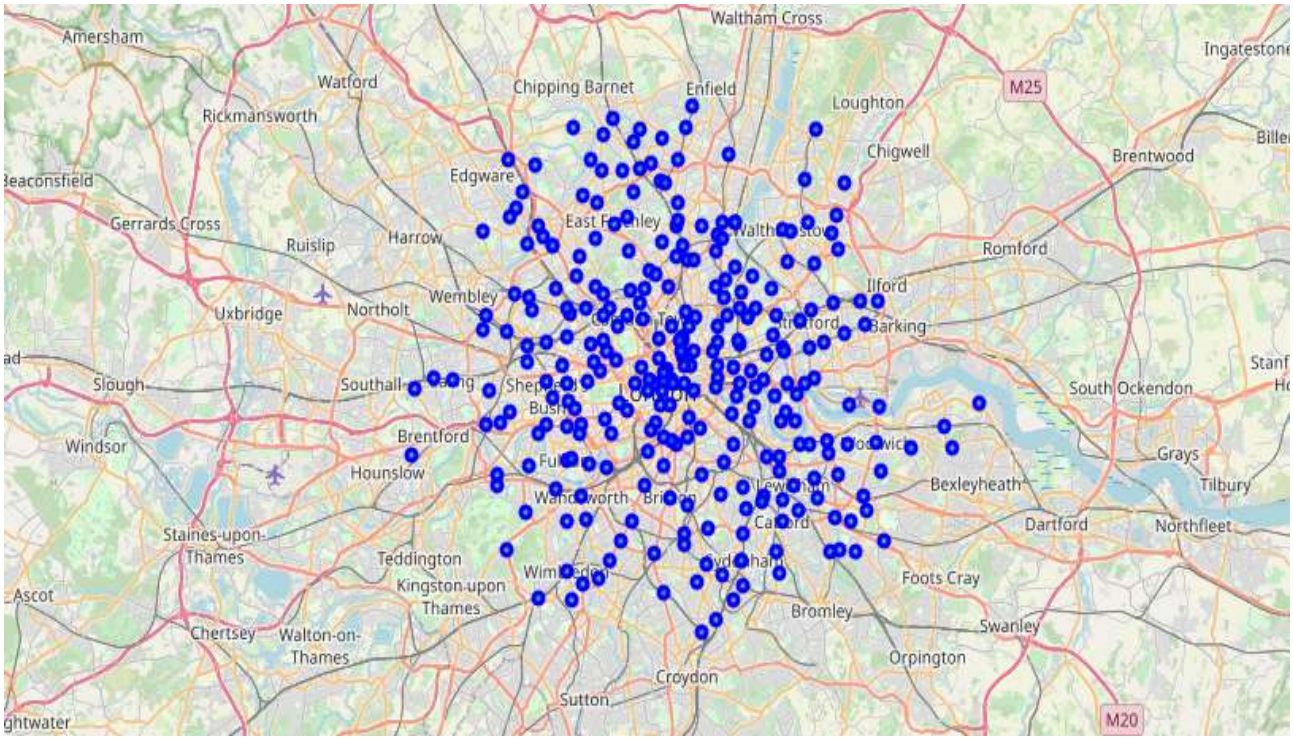

Figure 2: Map of London with geographical locations of areas that will be analyzed

Most gentrified boroughs

For this project we will analyse the most gentrified boroughs according to londonist.com. The list of gentrified boroughs is quite large, but to be safe we will use the five most gentrified boroughs in the London post town. These boroughs are: Kensington and Chelsea, Richmond upon Thames, Westminster, Hammersmith and Fulham and Merton.

In order to illustrate where these boroughs are located, the GEOJSON file of London boroughs was split up in a json file with the five “gentrified” boroughs and a json file with the rest of the boroughs. A map of London with the five most gentrified boroughs according to londonist.com in red is given in Figure 3.

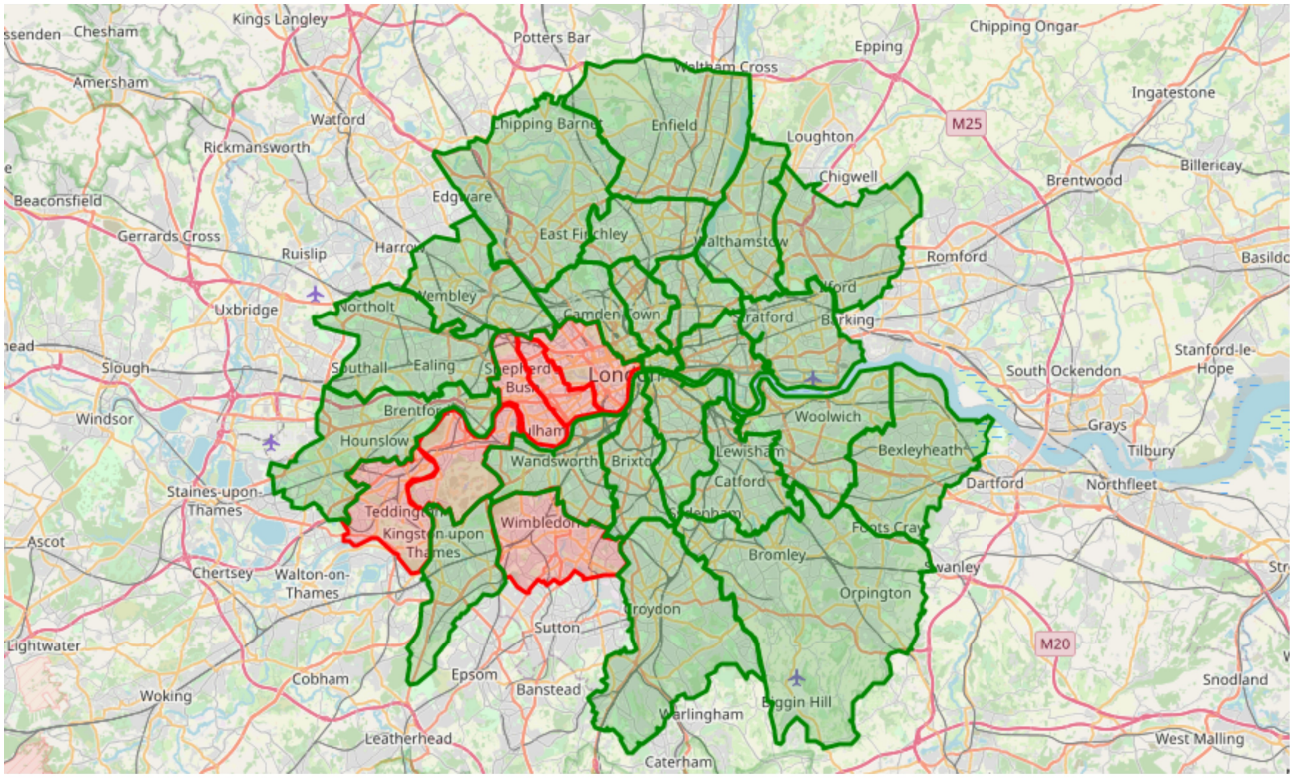

Figure 3: London boroughs with the most gentrified boroughs over the past years given in red

Venue data using Foursquare API

The Foursquare API was used in order to retrieve venue data for venues located in the areas given in Figure 2. For every area a call was made exploring the first 100 venues in a radius of 250 m (however, no areas returned more than 68 venues). The category and locations of all venues are stored in a dataframe.

	Area	Borough	Latitude	Longitude	Area Latitude	Area Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Abbey Wood	Bexley	51.49106	0.121216	51.49106	0.121216	Sainsbury's	51.492824	0.120724	Supermarket
1	Abbey Wood	Bexley	51.49106	0.121216	51.49106	0.121216	Abbey Wood Railway Station (ABW)	51.490825	0.123432	Train Station
2	Abbey Wood	Bexley	51.49106	0.121216	51.49106	0.121216	The Abbey Arms	51.490693	0.121182	Pub
3	Abbey Wood	Bexley	51.49106	0.121216	51.49106	0.121216	Platform 1	51.491023	0.119491	Platform
4	Abbey Wood	Bexley	51.49106	0.121216	51.49106	0.121216	Costcutter	51.491287	0.120938	Convenience Store
5	Abbey Wood	Bexley	51.49106	0.121216	51.49106	0.121216	Bean @ Work	51.491172	0.120649	Coffee Shop
6	Acton	Ealing	51.50814	-0.273261	51.50814	-0.273261	London Star Hotel	51.509624	-0.272456	Hotel
7	Acton	Ealing	51.50814	-0.273261	51.50814	-0.273261	Dragonfly Brewery at George & Dragon	51.507378	-0.271702	Brewery
8	Acton	Ealing	51.50814	-0.273261	51.50814	-0.273261	The Aeronaut	51.508376	-0.275216	Pub
9	Acton	Ealing	51.50814	-0.273261	51.50814	-0.273261	Amigo's Peri Peri	51.508396	-0.274561	Fast Food Restaurant

Figure 4: Snippet of dataframe with venues in areas

Real estate prices of London

The final piece of data that was used are the London real estate prices. A list of 2016 UK real estate prices in pounds per square meter was taken from the UK National Archives. This data set includes all boroughs in the UK.

	Borough	Price per m2
0	Hartlepool	987
1	Middlesbrough	1120
2	Redcar and Cleveland	1182
3	Stockton-on-Tees	1254
4	Darlington	1260
5	Halton	1339
6	Warrington	1750
7	Blackburn with Darwen	1053
8	Blackpool	1107
9	Kingston upon Hull, City of	1127

Figure 5: Real estate prices in pounds per square meter in the UK (2016)