Coursera Capstone Project
Battle of the neighborhoods – London

April 2020 | Martin Manjolo

Introduction

Background

- London is one of the world's largest tech centres and a sought-after destination for millennials
- London offers several opportunities for growth and is renowned for its high standard of life
- Challenging for prospective immigrants to find a and buy a home

Problem

- Goal is to find the right neighbourhood to buy a home within a given budget
- Focus on potential homeowners looking to buy a home near Central London

Audience

- Young millennials moving to the UK with a moderate-income base
- Looking to buy a 2 bedroomed flat in or around the Greater London Area
- Easy access to essential amenities and venues

Data sources and criteria

Sources

- 2020 data on London properties and relative price data is extracted from the HM Land registry available here: https://www.gov.uk/government/statistical-data-sets/price-paid-data-downloads

- Geographical coordinate data provided by FreeMapTools. UK postal code and geographical data is available here: https://www.freemaptools.com/download-uk-postcode-lat-lng.htm

- Foursquare data to discover recommended amenities and venues in the neighbourhoods that will influence our decision are:

Criteria

- A 2 bedroomed apartment or flat as close as possible to Central London

- Price range between 500,000 GBP and 510,000 GBP

- Venues and essential amenities should be located within 500 metres of the property

Data Cleaning

- 70,376 samples and 19 features emerged from data.

- Only used street names, prices, post code and geographical coordinates

- Filtering down on price range left use with 123 samples and 4 features

The resulting map of selected neighbourhoods after data cleaning
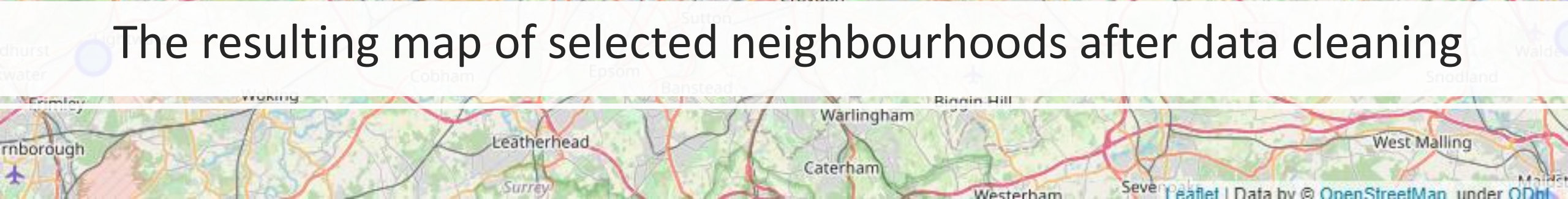
# Modelling and visualization using K-Means Clustering and Foursquare API

Foursquare API
- Employed the foursquare API to get information on venues around our target neighbourhoods based on amenities and essential facilities such as schools, hospitals and grocery stores.
- Limited our range to only include the top 100 venues within 500 metres of our neighbourhoods
- Analysed data and grouped the venues by neighbourhood and by taking the mean of the frequency of occurrence of each category
- We then selected the top 10 most common venues in each neighbourhood.

Clustering
- Clustered the neighbourhoods into 5 clusters
- Created a dataframe that included the cluster as well as the top 10 venues for each neighbourhood.

Selected neighbourhood clusters
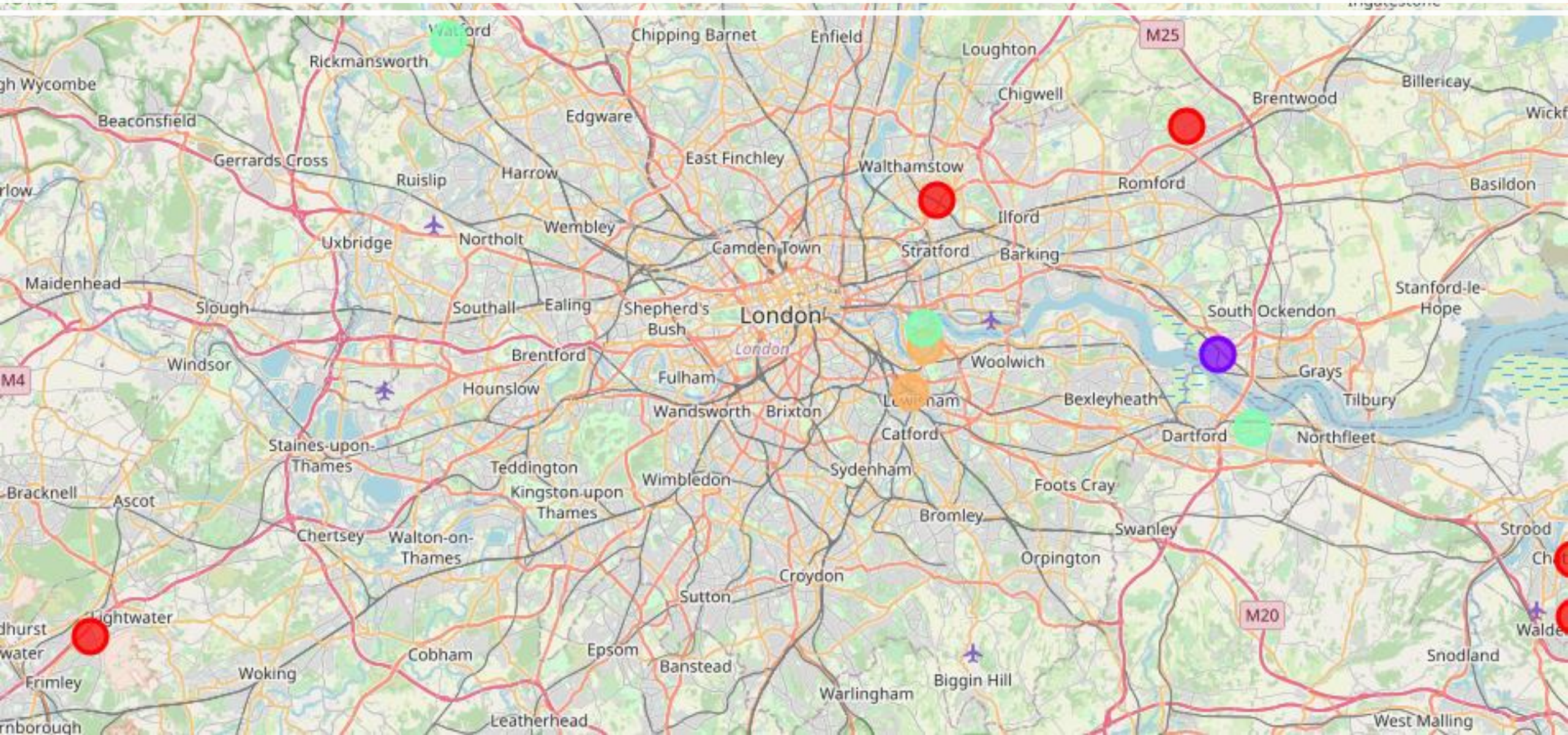
Results and Discussion

- Analysis shows that at our price point, there are more options for homes further outside of Central London.
- Fewer clusters closer to Central London, indicative that this market segment market is highly competitive and that properties rarely become available for sale
- The following streets and neighbourhoods are our top picks, offering the best combination of price, location and amenities:
  - London
  - Manchester
  - Wallington
  - South Croydon

Conclusion

- Project identified neighbourhoods near Central London that properties within our set price budget and had the right blend of amenities and venues

- Identified 4 neighbourhoods that meet the basic criteria.

- Final decision on the right neighbourhood will be made by prospective homeowners, considering the characteristics of the neighbourhoods, proximity to favourite venues and amenities, and ease of travel to and from Central London