



# **London Neighbourhoods by Most Common Businesses and Average Income**

A Clustering Analysis



# Business Problem

- London is a very touristic city with a vast offer of businesses and venues
- New Bakery to be inaugurated in one of London Zone 1 neighbourhoods in an above average income area and near to Hotels and touristic spots
- Needs to assess the profile of the neighbourhood and give some insights about the potential businesses and customers



# Data Acquisition and Cleaning

- Online database combines data from the Royal Mail, OS and National Statistics
- Foursquare developer API used to explore venues around each neighbourhood
- Postcodes that are not active are removed and the data is restricted to London Zone 1
- Check for missing data
- Select data and group by Postcode District and Ward around the mean coordinates of the locations

# Exploring Location Data

Location data from Foursquare API was used to find most common venues in London Zone 1 neighbourhoods.



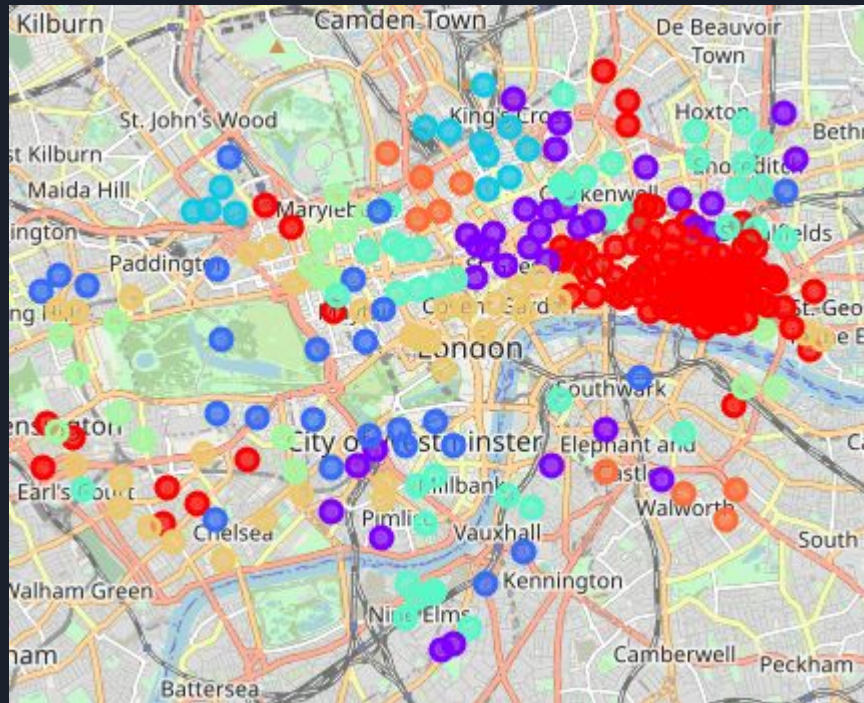


# Using k-Means Clustering

- Originally, Inner London is divided into 8 Postcode Districts
- Classify each neighbourhood into one of 8 clusters
- Average income and the 10 most common venues are used as input for the sklearn package implementation of k-means clustering

# Results

Clusters are more consistent in the City of London area and become more sparse in other regions of London Zone 1





# Conclusions

- Clustering analysis found the group of neighbourhoods with the characteristics the customer wanted, most around Hyde Park and Kensington regions
- Clusters give an insight of the most common businesses found in different levels of average income in London Zone 1
- Although common sense would probably indicate the region of the City of London as the most appropriate one, the analysis found a different solution