Birmingham Neighbourhoods

Applied Data Science Capstone Presentation

Prepared By,
SHOAIB UR REHMAN KHAN

Introduction

Business Problem

- Location to opening a new restaurant in Birmingham is the decision we have to make here.
- The objective of this capstone project is to analyze and select the best locations in the city of Birmingham, England to open new restaurant.
- The business problem is: In the city of Birmingham, England, if someone is looking to open a restaurant, where would you recommend that they open it?
- Using data science methodology and machine learning clustering this project aims to provide solut ions to answer the business problem

Data

Data Required

- List of all neighbourhod of Birmingham, England.
- Latitude and Longitude coordinates of the neighbourhoods.
- Venue Data, which is related to restaurants.

Data

Data Sources

- Wikipedia page which contains list of neighborhoods of Birmingham (https://en.wikipedia.org/wiki/B_postcode_area)
- Geocoder package for latitude and longitude coordinates of the neighbourhood.
- Foursquare API for Venue Data to explore the neighbourhoods.

Methodology

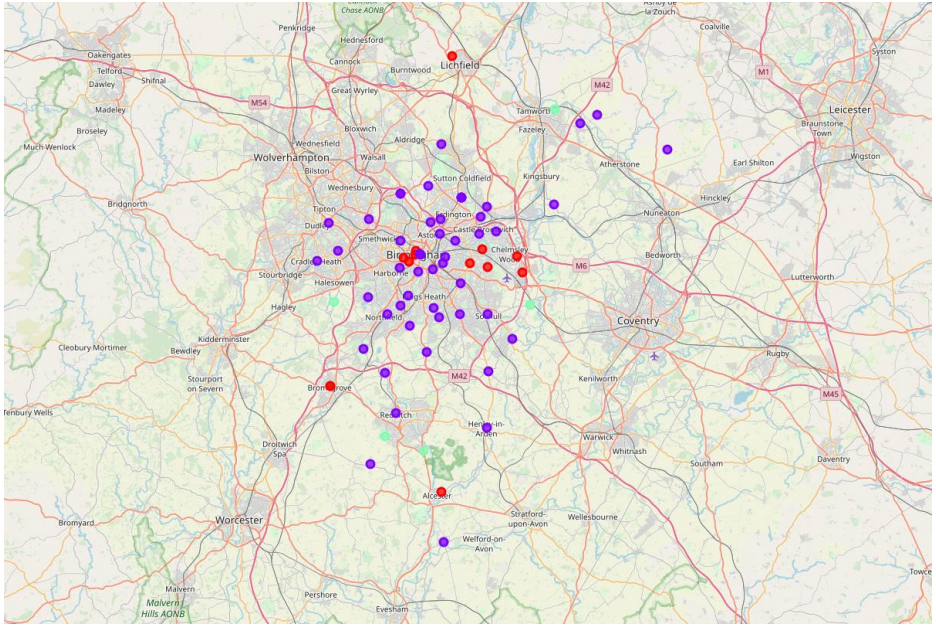
Exploratory Data Analysis & k-means Clustering

- Web scraping Wikipedia Page of neighbourhoods.
- Obtain latitude and longitude coordinates using Geocoder package.
- Use Foursquare API to get venue data.
- Check the returned venues of each neighbourhoods and examine the unique categories.
- Filter venue category by restaurants.
- Perform clustering on the data by using k-means clustering.
- Visualize the clusters in a map using folium.

Results

Categorized the neighbourhoods into three clusters

- Cluster 0: Neighbourhoods with moderate number of restaurants.
- Cluster 1: Neighbourhoods with low number to no existence of restaurants.
- Cluster 2: Neighbourhoods with high concentration of restaurants.



Discussion

- Most of the restaurant are found in the center of the Birmingham City.
- Highest number in cluster 2 and moderate number in cluster 0.
- Cluster 1 has very low number to totally no restaurants in the neighborhoods.
- Recommends on these findings to open new restaurant in neighborhoods in cluster 1.
- Little competition in cluster 0 and avoid cluster 2.

Recommendations

- We only considered one factor i.e. frequency of occurrence of restaurants.
- There is other factors like population and income of residents that could influence the location of restaurant.
- Can revise this methodology to determine the preferred locations to open a new restaurants.
- In addition, Use of paid Foursquare Developer account to bypass these limitations and obtain more results.

Conclusion

- Answer to business problem: The neighbourhoods in cluster 1 are the most preferred locations to open a restaurant.

- Findings of this project will help the relevant stakeholders to capitalize on the opportunities on high potential locations while avoiding overcrowded areas in their decisions to open a new restaurant

Thank You!