



Starting A New Restaurant In London By Analyzing Neighbourhoods

Bingqing He

IBM Applied Data Science Capstone



Introduction

- London: capital and largest city of England and the United Kingdom
- Considerable influence upon the arts, commerce, education, entertainment, fashion, finance, healthcare, media, professional services, research and development, tourism and transportation
- The third-most populous city in Europe



Problem

- Start a new restaurant
- Location and what kind of restaurant
- Client: investors and local residents



Data

- Neighbourhoods of London
 - Wikipedia Page
- Geographical Coordinates
 - GeoPy library in Python
- Venue
 - FourSquare API



Methodology

- Feature Extraction
 - Finding hot spot

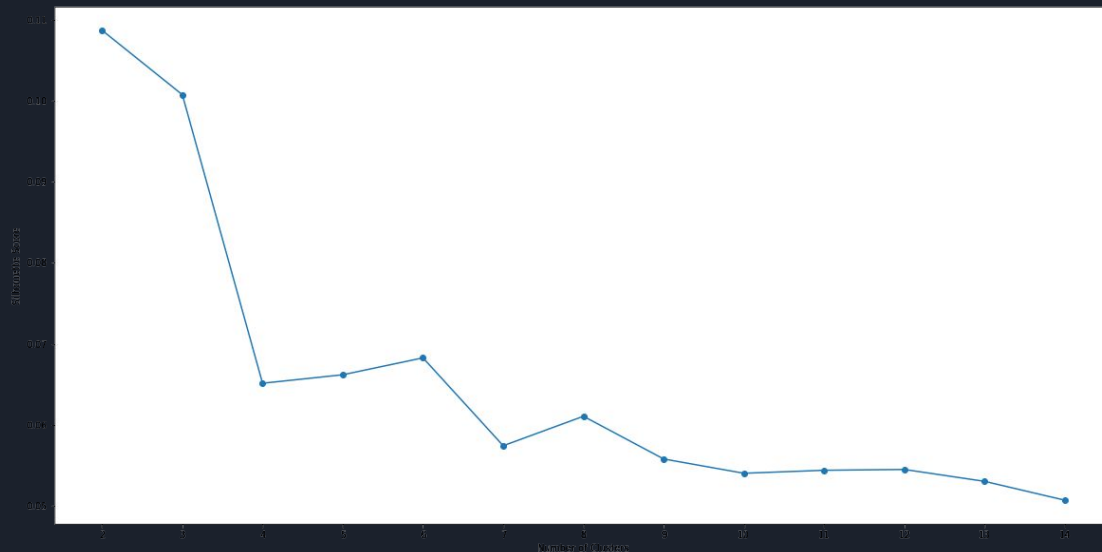
	Neighbourhood	Accessories Store	Afghan Restaurant	African Restaurant	Airport	Airport Lounge
0	Abbey Wood	0	0	0	0	0
1	Abbey Wood	0	0	0	0	0
2	Abbey Wood	0	0	0	0	0
3	Abbey Wood	0	0	0	0	0
4	Abbey Wood	1	0	0	0	0

5 rows x 398 columns



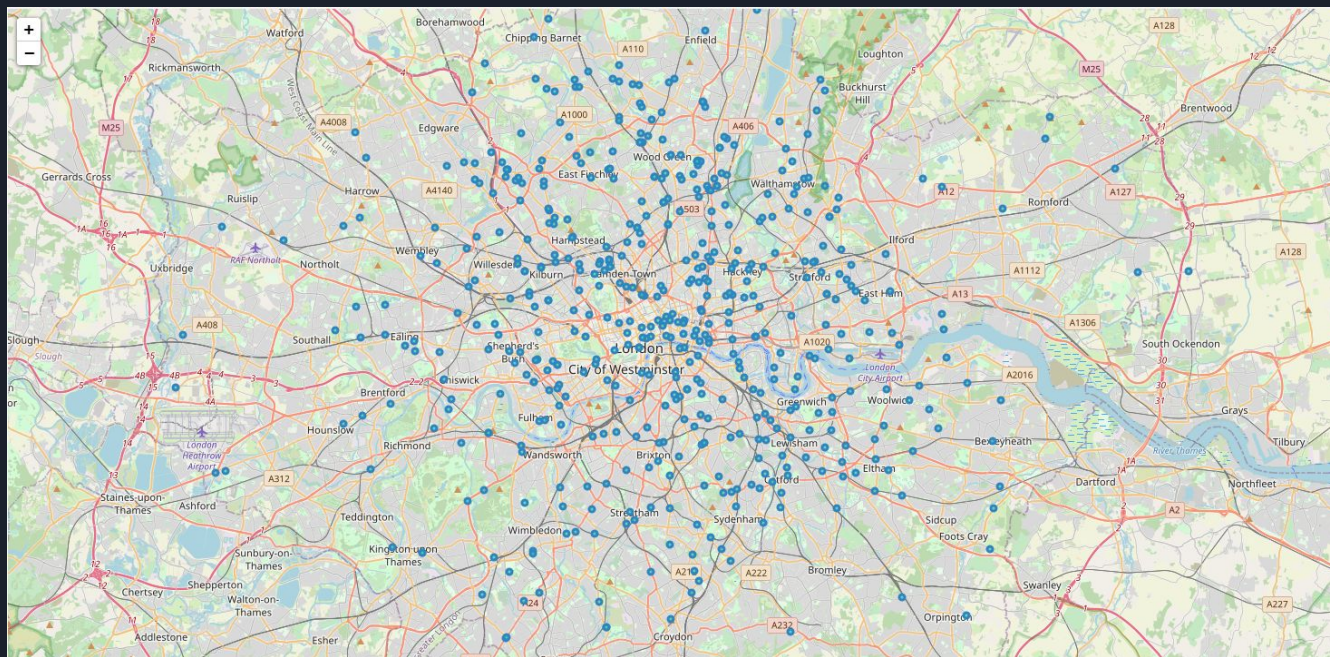
Methodology

- Machine Learning
 - K-Means Clustering



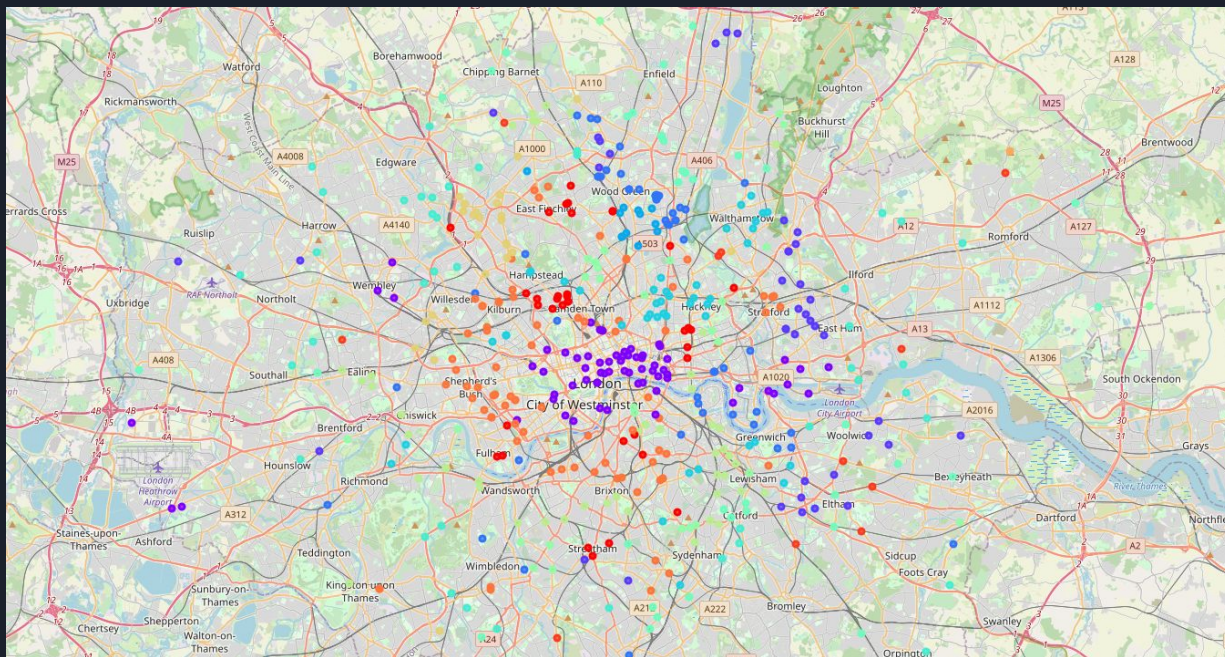
Plotting

- Folium in Python



Results

- Clusters visualization





Discussion

- The K-Means model works well and successfully
- Different clusters have different most common venues
- Consider more factors like parking and transportation



Conclusion

- Python's built-in libraries such as Folium and GeoPy, as well as BeautifulSoup API
- Data analysis and machine learning techniques
- Business data analysis includes the activities to make strategic decisions, achieve major goals and solve complex problems