ANALYSING CLUSTERS OF NEIGHBOURHOODS OF MOSCOW TO IDENTIFY POTENTIAL AREAS OF ESTABLISHING JAPANESE RESTAURANT BUSINESS

SAKET MISRA
SHAHHEED SUKHDEV COLLEGE OF BUSINESS STUDIES
UNIVERSITY OF DELHI

BUSINESS PROBLEM

Given a Business group possibly an experienced restaurant chain owner or a business group , venturing into the restaurant business with being new to having a few years of experience in the same , which areas (Neighborhoods) of Moscow could prove them to be good potential hotspots for establishing their restaurant setup , keeping in the mind the heavy competition of other similar or different businesses already established across the area , by recommending them the ideal cluster(s) in the city for them to explore and venture into for setting up their stand-alone restaurant and/or restaurant chain in the future.

DATA

Data Required :

1. List of Neighbourhoods of Moscow
2. Coordinates of the corresponding Neighbourhoods
3. Venue data corresponding to venues in the close proximity

Data Sources :

1. Neighbourhoods of Moscow : https://en.wikipedia.org/wiki/Category:Districts_of_Moscow
2. Geopy.geocoders library for Neighbourhood Coordinates
3. Foursquare API calls for venue data

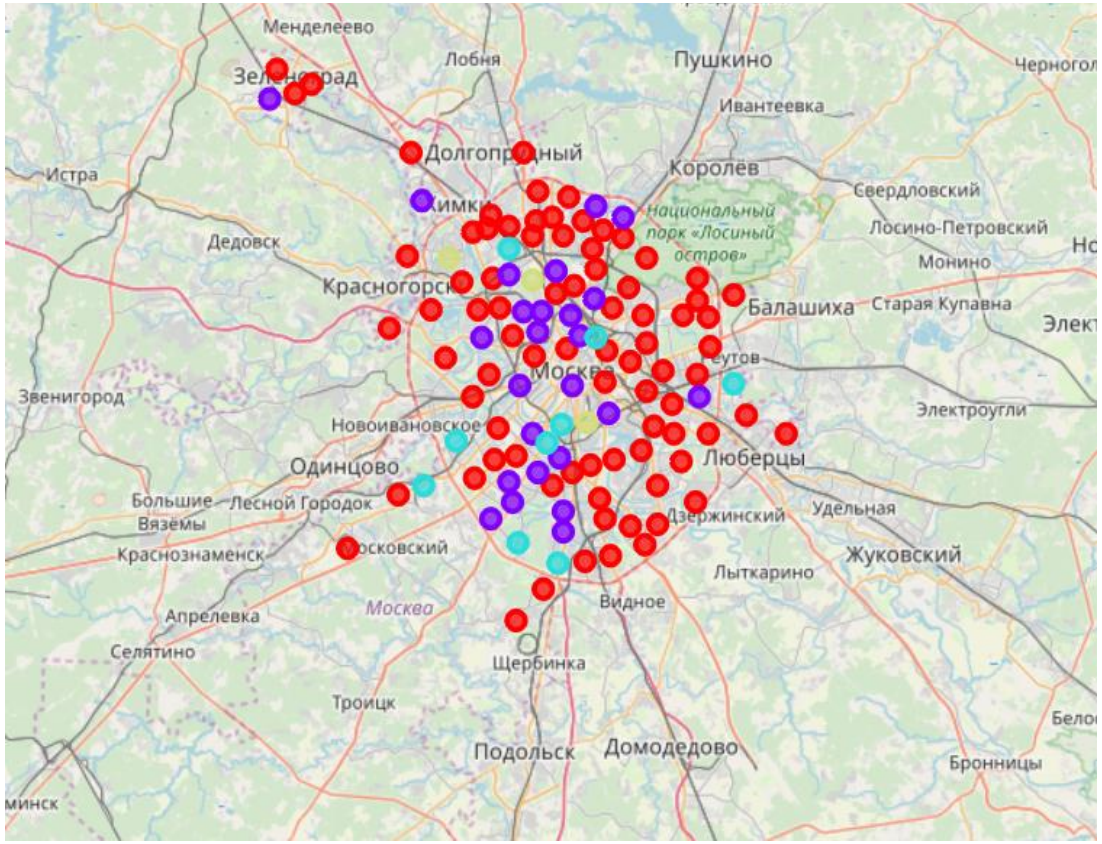
METHODOLOGY

1. Web scrapping list of neighbourhoods of Moscow from Wikipedia
2. Gathering co-ordinates (latitudes and longitudes) of neighbourhoods using geopy.geocoders library
3. Visualising the neighbourhoods on the map of Moscow using folium library
4. Gathering venue related data by making foursquare API calls for the respective neighbourhoods
5. Cleansing the data , by removing unnecessary features and shaping the data as per requirements (retaining only data pertaining to Japanese restaurant as venue category)
6. One hot encoding the data according to venue categories , and calculating aggregate means by grouping with neighbourhoods
7. Fitting the Kmeans clustering algorithm on our dataset by initialising K=4 and gathering labels pertaining to our dataset
8. Visualising the geographical span , and the magnitude of neighbourhoods covered under various clusters to draw any inferences on the obtained information.

RESULTS

1. Cluster 0 and Cluster 3 (yellow and red) cover least number of neighbourhoods
2. Cluster 1 and 2 predominate the geographical span of Moscow
3. Cluster 0 and 3 are geographically positioned on the outskirts of Moscow
4. Cluster 1 (Red) covers almost entire Moscow , while Cluster 2 (purple)

   covers most of the central moscow



DISCUSSION , CONCLUSIONS AND RECOMMENDATIONS

As it is clearly evident that the clusters 1 and 2 dominate in terms of coverage and geographical span of Moscow , and since they are situated mostly in the interiors of Moscow , indicate that , Central and near-Central Moscow has well established Japanese restaurant businesses. These are maybe the businesses that have been into place for years , and hence leaving little to no room and/or rising property prices pushing other newbies to the outskirts. The recommendation would be to setup a restaurant chain in cluster 0 as it only covers 7 neighbourhoods or cluster 3 which covers around 17 neighbourhoods. Since these clusters have less frequency of Japanese restaurants as compared to their counterparts , which have heavy concentration of Japanese restaurants , and rather suggest a mushrooming scenario in cluster 1 and 2 which would only lead to cut-throat competition , and pose difficulties in terms of profitability and drawing attention , on the other hand cluster 3 which contains a decent coverage of 17 neighbourhoods suggests that it balances the trade-off between less population coverage (assuming near equal population densities) and little to no competition and high population with cut-throat competition , making it an ideal cluster to establish the restaurant business.

THANK YOU !

This has been an absolutely enriching and a great learning experience , throughout the IBM data science certification course , It wouldn't have been possible without the indepth yet simplistic explanatory videos by the instructors of this course and my peers , who helped me get better with my assignments , and truly made me realise the importance of peer assessment and collaborative effort to gain this knowledge that I had via this course !