Opening a shopping Mall in Tehran

November 2020

Introduction

Malls can use big data, machine learning and AI to derive meaningful insights to minimize operational costs, build better customer engagement, explore new avenues for revenue, enable tenants to boost productivity, and more

Many companies are digging in and finding innovative ways to use predictive analytics to increase profitability at some of the biggest shopping centers in the world. Location analytics firms working in the retail space are providing mall owners with unprecedented information about consumer movements and behaviors to drive decisions and build marketing strategies.

Business Problem

There are many shopping malls in the city of Tehran, Iran and many more are being built. The location of shopping mall is one of the most significant factor that will determine whether the mall will be successful or not.

The purpose of this project is to analyze Tehran location-wise and make decision regarding the best place to open a new shopping mall. Using Clustering Machine learning method, we would like to answer a big question: Where would be the place to open a shopping mall in Tehran to increase the chance of business success?

Audience of this Project

Investors, retailers and property developers are the main audience of this project. They care about this project since they can increase their profit by finding the best place to open a mall

Data

1.List of neighborhoods in Tehran is the main data that we need. WikiPedia page has list of 71 neighborhoods in Tehran. We extract the data from the Wikipedia page by using Requests and BeautifulSoup Packages.

2. In order to get the venue data and plot the map, the Latitude and longitude coordinates are required. FourSquare API helps us to get access to this data. FourSquare API provides different categories of the venue data and we use shopping mall category to help us solve the business problem. Data cleaning, data wrangling, K-means clustering and map visualization with the help of Folium are other steps that required for this project

Method

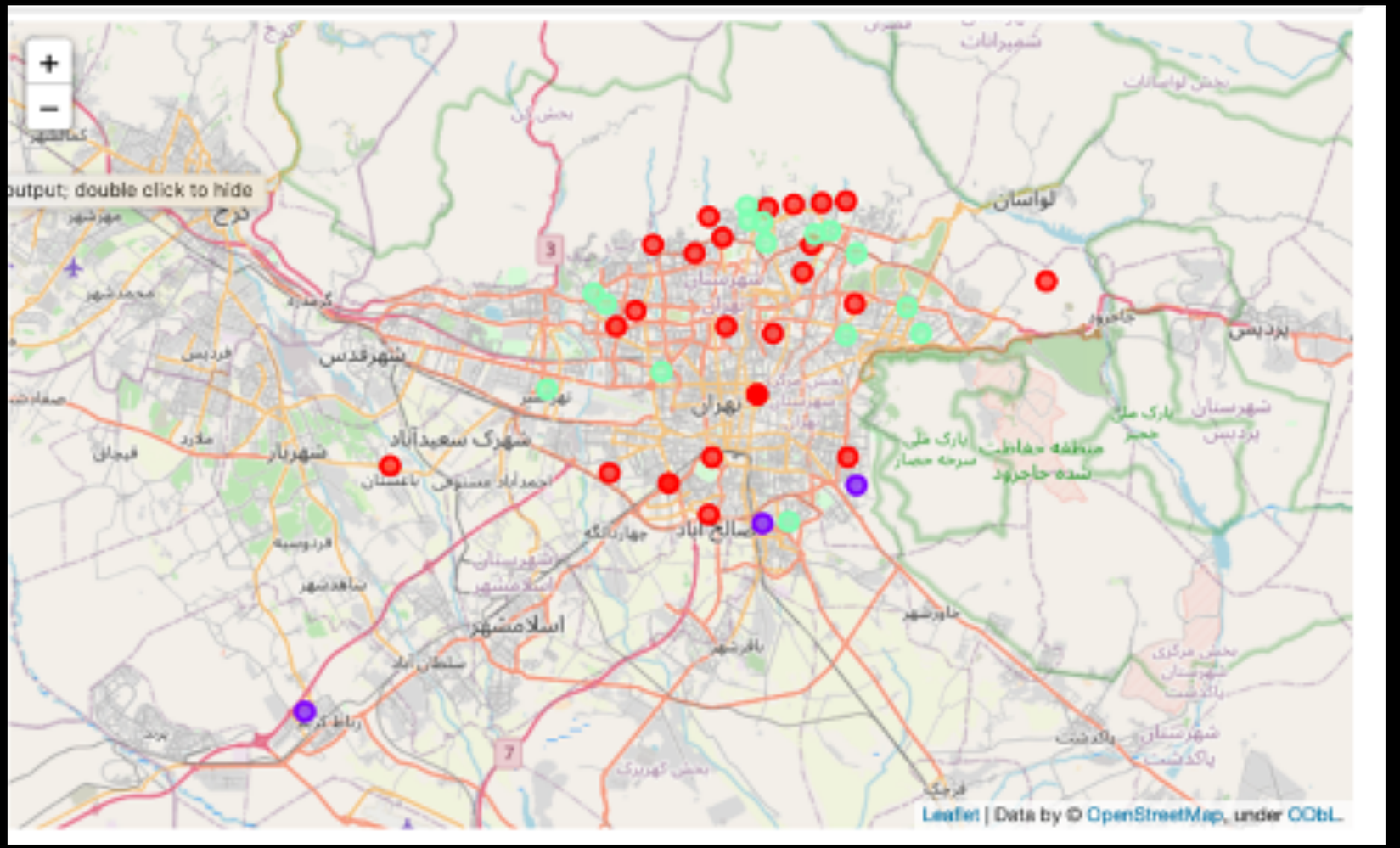
1. Wikipedia page https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Tehran
   Was used to get the list of neighborhoods in Tehran. In order to extract the list of locations, I used REQUEST and BEAUTIFULSOUP packages. For using FourSquare API , geographical coordinates (Latitude/Longtitude) are required. Therefore, GEOCORDER package was used to convert addresses into geographical coordinates.
1. By using the Pandas libraries, data was populated as a form of dataframe and then FOLIUM package was used to visualize the map of the neighbourhood.
2. To get the top 100 venues in Tehran, FourSquare API was used and geographical coordinates of the Tehran Neighborhoods was passed in.
3. We analyzed each neighborhood by grouping them to be able to filter the available venues.
4. By using the K-means Clustering, we clustered the data. By using 3 Clusters (K=3), We identified what neighborhoods have higher shopping malls concentrations. Based on the occurrence number, we can answer the main question: what neighborhood is the best for opening a new shopping mall

Result

Based on the clustering analysis, we defined 3 categories:

Category 1) Neighborhood with low number of shopping malls
Category 2) Neighborhood with average number of shopping malls
Category 3) Neighborhood with high number of shopping malls

The map below shows all 3 categories (clusters) in Tehran:


+
−
output; double click to hide
لواسان
شهرقدس
تهران
شهرک سعیدآباد
شهریار
صالح آباد
اسلامشهر
Leaflet | Data by © OpenStreetMap, under ODbL.

Discussion

Based on the 3 defined clusters and as visualized in the above map, It is clear that the south of Tehran has very low number of shopping malls (Purple dots). This shows a great opportunity to pen mail in this neighborhood since there is not much competition here. Therefore, based on our project finding, we recommend property developers and retailors to consider south of Tehran for opening a new shopping mall.

Conclusion

Based on occurrence of shopping malls we could answer the main question of property developers. However, other factors such as crime rate, local population, local resident income, … are other important factors that have potential effect on the result and the final decision. For future project, a data with all relevant variables can be used to increase the validity of the analysis.

•