



Venues data analysis in Tehran

Finding the best location to open a café or restaurant in Tehran, Iran

IBM Coursera Data Science Capstone

Authored by: Sahand Niasti

coursera

IBM

Tehran is one of the largest metropolises in the world where over **8.7 million** people live in city and around 15 million people in the larger metropolitan area of Greater Tehran. This city is the most populous city in Iran and Western Asia [1], and has the second-largest metropolitan area in the Middle East. So, it would not be shocking if you know that this city has a population density of **16,279** people per square kilometer [2]. As a resident of this city, I decided to use Tehran in my project. The city is divided into 22 districts in total. All these information tells us that this city has quite an intertwined and mixed structure.

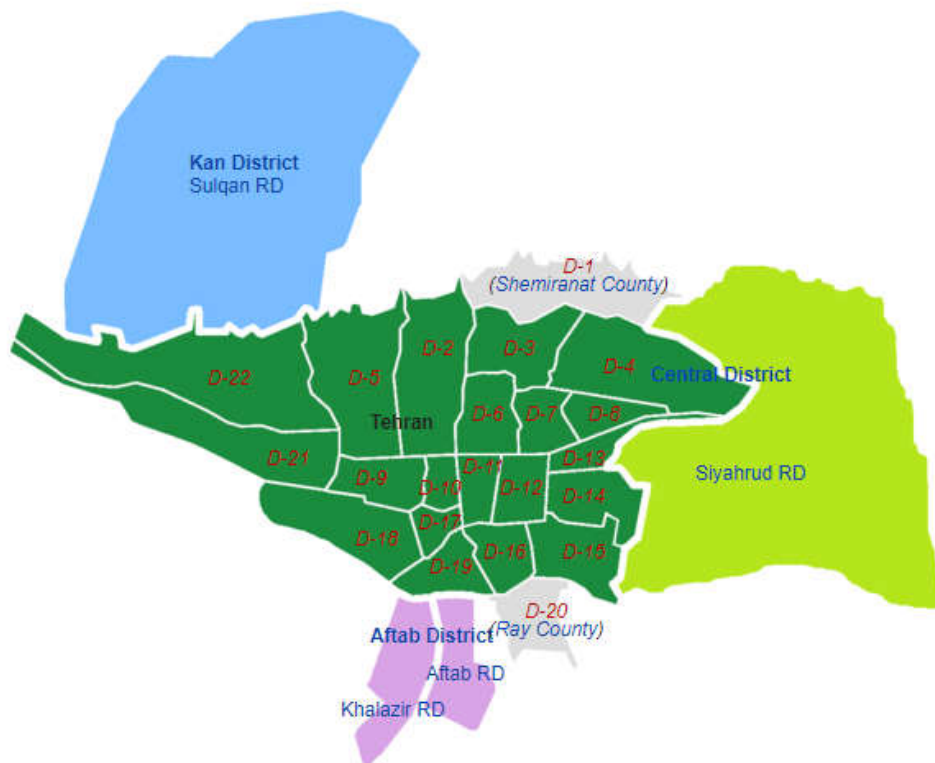


Figure 1. Tehran boroughs distribution in a map

Let say if you are a business manager who have never lived in Tehran and sees this populated city as a great opportunity to earn money. That person tends to invest in a form of café or restaurant, and our aim is to decide where or which neighborhoods should be selected to open such facility. In order to answer this question, you have to build a model get some recommendations where to start your business.

Business problem

The objective of this capstone project is to analyze and select the best locations in the city of Tehran, Iran to open a café or restaurant. Using data science methodology and machine learning techniques like clustering, and taking into account the fact that our business owner prefer to choose a district according to the social places density, leads us to the aim to provide solutions to this business question: In the city of Tehran , if a property developer is looking to open a new café or restaurant, where would you recommend that they open it?

Data Description

To solve the problem, we need the List of Boroughs and neighborhoods in Tehran with their Latitude and longitude coordinates. This defines the scope of the project which is confined to the city of Tehran, help to plot the map and eventually aid to get the venue data of the capital city of Iran. Sources of data and methods to extract them would be:

- Foursquare: It is a local search-and-discovery service which provides information on different types of entertainment, drinking and dining venues. Foursquare has an API that can be used to query their database and find information related to the venues, such as location, overall category, reviews and tips. [3]
- Wikipedia: There are not too many public data related to demographic and social parameters for the city of Tehran. Therefore, I decided to trust on this page and extract required information, with the help of Python requests and beautifulsoup packages. [4]
- Google Map: I also used ‘Search Nearby’ option to get the center coordinates of each Borough. [5]

[1] https://en.wikipedia.org/wiki/List_of_metropolitan_areas_in_Asia

[2] <https://en.wikipedia.org/wiki/Tehran>

[3] <https://developer.foursquare.com/>

[4] https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Tehran

[5] <https://www.google.com/maps>