

COURSERA

IBM Applied Data Science Capstone Project

Finding the best Location for Gas-Station in Dubai, UAE.

By

Nauman Mansuri

December, 2019



Introduction:-

With the booming automobile industry fuel consumption have tremendously increased to fulfil the need of commute. As a result provide great opportunity for oil company so build there gas station in Dubai city of UAE. In order to satisfy Endless fuel need. The company can earn a huge profit but it is not as easy at it seems location of Gas-station is to important before building. Not only location but there are several factors that affect the sales growth of a gas station Which needs to be taken to consideration.

Business Problem:-

The objective of this project is to select best location in Dubai, UAE to open a new Gas station using Data science technique and Machine learning algorithms like clustering which will be useful to provide solution for our Business problem.so if company wants to open their gas station in Dubai, UAE where you would advise to open?

Target Audience:-

This project is useful for the Oil companies and investors which are ready to invest in opening up a Gas-Station in the city of UAE, Dubai. This project is all time important for selecting the best geographical location according to Business need. As dubai has highly Concentrated Gas-Station in various areas and imbalance caused due to it will increase A chance to open up at a target location in order to hold strong advantage of it. According to report of gulf news Musaffah a area in UAE hits shortage of gas-station. Because of highly concentration to specific areas.

Data:-

Solving the problem will need following data

1. List of neighbourhoods in Dubai.
2. Latitude and longitude coordinates of those neighbourhoods. This is required in order to plot the map and also to get the venue data.
3. Venue data, particularly data related to Gas-Station. We will use this data to perform Kmeans clustering on the neighbourhoods.

Source: (https://en.wikipedia.org/wiki/List_of_communities_in_Dubai)

contains a list of neighbourhoods in Dubai with a total of 116 neighbourhoods. We will use web scraping techniques to extract the data from the Wikipedia page, with the help of Python requests and BeautifulSoup packages. Then we will get the geographical coordinates of the neighbourhoods using python Geocoder package which will give us the latitude and longitude coordinates of the neighbourhoods.

we will use Foursquare API to get the venue data for those neighbourhoods.. Foursquare API will provide many categories of the venue data, we are particularly interested in the Gas-station category in order to help us to solve the business problem put forward.

This is a project that will make use of many data science skills ,Machine learning , from web scraping (Wikipedia), working with API (Foursquare), data cleaning, data wrangling, to machine learning (K-means clustering) and map visualization (Folium), In the next section, we will present the Methodology section where we will discuss the steps taken in this project, the data analysis that we did and the machine learning technique that was used. we would analyse it and then select nest cluster and location to achieve our target.