Introduction / Business Problem

New York- A city Never Sleeps.

Visiting New York City on a budget might sound a bit far-fetched for some travelers, but with a little careful planning, you can have a successful trip to NYC on any budget.

While New York City is known to be one of the world's most expensive cities, there are plenty of ways to have an incredible trip within your budget. One of my favorite ways to explore city, venues, budget hotels is always through checking with Foursquare.

This project is aimed to make your travel experience to New York Simple by Providing all nearby locations, local areas, hotels, best places to visit in New York with help of Foursquare datasets

In this project, we will try to build a feature for recommendation engine which might be used by realtor for their web portal and also mobile platform to send personalized recommendations to its subscribers. It could be self-trained bot / personal assist with realtor features which specializes in showing personalized recommendations based on individuals preferences as everyone's requirements are not the same.

Data

Now that we have understood the business requirements. It is time to analyze and gather data for it. We will be using New York Municipalities data. We have collected data from valid sources, for our analysis and performed preliminary analysis for better understanding the data and preparing it for modelling. After the data is standardized, the qualified data is ready to be processed. In order to inspect the data, we will use the folium library to extract the map of NewYork and visualize the locations of the venues on the map.
We will query for each of their geo locations using geocoder library and venues from FourSquare API. For few of the locations, foursquare doesn't have data, we shall drop such rows as they are not useful for our analysis.
Based on the amenities provided by foursquare in the neighbourhood, cluster the nearby venues using k-means algorithm to find out the top ten amenities available for each location and extract those features. Based on the amenities cluster the neighborhoods which have similar characteristics.
Using this data will allow tourists to easily decide where to go when they are in a specific city.

Methodology:-

I used K-Means clustering algorithm to cluster based on the amenities, cluster the neighborhoods which have similar characteristics.
As we know user preferences are taken into consideration such as Banks, Shopping Malls, Bus Station, Grocery, Resturants, Coffee Shops, Deli Foods/ Bakeries etc ... Find out which ones are best for travel

Results:-

Based on the amenities available in each area, the city-county combinations are divided into 4 clusters. Each cluster has a unique amenity.

Discussion:-

Without a lot more work in the initial data exploration and methodology phase, it is not easy to figure out the best spot. However, we can see that Manhattan and Brooklyn are the famous travel spots

Conclusion:-

After examination of the clusters we discovered that Manhattan and Brooklyn are the famous travel spots