**CAPSTONE PROJECT REPORT**

1. **Introduction**
   1. Background

New York is a city which attracts tourists from all over the world. Its very famous for its variety of museums. It also has many restaurants serving cuisines from all over the world. Public transport is also quite good in New York. It includes buses, metro trains, taxis and of course a lot of people walk through the city of New York

* 1. Problem

Indians love to eat their native food wherever and whenever they travel abroad. An Indian Tourist visiting New York always has a problem looking for places that serve Indian Food.

* 1. Interest

In this project, the concerned stake holder is an Indian tourist who is staying at the Marrakech Hotel at New York. He/she is looking for Indian eateries nearby. He is also interested in knowing about the nearby Art/History Museums found close to the Hotel. So a map which shows him/her Indian restaurants and Art/History Museums close to his Hotel along with their rating and radial distance from the Hotel would help him plan his stay quite well

1. **Data Sourcing and Cleaning**

For this project, the data needed would be sourced using FourSquare Application.

* 1. Part A: A search query for “Indian restaurant” close to Marrakech Hotel would be punched in. The result would be a json file. The json file is then cleaned and structured into a pandas dataframe. Columns like restaurant name, id, latitude, longitude, address, radial distance are retained. The rest of information are filtered out. Then using each id of the different restaurants, their ratings are retrieved and appended into the original pandas dataframe (ny\_dataframe\_filtered).
  2. Part B: A search query for “Art/History Museums” close to Marrakech hotel would be punched in. The result would be a json file. The json file is then cleaned and structured into a pandas dataframe. Columns like restaurant name, id, latitude, longitude, address, radial distance are retained. The rest of information are filtered out. Then using each id of the different restaurants, their ratings are retrieved and appended into the original pandas dataframe ( nearby\_venues\_filtered).

1. **Methodology**

Once the two dataframes are ready, then using folium library, 3 maps are generated:

1. with Marrakech Hotel as a green circle surrounded by smaller blue circles denoting Indian restaurants
2. with Marrakech Hotel as a green circle surrounded by bigger red circles denoting Art/History Museums.
3. with Marrakech Hotel as a green circle surrounded by smaller blue circles denoting Indian restaurants and bigger red circles denoting Art/History Museums.

Now we use **KMeans Clustering method** to cluster the restaurants and museums based on radial distances from the Marrakech hotel where the concerned tourist is residing.

1. **Results**

A Final map of clustered group (*based on radial distance from Marrakech Hotel* )of Indian restaurants and Art/History Museums is displayed. The Indian restaurants as well as the Museums are color coded as per clusters which makes it easy for tourist to identify them. The pop label for each marker will also give the ratings and radial distance of the particular Indian restaurant and Museum.

1. **Discussion**

With the given map just displayed, the tourist will have a fair idea which are the places of his interest closer to his stay and which, further. Depending on what he sees on the map, he can plan his trip for the day.

1. **Conclusion**

In this way, Indian tourist staying at the Marrakech hotel would be in a position to plan his trip based on how far the Indian restaurants and museums are to his place of stay. Another feature that can be added on this code would be to generate a map with clustered group of Indian restaurants and Museums based on “ratings” rather than “distance”. This would give a color coded feature to the map and help the tourist in choosing only those places above a certain rating (or a list of color codes as would be in the map). That would narrow down the tourist to see only those places that other people have given better rating.