PROJECT ON BUSINESS SETUP FOR AN INDIAN RESTAURANT IN DUBAI

Annette Nayana Palanna

1. Introduction

1.1 Background

Dubai is an emirate in the United Arab Emirates. It is a city of skyscrapers, ports, and beaches. Dubai is called one of the top tourist destinations in the world. Dubai Metro is automated and driverless and is recognized as the third largest in the world. Dubai Metro Line, which runs like an artery through the heart of Dubai, was the longest single driverless Metro line in the world with a route of 75Km as 2016 recognized by Guinness world records [5].

Dubai has fastest growing population in the world and is home to an exciting, diverse, multi-cultural blend of young, dynamic, professional people all enjoying the unrivalled quality of life in Dubai. Dubai is a crime-free place where administrative efficiency and openness to business have encouraged astounding growth. Favorable business conditions and shrewd investment decisions have led Dubai to become the best destination for investment [4].

1.2 Problem

This study is focused to detect a prime location in Dubai to set up an Indian cuisine restaurant, particularly walking distance from the Metro station. Our study is based on the following attributes in vicinity of each Metro station as mentioned below.

- Number of Indian restaurants, other restaurants and Hotels
- Population density
- House price index
- Tourist attractions

1.3 Interest

Dubai is a well-known tourist destination and hence hospitality industry has become the most prominent sector for investment. Therefore, this project is introduced to detect the top prime locations for our stakeholders to setup an Indian restaurant at prime location in Metro vicinity.

2. Data Acquisition and Cleaning

2.1 Data Sources

The data employed for analysis in this project is collected from Bayanat UAE website [1]. This data comprises of list of Metro stations, line they belong to and its responding coordinates in the Comas Separated Values (.csv) format.

2.2 Data Cleaning

Dubai Metro has 50 Metro stations in total that fall under 2 lines – Red line and Green line. The red line has 29 stations and 19 stations on green line having two intersections on Burjuman Metro station and Union Metro station [5].

The data collected has 2 intersections falling under both red line and green line category, which is identified as duplicate. The study is based on unique Metro stations irrespective of their lines, therefore duplicate is discarded. The updated data set contains 48 Metro stations and column names are made readable.

3. Methodology

In this project we will direct our efforts on detecting areas of Metro stations that have low restaurant density, particularly those with low number of Indian restaurants. We will limit our analysis within the radius of 500m from the Metro stations.

3.1 Explore stations in Dubai Metro

The figure 1 displays 48 Metro stations on a folium map. The coordinates of Dubai city are extracted using Nominatim geo location service from geopy library, where each Metro station is superimposed on top of folium map.

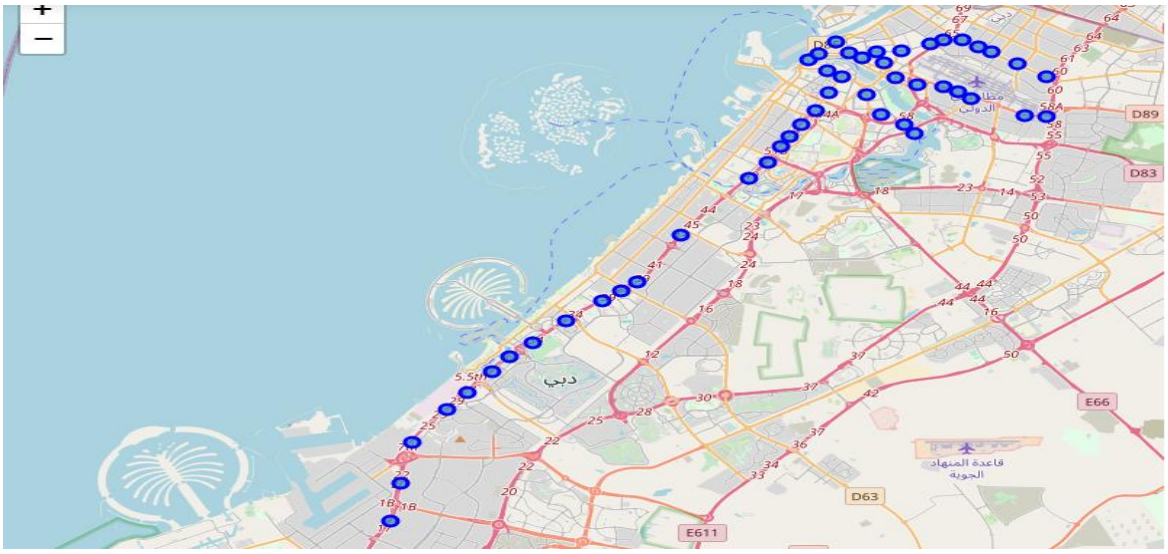

Figure 1: Dubai Metro stations

To extract relevant features from each station, we define Foursquare credentials and versions. From Foursquare API service name, location and category of venues associated with each Metro station within the range of 500meters from the stations are obtained. The resulting data frame has 1544 venues and 221 unique categories.

The figure 2 interprets the total number of venues corresponding to each Metro station.

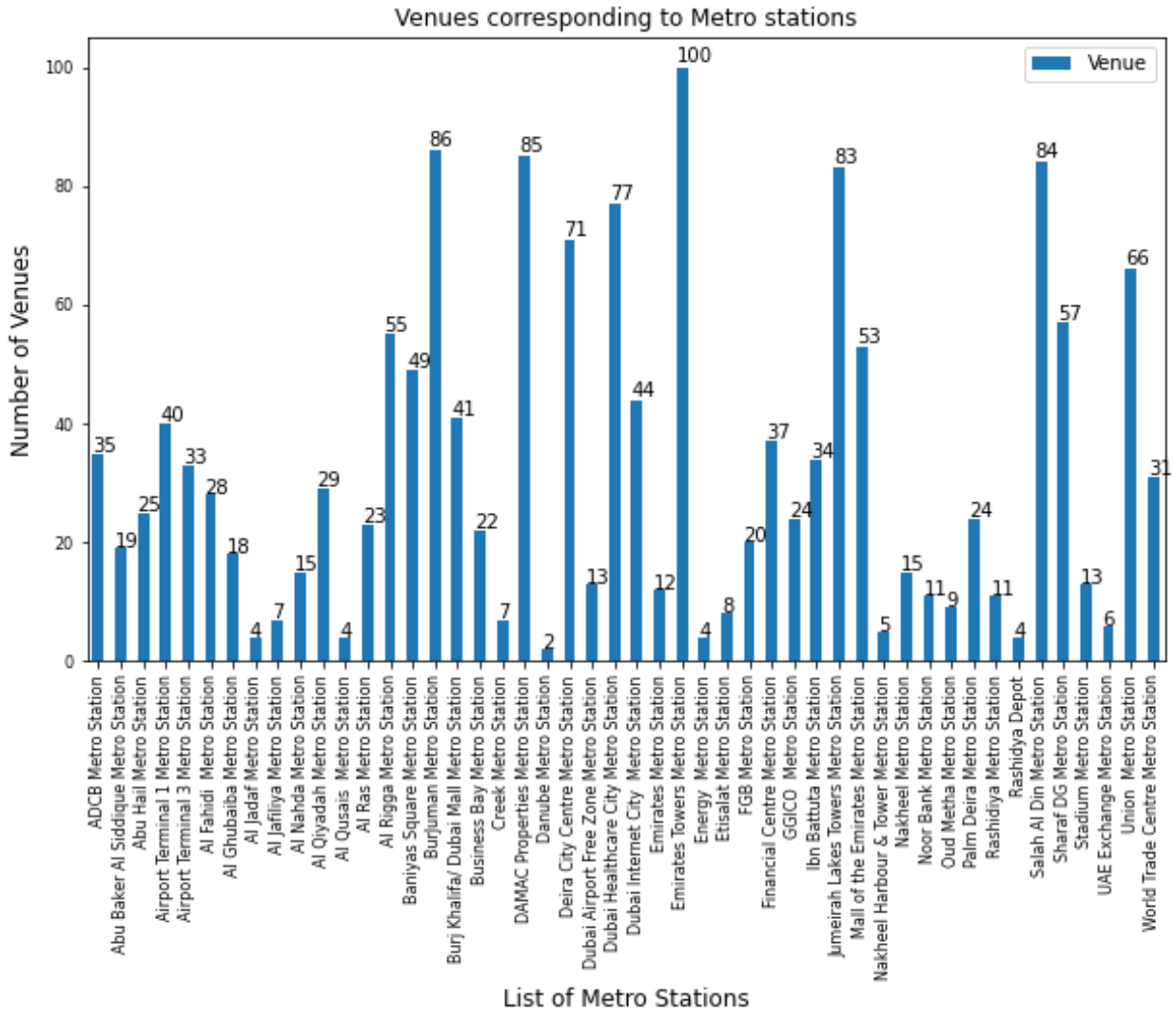

Figure 2

From the above figure we infer Emirates Tower Metro station has 100 venues, DAMAC Metro station has 87 venues and so on.

It is observed that several Metro stations have lesser venues, which makes the analysis difficult because these stations lack tourist spots, distant from city limit and are most likely unique when employing clustering algorithm. Therefore Metro stations with less than 25 venues are discarded. The resulting dataframe contains 23 popular Metro stations.

3.2 Analyze each Metro station

To employ machine learning algorithm for analysis categorical variables of the dataframe are to be converted to Machine Language. Therefore one-hot encoding is used to obtain numerical data. Further mean value is calculated on data and group based on station name. The resulting dataframe consist of 196 venues.

The figure 3 shows top ten venues of five Metro stations.

	Station Name	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	ADCB Metro Station	Indian Restaurant	Hotel	Asian Restaurant	Fast Food Restaurant	Filipino Restaurant	Pakistani Restaurant	Chinese Restaurant	Restaurant	Convenience Store	Department Store
1	Abu Hail Metro Station	Café	Indian Restaurant	Clothing Store	Hotel	Baby Store	Pizza Place	Convenience Store	Fast Food Restaurant	Bus Station	Burrito Place
2	Airport Terminal 1 Metro Station	Airport Terminal	Airport Service	Rental Car Location	Airport Lounge	Coffee Shop	Breakfast Spot	Sandwich Place	Smoke Shop	Food Court	Mobile Phone Shop
3	Airport Terminal 3 Metro Station	Airport Lounge	Coffee Shop	Duty-free Shop	Spa	Hotel	Juice Bar	Clothing Store	Smoke Shop	Burger Joint	Sandwich Place
4	Al Fahidi Metro Station	Indian Restaurant	Asian Restaurant	Pizza Place	North Indian Restaurant	Sporting Goods Shop	Camera Store	Chinese Restaurant	Coffee Shop	Pub	Persian Restaurant

Figure 3

3.3 K-Means Clustering

The clustering algorithm employed in this project is one of the most popular unsupervised machine learning algorithms. This algorithm is iterative and uses Euclidean-based similarity measures to cluster the dataset into 2 clusters. In this project value of K is obtained from the Elbow method which provides the optimal value of K for the data of our interest. [Figure 4]

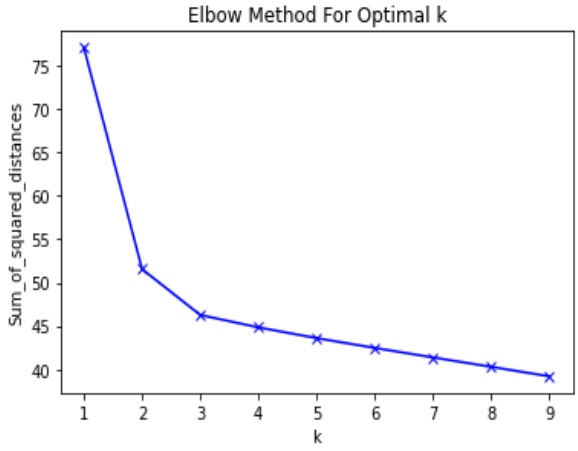

Figure 4

Figure 5 shows the clusters in two different colors corresponding to Metro stations superimposed on top of folium map. Cluster 0 is represented in red and has 5 stations whereas Cluster 1 in purple and has 18 stations.

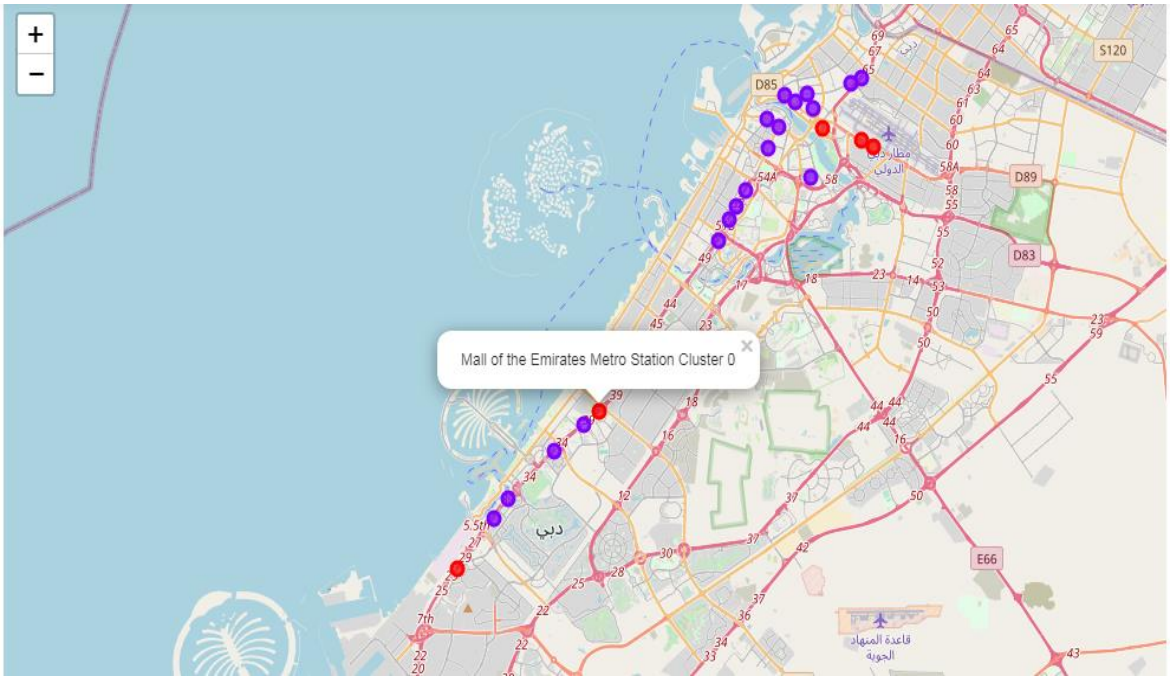

Figure 5

The catplot visualization shown in figure 6 depicts the cluster labels associated with most common venue pertaining to each Metro station. Cluster 0 comprises of arbitrary venues which is out of scope. It is observed that hotel, restaurants, coffee shop and café categories fall under Cluster 1 which is the prominent feature in this study.

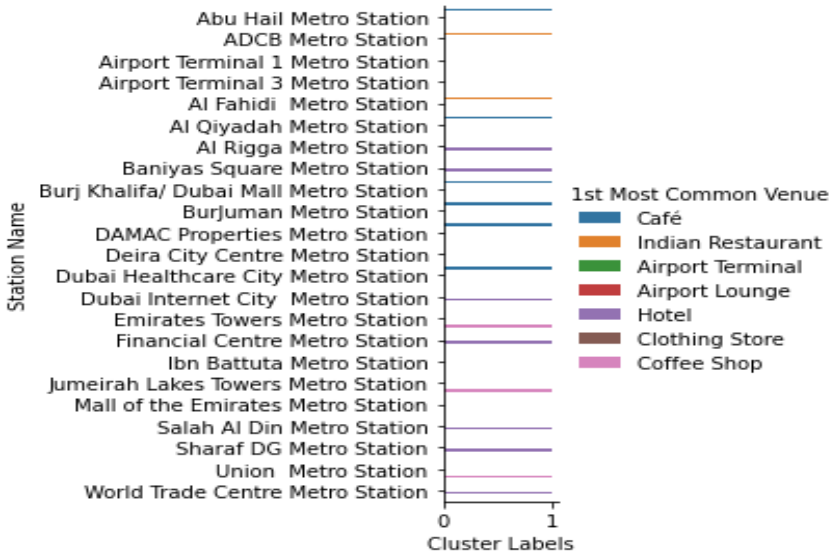

Figure 6

As our interest fall under venues in restaurants/hotel categories all other venues namely coffee shops, pizza places, bakeries etc. are not considered since they are not direct competitors. Therefore, analysis will be performed on restaurants/Hotels categories alone.

Further, the number of Indian cuisine restaurants is compared with other restaurants and hotels (including Indian cuisine) pertaining to each metro station from Cluster 1. From the figure 7, it is observed Al Qiyada Metro Station, Al Rigga Metro station and Burj Khalifa Metro station has zero Indian restaurants in 500 meters vicinity. Whereas DAMAC Metro station has highest number of restaurants and Abu Hail Metro station has 6 of them.

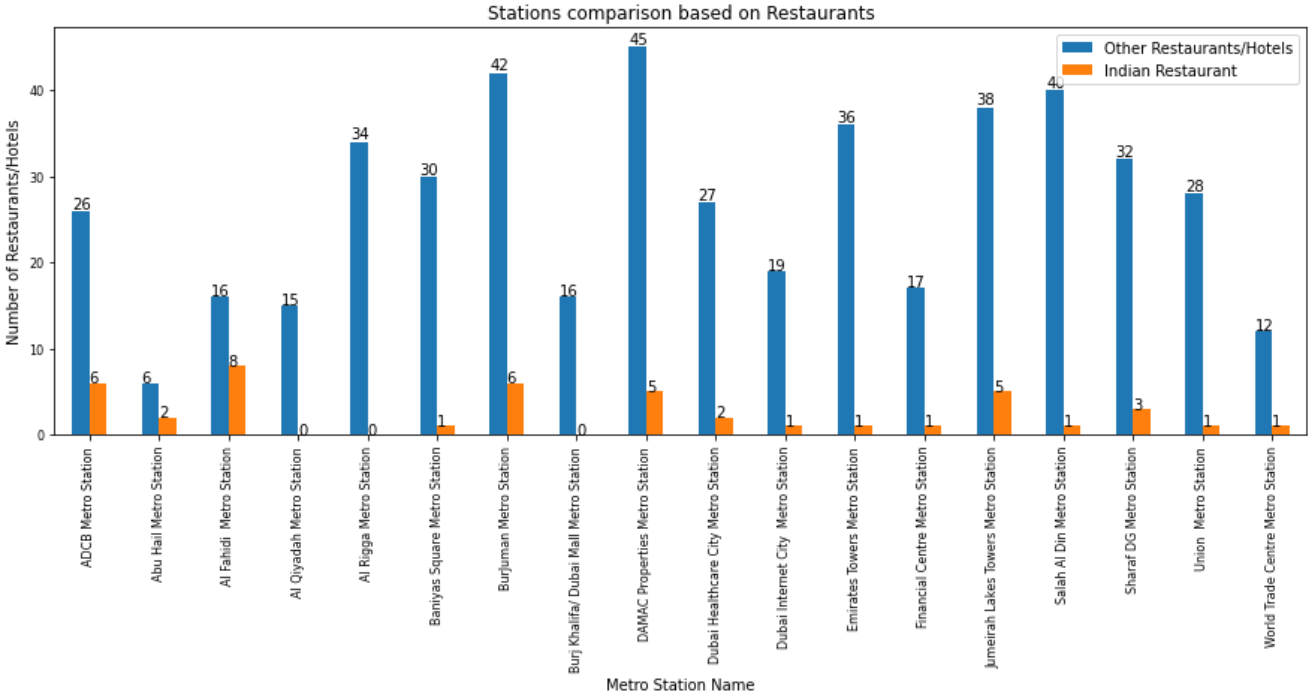

Figure 7

In order to draw a conclusion for optimal location for the Business, let us visualize the intensity of Indian restaurant density with the Heat map shown in figure8. In this case, the heatmap represents matrix wherein each column represent the total number of restaurants/hotels corresponding to each Metro stations represented in the row. Each square of the heatmap represents a cell which displays density of Indian restaurants ranging from low to high density value.

This study is compared with House price index [2], Population density [3], Tourist attractions [6] and from the Number of Indian restaurants, other Restaurants and Hotels.

From the below figure, we observe that Al Qiyadah Metro station, Al Rigga Metro station and Burj Khalifa Metro station has lower density of Indian restaurants making it an ideal location to set up the business. However, there are various factors that can influence the decision. Let us analyze each station in this zone from a business perspective. We have noticed that Burj Khalifa Metro station is Dubai's downtown with prominent attractions making it one of the top tourist destinations in the world. The

House price index is extremely high for this locality, involving premium cost for business setup. Al Rigga Metro station has high population density with low house price index, making it most suitable for the business. Whereas Al Qiyadah Metro station has low House price index, mostly industrial area and located far away from the city limits making it not a great location for the business.

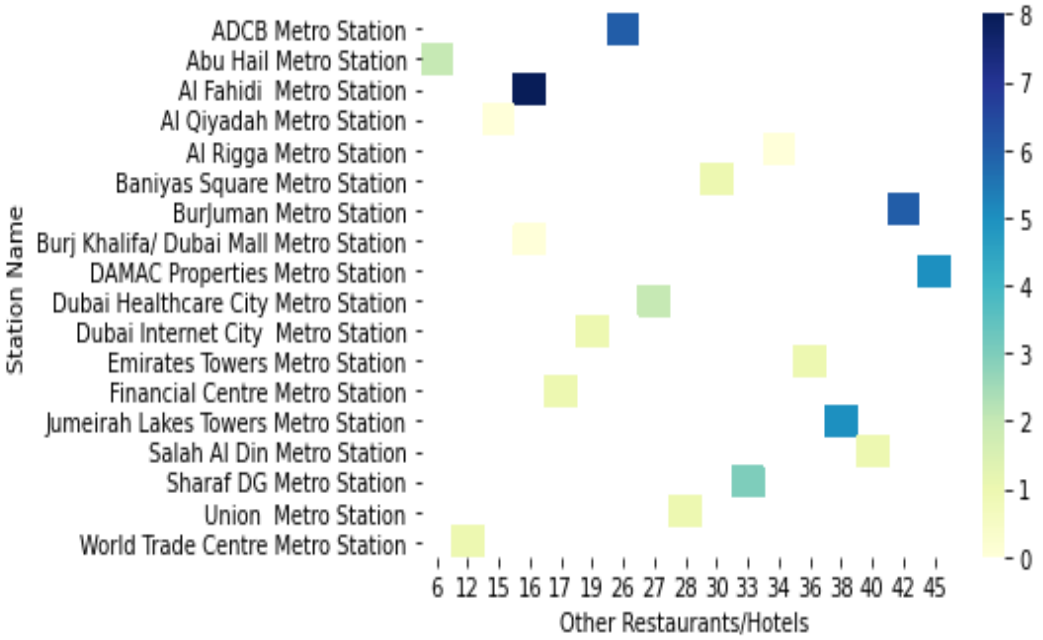

Figure 8

The heat map also reveals stations such as Baniyas Square Metro station, Dubai Internet City Metro station, Emirates Tower Metro station, Financial Centre Metro station, Salah Al Din Metro station, Union Metro station and World Trade Centre Metro station has lower density of Indian restaurants making it a second ideal locations. Let us analyze each station in this zone from a business perspective. Baniyas Square has a lot of tourist attraction with higher population density and low House price index making an ideal location for Business. Whereas Emirates Tower, Financial Centre and World Trade Centre fall under one stretch in metro line closer to downtown area having high House price index. Thus population density is lower in these regions making it less feasible area of interest for business. Dubai Internet City is dedicated to Information Technology sectors having low population density therefore not an ideal location for restaurant business. We have noticed mostly coffee shops and café category in this region. Salah Al Din Metro station and Union Metro come under same metro line and are located less than 1 km radius. These two stations have high density of population with mid-level House price index making it significant and affordable location with plenty of tourist attractions in the vicinity.

Abu Hail Metro station, Dubai Healthcare City Metro station and Sharaf DG Metro station falls under third optimal location with mid-density of Indian restaurant in the region. Let us analyze each station in this zone from a business perspective. Abu Hail is densely populated with low income households having low House price index. We have found that this area is not a tourist spot, thus making it an ideal location for a low budget business setup. Dubai Healthcare city is dedicated healthcare freezone area with less

tourist spots. However, population density in this area is very low due to high House pricing index. Sharaf DG has high population density having residential area with reasonable House price index. This could be an ideal location for setting up the business.

All other stations shown in the hottest region in the heatmap with highest density of Indian restaurants will not be the best suit for business.

4. Results and Discussions

The dataset for Dubai Metro, both Red and Green lines had 1544 venues belonging to 221 unique categories. The venues were selected within a range of 500 meter radius from each metro station which was retrieved using Foursquare geopy service. By employing clustering algorithm, Metro stations were clustered to obtain zones of interest. Since our objective is to find the best location for an Indian cuisine restaurant, we have chosen cluster belonging to 'food' venue category for this research. This primarily contains hotels, restaurants, cafes and coffee shops. As per the study, large numbers of restaurant were detected in red line Metro and few restaurants in the green line Metro. It is observed that several Metro stations have lesser venues, which makes the analysis difficult because these stations lack tourist spots, distant from city limit and are most likely unique when employing clustering algorithm. Therefore Metro stations with less than 25 venues were discarded.

This study intends to find the optimal location for restaurant business. The following category like coffee shops, café's, Pizza shops are not considered as our direct competitors therefore these were eliminated from the research and had selected only "other restaurants/hotels" category for the analysis. This enables the research to locate the optimal location to setup an Indian restaurant. Further we perform analysis based on popular tourist attractions, population density and House price index for each Metro station, which has low Indian restaurant density as depicted from the heatmap. The entire hot region in the heatmap corresponding to high Indian restaurant density was withdrawn from the research.

To draw the inference, recommended zones must be considered as a jump off point for a detailed study as it will eventually summarize the location where relevant factors are taken into consideration with less/no competition.

5. Conclusion

The objective of this project was to detect the optimal location for Indian cuisine restaurants within 500 meter radius with less or no competition. By considering restaurant density distribution from heatmap, we have identified several potential candidate stations by looking at Population density, popular tourist attractions and House price index with no/less competitions which will aid stakeholders in tapering the search for the most favorable business investment. The final decision in deriving on optimal location is solely based on their preferences.

6. References

1. http://data.bayanat.ae/en_GB/dataset/metro-stations/resource/666412a8-3383-47df-b521-c70150698e6a
2. https://www.propertyfinder.ae/blog/infographic-dubai-metro-rent-map/
3. https://www.dsc.gov.ae/en-us
4. https://en.wikipedia.org/wiki/Dubai
5. https://en.wikipedia.org/wiki/Dubai_Metro
6. https://www.visitdubai.com/en