BUSINESS OPPORTUNITIES IN TRIVANDRUM, INDIA

Coursera Capstone Project, IBM Data science Certification

G. HARIKRISHNAN

INTRODUCTION

A movie is best enjoyed when seen in a theatre. Partially due to the theatre ergonomics and partially due to that disturbance free environment. A theatre is where individuals rush to during free time or for relaxation. Coming to theatres , almost all shopping malls have a theatre and multiplexes are a common sight. Film industry is always on the run and theatre form a crucial role in the revenue gains of a movie. The more interesting the film is , the more people who comes to see it. A person Starting a theatre is a highly profitable task given that the theatre is located in an optimum position.

Just like a movie, food also plays a major part in our lives. Most of us enjoy the restaurant food as well as the ambience it offers. The restaurant is also a major relief for the working-class people as a means to energize them in their busy schedule. Starting a restaurant may seem like an easy task but to choose the right restaurant at an optimum place to maximize the sales is tough task. Hence, we will be seeing this also.

 The location is one of the key decisions to be taken while choosing a place for this large budget tasks. This can be done by analysing the data of the city.

Business Opportunities

The project aims to utilize all the data science techniques learned so far to analyse the data of a city to come to a location as to where it would be optimum to start a Movie theatre or a restaurant in Trivandrum. This will be useful for business men looking for a location to start off the business.

Target audience

This project is oriented for business men who intend to start off a new business in the entertainment industry. This will also be good for the people returning from the foreign countries due to the recent pandemic to find a new place to invest their money in.

DATA

Like the oil for a generator, data is the essential commodity of prime importance for a data analysis. So, for the project we require the list of regions in the city, the latitude and longitude of all these places, the various venues in these localities. With these data we can come to the analysis.

Source of Data

The list of locations In the city is available through the Wikipedia page (https://en.wikipedia.org/wiki/Category:Suburbs_of_Thiruvananthapuram) . This can be scrapped for data of these locations. Then using geocoder, we will be finding the latitudes and longitudes of these places and storing them. Finally using these position vectors and foursquare api we will get the location data of all these places and their corresponding business.

This acquired data will first be used for location of theatre and then for restaurants. For both these cases we will be dividing the data into useable columns and then clustering them. These clusters will then be analysed to find the best spots. The best spots can be chosen by the data in conjunction with the nearness of the location to certain point of interests

METHODOLOGY

For this project we are utilizing Jupyter notebook, a graphical interface for python and similar languages. Jupyter lab allows us to execute a small portion of code and see the output of that process rather than the final result.

First of all, we import all the python run time libraries like Numpy and pandas. We also import scikit learn for the clustering and folium and geocoder for the visual analysis. Using Beautifulsoup we scrap a Wikipedia page (https://en.wikipedia.org/wiki/Category:Suburbs_of_Thiruvananthapuram) for the locations in the city. The data is stored in a data frame. Using geocoder, a utility we find the latitudes and longitudes of all these localities. We use an ArcGIS instead of Google for acquiring the Coordinates. Using map visualisation, we remove those places which are wrongly pointed and those which are overlapping. These points are dropped from the data frame.

Using foursquare API, we then find the venues nearby the localities, Here I have placed a limit of 3000 (3 KM).This is a random assumption that a person won't be willing to look for food beyond 3 KM radius of his house. The foursquare API requires a developer account with a client id and secret key. While passing all these credentials we will get all the area details with foursquare. Using dummy variable can see what all are available. We the take a frequency of these in the areas and take a group it by each neighbourhood.

Theatre

For the theatre , from the frequency data frame we choose the columns that are the shows the frequency of theatres. Here chose different columns as all those will indirectly mean that a theatre is present in that location. These are then made into a separate data frame. Using K means method of unsupervised clustering, we then split this into 4 frames depending on the values of these columns. The four clusters are marked in the data frame along with the location coordinate of the place, a map is visualized using folium. Using the different colour in the map we can observe what is the cluster of each points in the map. Also, the map allows us to get insight of the locality where the theatres are more, the geo economic condition of that are. The finally we separate the data frames individually to know the underlying information of all cluster.

Restaurants

Just like in the case of theatres, we made a separate data frame from the frequency data frame. Then using K means clustering we divided the area into 3 individual locations. The clusters along with the location of each clusters are then utilised to forma a map using folium. The observations are done here also of the locality . The clusters are then individually analysed to come to a conclusion.

RESULTS

1. Theatre

It can be observed that the cluster 1 has high concentration of movie theatres, this can be seen in the map also that these are the sites in the heart of the city, hence this also points out that the city centre is having sufficient theatres. Coming to cluster 2,it lies far from city centre but the cluster has seen that here are theatres but low in number cluster 3 lies throughout the city, has low to none of the amount available theatres. Cluster 4 is similar to cluster 2 has multiplexes. From a business point of view, it is better to open a theatre in the places marked in cluster 3

Image.1 Theatre Clusters

2. Restaurants

Here the 3 clusters are cluster 1 has large concentration of restaurants hence in cluster 1 we should not open a restaurant, in cluster 2 we see that neither restaurant nor fast food restaurants are common, means they are rare, these even include the city centre. Cluster 3 has good number of restaurants and fast food chains but lower than cluster 1. So best option will be to open restaurants in areas in cluster 2.
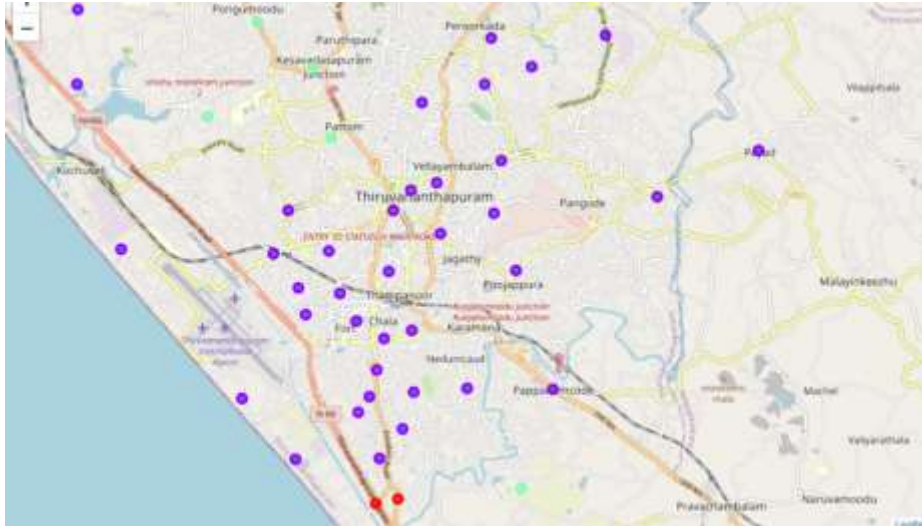
Image.2 Restaurant Cluster

DISCUSSION AND FUTURE SCOPE

The above analysis is purely done on the basis of the concentration of similar location in a locality and not any other factor. The information of each localities taken from foursquare only and there may be missing data. The above analysis is to give an idea of the business opportunities in the city.

This project can be further implemented in may location and including other factors like population density, near by land marks etc. to increase the accuracy of the analysis. The same project can be used for finding business opportunities in different sectors in the same city.

CONCLUSION

From this project we have initially declared a problem statement of business value. Then using the data science methodologies, acquired , sorted and analysed the data for finding the solution of the problem statement. This analysis will prove good for any business man who is interested in starting or acquiring business in the area. To directly answer the problem statement, I would like to suggest cluster 3 in the analysis for starting a Theatre business and cluster 2 of the analysis of restaurants for starting a restaurant business.

REFFERENCES

1. Location information from Wikipedia page

https://en.wikipedia.org/wiki/Category:Suburbs_of_Thiruvananthapuram

2. Foursquare

https://foursquare.com/developers/apps