**THE BATTLE OF NEIGHBORHOODS**

**SECTION 1-INTRODUCTION**

**BACKGROUND**

In this section, I will define my idea of choosing this project, where I leverage the Four square location data to solve the imagined business opportunity.

There are 100’s and 1000’s of sites on Internet including Foursquare that tells us about places to go, things to see, restaurants to eat but when it comes to finding out all the information about a city, it requires a lot of work.

While choosing an appropriate location to open a restaurant, it’s important to understand what to look out for. The following factors must be considered for choosing just the right spot.

1-Visibility

If the restaurant is located in a side street with a little foot traffic, it’s very difficult to attract customers. To determine visibility, look at foot and car traffic patterns.

2-Crime Rates

Crime rates are unglamorous considerations, but if the restaurant is located in a crime-laden area, are the target customers going to visit? High crime rates can make potential customers uncomfortable, and if they feel they’ll be mugged walking to their cars, it will only drive away business, no matter how legendary it is.

3-Surrounding Business and Competitor Analysis

Research regarding surrounding businesses is a must. Are they doing well? Is the area affluent? Is there enough room for our business? Also, we’ll want to know what types of restaurants do well in the area; however, we don’t want to open a pizzeria if there are four in the area. Areas can only support so many of the same type of restaurant. What will distinguish any new restaurant is excellent service and consistently wonderful food.

4-Accessibility

Some restaurants find success in just-off-the-highway locations, or located near exits for those interstate travellers who need a bite to eat but don’t want the usual fast-food restaurant. We’d have to keep peak times in mind for these kinds of locations, as well as customer demographic.

5-Affordability

Cost is always a bottom-line consideration for any business. If the rent or purchase of the space is more than we’ll earn in each month in profits, that location is not feasible at that time. However, if we know that we’ll generate business from that location, then it can be considered. Although some risks do pay off, we don’t want to be at the point where we’re struggling to cover basic costs.

6-Safety

Once chosen, keep in mind that OSHA reports that slips, trips, and falls are the most common workplace accident, accounting for 15 percent of all accidental deaths, and they are more prevalent in the restaurant industry. It’s important to install slip-resistant flooring, have slip-resistant mats in the kitchen spaces, and provide the workers with highly-rated safe, slip-resistant footwear to help prevent those preventable slips and trips. Slips and trips could cost the company more in fines than renting a space, so make sure to consider all angles of safety when choosing the best space. If the space has its own equipment, make sure that all fryers are safe, ice machines aren’t leaking, and there aren’t any trip hazards in the kitchen.

**BUSINESS PROBLEM**

In this scenario, it is urgent to adopt machine learning tools in order to assist homebuyers in US to make wise and effective decisions. As a result, the business problem we are currently posing is: how could we provide support to homebuyers in USA to purchase a suitable real estate in this uncertain economic and financial scenario?

**PROJECT IDEA**

My idea for the Capstone Project is to show that when driven by venue and location data from Four Square, it is possible to present the cautious and nervous traveller/migrant with a list of attractions to visit supplemented with the graphics showing the chances of opening a successful restaurant in any particular area.

The Approach is as follows:

1. The travellers decide on a city location [in this case Chicago].
2. The Fore Square website is scrapped for the top venues in the city.
3. From this list of top venues the list is augmented with additional geographical data.
4. Using this additional geographical data the top nearby restaurants are selected.
5. A map is presented to the to the traveller showing the selected venues and the number of restaurants already being there in that area.

**Who is this solution targeted at?**

I believe this is a relevant challenge with valid questions for anyone moving to other large city in US, EU or Asia. The same methodology can be applied in accordance to demands as applicable. This case is also applicable for anyone interested in exploring starting or locating a new business in any city. Lastly, it can also serve as a good practical exercise to develop Data Science skills.

There are many data science aspect of this project including:

1. Data Acquisition
2. Data Cleansing
3. Data Analysis
4. Machine Learning
5. Prediction

**SECTION 2- DATA SECTION**

**REQUIREMENTS**

To explore and target recommended locations across different venues according to the presence of amenities and essential facilities, we will access data through Four Square API interface and arrange them as a data frame for visualization

For this the following data is required:

1. Geographical location of Lawrence, Chicago.
2. Four Square API to get the locations of all the nearby restaurants.
3. Folium Library to visualize the results.

**METHODOLOGY**

The Methodology section will describe the main components of our analysis and predication system. The Methodology section comprises four stages:

1. Collect Inspection Data

2. Explore and Understand Data

3. Data preparation and preprocessing

4. Modeling

**1-COLLECT INSPECTION DATA**

First, we have imported all the necessary libraries and then the geolocator is used to get the latitude and longitude of West Lawrence Avenue, Chicago. To look for similar restaurants in the neighborhood of West Lawrence, Four Square API is used and the resulting json files are converted to a data frame.

Before using data, we will have to explore and understand it.

#### 2. EXPLORE AND UNDERSTAND DATA

#### We read the dataset that we collected from the FourSquare API into a pandas' data frame and display the first five rows of it using head() function.

#### We filtered the dataframe to retain only the relevant attributes like name, address, latitude and longitude. Then, using a function get\_category\_type() we retrieved the categories of all the restaurants.

#### All the restaurants are then plotted on the map using folium library.

**3-DATA PREPERATION AND PRE-PROCESSING**

At this stage, we prepare our dataset for the modeling process, opting for the most suitable machine learning algorithm for our scope. Accordingly, we perform the following steps:

* Rename the column name.
* Select only the location or id data.
* Read the restaurant-wise coordinates into a data frame, eliminating recurring word Location from individual names.
* Deal with the missing Values
* Normalize the data.
* Plot the data using Choroplath maps.

**MODELING**

After exploring the dataset and gaining insights into it, we are ready to use the clustering methodology to analyze real estates. We will use the k-means clustering technique as it is fast and efficient in terms of computational cost, is highly flexible to account for mutations in real estate market in US and is accurate.

* Retrieve all the numerical values from the table like address, latitude, longitude.
* Normalize the retrieved values and make their 3 clusters.
* Retrieve the names of the venues of all the 3 clusters.
* Append the label assigned to each row in the table.
* Plot all the clusters on Chicago map.

**SECTION 3- RESULT AND DISCUSSION**

The visitor can now decide the venue for his business by considering the number of restaurants in that area. He should consider the area with fewer number of restaurants, as it would increase his chances of earning profit. If not satisfied, and want to explore more venues, then again Foursquare API can be used in the similar manner to retrieve the details of all the venues of Chicago.

Although all of the goals of this project were met there is definitely room for further improvement and development as noted below. However, the goals of the project were met and, with some more work, could easily be developed into a fully pledged application that could support the cautious traveller in an unknown location.

The following are suggestions how this project could be further developed:

1-Best time to visit each venue

2-Suggestions for morning, afternoon, evening and night time

3-Daily itineraries

4-Route planning and transportation

5-Time lapse of the crime in the area of the venue

**SECTION 4-CONCLUSION**

The competitor analysis can also be refined by segregating the restaurants in North Beach into different categories. This information is extremely useful because we certainly don’t want too many competitors in the same sector. While the approach discussed here is primitive, it nevertheless showcases the usefulness of data analysis!

I feel rewarded with the efforts, time and money spent. I believe this course with all the topics covered is well worthy of appreciation. This project has shown me a practical application to resolve a real situation that has impacting personal and financial impact using Data Science tools.

The mapping with Folium is a very powerful technique to consolidate information and make the analysis and decision thoroughly and with confidence. I would recommend for use in similar situations.[**¶**](https://render.githubusercontent.com/view/ipynb?commit=59c42f55f76f18b565621714064716ed8df8bec7&enc_url=68747470733a2f2f7261772e67697468756275736572636f6e74656e742e636f6d2f676e617669613030372f436f7572736572615f43617073746f6e652f353963343266353566373666313862353635363231373134303634373136656438646638626563372f43617073746f6e652532302d253230546865253230426174746c652532306f662532304e65696768626f72686f6f64732532302d253230474e2532302d25323046696e616c2e6970796e62&nwo=gnavia007%2FCoursera_Capstone&path=Capstone+-+The+Battle+of+Neighborhoods+-+GN+-+Final.ipynb&repository_id=157266007&repository_type=Repository#The-mapping-with-Folium-is-a-very-powerful-technique-to-consolidate-information-and-make-the-analysis-and-decision-thoroughly-and-with-confidence.-I-would-recommend-for-use-in-similar-situations.)