INTRODUCTION

Chicago is the **most populous city** in the U.S. state of Illinois and the third most populous city in the United States.

Located on the shores of freshwater Lake Michigan, Chicago is an international hub for finance, culture, commerce, industry, education, technology, telecommunications, and transportation.

Chicago's 58 million domestic and international visitors in 2018 made it the second most visited city in the nation, not far behind New York City's 65 million visitors in 2018. Thus, business opportunities are vast in the city and so is the competition. Many big players have been attracted to this city in all areas of business. As is it a highly developed city, the cost of doing the business is huge and risk is high. Thus, any new business venture or expansion needs to be analyzed and studied carefully. This will ensure good understanding of the business environment and strategical planning to reduce the risk factor and increase returns.

BUSINESS PROBLEM

As a business and tourist hub, the city is famous for its cuisine as well. From street food to high end Michelin star restaurants, you will find everything in Chicago. Various well-known chefs have their restaurants there. In 2003, Robb Report named Chicago the country's "most exceptional dining destination".
The diversity in people settled here is reflected in its food as well. Among various international cuisines, Mediterranean food has huge fan following. There are many Mediterranean restaurants among various parts of the city. Daily new ones are opening but the number of them shutting down is also big. So, what differentiates a successful one from the unsuccessful. Its not always the quality of food. Some times many small things have to be considered to make a project work.
Find a suitable location.
The area where there are least number of Mediterranean restaurants will be the ideal one to open a new restaurant. Also, the neighborhood should be either middle east origin or the people should be interested in less sophisticated setting but more flavorful food.
A very high end street is not a suitable area for such eatery. Strategical planning is must. Various factors need to be studied before giving the conclusions.

Target Audience:

The objective is to locate and recommend to the management the best neighborhood to open a Mediterranean restaurant in Chicago.

DATA ACQUISITION

This demonstration will make use of the following data sources:

Chicago Neighborhood Dataset

Data will be retrieved from Chicago (Boundaries – Neighborhoods) open from https://data.cityofchicago.org website.

Neighborhood boundaries in Chicago, as developed by the Office of Tourism. These boundaries are approximate and names are not official. The data can be viewed on the Chicago Data Portal with a web browser. However, to view or use the files outside of a web browser, you will need to use compression software and special GIS software, such as ESRI ArcGIS (shapefile) or Google Earth (KML or KMZ), is required.

Chicago Neighborhood location Information

Data coordinates of Neighborhood will be retrieved using GeoPy Geocoder Library.

Chicago Top Venue Recommendations from FourSquare API

(FourSquare website: www.foursquare.com)

I will be using the FourSquare API to explore neighborhoods in selected towns in Chicago. The Foursquare explore function will be used to get the most common venue categories in each neighborhood, and then use this feature to group the neighborhoods into clusters. The following information are retrieved on the first query:

- Venue ID
- Venue Name
- Coordinates : Latitude and Longitude
- Category Name

METHODOLOGY

Chicago Towns List : The source data contains Neighborhoods of City of Chicago. Data Cleanup and re-grouping. The retrieved table contains some un-wanted entries and needs some cleanup. The following tasks will be performed:

- Drop/ignore cells with missing data.
- Use most current data record.
- Fix data types.
- Adding geographical coordinates of each town location

1.Retrieve town coordinates.

GeoPy library was be used to retrieve the coordinates (latitude and longitude of each town centers. For this exercise, The town coordinates will be used in retrieval of Foursquare API location data.

Getting Geo-Location of the Neighborhoods

```python
appended_data = []
for row in data_chicago.itertuples():
    neigh=row.PRI_NEIGH+',Chicago,IL'

    location = geolocator.geocode(neigh)
    if location is not None:
        print (neigh)
        data_chic_latlong = {
                'Neighborhood':row.PRI_NEIGH,
                'Latitude':  location.latitude,
                'Longitude':location.longitude,
                    }

        appended_data.append(data_chic_latlong)
# see pd.concat documentation for more info
df=pd.DataFrame(appended_data)
# write DataFrame to an excel sheet
df.to_excel('appended.xlsx')
print(df)
```

2. Segmenting and Clustering Towns in Chicago:

Retrieving FourSquare Places of interest.

Using the Foursquare API, the explore API function was be used to get the most common venue categories in each neighborhood, and then used this feature to group the neighborhoods into clusters. The k-means clustering algorithm was used for the analysis. Fnally, the Folium library is used to visualize the recommended neighborhoods and their

emerging clusters. In the ipynb notebook, the function getNearbyVenues extracts the following information for the dataframe it generates:

- Venue ID
- Venue Name
- Coordinates : Latitude and Longitude
- Category Name

The function getVenuesByCategory performs the following:

1. Category based venue search to simulate user venue searches based on certain places of interest. This search extracts the following information:
   - Venue ID
   - Venue Name
   - Coordinates : Latitude and Longitude
   - Category Name
2. For each retrieved venueID, retrieve the venues category rating.
   This Foursquare search is expected to collect venues in the following category:
   - Mediterranean Restaurants

   I used the FourSquare API to retrieve venue scores of locations. Note that there is max query limit of 50 in FourSquare API for free subscription.

   Data cleanup un-needed entries
   - Eliminate possible venue duplicates.
   - Improve the quality of our venue selection by removing venues with no ratings or 0.0

TOWN	SCORE															
	5.4	5.8	5.9	6.5	6.6	6.8	7	7.1	7.3	7.6	7.7	7.9	8.3	8.6	8.7	8.9
Andersonville											1		1			
Bucktown														1	1	
East Village					1		1			1						
Grant Park					1		1			1						
Greektown	1			1						1						
Hyde Park								1								
Lincoln Park												1				
Little Village					1		1			1						
Loop					1		1			1						
Magnificent Mile									1					1		1
O'Hare			1													
Printers Row						1		1								
Streeterville									1							
United Center												1				
Uptown		1														
West Loop					1		1			1						
Grand Total	1	1	1	1	5	1	5	2	2	6	1	2	1	2	1	1

What are the top 10 Venue Recommendation Scores ?

SCORE
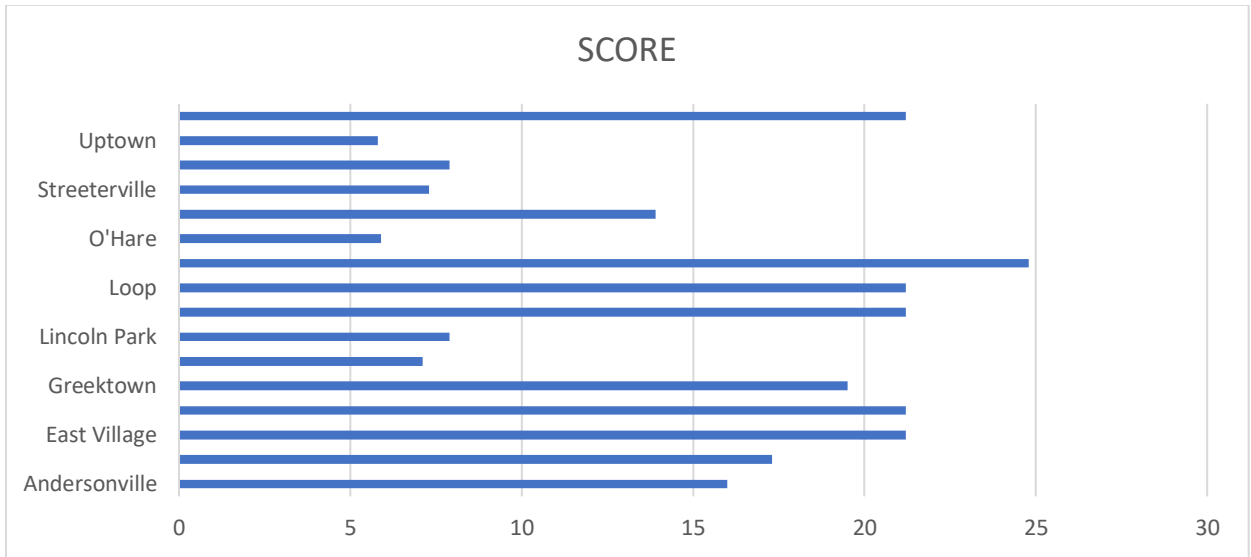
Analyze Chicago Town nearby recommended venues

Technique : One Hot Encoding

```python
# one hot encoding
chicago_onehot = pd.get_dummies(chicago_town_venues[['category']], prefix="", prefix_sep="")

# add neighborhood column back to dataframe
chicago_onehot['Neighborhood'] = chicago_town_venues['Town']

# move neighborhood column to the first column
fixed_columns = [chicago_onehot.columns[-1]] + list(chicago_onehot.columns[:-1])
chicago_onehot = chicago_onehot[fixed_columns]

chicago_onehot.head()
```

	Neighborhood	Falafel Restaurants	Food Stands	Greek Restaurants	Mediterranean Restaurants	Middle Eastern Restaurants	Moroccan Restaurants	Tapas Restaurants
0	Printers Row	0	0	0	0	0	0	1
1	Printers Row	0	0	0	1	0	0	0
2	Printers Row	0	0	0	1	0	0	0
3	Printers Row	0	0	0	1	0	0	0
4	Printers Row	0	0	0	1	0	0	0

Categorized Result

	Neighborhood	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
0	Albany Park	Mediterranean Restaurants	Tapas Restaurants	Moroccan Restaurants	Middle Eastern Restaurants	Greek Restaurants	Food Stands	Falafel Restaurants
1	Andersonville	Middle Eastern Restaurants	Mediterranean Restaurants	Tapas Restaurants	Moroccan Restaurants	Greek Restaurants	Food Stands	Falafel Restaurants
2	Belmont Cragin	Mediterranean Restaurants	Tapas Restaurants	Moroccan Restaurants	Middle Eastern Restaurants	Greek Restaurants	Food Stands	Falafel Restaurants
3	Boystown	Mediterranean Restaurants	Tapas Restaurants	Moroccan Restaurants	Middle Eastern Restaurants	Greek Restaurants	Food Stands	Falafel Restaurants
4	Bridgeport	Mediterranean Restaurants	Tapas Restaurants	Moroccan Restaurants	Middle Eastern Restaurants	Greek Restaurants	Food Stands	Falafel Restaurants

RESULTS : k-means Cluster Results

Clustered results for k-means to cluster with 5 clusters.

Cluster 2 and 4
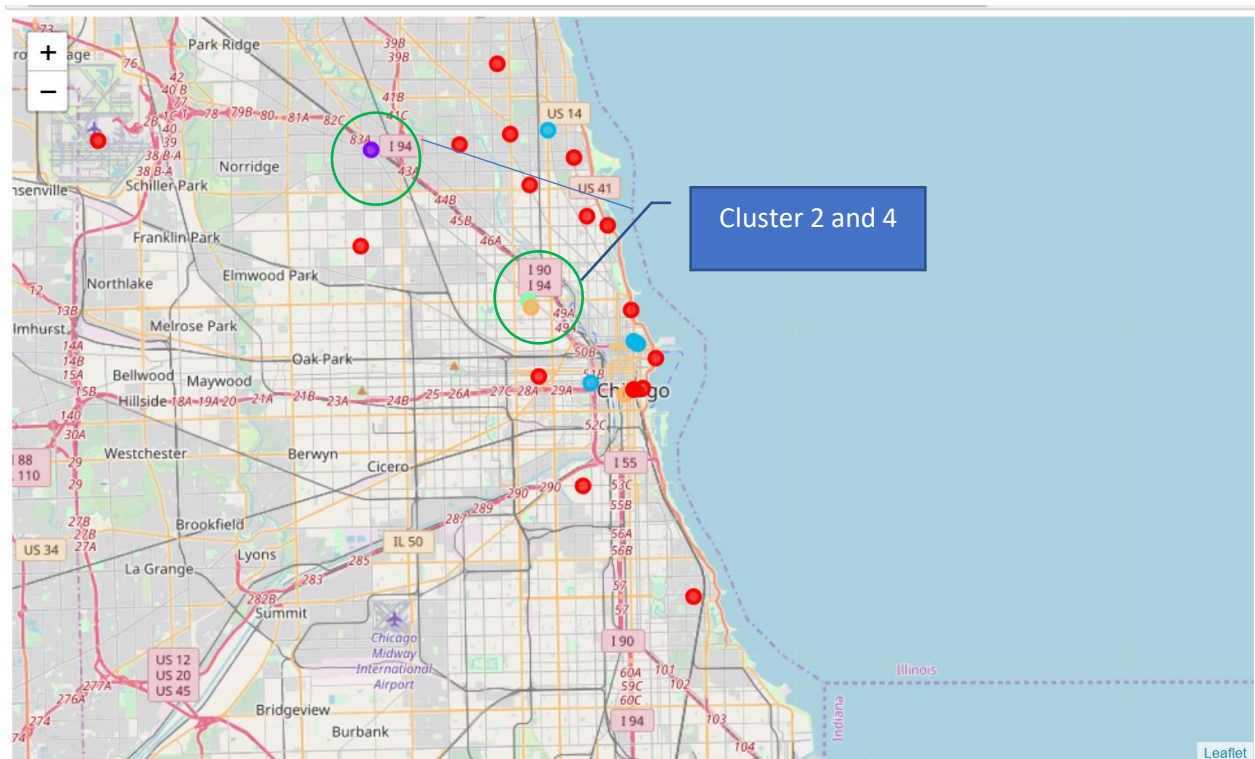
Discussion and Conclusion

On this notebook, Analysis of best neighborhood to open a Mediterranean food joint has been presented. Based on the number of Mediterranean food shops(restaurants, kiosks or takeaway)in the area recommendations are provided for a suitable place for opening a new one. Since Chicago is a big city and an expensive one too, narrowing down the areas with less such joints and rating should be a good insight for the business.

Using Foursquare API, we have collected a good amount of data regarding the food preferences in the area. The sourcing from FourSquare has also its limitations. The list of people with food preferences is not an exhaustive one. The locations of currently running Mediterranean food joints is extensive though. Sourcing from the venue recommendations from FourSquare has its limitation, The list of venues is not exhaustive list of all the available venues is the area. Furthermore, not all the venues found in the the area has a stored ratings. For this reason, the number of analyzed venues are only about 50% of all the available venues initially collected. The results therefore may significantly change, when more information are collected on those with missing data.

The generated clusters from our results shows Cluster 2 and Cluster 4 are best places to open the Mediterranean Restaurant. The results shows that while Jefferson Part and BuckTown neighborhoods can be best suited to open a restaurant with the available data.