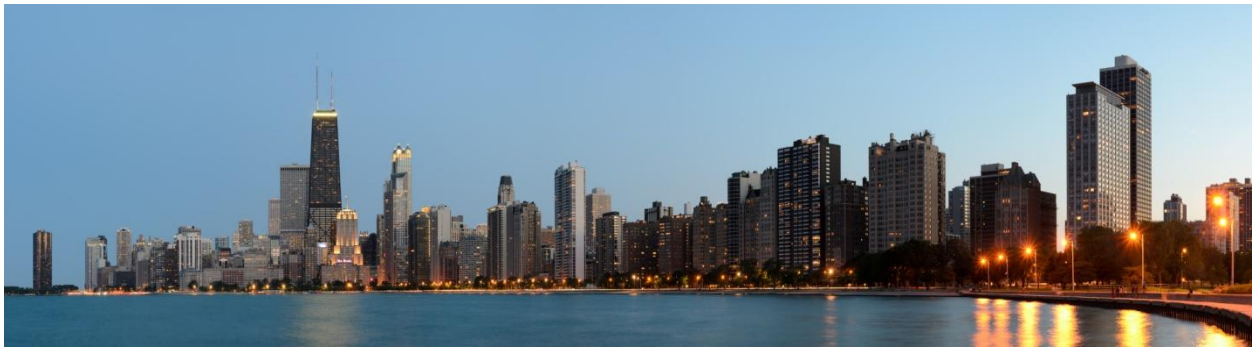
I. Table of Contents



City of Chicago

II. Introduction:

Located on the shores of freshwater Lake Michigan Chicago, officially the City of Chicago, is the 3rd most populous city in the United States. With an estimated population of 2,693,976 in 2019, it is also the most populous city in the Midwestern United States. Chicago is the principal city of the Chicago metropolitan area, the Combined Statistical Area (almost 10 million residents) often called Chicagoland. It constitutes the third most populous urban area in the United States after New York City and Los Angeles.

Chicago is an international hub for finance, culture, commerce, industry, education, technology, telecommunications, and transportation. It is the site of the creation of the first standardized futures contracts, issued by the Chicago Board of Trade, which today is part of the largest and most diverse derivatives market in the world, generating 20% of all volume in commodities and financial futures alone. O'Hare International Airport is routinely ranked among the world's top six busiest airports according to tracked data by the Airports Council International. The region also has the largest number of federal highways and is the nation's railroad hub.
Several Fortune 500 companies, Including Allstate, Boeing, Caterpillar, Exelon, Kraft Heinz, McDonald's, Mondelez International, Sears, United Airlines Holdings, US Foods and Walgreens.

Chicago's 58 million tourist visitors in 2018 set a new record. Landmarks in the city include Millennium Park, Navy Pier, the Magnificent Mile, the Art Institute of Chicago, Museum Campus, the Willis (Sears) Tower, Grant Park, the Museum of Science and Industry, and Lincoln Park Zoo. Chicago's culture includes the visual arts, literature, film, theatre, comedy (especially improvisational comedy), food, and music, particularly jazz, blues, soul, hip-hop, gospel, and electronic dance music including house music. Chicago is home to all major-league sports teams including Chicago Cubs baseball, Bulls Basket ball, Bears football, Black hawks ice hokey. Of the area's many colleges and universities, the University of Chicago, Northwestern University, and the University of Illinois at Chicago are classified as "highest research" doctoral universities.

III. Business problem

Chicago's world famous music events, festivals, major-league and collage sports events, several industry conventions, Annual Auto show, magnificent mile shopping etc., draw lot of domestic and foreign tourist. The combination of 58 million tourist (yr 2018) visitors and 10 million residents opens new opportunities for new Latin American restaurants. Finding good locations is key to the success of restaurant Industry. Purpose of the project is to forecasts new locations for the business.

IV. Data

A. Chicago Neighborhoods:

Unlike New York city, city of Chicago does not have boroughs. Chicago is divided into community areas and neighborhoods. Chicago is divided into 77 community areas for statistical and planning purposes. Census data and other statistics are tied to the areas.
The areas are distinct from but, related to the more numerous neighborhoods of Chicago; an area often corresponds to a neighborhood or encompasses several neighborhoods. Project explores the data from the view of neighborhoods. File of community areas and neighborhoods is stored in a google cloud storage bucket.

B. GeoPy:

Geopy is used for by python developers to locate the coordinates of addresses, cities, countries, and landmarks across the globe using third-party geocoders and other data sources. Project uses geopy.geocoders for getting latitude and longitude coordinates. Using the list of neighboroods, we get the latitude and longitude coordinates.

C. Foursquare api:

This project uses Foursquare api for a getting list of nearby venues by neighborhoods. Foursquare is a location data provider of information about all venues and events within an area of interest (in this case neighborhoods). Information includes venue names, locations, menus and even photos. As such, the foursquare location platform will be used as the data source since all the required information can be obtained through this API.

Using the neighborhoods latitude and longitude coordinates, connect to the Foursquare api and gather information about venues in each neighborhood. The search radius set to 100 meters for each neighborhood,

The data retrieved from Foursquare contained information of venues within a specified distance of the longitude and latitude of the postcodes. The information obtained per venue as follows:

1. Neighborhood
2. Neighborhood Latitude
3. Neighborhood Longitude
4. Venue
5. Name of the venue e.g. the name of a store or restaurant
6. Venue Latitude
7. Venue Longitude
8. Venue Category

D. Folium

Folium is used for mapping Chicago and neighbor hoods. Folium makes it easy to visualize data that's been manipulated in Python on an interactive leaflet map

.

E. Machine Learning K-Means Clustering

K-Means Clustering is an unsupervised machine learning algorithm. K-Means is used for classification and assigning groups.

V. Data Analysis Process

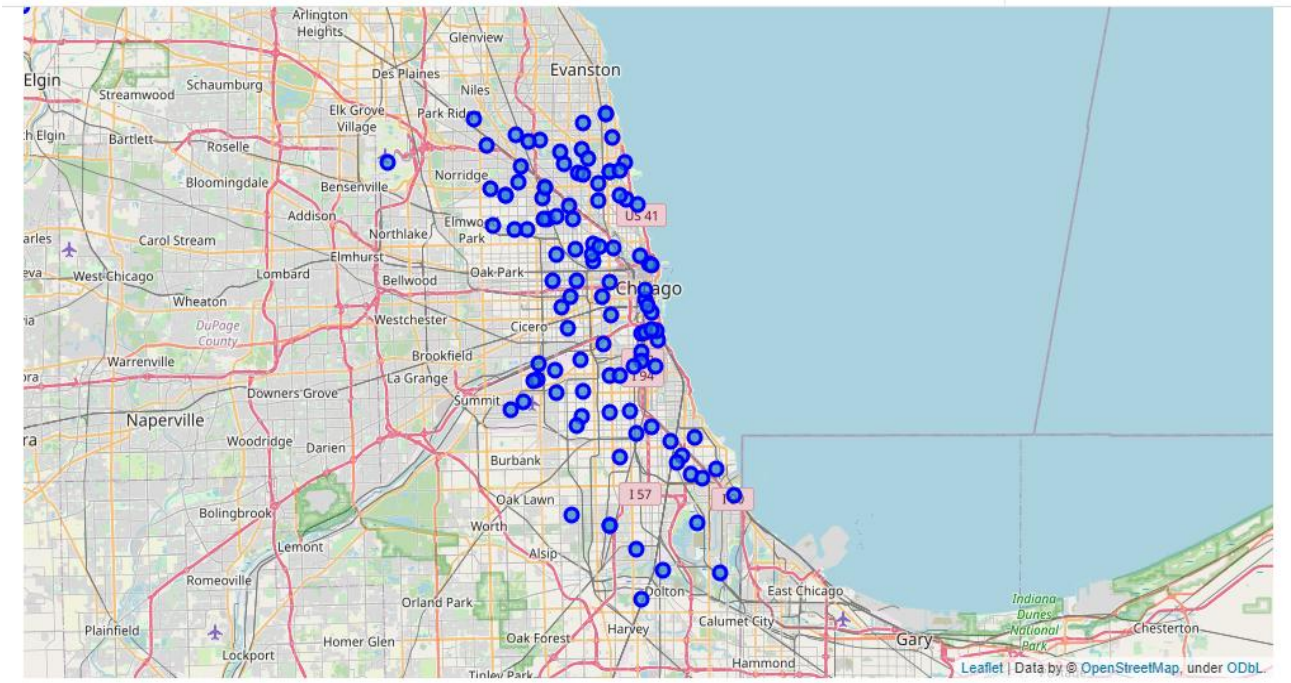
A. Neighborhoods, latitude and longitude

First step get the latitude and longitude of all 77 Neighborhoods in Chicago, IL.

	Community_Area	Neighborhood	Latitude	Longitude
0	Albany Park	Albany Park, IL	41.971937	-87.716174
1	Albany Park	Mayfair, IL	51.511087	-0.147058
2	Albany Park	North Mayfair, IL	33.462919	-111.755610
3	Albany Park	Ravenswood Manor, IL	41.964622	-87.701380
4	Archer Heights	Archer Heights, IL	41.811422	-87.726165
5	Armour Square	Armour Square, IL	41.840231	-87.632986

B. Map Chicago and all Neighborhoods using folium

Map the neighborhoods of Chicago, IL



C. Neighborhoods and venues latitude and longitude

Get the Neighborhood venues, latitude and longitude and create a new dataframe.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Albany Park, IL	41.971937	-87.716174	Chicago Produce	41.970553	-87.716327	Grocery Store
1	Albany Park, IL	41.971937	-87.716174	Cairo Nights Hookah Lounge	41.975776	-87.715547	Hookah Bar
2	Albany Park, IL	41.971937	-87.716174	Nighthawk	41.967974	-87.713415	Cocktail Bar
3	Albany Park, IL	41.971937	-87.716174	Peking Mandarin Resturant	41.968292	-87.715783	Chinese Restaurant
4	Albany Park, IL	41.971937	-87.716174	Popeyes Louisiana Kitchen	41.968756	-87.713019	Fried Chicken Joint
5	Albany Park, IL	41.971937	-87.716174	Hiromi's Oriental Restaurant	41.968144	-87.718719	Karaoke Bar
6	Albany Park, IL	41.971937	-87.716174	Markellos Baking Company	41.968602	-87.716607	Bakery
7	Albany Park, IL	41.971937	-87.716174	Banpojung	41.975707	-87.715609	Korean Restaurant
8	Albany Park, IL	41.971937	-87.716174	Subway	41.968748	-87.712861	Sandwich Place
9	Albany Park, IL	41.971937	-87.716174	T-Mobile	41.968751	-87.713158	Mobile Phone Shop
10	Albany Park, IL	41.971937	-87.716174	Dunkin'	41.968255	-87.712964	Donut Shop

D. Neighborhood and top 10 venues

Create a dataframe of top 10 venues for each Neighborhood.

	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	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Albany Park, IL	Sandwich Place	Grocery Store	Bakery	Mobile Phone Shop	Hookah Bar	Donut Shop	Chinese Restaurant	Latin American Restaurant	Cocktail Bar	Korean Restaurant
1	Altgeld Gardens, IL	Food	Grocery Store	Park	Dry Cleaner	Eastern European Restaurant	Electronics Store	Empanada Restaurant	English Restaurant	Escape Room	Ethiopian Restaurant
2	Andersonville, IL	History Museum	American Restaurant	Travel & Transport	Yoga Studio	Fabric Shop	Eastern European Restaurant	Electronics Store	Empanada Restaurant	English Restaurant	Escape Room
3	Archer Heights, IL	Mexican Restaurant	Grocery Store	Mobile Phone Shop	Bank	Candy Store	Coffee Shop	Big Box Store	Gym / Fitness Center	Gas Station	Sandwich Place
4	Armour Square, IL	Chinese Restaurant	Cosmetics Shop	Italian Restaurant	Grocery Store	Gas Station	Asian Restaurant	Café	Sandwich Place	Hot Dog Joint	Ice Cream Shop
5	Ashburn, IL	Lake	Yoga Studio	Dry Cleaner	Eastern European Restaurant	Electronics Store	Empanada Restaurant	English Restaurant	Escape Room	Ethiopian Restaurant	Event Service

VI. Machine Learning - K-Means Clustering

Clustering Approach: For comparing and exploring neighborhoods, segment and group neighborhoods into clusters to find similar clusters of neighborhoods. Creating data clusters, which is a form of unsupervised machine learning; k-means clustering algorithm.

K-Means Clustering unsupervised machine learning algorithm is used for data analysis. Number of clusters is set to 5 for this analysis.

A. Neighborhood, latitude, longitude, labels, and 10 venues

Community_Area	Neighborhood	Latitude	Longitude	Cluster Labels	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
Albany Park	Albany Park, IL	41.971937	-87.716174	1.0	Sandwich Place	Grocery Store	Bakery	Mobile Phone Shop	Hookah Bar	Donut Shop	Chinese Restaurant	Latin American Restaurant	Cocktail Bar	Korean Restaurant
Albany Park	Mayfair, IL	51.511087	-0.147058	1.0	Hotel	Boutique	French Restaurant	Art Gallery	Clothing Store	Seafood Restaurant	Steakhouse	Lounge	Café	Park
Albany Park	North Mayfair, IL	33.462919	-111.755610	1.0	Pool	Yoga Studio	Fabric Shop	Eastern European Restaurant	Electronics Store	Empanada Restaurant	English Restaurant	Escape Room	Ethiopian Restaurant	Event Service
Albany Park	Ravenswood Manor, IL	41.964622	-87.701380	1.0	Train Station	Playground	Video Game Store	Mexican Restaurant	Convenience Store	Brewery	Museum	Garden	Park	Indoor Play Area
Archer Heights	Archer Heights, IL	41.811422	-87.726165	1.0	Mexican Restaurant	Grocery Store	Mobile Phone Shop	Bank	Candy Store	Coffee Shop	Big Box Store	Gym / Fitness Center	Gas Station	Sandwich Place

B. Cluster 1

Neighborhood and top 10 venues

	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	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
52	Groveland Park, IL	Convenience Store	Music Venue	Falafel Restaurant	Electronics Store	Empanada Restaurant	English Restaurant	Escape Room	Ethiopian Restaurant	Event Service	Event Space
68	Edgewater Beach, IL	Food	Bar	Dumpling Restaurant	Eastern European Restaurant	Electronics Store	Empanada Restaurant	English Restaurant	Escape Room	Ethiopian Restaurant	Event Service
170	North Lawndale, IL	Concert Hall	Athletics & Sports	Food	Convenience Store	Hostel	Dry Cleaner	Eastern European Restaurant	Electronics Store	Empanada Restaurant	English Restaurant
176	Norwood Park West, IL	Convenience Store	Cosmetics Shop	Falafel Restaurant	Electronics Store	Empanada Restaurant	English Restaurant	Escape Room	Ethiopian Restaurant	Event Service	Event Space
177	Old Norwood, IL	Food	Spa	Convention Center	Eastern European Restaurant	Empanada Restaurant	English Restaurant	Escape Room	Ethiopian Restaurant	Event Service	Event Space

C. Cluster 2

Neighborhood and top 10 venues

	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	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Albany Park, IL	Sandwich Place	Grocery Store	Bakery	Mobile Phone Shop	Hookah Bar	Donut Shop	Chinese Restaurant	Latin American Restaurant	Cocktail Bar	Korean Restaurant
1	Mayfair, IL	Hotel	Boutique	French Restaurant	Art Gallery	Clothing Store	Seafood Restaurant	Steakhouse	Lounge	Café	Park
2	North Mayfair, IL	Pool	Yoga Studio	Fabric Shop	Eastern European Restaurant	Electronics Store	Empanada Restaurant	English Restaurant	Escape Room	Ethiopian Restaurant	Event Service
3	Ravenswood Manor, IL	Train Station	Playground	Video Game Store	Mexican Restaurant	Convenience Store	Brewery	Museum	Garden	Park	Indoor Play Area
4	Archer Heights, IL	Mexican Restaurant	Grocery Store	Mobile Phone Shop	Bank	Candy Store	Coffee Shop	Big Box Store	Gym / Fitness Center	Gas Station	Sandwich Place
...	...	...	...	...	...	...	...	...	...	...	...
235	Ukrainian Village, IL	Coffee Shop	Pub	Dive Bar	Ukrainian Restaurant	Eastern European Restaurant	Bakery	Grocery Store	Bar	Art Museum	Deli / Bodega
236	West Town, IL	Dive Bar	Grocery Store	Mexican Restaurant	Bar	Pet Store	Toy / Game Store	Record Shop	Chinese Restaurant	Bank	Taco Place

D. Cluster 3

Neighborhood and top 10 venues

	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	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
37	Burnside, IL	Hotel	Yoga Studio	Fabric Shop	Eastern European Restaurant	Electronics Store	Empanada Restaurant	English Restaurant	Escape Room	Ethiopian Restaurant	Event Service
70	Lakewood / Balmoral, IL	ATM	Hotel	Yoga Studio	Fabric Shop	Eastern European Restaurant	Electronics Store	Empanada Restaurant	English Restaurant	Escape Room	Ethiopian Restaurant

E. Cluster 4

Neighborhood and top 10 venues

	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	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
136	Beverly Woods, IL	Coffee Shop	Yoga Studio	Fabric Shop	Eastern European Restaurant	Electronics Store	Empanada Restaurant	English Restaurant	Escape Room	Ethiopian Restaurant	Event Service
171	Hollywood Park, IL	Coffee Shop	Yoga Studio	Fabric Shop	Eastern European Restaurant	Electronics Store	Empanada Restaurant	English Restaurant	Escape Room	Ethiopian Restaurant	Event Service

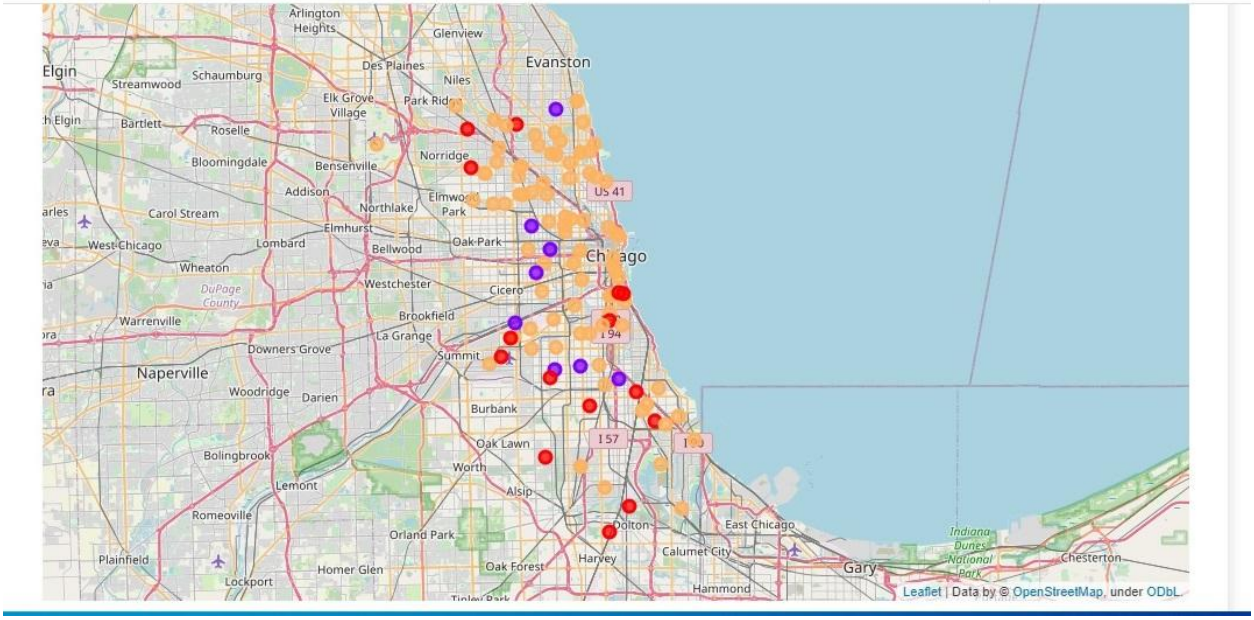
F. Cluster 5

Neighborhood and top 10 venues

	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	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
15	Auburn Gresham, IL	Park	Food	Discount Store	Basketball Court	Yoga Studio	Exhibit	Empanada Restaurant	English Restaurant	Escape Room	Ethiopian Restaurant
21	West Humboldt Park, IL	Park	Lake	Baseball Field	Café	Plaza	Food Truck	History Museum	Museum	Beach	Yoga Studio
38	Calumet Heights, IL	Bus Station	Gym / Fitness Center	Park	Deli / Bodega	Yoga Studio	Electronics Store	Empanada Restaurant	English Restaurant	Escape Room	Ethiopian Restaurant
46	Marquette Park, IL	Park	Soccer Field	Liquor Store	Yoga Studio	Exhibit	Electronics Store	Empanada Restaurant	English Restaurant	Escape Room	Ethiopian Restaurant
49	Clearing West, IL	Airport Terminal	Airport	Park	Yoga Studio	Fabric Shop	Electronics Store	Empanada Restaurant	English Restaurant	Escape Room	Ethiopian Restaurant
54	Prairie Shores, IL	Park	Bus Station	Train Station	Gym / Fitness Center	Shopping Mall	Event Space	Eastern European Restaurant	Electronics Store	Empanada Restaurant	English Restaurant
55	South Commons, IL	Deli / Bodega	Bus Station	Park	Gym / Fitness Center	Rental Car Location	Eastern European Restaurant	Electronics Store	Empanada Restaurant	English Restaurant	Escape Room

VII. Results

A. Final map of Chicago and Clusters (neighborhoods)



VIII. Final Discussion

Question we tried to address: Objective of this project is to find a city with population and many diverse venues for locating new Latin American restaurants. Chicago the 3rd most populace city in USA was selected for this study. Chicago also receives 58 million tourists annually. Data is analyzed by neighborhoods and venues.

Machine learning K-Means clustering algorithm is used for clustering neighborhood and venues. The Number of clusters is set to 5 for this study.

IX. Conclusion

Chicago has a large tourist and domestic population, there are many venues and attractions in and around Chicago. It's also a Midwest transportation hub. Chicago Airports are well connected to all International cities. Chicago has the resources to support new restaurants.

X. Additional studies

The Covid-19 pandemic has impacted the global hospitality industry. Additional study may required to analyze the impact on Chicago restaurants. Food and Beverage industry forecasts to recover in the 3rd quarter of 2021.

XI. References

A. Data sources:

- Wikipedia – list of neighborhoods by community area Chicago
- City of Chicago portal
- geopy.geocoders for latitude and longitude
- Foursquare api for list of venues

B. Libraries used for development

- Pandas: For creating and manipulating data frames.
- Numpy: Powerful n-dimensional arrays. Numerical computing tools.
- Folium: folium makes it easy to visualize data that's been manipulated in Python on an interactive leaflet map. It enables both the binding of data to a map for choropleth visualizations as well as passing rich vector/raster/HTML visualizations as markers on the map..
- Scikit Learn: Simple and efficient tools for predictive data analysis, used for k-means clustering.
- GeoPy Geocoder: for retrieving Location Data.
- Matplotlib: Python Plotting Module.