Pick a suitable location for a restaurant in New York state.

Kanna Venkata Raghavendra Karthik

January 15, 2021

1. Introduction

1.1 Background

People love to eat good food and different cuisine restaurants have provided them to taste various delicacies. Restaurants have always played a crucial role in business, social life of the society. Many personal and professional events happen at restaurants. But a restaurant's investment and maintenance costs would be high. So every restaurant owner wants to get a high return on their investments. The taste of food, ambience and the location of the restaurant are few major contributing factors for a successful restaurant.

1.2 Problem

Consider the following hypothetical scenario :

‘Resto-grand’ a successful multi-chained restaurant company has its brands around Europe, India, California. They wanted to start their restaurants in New York state and required some suggestions on where to open them.

The aim of the project is to find suitable locations for the restaurants in New York state.

1.3 Audience

Here Resto-grand's business owners to whom we are preparing the report is the target audience. But this report can be used by other business people who want to find a suitable location to open a restaurant in New York state.

2. Data

To tackle the above mentioned problem, we need to have the dataset that contains :

- Cities present in New York State.
- Latitude and Longitude locations of the cities of New York state.
- Top venues present in each city of New York state.

2.1 Data Sources

- List of cities present in New York state can be obtained from wikipedia's page : (https://en.wikipedia.org/wiki/List_of_cities_in_New_York)
- Latitude and Longitude locations of them can be found using python's geopy package.
- We will use foursquare.com's API to find the most common venues present at given geographical coordinates.

2.2 Description of data

The below output is obtained after cleaning the data acquired from wikipedia's page and using python's geopy package. Each row represents a city in New York state, its county, population and geographical coordinates.

```
df.head()
```

Out[11]:

	City	County	Population	Latitude	Longitude
0	Albany	Albany	97856	42.651167	-73.754968
1	Amsterdam	Montgomery	18620	42.953685	-74.219581
2	Auburn	Cayuga	27687	42.932020	-76.567203
3	Batavia	Genesee	15465	42.998014	-78.187551
4	Beacon	Dutchess	15541	41.504879	-73.969682

We will use foursquare's API to find the most common venues present around the 30 km radius of each city of New York.

3. Methodology

The first step is to collect data about cities of New York state from the wikipedia page at the following link (https://en.wikipedia.org/wiki/List_of_cities_in_New_York). There are 62 cities in New York state.

	City	County	Population [1][2](2010 census)	Incorporationdate	FIPS code(subdivision)	FIPS code (place)
0	Albany	Albany	97856	1686	3600101000	3601000
1	Amsterdam	Montgomery	18620	1830	3605702066	3602066
2	Auburn	Cayuga	27687	1848	3601103078	3603078
3	Batavia	Genesee	15465	1915	3603704715	3604715
4	Beacon	Dutchess	15541	1913	3602705100	3605100
...	...	...	...	...	...	...
57	Utica	Oneida	62235	1832	3606576540	3676540
58	Watertown	Jefferson	27023	1869	3604578608	3678608
59	Watervliet	Albany	10254	1896	3600178674	3678674
60	White Plains	Westchester	56853	1916	3611981677	3681677
61	Yonkers	Westchester	195976	1872	3611984000	3684000

62 rows × 6 columns

3.1 Data cleaning and exploration

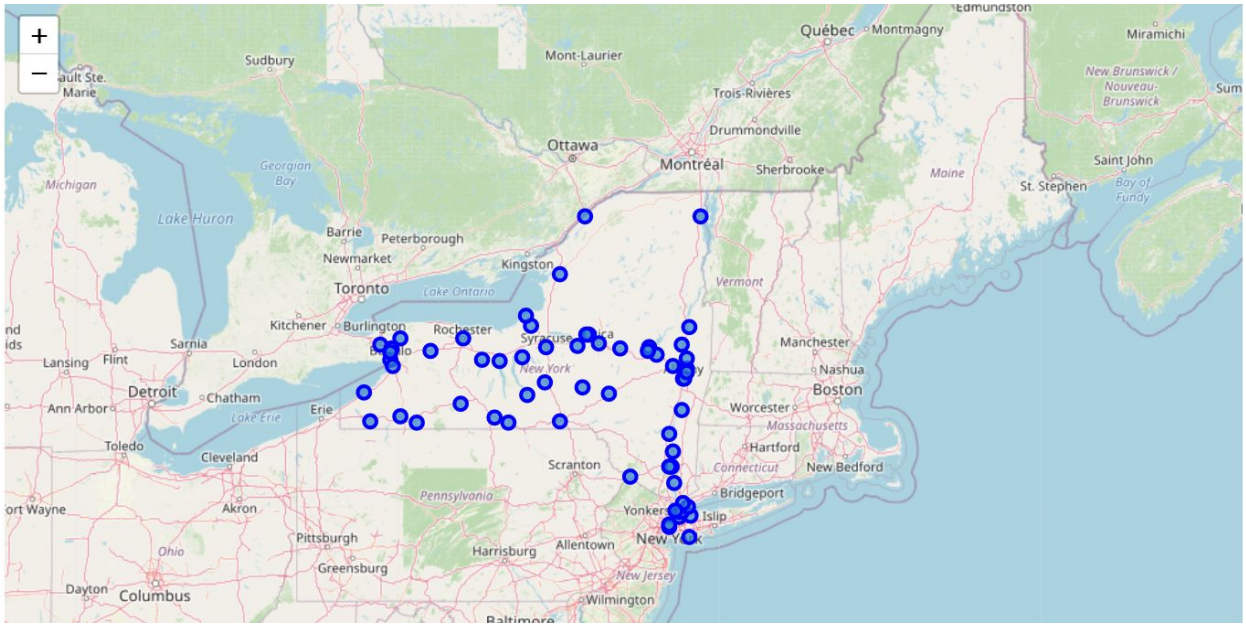
Then we tidy up the data to contain city, county, population columns. We will use this data to find the geographical coordinates of each city using python's geopy package and the top 5 rows of data is as follows -

```
df.head()
```

Out[11]:

	City	County	Population	Latitude	Longitude
0	Albany	Albany	97856	42.651167	-73.754968
1	Amsterdam	Montgomery	18620	42.953685	-74.219581
2	Auburn	Cayuga	27687	42.932020	-76.567203
3	Batavia	Genesee	15465	42.998014	-78.187551
4	Beacon	Dutchess	15541	41.504879	-73.969682

Using folium, a map rendering software we shall just see the location of cities on the world map. The blue points represent our 62 cities in New York state.



Later we use foursquare's api to find the list of venues around a city ('Amsterdam') to see where we get necessary venues in that city. We will restrict the limit of venues obtained to 100 and which are present around 30 km radius of the city. Once we receive the response in json format, we will tidy it up and represent it in a dataset having venue's name, its category and its geographical coordinates.

The top 5 rows of the dataset is shown below -

Out[23]:

	name	categories	lat	lng
0	Bosco's Restaurant & Bar	Italian Restaurant	42.945115	-74.200738
1	Recovery Sports Grill	American Restaurant	42.958699	-74.185112
2	Full House Buffet	Chinese Restaurant	42.968892	-74.184887
3	Ruby Tuesday	American Restaurant	42.966054	-74.186193
4	Moe's Southwest Grill	Mexican Restaurant	42.964819	-74.187859

```
In [24]: print('{} venues were returned by Foursquare.'.format(nearby_venues.shape[0]))
```

```
100 venues were returned by Foursquare.
```

```
In [25]: print ('{} unique categories in Amsterdam.'.format(nearby_venues['categories'].value_counts().shape[0]))
```

```
44 unique categories in Amsterdam.
```

As we can see, we got 100 venues with 44 unique categories in Amsterdam. It contains restaurants as well.

We generalize this and find the venues around 30km in each city of New York state and create a dataframe which consists of venue's name,category,geographical location along with to which city it belongs and city's geographical location. The top 5 rows are shown below :

```
In [51]: newyork_venues.head()
```

Out[51]:

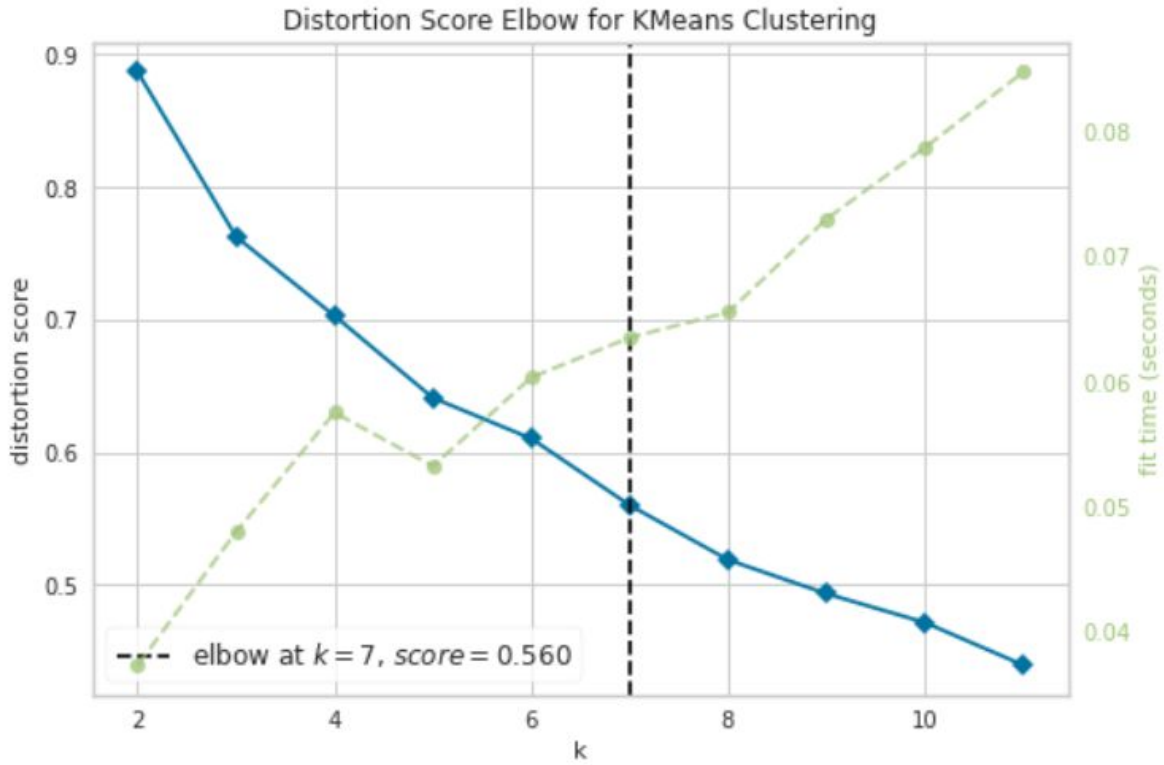
	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Albany	42.651167	-73.754968	Renaissance Albany Hotel	42.650625	-73.755687	Hotel
1	Albany	42.651167	-73.754968	Iron Gate Cafe	42.655974	-73.762504	Café
2	Albany	42.651167	-73.754968	City Beer Hall	42.649660	-73.754787	Pub
3	Albany	42.651167	-73.754968	Palace Theatre	42.654736	-73.750192	Theater
4	Albany	42.651167	-73.754968	The Olde English Pub & Pantry	42.653958	-73.748563	Pub

Later we use the one-hot encoding technique on the 'Venue Category' column to find the unique venues present in the city / neighborhood and then calculate the frequencies of venues present in each city. Finally we obtain a dataset each row containing the city name and the top 10 most common venues in that city. A portion of it can be seen below -

	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	Coffee Shop	Ice Cream Shop	Grocery Store	Pub	Brewery	Pizza Place	Sandwich Place	Bakery	Burger Joint	American Restaurant
1	Amsterdam	Convenience Store	American Restaurant	Italian Restaurant	Ice Cream Shop	Mexican Restaurant	Restaurant	Pub	Farm	Donut Shop	Pizza Place
2	Auburn	American Restaurant	Restaurant	Café	Ice Cream Shop	Bar	Pub	Park	Pizza Place	Winery	Farm
3	Batavia	American Restaurant	Pizza Place	Ice Cream Shop	Diner	Bar	Deli / Bodega	Donut Shop	Supermarket	Coffee Shop	Department Store
4	Beacon	Trail	Bar	Ice Cream Shop	American Restaurant	Bookstore	Brewery	Liquor Store	Coffee Shop	Bakery	State / Provincial Park
5	Binghamton	American Restaurant	Italian Restaurant	Pizza Place	Bar	Brewery	Diner	Hardware Store	Coffee Shop	Park	Pharmacy
6	Buffalo	Park	Supermarket	Grocery Store	Coffee Shop	Café	American Restaurant	Wine Shop	Pizza Place	Burger Joint	Ice Cream Shop

3.2 Machine Learning method

As we have data available for our machine learning model we shall proceed this step by using an unsupervised algorithm as we don't have any labelled data. We use the KMeans Clustering algorithm that arranges the similar cities into a cluster and dissimilar cities in different clusters. For this algorithm we need to pass a parameter called number of clusters. The optimal value of this can be found by using the KElbowVisualizer module in the yellowbrick package.



From the above graph, we can see the value of k = 7. So, we will input this value as the number of clusters into KMeans algorithm and fit the dataset containing the top 10 most common venues in each city. When done, we add the cluster labels to the dataset. The first 5 rows of the table are shown below -

```
In [71]: neighborhoods_venues_sorted_w_clusters.head()
```

Out[71]:

	Cluster Labels	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	4	Albany	Coffee Shop	Ice Cream Shop	Grocery Store	Pub	Brewery	Pizza Place	Sandwich Place	Bakery	Burger Joint	American Restaurant
1	5	Amsterdam	Convenience Store	American Restaurant	Italian Restaurant	Ice Cream Shop	Mexican Restaurant	Restaurant	Pub	Farm	Donut Shop	Pizza Place
2	1	Auburn	American Restaurant	Restaurant	Café	Ice Cream Shop	Bar	Pub	Park	Pizza Place	Winery	Farm
3	2	Batavia	American Restaurant	Pizza Place	Ice Cream Shop	Diner	Bar	Deli / Bodega	Donut Shop	Supermarket	Coffee Shop	Department Store
4	1	Beacon	Trail	Bar	Ice Cream Shop	American Restaurant	Bookstore	Brewery	Liquor Store	Coffee Shop	Bakery	State / Provincial Park

4. Results

We shall merge the above dataset with the dataset which has the city's geographical details. We can slice the data into 7 different clusters as per the 7 cluster labels(0,1,2,3,4,5,6).

4.1 Cluster - 0

Cluster 0 contains cities where parks, bakeries, gyms are priority in the top most venues.

	City	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
29	Middletown	40.839722	-73.831667	0	Park	Bakery	Gym	Theater	Plaza	Gourmet Shop	Pizza Place	Museum	Cocktail Bar	Gym / Fitness Center
30	Mount Vernon	40.912582	-73.837079	0	Park	Bakery	Gym	Theater	Pizza Place	Museum	Plaza	Gourmet Shop	Butcher	Garden
31	New Rochelle	40.911539	-73.782636	0	Park	Bakery	Gym	Boxing Gym	Theater	Garden	Exhibit	Museum	Art Museum	Pizza Place
32	New York	40.712728	-74.006015	0	Park	Ice Cream Shop	Bookstore	Gourmet Shop	Bakery	Movie Theater	Scenic Lookout	Pier	Cycle Studio	Thai Restaurant
56	Troy	40.740256	-74.005535	0	Park	Gourmet Shop	Bakery	Bookstore	Gym	Art Gallery	Theater	Seafood Restaurant	Taco Place	Scenic Lookout
61	Yonkers	40.931210	-73.898747	0	Park	Bakery	Gym	Theater	Gourmet Shop	Plaza	Museum	Pizza Place	Performing Arts Venue	State / Provincial Park

4.2 Cluster - 1

Cluster 1 contains the cities where cafes, bars, restaurants, ice cream shops are priority in the top most venues.

	City	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
2	Auburn	42.932020	-76.567203	1	American Restaurant	Restaurant	Café	Ice Cream Shop	Bar	Pub	Park	Pizza Place	Winery	Farm
4	Beacon	41.504879	-73.969682	1	Trail	Bar	Ice Cream Shop	American Restaurant	Bookstore	Brewery	Liquor Store	Coffee Shop	Bakery	State / Provincial Park
6	Buffalo	42.886717	-78.878392	1	Park	Supermarket	Grocery Store	Coffee Shop	Café	American Restaurant	Wine Shop	Pizza Place	Burger Joint	Ice Cream Shop
14	Geneva	42.869027	-76.978612	1	Vineyard	Winery	American Restaurant	Bar	Brewery	Café	Ice Cream Shop	Restaurant	Coffee Shop	Pizza Place
19	Hudson	42.252865	-73.790959	1	American Restaurant	Café	Pizza Place	Brewery	Diner	Farm	Restaurant	Trail	Bar	Furniture / Home Store
20	Ithaca	42.439604	-76.496802	1	Ice Cream Shop	American Restaurant	Park	State / Provincial Park	Coffee Shop	Brewery	Sandwich Place	Bakery	Trail	Café
23	Kingston	41.928781	-74.002370	1	Café	American Restaurant	Pizza Place	Coffee Shop	Trail	Ice Cream Shop	Restaurant	Park	Bakery	Cheese Shop
24	Lackawanna	42.819090	-78.833717	1	Supermarket	American Restaurant	Café	Coffee Shop	Ice Cream Shop	Gym	Beer Store	Wine Shop	Park	Grocery Store
26	Lockport	43.168710	-78.696567	1	Park	Coffee Shop	Hot Dog Joint	Café	Mexican Restaurant	Supermarket	Pizza Place	American Restaurant	Wine Shop	Bar
33	Newburgh	41.503427	-74.010418	1	Trail	Bar	American Restaurant	Bookstore	Ice Cream Shop	Brewery	Pizza Place	Liquor Store	Coffee Shop	Bakery
34	Niagara Falls	43.084360	-79.061469	1	Park	Coffee Shop	Italian Restaurant	American Restaurant	Hotel	Scenic Lookout	State / Provincial Park	Waterfall	Ice Cream Shop	Restaurant
35	North Tonawanda	43.040774	-78.866053	1	Supermarket	Grocery Store	Café	Park	American Restaurant	Pizza Place	Burger Joint	Greek Restaurant	Waterfall	Deli / Bodega
42	Peekskill	41.289811	-73.920492	1	Trail	Ice Cream Shop	Scenic Lookout	Park	Pizza Place	State / Provincial Park	American Restaurant	Brewery	Restaurant	Farm
44	Port Jervis	41.375094	-74.692663	1	Ice Cream Shop	American Restaurant	Bagel Shop	Brewery	Golf Course	Pizza Place	Deli / Bodega	Diner	Mexican Restaurant	Campground
45	Poughkeepsie	41.706554	-73.928367	1	Trail	Pizza Place	American Restaurant	Convenience Store	Brewery	Bakery	New American Restaurant	Diner	Coffee Shop	Café
55	Tonawanda	42.993582	-78.881904	1	American Restaurant	Park	Supermarket	Pizza Place	Café	Coffee Shop	Grocery Store	Greek Restaurant	Gourmet Shop	Deli / Bodega

4.3 Cluster - 2

Cluster 2 contains the cities where many restaurants, pizza places and bars are priority in the top most venues.

	City	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
3	Batavia	42.998014	-78.187551	2	American Restaurant	Pizza Place	Ice Cream Shop	Diner	Bar	Deli / Bodega	Donut Shop	Supermarket	Coffee Shop	Department Store
5	Binghamton	42.096968	-75.914341	2	American Restaurant	Italian Restaurant	Pizza Place	Bar	Brewery	Diner	Hardware Store	Coffee Shop	Park	Pharmacy
9	Corning	42.143526	-77.054341	2	American Restaurant	Italian Restaurant	Pizza Place	Park	Ice Cream Shop	Café	Discount Store	Supermarket	Trail	Restaurant
11	Dunkirk	42.479502	-79.333932	2	American Restaurant	Pharmacy	Discount Store	Bar	Ice Cream Shop	Beach	Pizza Place	Convenience Store	Coffee Shop	Liquor Store
12	Elmira	42.089796	-76.807734	2	American Restaurant	Pizza Place	Italian Restaurant	Park	Sandwich Place	Bakery	Discount Store	Supermarket	Coffee Shop	Ice Cream Shop
13	Fulton	43.322846	-76.417159	2	Pizza Place	American Restaurant	Coffee Shop	Supermarket	Bar	Mexican Restaurant	Ice Cream Shop	Italian Restaurant	Café	Park
39	Oneida	43.214405	-75.403915	2	American Restaurant	Pizza Place	Italian Restaurant	Bar	Brewery	Japanese Restaurant	Café	Donut Shop	Farmers Market	Supermarket
40	Oneonta	42.453492	-75.062953	2	Italian Restaurant	American Restaurant	Café	Pizza Place	Brewery	Pharmacy	Bakery	Hotel	Discount Store	Coffee Shop
41	Oswego	43.454728	-76.509597	2	American Restaurant	Pizza Place	Sandwich Place	Bar	Coffee Shop	Donut Shop	Diner	Mexican Restaurant	Ice Cream Shop	Discount Store
43	Plattsburgh	44.692820	-73.455620	2	American Restaurant	Pizza Place	Sandwich Place	Gas Station	Restaurant	Diner	Bagel Shop	Supermarket	Cosmetics Shop	Bar
48	Rome	43.212847	-75.455730	2	American Restaurant	Pizza Place	Italian Restaurant	Bar	Donut Shop	Brewery	Gas Station	Farmers Market	Coffee Shop	Fast Food Restaurant
50	Salamanca	42.157841	-78.715031	2	Sandwich Place	Bar	Pizza Place	American Restaurant	Café	Discount Store	Ski Area	Brewery	Coffee Shop	Steakhouse
53	Sherrill	43.073904	-75.598297	2	Pizza Place	American Restaurant	Italian Restaurant	Café	Bar	Hotel	Farmers Market	Grocery Store	Sandwich Place	Golf Course
54	Syracuse	43.048122	-76.147424	2	Pizza Place	American Restaurant	Supermarket	Grocery Store	Coffee Shop	Italian Restaurant	Diner	Plaza	Park	Department Store
57	Utica	43.100903	-75.232664	2	Pizza Place	Italian Restaurant	American Restaurant	Bar	Brewery	Department Store	Deli / Bodega	Grocery Store	Diner	Farmers Market

4.4 Cluster - 3

Cluster 3 contains the cities where discount stores, sandwich and pizza places are priority in the top most venues.

	City	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
18	Hornell	42.327848	-77.661102	3	Discount Store	Sandwich Place	Pizza Place	Pharmacy	Grocery Store	Bar	Convenience Store	Campground	Café	American Restaurant
36	Norwich	42.531184	-75.523515	3	Discount Store	Pizza Place	Sandwich Place	Grocery Store	Golf Course	Donut Shop	Italian Restaurant	Pharmacy	Post Office	Rest Area
38	Olean	42.077478	-78.429861	3	Sandwich Place	Pizza Place	Discount Store	Bar	Café	American Restaurant	Coffee Shop	Steakhouse	Brewery	Convenience Store

4.5 Cluster - 4

Cluster 4 contains the cities where coffee shops, pizza places , ice cream shops are priority in the top most venues.

	City	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
0	Albany	42.651167	-73.754968	4	Coffee Shop	Ice Cream Shop	Grocery Store	Pub	Brewery	Pizza Place	Sandwich Place	Bakery	Burger Joint	American Restaurant
7	Canandaigua	42.885858	-77.279976	4	Ice Cream Shop	Bar	Brewery	Coffee Shop	Pizza Place	Italian Restaurant	Park	Pub	Breakfast Spot	Bakery
8	Cohoes	42.774245	-73.700119	4	Coffee Shop	Pizza Place	Grocery Store	Brewery	Pub	Convenience Store	Steakhouse	Ice Cream Shop	Bakery	Mexican Restaurant
10	Cortland	42.601181	-76.180484	4	Coffee Shop	Pizza Place	Bar	American Restaurant	Ice Cream Shop	Diner	Ski Area	Liquor Store	Farmers Market	Sports Bar
16	Glens Falls	43.309941	-73.644447	4	American Restaurant	Coffee Shop	Hotel	Convenience Store	Ice Cream Shop	Brewery	Diner	Café	Supermarket	Pizza Place
21	Jamestown	42.097002	-79.235326	4	Coffee Shop	Bar	American Restaurant	Discount Store	Ice Cream Shop	Breakfast Spot	Café	Seafood Restaurant	Sandwich Place	Hotel
28	Mechanicville	42.902855	-73.687341	4	Coffee Shop	American Restaurant	Convenience Store	Pub	Bakery	Hotel	Italian Restaurant	Brewery	Mexican Restaurant	Tea Room
37	Ogdensburg	44.694285	-75.486374	4	Fast Food Restaurant	Coffee Shop	Pizza Place	Grocery Store	Discount Store	Restaurant	Sandwich Place	Pub	Bank	Gas Station
46	Rensselaer	42.642579	-73.742898	4	Coffee Shop	Ice Cream Shop	Pizza Place	Grocery Store	Pub	Brewery	Sandwich Place	Bakery	Burger Joint	Bookstore
47	Rochester	43.157285	-77.615214	4	Park	Pub	Bakery	Coffee Shop	Sandwich Place	Liquor Store	Mexican Restaurant	Café	Brewery	Supermarket
51	Saratoga Springs	43.082179	-73.785392	4	American Restaurant	Convenience Store	Coffee Shop	Hotel	Brewery	Italian Restaurant	Food & Drink Shop	Tea Room	Mexican Restaurant	Café
52	Schenectady	42.814243	-73.939569	4	Mexican Restaurant	Convenience Store	Pub	Bakery	Ice Cream Shop	Gym	Coffee Shop	Brewery	Hotel	Italian Restaurant
59	Watervliet	42.730078	-73.701230	4	Coffee Shop	Brewery	Ice Cream Shop	Convenience Store	Pizza Place	Pub	Grocery Store	Steakhouse	American Restaurant	Bakery

4.6 Cluster - 5

Cluster 5 contains the cities where convenience stores, sandwich places and restaurants are priority in the top most venues.

	City	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
1	Amsterdam	42.953685	-74.219581	5	Convenience Store	American Restaurant	Italian Restaurant	Ice Cream Shop	Mexican Restaurant	Restaurant	Pub	Farm	Donut Shop	Pizza Place
17	Gloversville	43.050509	-74.345724	5	Convenience Store	Sandwich Place	American Restaurant	Pharmacy	Pizza Place	Gas Station	Discount Store	Italian Restaurant	Ice Cream Shop	Restaurant
22	Johnstown	43.006869	-74.367644	5	Convenience Store	Sandwich Place	American Restaurant	Discount Store	Pharmacy	Gas Station	Italian Restaurant	Pizza Place	Ice Cream Shop	Supermarket
25	Little Falls	43.043404	-74.859596	5	Discount Store	Italian Restaurant	Gas Station	Convenience Store	American Restaurant	Fast Food Restaurant	Diner	Sandwich Place	Rest Area	Supermarket
58	Watertown	43.974784	-75.910757	5	Italian Restaurant	American Restaurant	Discount Store	Convenience Store	Grocery Store	Sandwich Place	Donut Shop	Pizza Place	Supermarket	Seafood Restaurant

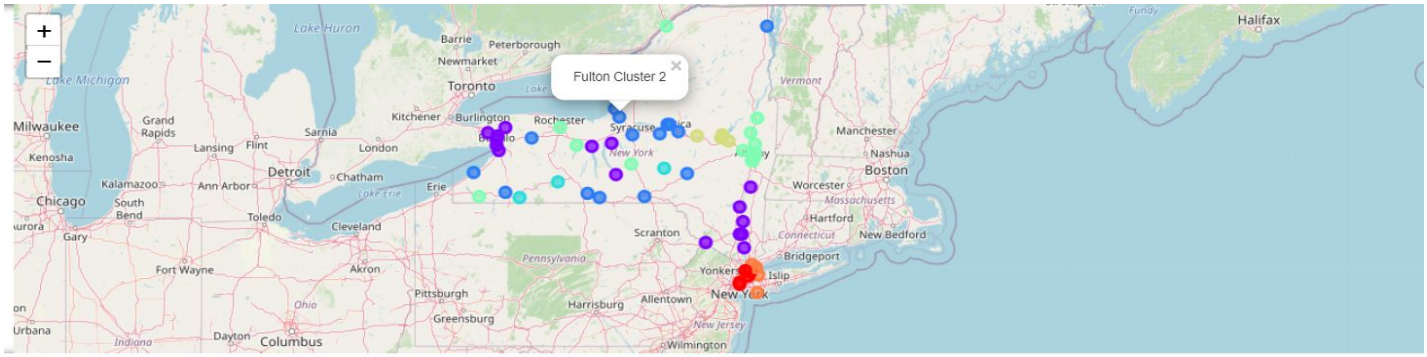
4.7 Cluster - 6

Cluster 6 contains the cities where pizza places and bakeries are priority in the top most venues.

	City	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
15	Glen Cove	40.862755	-73.633609	6	Pizza Place	Deli / Bodega	Bakery	Park	Italian Restaurant	Bagel Shop	Boxing Gym	Grocery Store	American Restaurant	Gym / Fitness Center
27	Long Beach	40.588512	-73.657861	6	Beach	Pizza Place	Bakery	Grocery Store	Bagel Shop	Park	Surf Spot	Deli / Bodega	Vegetarian / Vegan Restaurant	Boxing Gym
49	Rye	40.980821	-73.684294	6	Pizza Place	Park	Grocery Store	Italian Restaurant	Trail	American Restaurant	Gym / Fitness Center	Deli / Bodega	State / Provincial Park	Mexican Restaurant
60	White Plains	41.033986	-73.762910	6	Pizza Place	Grocery Store	Park	Italian Restaurant	Trail	Deli / Bodega	State / Provincial Park	Gym / Fitness Center	Brewery	Castle

5. Discussions and recommendations

Based on what we have learned about the clusters, we can advise the Resto-grand's business owners to consider the cities from Cluster-2 as the most suitable location for their restaurants. These are the cities that are well represented by the restaurants, other bar and pizza places. This shows there is very high demand in the locality for the restaurant and in turn can be successful provided proper ambience and taste of the food. All the blue colored circles represent this cluster.



6. Conclusion

In this project we discussed the process of solving a hypothetical real world scenario of finding suitable locations for opening a restaurant in New York state. The analysis is done mainly using the python and python libraries such as pandas, scikit-learn, folium, yellowbrick to name a few. We made use of the foursquare api to find the popular venues and its categories present in the city within a 30km radius. KMeans unsupervised machine learning algorithm was used to group these cities into clusters after knowing the top most venue category in each city.

7. References

The Jupyter Notebook used for this project can be found on GitHub.
https://github.com/kvr-karthik/Coursera_Capstone/blob/master/IBM%20Capstone%20Final%20Assignment.ipynb