

# **The Battle of Neighborhood**

## **-New Restaurant in Capital Region, NY**

### **Introduction**

ALBANY — The Capital Region is an oasis of population growth in upstate New York since 2010, according to the most recent estimates by the U.S. Census Bureau. The high-tech, health care and higher education sectors are feeding the Capital Region's expansion, according to local experts.

"The only area outside of New York City that has job growth is the Capital Region," said Mark Eagan, CEO of the Capital Region Chamber. "Our private sector job growth is above all the other metropolitan areas in the New York." This growth placed the Capital Region in the ranks of most counties across the nation that saw population growth, as foreign migration into our area offset the loss of those who moved elsewhere.

Given the economic and population growth in Albany, there's good opportunity to investigate new business, such as opening new restaurant. This project is trying to analyze the neighborhood data in Albany region and identifying candidate spots for new restaurant. We plan to identify locations where the place is close to popular venues, without too many existing competitions, and in a safe region.

### **Source of the data**

The criminal incident data for Albany NY is available in government website:  
<https://data.albanyny.gov/Public-Safety/APD-Arrests-Dataset-by-Neighborhood/7y34-47cz>

The location data for top venues in Albany can be found in foursquare location data

### **Methodology**

First, raw data of criminal incidents are downloaded and loaded in pandas dataframe. We will clean up the data such that only those major criminal types are kept for easy statistic analysis.

The criminal incidents are then plotted using folium map with labels.

To obtain the criminal frequency in different areas, a mark cluster object is defined in the dataframe, followed by looping through the dataframe and add each data point to the mark cluster, the incident rate for each area are identified.

After that, geolocator is used to identify the latitude and longitude of Albany, NY. Foursquare location data is used to identify the venues nearby the address.

Further, by analyzing and grouping the category items for those venues, the venues are separated into two groups: restaurants and non-restaurants

That information, together with criminal incident rate, are plotted together using folium map, and best locations for candidate restaurant can be found from the map.

## Results

First, criminal record are downloaded from government website as shown below:

Criminal Record downloaded from government Website

```
In [47]: df_incidents = pd.read_csv('https://data.albanyny.gov/api/views/7y34-47cz/rows.csv?accessType=DOWNLOAD')
print('Dataset downloaded and read into a pandas dataframe!')
```

Dataset downloaded and read into a pandas dataframe!

```
In [48]: df_incidents.head()
```

Out[48]:

	ArrestID	Incident	Arrest Date	Arrest Time	Age	Sex	Race	Ethnicity	Law	Crime LongID	Code Description	Attempt	Arrest type	Last 30 Days	City	State	Neighbo Assor
0	APD*190002373	19067785.0	09/27/2019 12:00:00 AM	11:45	50	F	O		H PL	PL 165.05 03 AM3	USE VEH BY TEMP CUSTD- NO CNSNT	C	Arrest Warrant	False	Albany	NY	Park
1	APD*190002376	19137995.0	09/27/2019 12:00:00 AM	20:59	48	F	B		N PL	PL 195.05 AM2	OBSTRUCT GOVERNMENTL ADMIN 2ND	C	Crime in Progress	False	Albany	NY	Dov
2	APD*190002378	19138000.0	09/27/2019 12:00:00 AM	21:23	29	M	I		N PL	PL 230.04 AM3	PATRONIZE PROSTITUTE 3RD	C	Crime in Progress	False	Albany	NY	Beve
3	APD*190002594	19153449.0	10/27/2019 12:00:00 AM	20:00	19	M	B		N PL	PL 265.03 03 CF2	CPW-2ND: LOADED FIREARM	C	Crime in Progress	True	Albany	NY	W
4	APD*190002375	19137884.0	09/27/2019 12:00:00 AM	17:59	59	M	B		N PL	PL 240.26 01 OV2	HARRASSMENT 2ND- PHY CONTACT	C	Complaint	False	Albany	NY	Dov

After that, manipulating the data to extract X, Y coordinate for major criminal types:

## Data frame manipulations

```
In [56]: df_incidents[['P','X','Y']] = df_incidents.NeighborhoodXY.str.split(" ",expand=True,)
```

```
In [57]: df_incidents['X'] = df_incidents['X'].str[1:]
df_incidents['Y'] = df_incidents['Y'].str[:-1]
df_incidents = df_incidents.drop(['P'], axis=1)
```

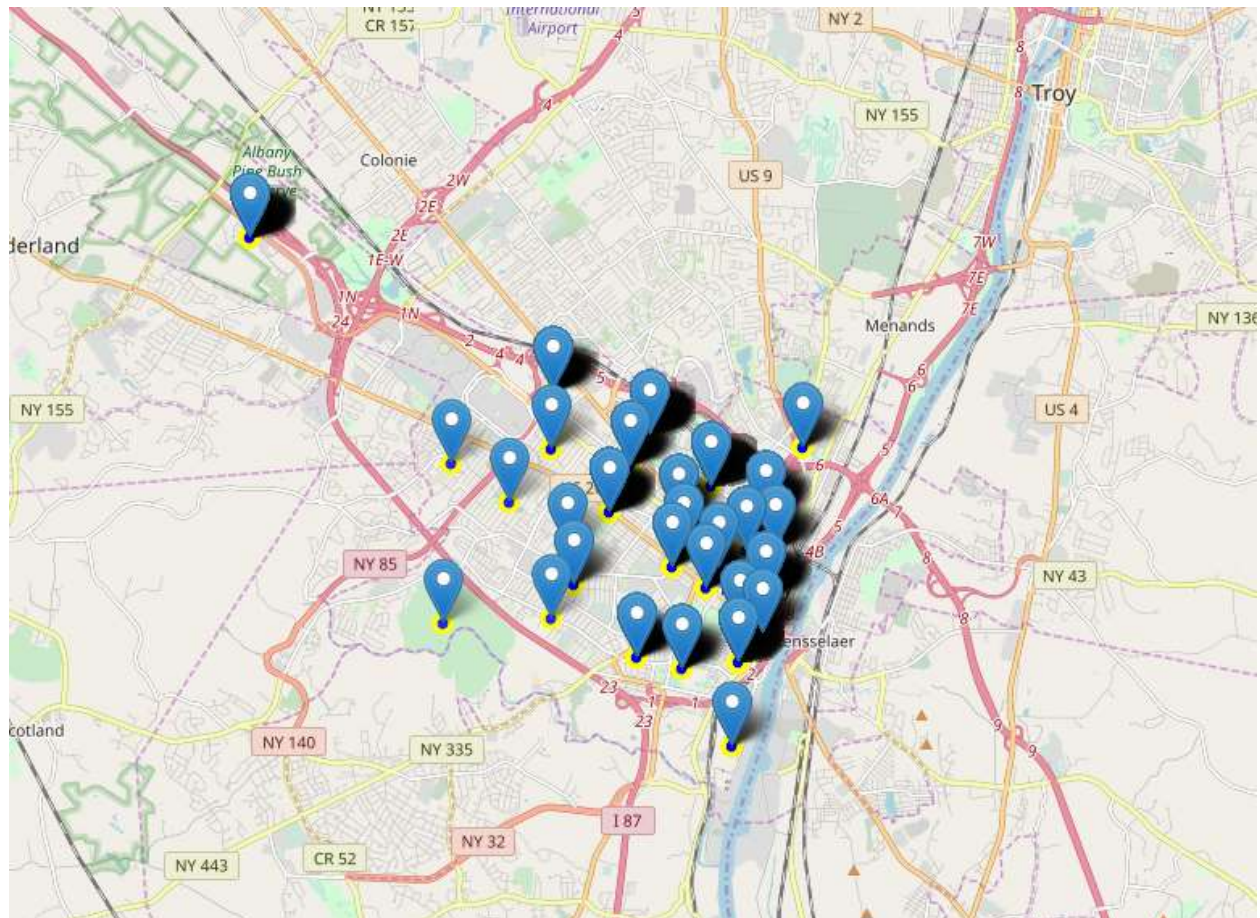
```
In [59]: df_incidents['X'] = df_incidents['X'].astype(float)
df_incidents['Y'] = df_incidents['Y'].astype(float)
df_incidents = df_incidents.dropna(subset=['X'])
df_incidents = df_incidents.dropna(subset=['Code Description'])
df_incidents = df_incidents.rename(columns={'Code Description':'Category'})
```

```
In [61]: df_incidents.head()
```

Out[61]:

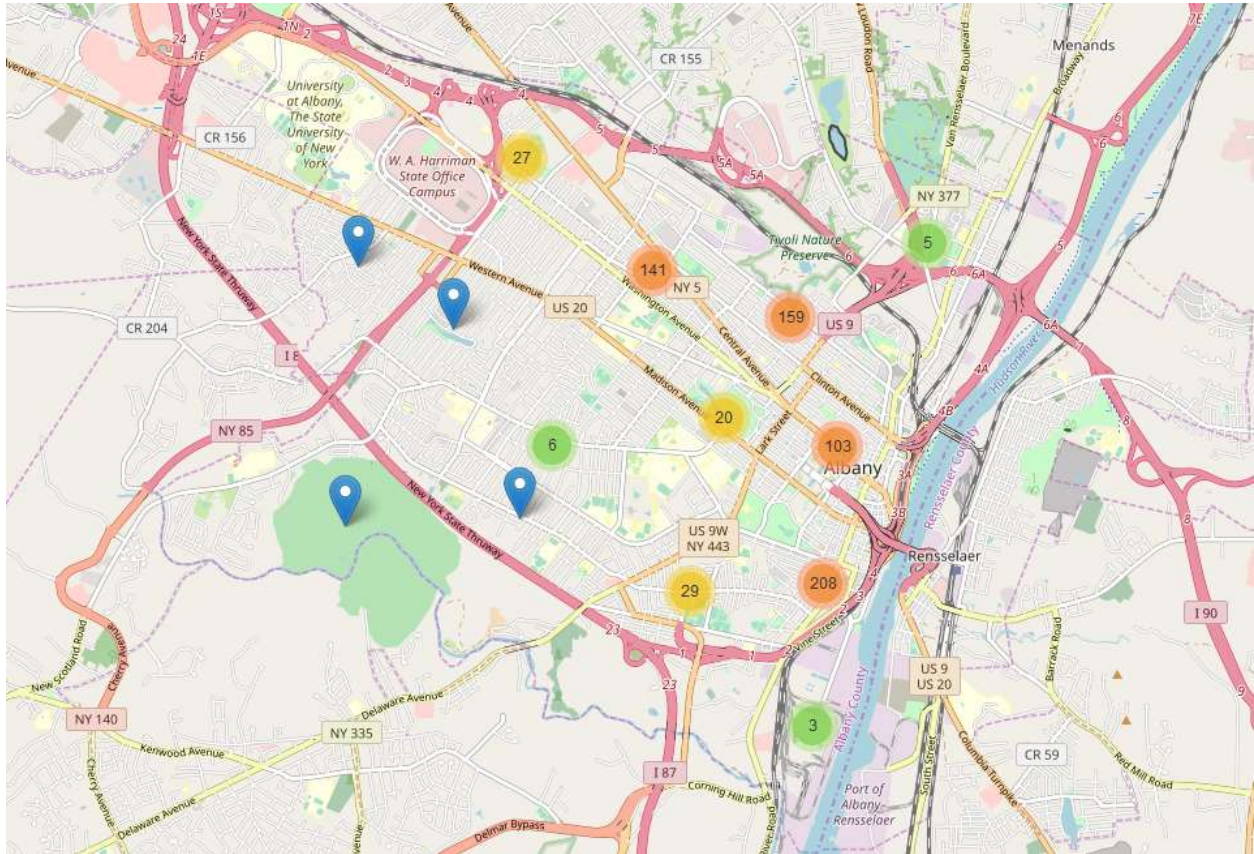
Crime Time	Age	Sex	Race	Ethnicity	Law	Crime LongID	Category	Attempt	Arrest type	Last 30 Days	City	State	Neighborhood Association	NeighborhoodXY	X	Y
7:59	59	M	B	N	PL	PL 240.26 01 OV2	HARRASSMENT 2ND- PHY CONTACT	C	Complaint	False	Albany	NY	Downtown	POINT (-73.7532379834 42.6495622808)	-73.753238	42.649562
7:19	22	F	W	N	PL	PL 220.03 AM7	CRIM POSS CONTRL SUBST	C	Crime in Progress	False	Albany	NY	West Hill	POINT (-73.764844247 42.6666911887)	-73.764844	42.666691
1:23	32	F	W	N	PL	PL 220.03 AM7	CRIM POSS CONTRL SUBST	C	Crime in Progress	False	Albany	NY	Beverwyck	POINT (-73.7822271336 42.6700629397)	-73.782227	42.670063
2:35	52	M	B	N	PL	PL 120.00 01 AM3	ASLT W/INT CAUSES PHYS INJURY	C	Complaint	False	Albany	NY	South End	POINT (-73.7592617522 42.6392747943)	-73.759262	42.639275

Then criminal incidents are plotted using folium map with labels for Albany region.

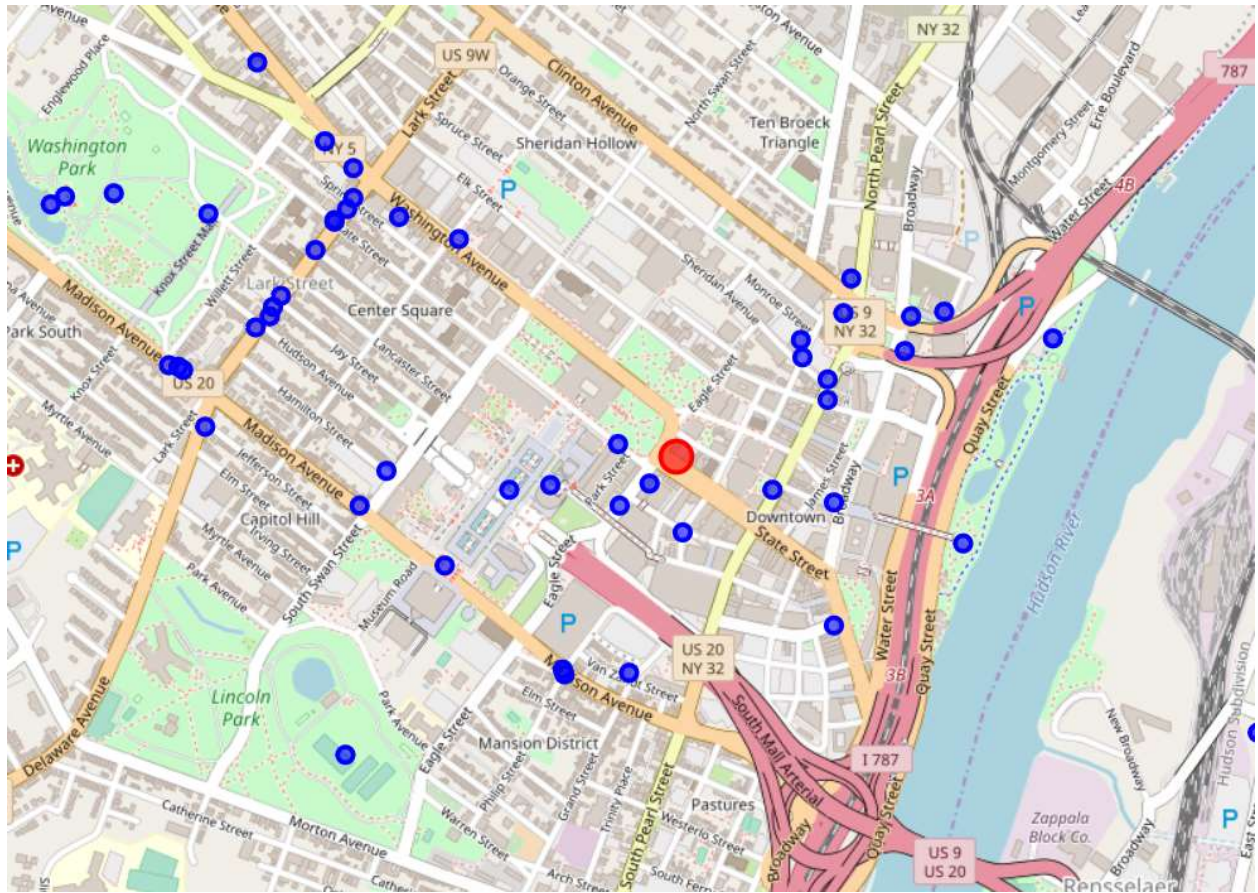




By adding incidents within nearby circle in cluster, incident rate is identified

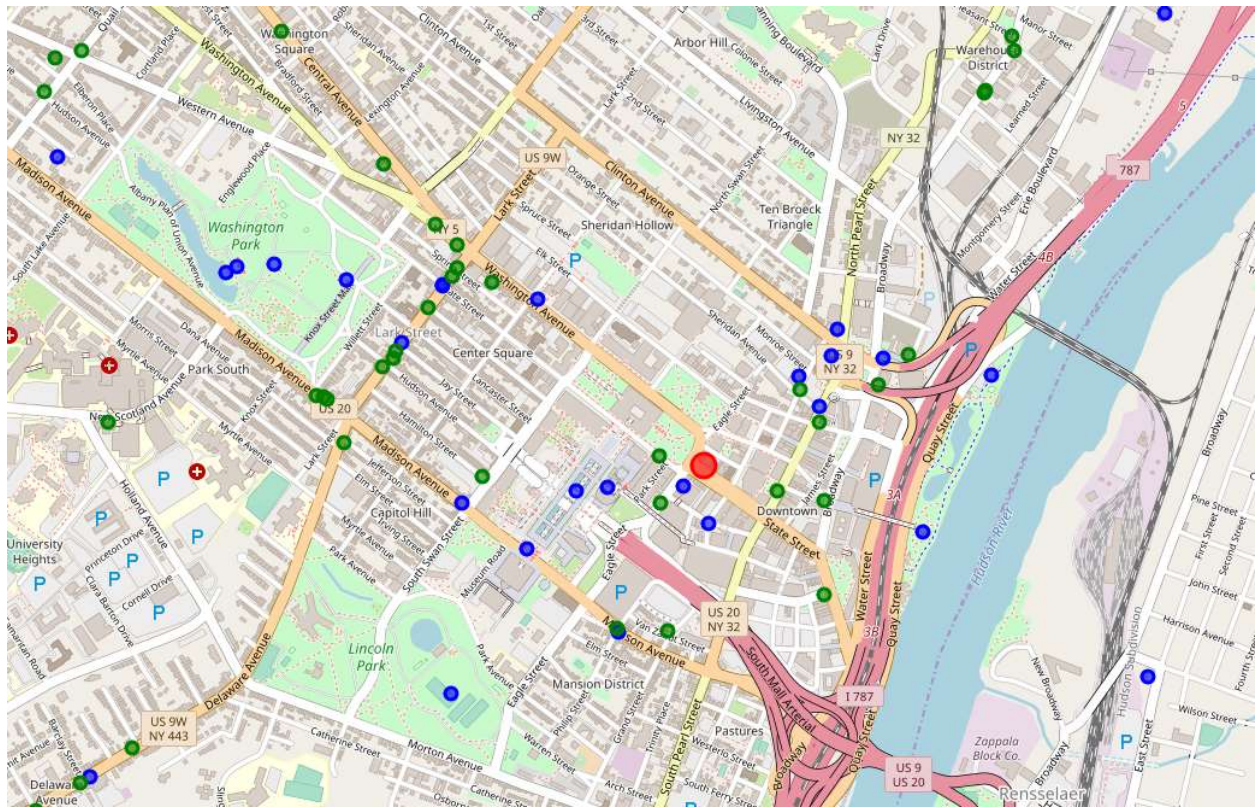


Next, using foursquare location data, top venues in Albany are identified:

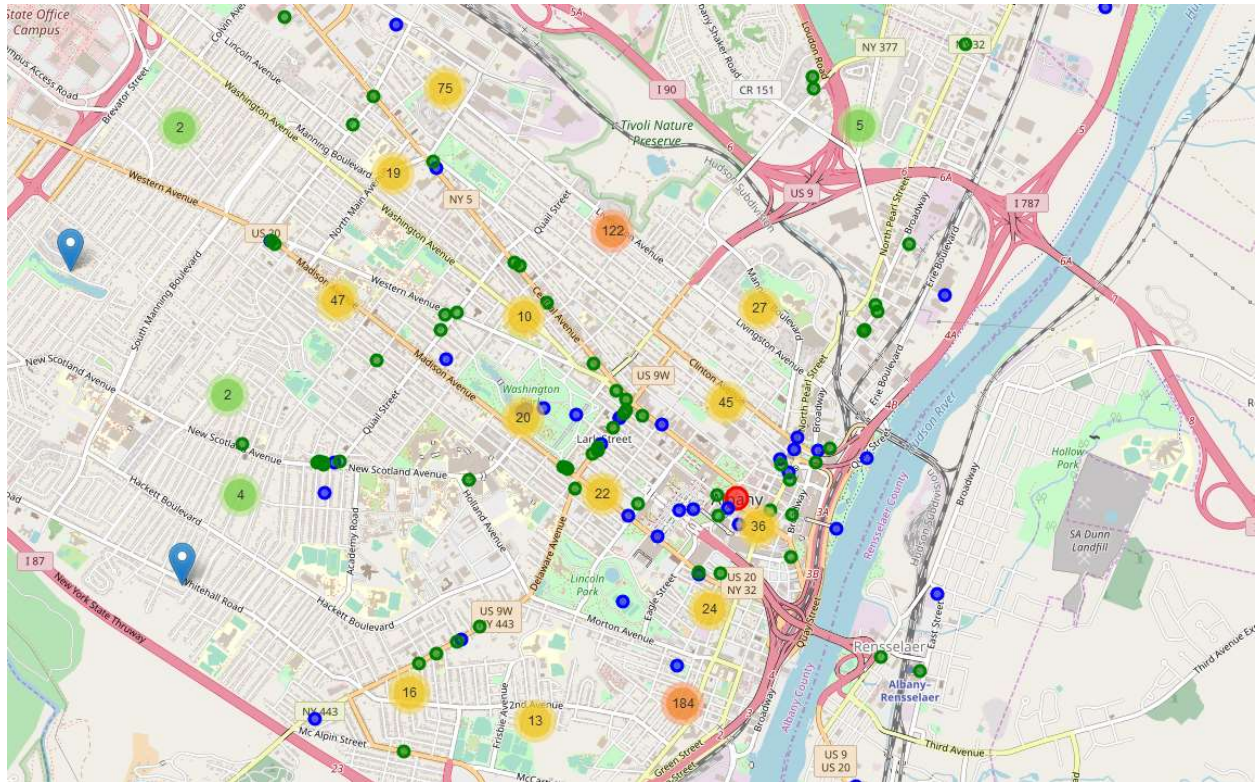


Further, those venues are be divided into non-restaurant and restaurant:





Finally, the venue data are plotted together with criminal incident rate:



## Discussion

- ❖ Our goal is to determine best location for a new restaurant. The ideal place of a new restaurant should have following criteria:
  - (1) It is close to hot venues
  - (2) Not many existing restaurants nearby to compete customers
  - (3) It is in a safe area.
- ❖ From the consolidate map, a few top candidate places for new restaurants are:
  - (1) Watervliet Avenue Extension
  - (2) Intersection of South Swan Street and Madison Avenue
  - (3) Intersection of Chapel Street and Clinton Avenue

## Conclusion

In this project, I analyzed criminal and venue data in Albany NY region, by finding the criminal rate in different clusters and separating top venues into non-restaurant and restaurant types, consolidate map is plotted and top candidate place for new restaurants are identified.