Chase Whitman IBM Applied Data Science Capstone September 1, 2020

# **Exploring Potential Restaurant Locations in Essex County, New York**

#### I. Introduction

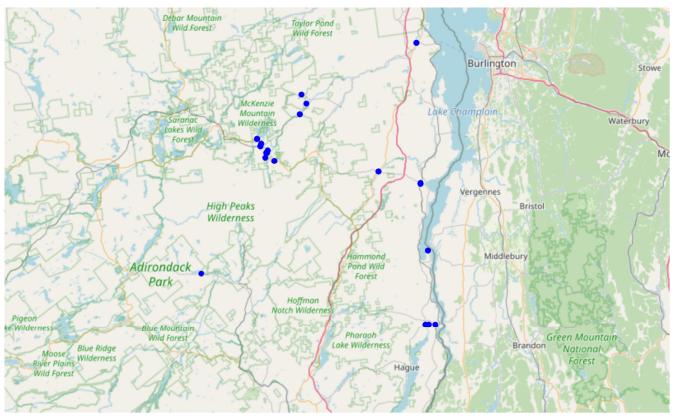
Essex County is a small county located in the northern region of New York, United States. The county was founded in 1799 and has many historical sites located within its boundaries. There are approximately 40,000 residents in the county. The goal for this project is to identify the best locations to open up a restaurant. The target audience would be aspiring restaurant owners or developers. The project will not focus on any specific type of restaurant, but rather provide ideal locations for any food establishment in general.

An individual or company looking to open up a restaurant in Essex County will care about this project because the population for the county is relatively small. Therefore, a successful restaurant will likely require consistent business from tourists and non-residents. The restaurant's location must be within a close vicinity of popular tourist destinations. A location within walking distance of the tourism site is the most ideal, as tourists can walk over to the restaurant immediately after visiting the attraction. Additionally, if they search for nearby restaurants from within the attraction, then the new restaurant will be listed as one of the first results and greatly increase its chance of being picked.

### II. Data

The State of New York provides a geographic data set of all tourist attractions in Essex County on the Data.gov website. The data set contains 24 tourist attractions along with their respective addresses and websites. The addresses must first be geocoded to retrieve their latitude and longitude coordinates. These coordinates will then serve as the center points for the Foursquare search area.

Foursquare location data will be used to explore the surrounding area. A search radius of 500 meters will be used because the average person can walk 500 meters in five to ten minutes. The Foursquare data can provide important information about existing venues near each of the tourist attractions. For example, the data can show the number of restaurants that already exist near each of the venues and provide insight on the amount of competition that would likely occur. The Foursquare data can also show other types of nearby venues that would be beneficial to a restaurant. A hotel, for example, could draw in additional customers, especially if the restaurant was within a reasonable walking distance.



Map 1: Tourist attractions in Essex County, NY

## III. Methodology

The first part of the project involved exploring the Essex County tourist attraction data set. The tourist attraction data set only contained 24 tourist sites. The type of sites included in the data set were museums, railroads, theaters, art centers, and more. All sites contained addresses and websites. The addresses had to be geocoded to obtain the coordinates. The Geopy module introduced in previous lessons only produced coordinates for approximately half of the sites. The Geocoder module with the ArcGIS provider was able to successfully geocode all of the locations.

The next step in the project involved using Foursquare location data to find all venues within a 500 meter radius of each tourist attraction. The search returned 209 venues. Examining the value counts for each location demonstrated clear trends of distribution. Most attractions had between one to seven venues nearby. There were three major outliers. These three tourist attractions were all within the Lake Placid center and were likely referencing the same venues. Without even researching the area, it was obvious that Lake Placid was the major shopping and entertainment center in the county.

After reviewing the different venue categories in the Foursquare search, the venue data was filtered into two new dataframes. One dataframe included all venues that likely served food. The second dataframe included all hotels and motels. There were 16 hotels and motels within a 500 meter radius of the tourist attractions. The venues that served food were divided into very specific venue categories, such as 'American Restaurant' and 'BBQ Joint'. In order to compile all food-serving venues into one dataframe, a new list was created that iterated through each venue category searching for keywords, like restaurant, deli, and diner. The final dataframe included 51 restaurants. The data was

then sorted and ranked into the five most common venues surrounding each of the nine attractions. Finally, the data was visualized on a map using the Folium package.

	Tourist Attraction	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Adirondack History Center Museum	American Restaurant	Steakhouse	BBQ Joint	Breakfast Spot	Burger Joint
1	Adirondack Scenic Railroad	Tapas Restaurant	American Restaurant	Southern / Soul Food Restaurant	Snack Place	Sandwich Place
2	Hancock House Museum	American Restaurant	Diner	Pizza Place	BBQ Joint	Breakfast Spot
3	High Falls Gorge	Diner	Tapas Restaurant	Steakhouse	BBQ Joint	Breakfast Spot
4	Lake Placid Center for the Arts	Deli / Bodega	Fast Food Restaurant	Mexican Restaurant	BBQ Joint	Burger Joint
5	Lake Placid Olympic Museum	American Restaurant	Italian Restaurant	Breakfast Spot	Southern / Soul Food Restaurant	Snack Place
6	Lake Placid Skating	American Restaurant	Italian Restaurant	Breakfast Spot	Southern / Soul Food Restaurant	Snack Place
7	Lake Placid-North Elba Historical Society	Tapas Restaurant	American Restaurant	Southern / Soul Food Restaurant	Snack Place	Sandwich Place
8	Ticonderoga Heritage Museum	American Restaurant	Diner	Sandwich Place	Pizza Place	Deli / Bodega

*Table 1: Five most common restaurant types per tourist attraction.* 

### IV. Results

Not all tourist attractions had restaurants within 500 meters. Only nine of the 24 attractions had venues that met the criteria. Of these nine attractions, the ones located in Lake Placid had the most restaurants. Similar to the restaurant venues, not all attractions had hotels or motels nearby. Only seven of the attractions had hotels within walking distance. Also, similar to the restaurant venues, the majority of the hotels and motels were within the Lake Placid area. In general, American restaurant types were the most common in the county. American restaurants were the first or second most common restaurant type in seven of the nine tourist attractions. Twelve out of 24, or 50%, of the tourist attractions did not have a single restaurant or hotel within walking distance.

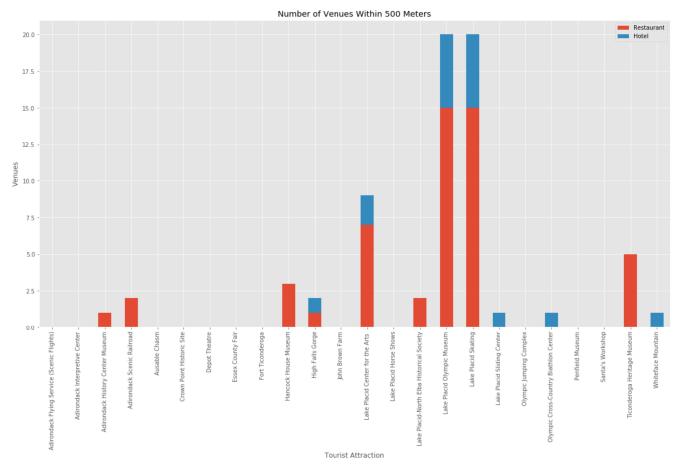
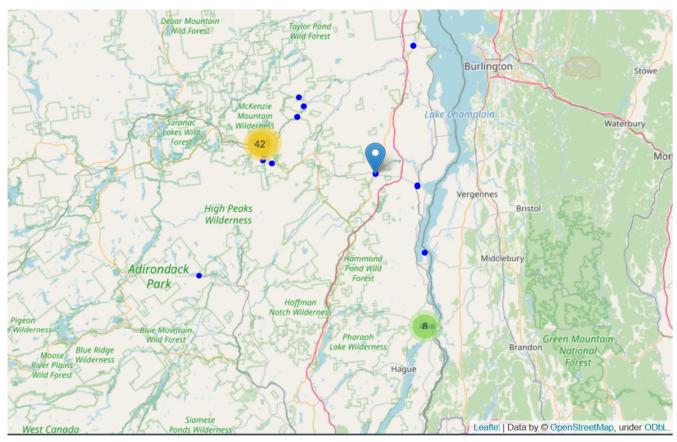


Figure 1: Number of restaurant and hotels per tourist attraction.



Map 2: Venue Clusters

#### V. Conclusion

The results of this project can aid aspiring restaurant owners in finding the most ideal location to open their restaurant. The best location will be dependent upon many factors, such as the owner's budget, the type of restaurant they wish to open, and how close they want to be to surrounding restaurants. A restaurant owner or analyst can use the table containing the most common venue types to determine whether their specific restaurant type would be a good fit for the area. For example, delis and fast food restaurants are the most common type near the Lake Placid Center for the Arts. This may be a prime location for a nicer, sit-down restaurant due to the lack of competition in the area. Furthermore, the data also showed that many tourist attractions did not have restaurants within walking distance. A restaurant owner can capitalize on this opportunity by opening up one in the area. Lake Placid had the most restaurants and hotels close to tourist attractions and may be the most popular area in the county for tourism. As a result, the area will likely be more competitive and expensive.