Hospitality and Food Industry Data Analysis of New Orleans

Alex Yu

December 11th 2020

1. Introduction

New Orleans is a city for great food and eats. According to the World Best Awards 2020 and Travel + Leisure voting it ranked among the top 3 destinations in the world for quality and affordable food. The city of New Orleans only has about 400,000 residents and 71 neighborhoods. However, this city is one of the busiest in terms of tourism and hospitality, each year raking in millions of visitors from all over the nation and sometimes even the world.

With tourism and hospitality being so valued in this city, it raises the question of where those business owners of restaurants or hotels and potential visitors may be attracted towards. Obviously not all locations even in the great city of New Orleans are made equal for this business. With this in consideration, we can map out the city of New Orleans and its corresponding neighborhoods to better inform our business owners and food/travel enthusiasts.

2. Data Source and Cleaning

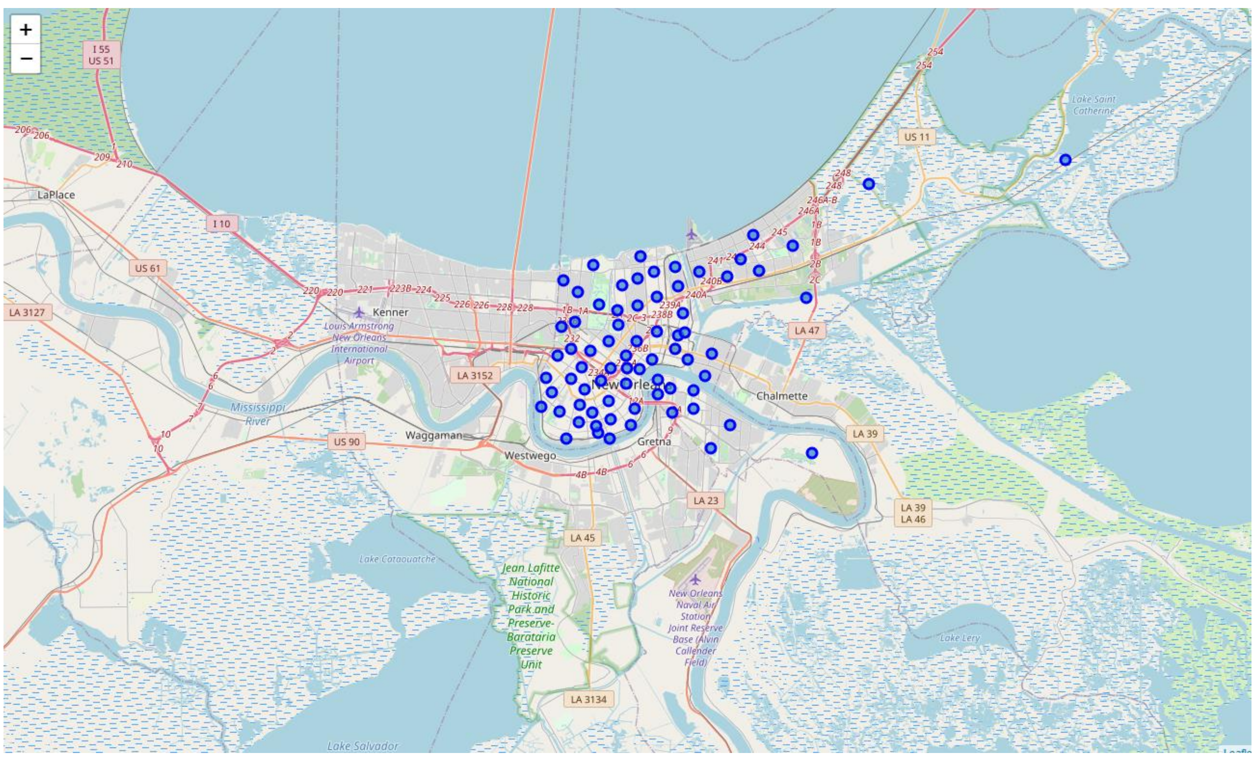
To address the above question and overarching business problem we require some fundamental locational data:

- First, retrieved from Wikipedia I have scraped neighborhoods of New Orleans and corresponding latitude and longitude figures.
- To help with clustering based on nearby venues I used the Foursquare API.

The dataset I was working with included three important features that would be critical in looking up and recording nearby venues. They included: Neighborhood, Latitude, and Longitude.

	Neighborhood	Longitude	Latitude
0	U.S. NAVAL BASE	-90.026093	29.946085
1	ALGIERS POINT	-90.051606	29.952462
2	WHITNEY	-90.042357	29.947200
3	AUDUBON	-90.121450	29.932994
4	OLD AURORA	-90.000000	29.924440

To ensure that the coordinates made qualitative sense with its nearby New Orleans setting I created a map primitive map using folium, marking each neighborhood with its reported coordinates.



Because of the due diligence in the web scraping phase, there was less preprocessing of the data in comparison to the lab and assignment of segmenting and clustering New York and Toronto. I was able to directly and iteratively work through the entire Neighborhood column and grab the corresponding venue values from the Foursquare API

Using my respective CLIENT_ID and CLIENT_SECRET with the Foursquare API, I was able to retrieve all of the local venues for the 67 neighborhoods of the dataset. Then using the groupby() and get_dummies() I was able to condense the items retrieved from the API to create a binary set of columns for each unique venue and condense the count among each unique venue for the different neighborhoods.

3. Methodology

The next step was to create a new dataframe that took the mean of the condensed frequencies to make the clustering easier and accurate. However, after completing this step and most of the cleaning I decided that in order to better address and analyze the hospitality and food industry locational data, I should limit the scope of the venues. Thus, I compiled a list of venues that specifically belonged to the mentioned industries and made those the possible venues to cluster by. The result of that dataframe is picture below:

```
venue_list = ['Neighborhood','American Restaurant','Asian Restaurant', 'BBQ Joint', 'Bagel Shop', 'Bakery', 'Bar',
              'Baseball Field', 'Baseball Stadium', 'Basketball Court', 'Bed & Breakfast','Beer Garden', 'Beer Store', 'Bistro', 'Breakfast Spot', 'Brewery', 'Burger Joint', 'Burrito Pl
             'Café', 'Cajun / Creole Restaurant', 'Candy Store', 'Caribbean Restaurant', 'Cheese Shop', 'Chinese Restaurant', 'Cocktail Bar', 'Coffee Shop', 'Comedy Club', 'Comfort Food
             'Deli / Bodega', 'Dessert Shop', 'Dive Bar', 'Donut Shop', 'Eastern European Restaurant', 'Ethiopian Restaurant', 'Farmers Market', 'Fast Food Restaurant', 'Flea Market', '
              'Food', 'Food & Drink Shop', 'Food Court', 'Food Truck', 'French Restaurant', 'Fried Chicken Joint', 'Gay Bar', 'General Entertainment', 'Gift Shop', 'Gourmet Shop', 'Groc
             'Halal Restaurant', 'Harbor / Marina', 'Hostel', 'Hot Dog Joint', 'Hotel', 'Hotel Bar', 'Hotel Pool', 'Ice Cream Shop', 'Indian Restaurant', 'Indie Theater', 'Irish Pub', '
             'Juice Bar', 'Karaoke Bar', 'Liquor Store', 'Lounge', 'Mediterranean Restaurant', "Men's Store", 'Mexican Restaurant', 'Middle Eastern Restaurant', 'Miscellaneous Shop',
             'Museum', 'Music Store', 'Music Venue', 'National Park', 'New American Restaurant', 'Nightclub', 'Nightlife Spot', 'Other Great Outdoors', 'Outdoors & Recreation', 'Park',
             'Pier', 'Pizza Place', 'Playground', 'Plaza', 'Pool', 'Pool Hall', 'Poutine Place', 'Pub', 'Public Art', 'Racetrack',
             'Recreation Center', 'Rental Car Location', 'Residential Building (Apartment / Condo)', 'Restaurant','Rock Club', 'Salad Place', 'Sandwich Place', 'Scenic Lookout', 'Seafoo
             'Smoothie Shop', 'Soup Place', 'Southern / Soul Food Restaurant', 'Spa', 'Spanish Restaurant', 'Speakeasy', 'Sports Bar', 'Stables', 'Steakhouse',
             'Sushi Restaurant', 'Taco Place', 'Tapas Restaurant', 'Tennis Court', 'Tex-Mex Restaurant', 'Thai Restaurant', 'Theater', 'Tourist Information Center',
             'Travel & Transport', 'Vegetarian / Vegan Restaurant', 'Vietnamese Restaurant','Wine Bar', 'Wine Shop', 'Winery', 'Wings Joint' ]
```

```
NO_grouped = NO_grouped[venue_list]
NO_grouped.head()
```

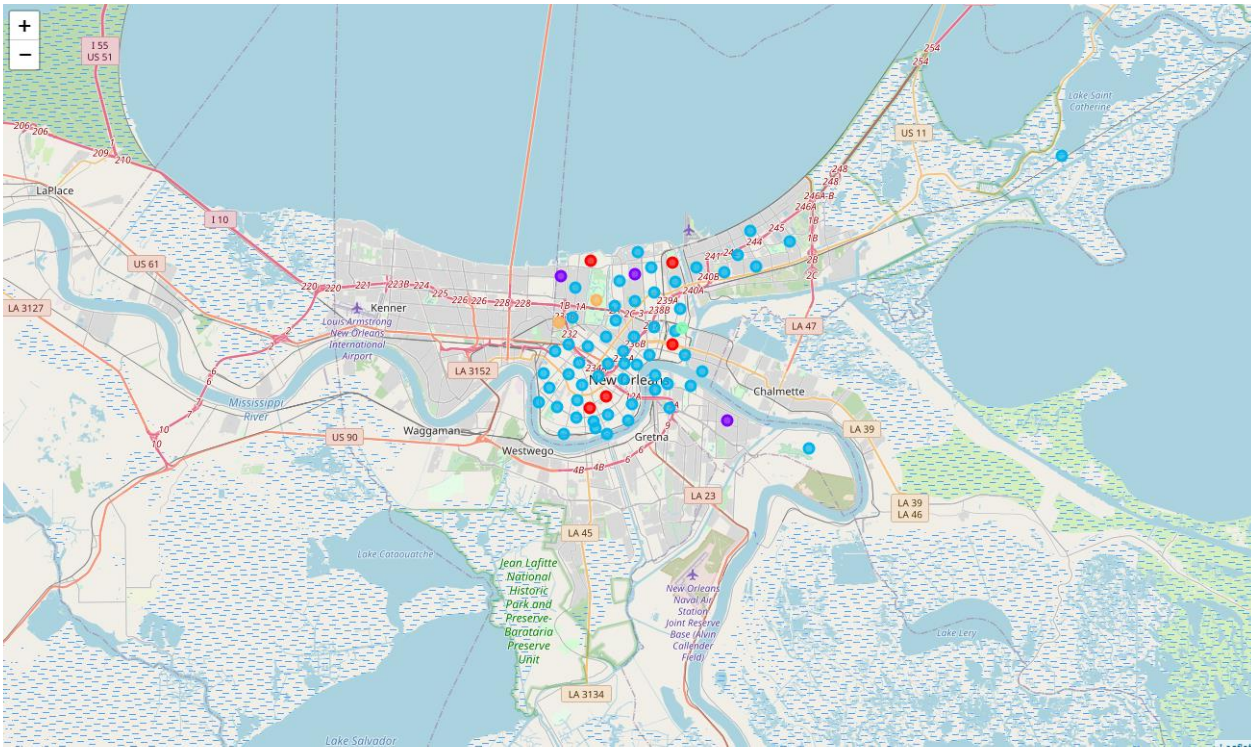
	Neighborhood	American Restaurant	Asian Restaurant	BBQ Joint	Bagel Shop	Bakery	Bar	Baseball Field	Baseball Stadium	Basketball Court	...	Thai Restaurant	Theater	Tourist Information Center	Travel & Transport	Vegetarian / Vegan Restaurant	Vietnamese Restaurant	Wine Bar	Wine Shop	Winery	Wings Joint
0	ALGIERS POINT	0.0	0.0	0.0	0.0	0.0	0.130435	0.0	0.0	0.0	...	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.043478
1	AUDUBON	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	...	0.0	0.083333	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000
2	B. W. COOPER	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	...	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000
3	BAYOU ST. JOHN	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	...	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000
4	BLACK PEARL	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	...	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000

5 rows × 127 columns

With this done I was almost done with the dataframe to be fitted on by the machine learning algorithm. The last thing to finish cleaning my data was to create a method and iteratively go through the row to compile the top 10 venues for each neighborhood. The final dataframe that is fitted on contained the features: Neighborhood, Latitude, Longitude, and 1st-10th most Common Venue.

Now it was time to use our machine learning model. I decided upon using an unsupervised machine learning algorithm of clustering as way of looking at similarities among clusters to hopefully infer about where businesses in the hospitality/food/tourism industry lie along the city of New Orleans. K Means Clustering, specifically is what I used, as it is a very robust and easy to implement cluster algorithm which is why I chose to use it in my project.

Using k=5 clusters I was able to fit it among my cleaned data and then used folium to create a clustered map that color coded the fitted clusters.



4. Results and Discussion

After running through K Means Clustering the model was able to cluster the 67 neighborhoods into 5 groups. Looking at the folium map, geographically we can see that there is some connection between the location of the neighborhood and it's cluster label. Those more inland corresponded to the blue marker while the one's bordering the coast had other colors.

I also grouped the dataframe by cluster and examine each of the cluster and the top venues that each neighborhood had. It is important to note that besides geographical locations the inland clusters also had more venues geared towards hospitality and social service likes hotels, clubs, and bars, while those on the coast had more restaurants and grocery stores or farmer's market.

Another key finding were that red markers were grouped by the large presence of parks and outdoor recreation which could be an ideal place for mobile modes of dining like food trucks.

Cluster 0 - RED

```
NO_merged.loc[NO_merged['Cluster Labels'] == 0, NO_merged.columns[list(range(9))]]
```

	Neighborhood	Longitude	Latitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
12	ST. CLAUDE	-90.039268	29.971814	0	Plaza	Park	Wings Joint	Japanese Restaurant	Ethiopian Restaurant
15	CENTRAL CITY	-90.086689	29.939465	0	Park	Fried Chicken Joint	Seafood Restaurant	Cajun / Creole Restaurant	Grocery Store
25	MILAN	-90.098362	29.932548	0	Wine Shop	Dive Bar	Park	Theater	Wings Joint
44	LAKESHORE - LAKE VISTA	-90.097847	30.023253	0	Harbor / Marina	Park	Wings Joint	General Entertainment	Ethiopian Restaurant
51	PONTCHARTRAIN PARK	-90.039310	30.021989	0	Park	Wings Joint	General Entertainment	Ethiopian Restaurant	Farmers Market

5. Conclusion

Overall I was able to get a better sense of how building and services of the food, tourism, and hospitality industry were laid out. More of the social attractions like bars, clubs, and parks were located more inland and a great place for aspiring owners to prop up their venues. As for more dining and especially of Creole/Haitian type, they were more likely to be found along the shores and in the non-blue clusters.

Moving forward more analysis can be done on possibly reviews of said tourist and restaurant attractions to get a better sense of optimal locations for touring, dining, or having fun for both owners and customers.