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.

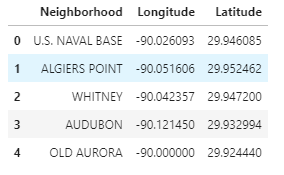
With tourism and hospitality being so valued in this city, it raises the question of where those business owners of restaurants or hotels and potential visitors may be attracted towards. Obviously not all locations even in the great city of New Orleans are made equal for this business. With this in consideration,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

1. **Data Source and Cleaning**

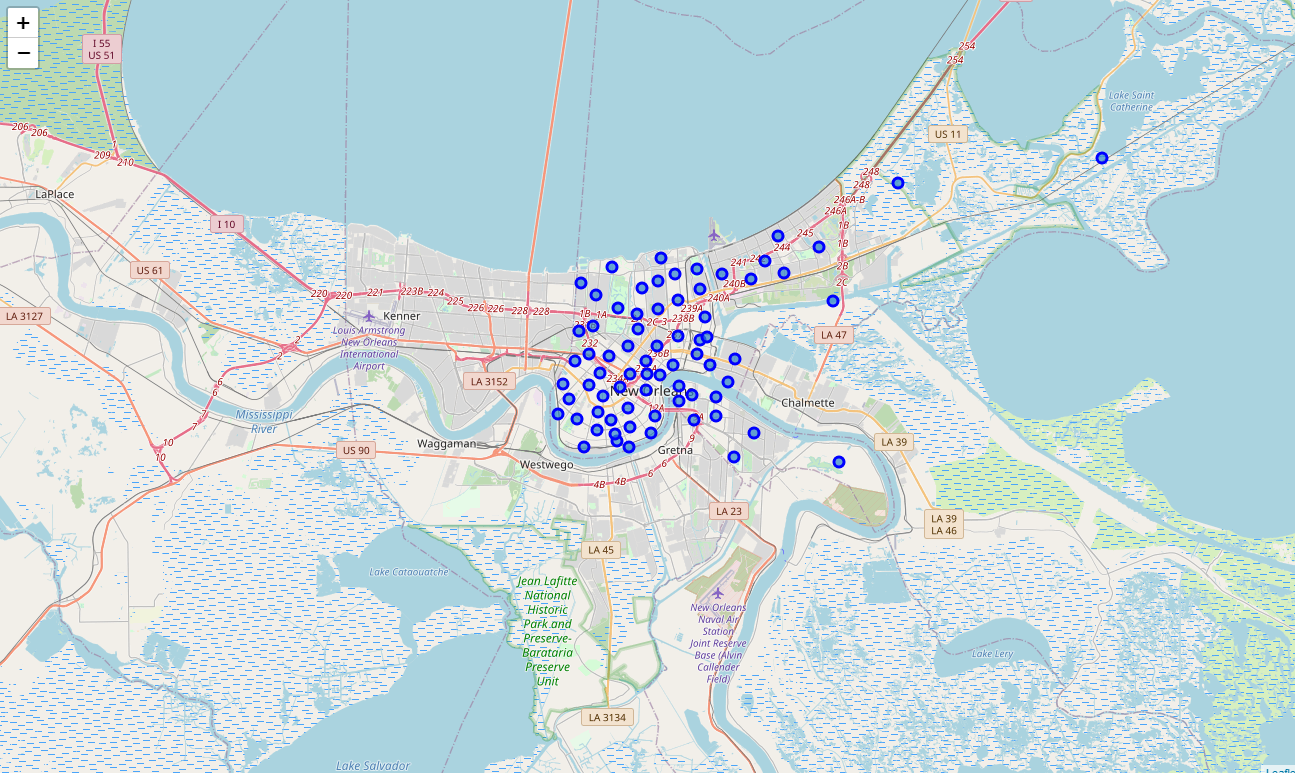
To address the above question and overarching business problem we require some fundamental locational data:

* 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.



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.

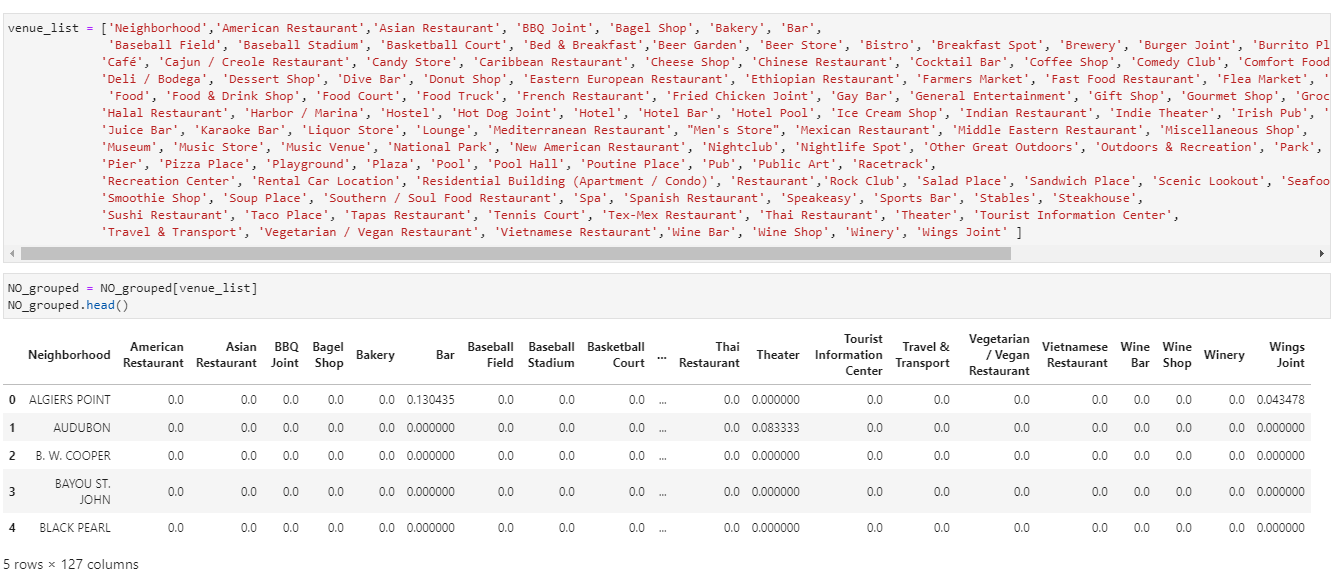


Because of the due diligence in the web scraping phase, there was less preprocessing of the data in comparison to the lab and assignment of segmenting and clustering New York and Toronto. I was able to directly and iteratively work through the entire Neighborhood column and grab the corresponding venue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

1. **Methodology**

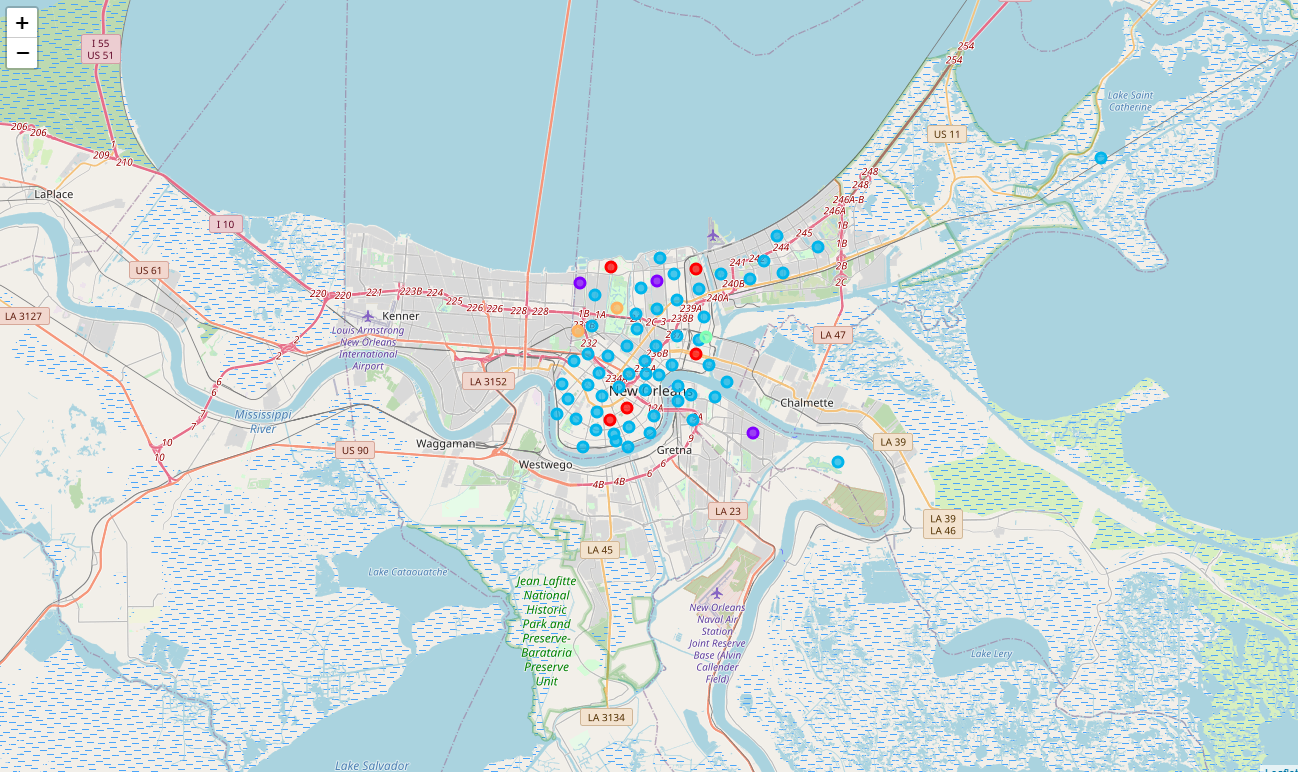
The next step was to create a new dataframe that took the mean of the condensed frequencies to make the clustering easier and accurate. However, after completing this step and most of the cleaning I decided that in order to better address and analyze the hospitality and food industry locational data, 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:



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.



1. **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.