

THE BATTLE OF NEIGHBOR HOODS

NEW YORK FOOD
INDUSTRY



DECIDING THE BEST PLACE TO OPEN A RESTAURANT

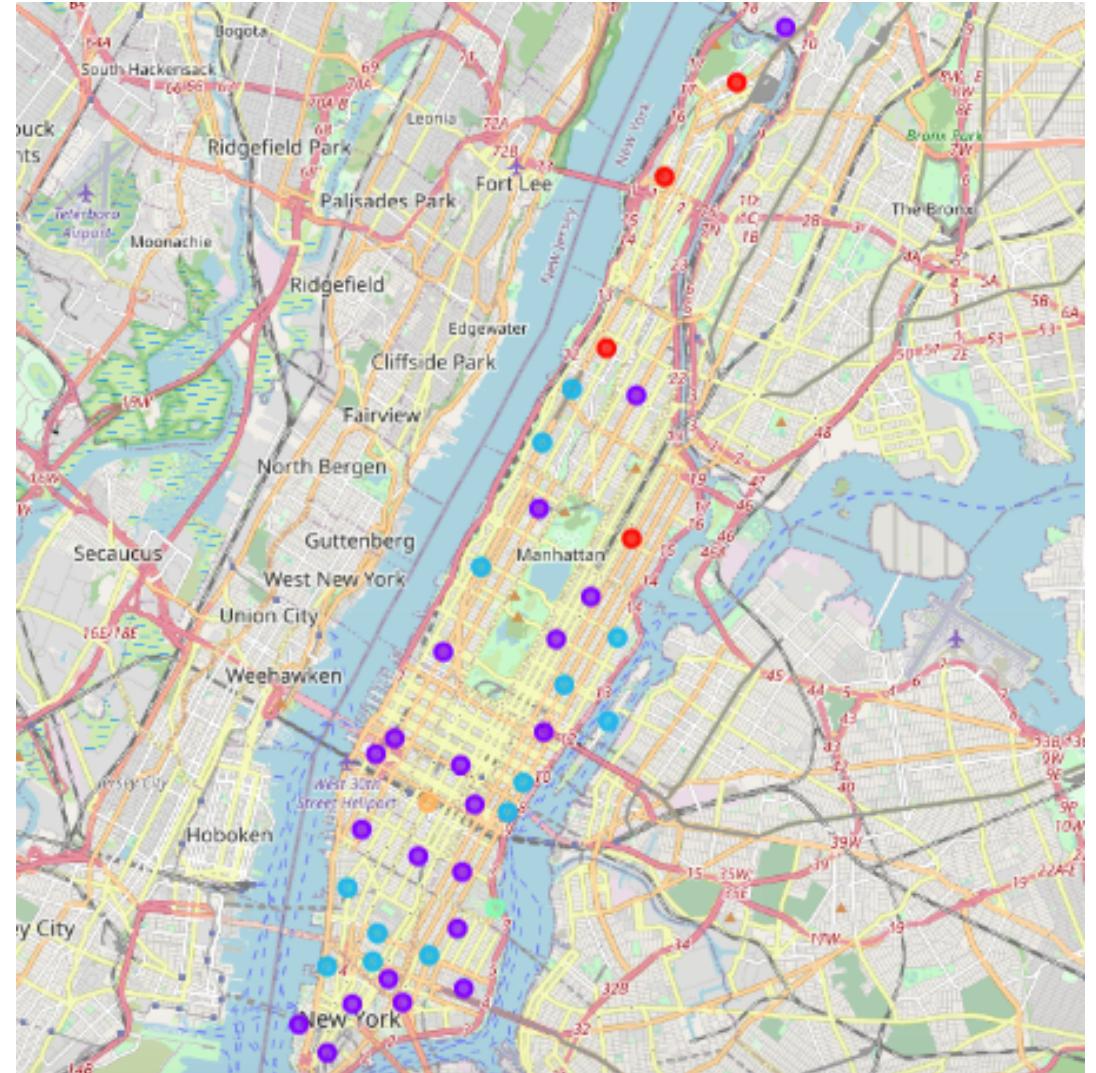
- As a new owner one of the most difficult things is to decide the place where you will open your next restaurant. You need to consider many factors that can impact your business in the future.
- The objective of this project is to provide information for a new restaurant owner about the neighborhoods in New York city, considering for each borough the preferences of the customers in each neighborhood.
- With this information, the new owner will be able to make a profitable decision as he would know if his type of restaurant has a good valoration on the neighborhood or not.

DATA ACQUISITION AND CLEANING

- For this project we use 2 different data sets:
 1. New York data provided by the course (Neighborhood and Borough data).
 2. New York top venues per neighborhood (Foursquare API).
- Once we collect the data, we processed it to generate the data frames we will need to run our models.

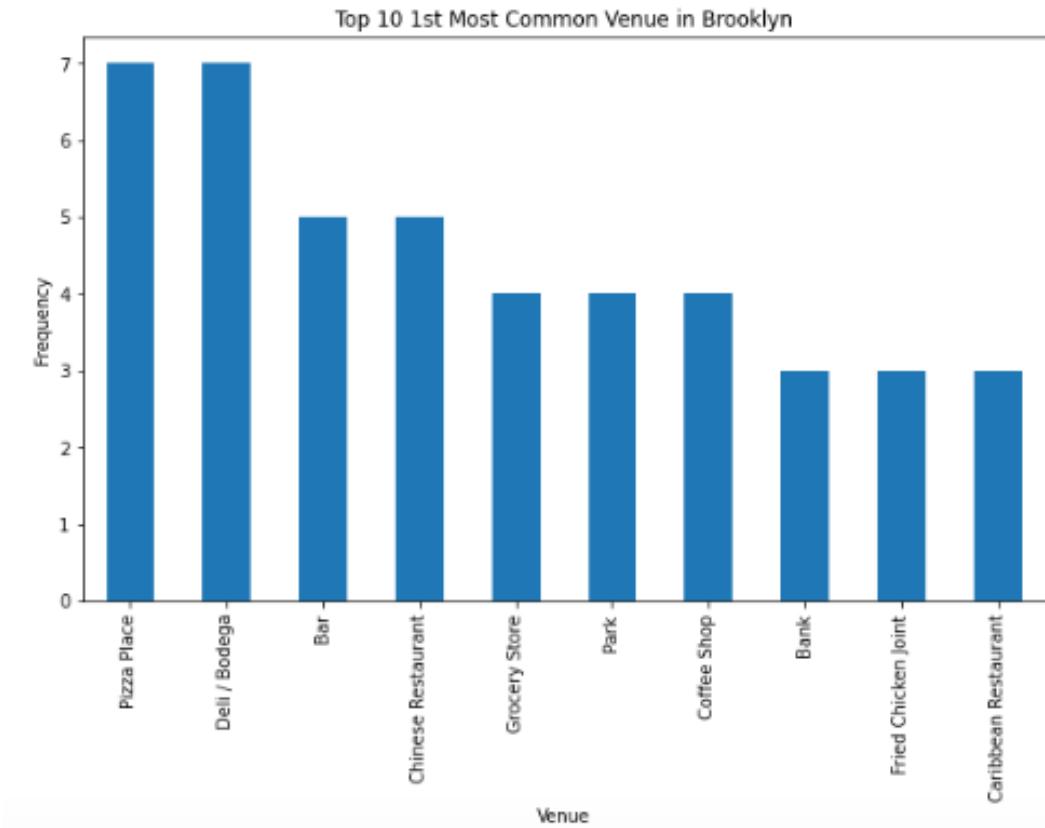
K-MEANS CLUSTER MODEL

- We use k-means cluster model to cluster our data for every borough depending on every neighborhood top 10 venues.
- We decide a value of $k = 5$ to generate 5 clusters considering that the model is an unsupervised model looking to minimize intra-cluster distance and maximize inter-cluster distance.
- With the cluster results we analyze how each cluster was define considering the top 10 venues and realize which one would be better for different types of venues to open a new store/restaurant.



TOP 10 MOST COMMON VENUES PER BOROUGH

- We extract the top 10 most common venues for each borough, so we can decide in which borough would be better which type of restaurant.
- For example, looking to Brooklyn bar chart, we can realize that pizza places, Chinese restaurants and bars are consider as the most popular places for their costumers.



CONCLUSIONS

- As we can see with the results, using data science tools to make company decisions may impact the future of our business.
- It is very important to consider all the data that we have of the industry before taking any decision.