Site recommendation and competitor analysis for Latin Restaurant

Data Science capstone project

Table of contents

Introduction

Background

Business problem

Target Audience

Data Section and analysis

Data Sources

Data cleaning process

Basic exploration and modeling

Results and Discussion

Conclusions

Recommendations

Introduction-Problem

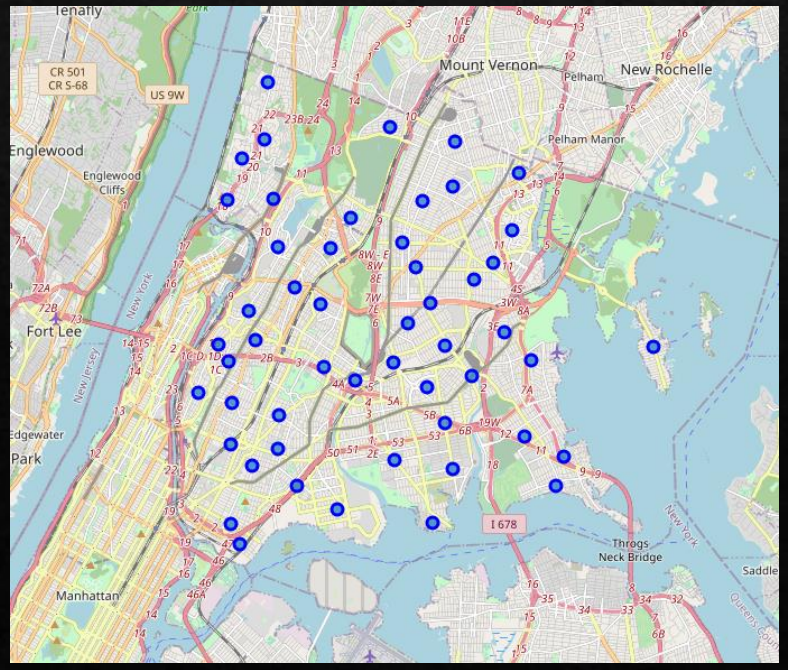
- **The Bronx is the only New York borough with a Hispanic majority**. At the 2010 Census, 53.5% of Bronx's population was of Hispanic, Latino, or Spanish origin (they may be of any race).
- Restaurants and Culinary habits have always been one of the favorite ways to socialize and meet each other.
- My project it's a about determining **best restaurant location and segmenting** competitors in Bronx city to find **low Latin restaurant desensity** by means of data science **identify best location** for opening a new restautant.
- My target Audience:
  - Restaurant groups, retail chains, and other enterprises can use this work and use POI data to help solve problems such as site selection, competitive analysis, and much more.

Data source and cleanning

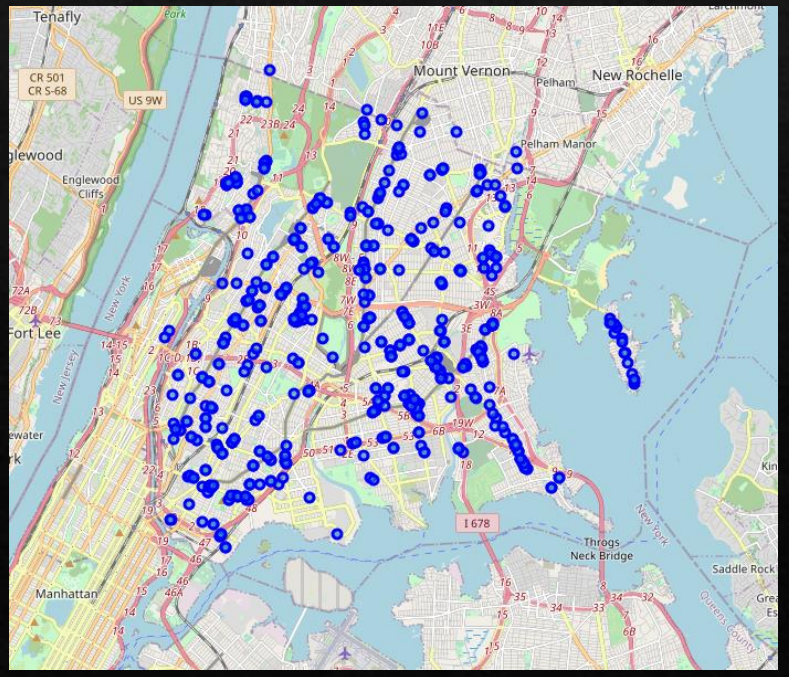
- The restaurant data used in this study was pulled from **Foursquare's Places Database** - a comprehensive dataset spanning 190+ countries and 50 territories. The data set is updated continuously and published daily as our world is constantly evolving.
- New York data location file with neighborhoods and boroughs.
- Sliced and merged both source to create a data frame with restaurants categories with 607 rows and 7 features in total.

Data Analysis | Bronx venue location

- Map shows 52 neighborhoods from Bronx borough.

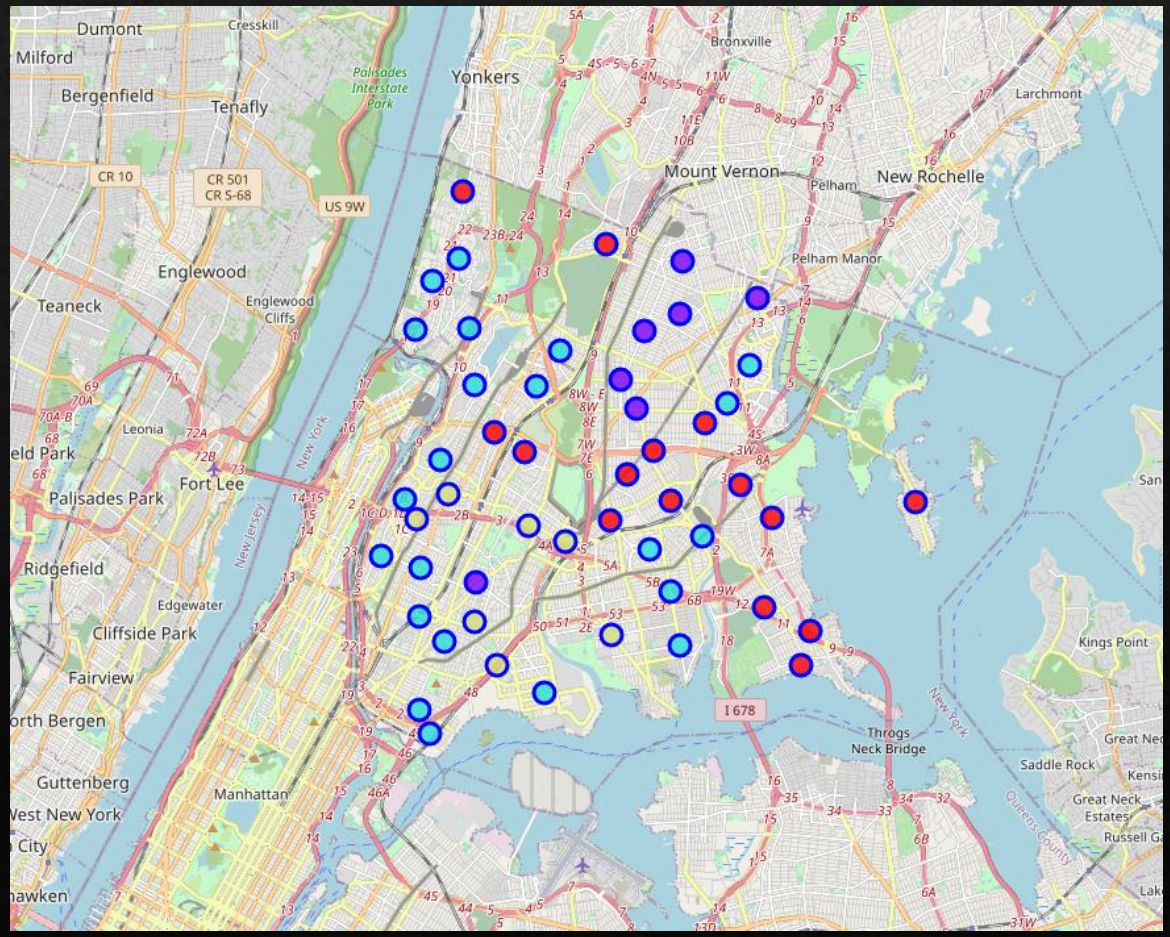


- After merging with Four Square venues location data, we obtained 607 venues as shown in the map bellow



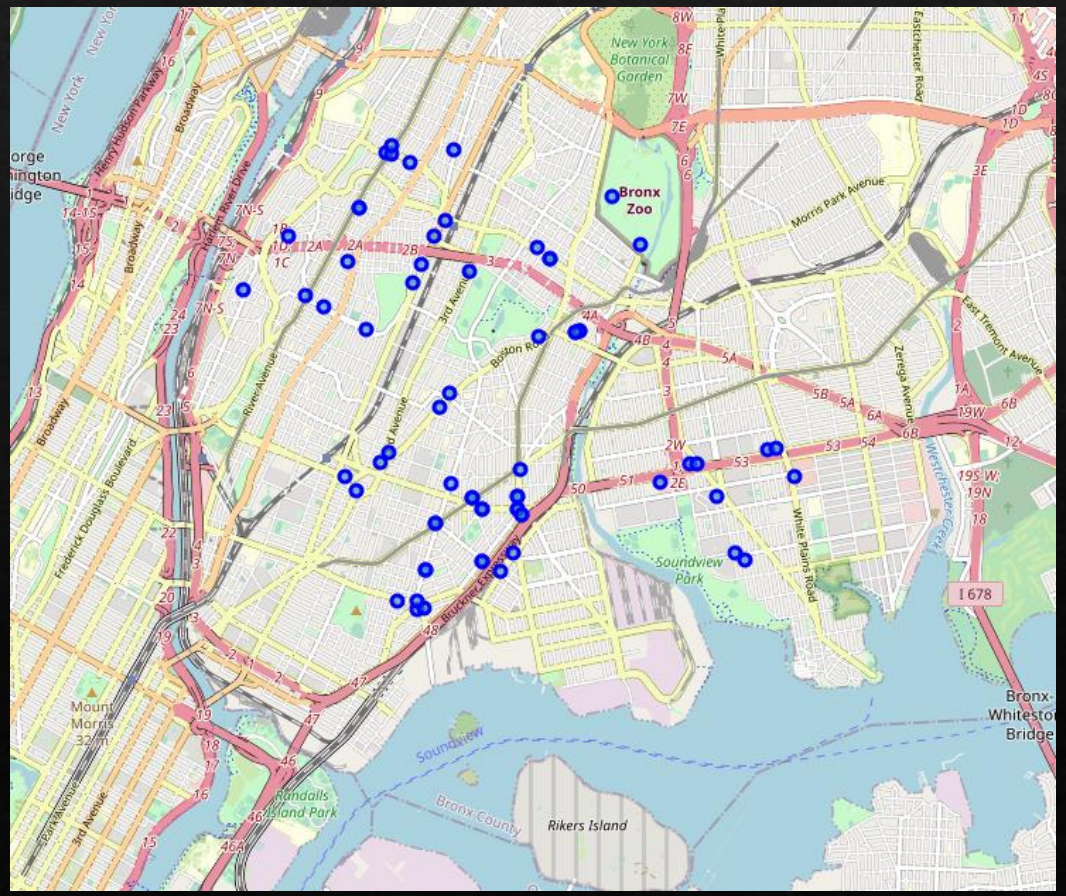
Clustering neighborhoods

- After running K-means clustering machine learning method, neighborhoods are divided into four groups.
- Maps shows color coded cluster, based on top 10 most common restaurant category venue.
- We can figure out **Cluster three** it's the most densed based on Restarant count, that represent our competitor. those neighborhood are not the best site selection site just taking into consideration venues similarity.
- **Cluster four(light green)** seems to be the best Neighborhood locations to be considered to open a new restaurant.



Neighborhood selection

- Highest concentration of restaurants was detected south and and west from center, so we focused our attention to areas south, corresponding to boroughs **East Tremon, Mount Hope, Mount Eden, Morrisania, Longwood**



Results and Discussion

- Neighborhood segmentation shows although there is a great number of restaurants in Bronx 596 after slicing first data location, we find low restaurant density in the north region.
- Highest concentration of restaurants was detected south and and west from center, so we focused our attention to areas south, corresponding to boroughs **East Tremon, Mount Hope, Mount Eden, Morrisania, Longwood**.
- Result of all this is four cluster being cluster number one containing largest number of 1srt Most common venue of Caribbean restaurant, and cluster number four having the lowest 1srt Most common venue of Caribbean restaurant

Conclusion

- Final decission on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.