
TOP TOURIST DESTINATIONS IN QUEENS, NEW YORK



AUGUST 12, 2019

IBM DATA SCIENCE CAPSTONE PROJECT

Authored by: OLUSOLA ADEKUNLE

1. INTRODUCTION

Founded in 1683, Queens is the one of the lively, vibrant and largest borough in New York. Queens Borough is located on Long Island cross the East River from Manhattan with notable tourist attractions like; Flushing Meadows-Corona Park or the Queens Botanical Garden, the park's Queens Museum is known for the "Panorama," a building-for-building model of New York City, Nearby Citi Field is the stadium of pro baseball team, the Mets and many more. Queens is the easternmost of the other five boroughs in New York, is adjacent to the borough of Brooklyn at the southwestern end of Long Island. Due to her uniqueness and beauty, Queens has been unarguably one of New York's best boroughs for new residents to make their homes and also a scenery tourist destination. Many parts of Queens have a more suburban feel and offer the perfect retreat from the bright lights and bustling streets of nearby Manhattan.

BUSINESS PROBLEM

This project evaluates the neighborhoods in Queens for people intending to visit; either to start a business or for pleasure. More importantly I try to highlights peculiarity of the neighborhoods and opportunities available for intending business owners and pleasure seekers.

Over the years, there have been increasing number of people visiting the borough and realizing what a great destination it is. There's history, culture, views, and food without the high prices as seen in other New York borough and this project will make every visit a memorable one.

DATA

While exploring the diversity of the Queens, I required

1. Neighborhoods in Queens. This define the scope of the project.
2. Latitude and Longitude coordinates of those neighborhood. This is required to plot maps and provide venue data.
3. Queens public spatial data

Data Sources

Queens data shows list of Queens neighborhood together with spatial locations
<https://geo.nyu.edu/catalog/nyu-2451-34572>

In addition, Foursquare API will be used to get the venue and venue category data of the Queens neighborhood. Foursquare has one of the largest databases for millions of locations and is used by thousands of developers. Foursquare provides many more categories of venue data but for this project, my interest is in top destinations across Queens in New York.

METHODOLOGY

The data has about 81 instances of Queens neighborhood including the latitude and longitude data. The data also confirmed that New York has 306 neighborhoods and five boroughs. In order to segments the neighborhoods and explore them, I needed essentially a dataset that contains the Queens borough and the neighborhoods that exist in the borough. as well as the latitude and longitude coordinates of each neighborhood.

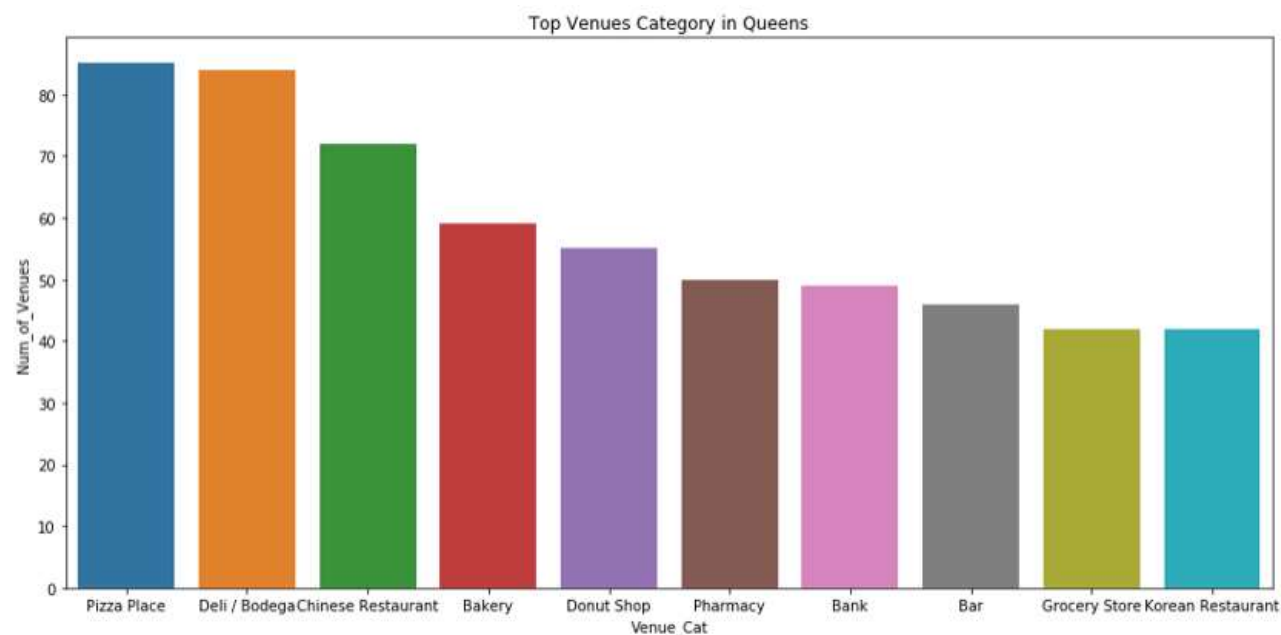
The data which is store in a key value json format. All relevant data apparently were in the feature key which is basically a list of the neighborhood. The step was to define a new variable call “Neighbourhoods_data” that include the data. Next step, was to transform the data of nested python dictionaries into a pandas dataframe, defining the empty columns and loop through the data and fill the data frame one row at time.

After successfully extracted the Queens dataset, next step was to use the Foursquare API to get all the top venues within 500m radius. The Foursquare dataset consist 100 unique venues. Each Queens neighborhood shows their uniqueness in terms of the venue category associated with it.

Now, let us examine each cluster and determine the discriminating venue categories that distinguish each cluster. Based on the defining categories, you can then assign a name to each cluster.

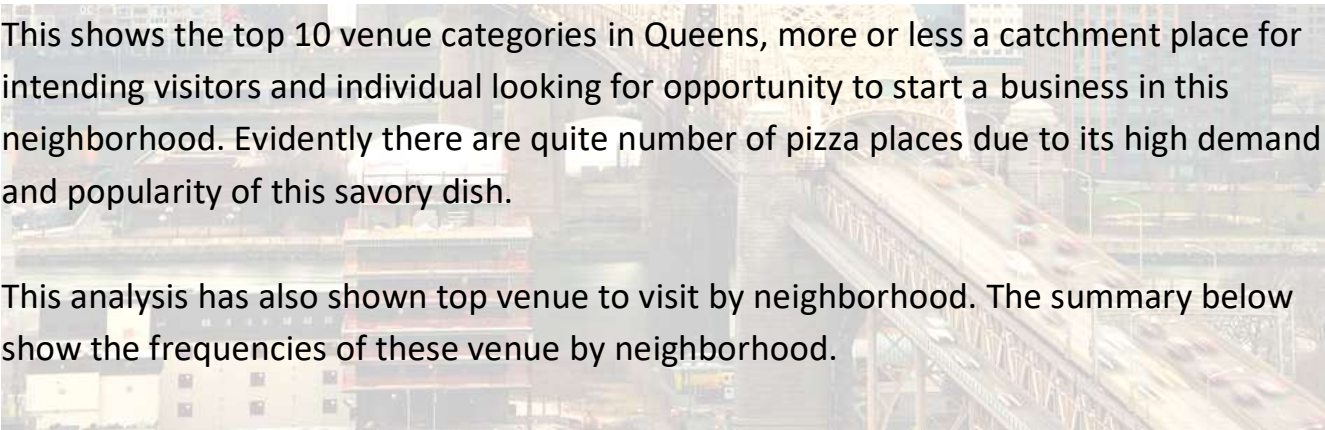
RESULTS

Visual exploratory analysis of the venue category in Queens neighborhood is shown below.



This shows the top 10 venue categories in Queens, more or less a catchment place for intending visitors and individual looking for opportunity to start a business in this neighborhood. Evidently there are quite number of pizza places due to its high demand and popularity of this savory dish.

This analysis has also shown top venue to visit by neighborhood. The summary below show the frequencies of these venue by neighborhood.



----Brookville----			----East Elmhurst----			----Far Rockaway----		
	venue	freq		venue	freq		venue	freq
0	Deli / Bodega	1.0	0	Donut Shop	0.22	0	Deli / Bodega	0.11
1	Yoga Studio	0.0	1	Ice Cream Shop	0.11	1	Pizza Place	0.09
2	Nail Salon	0.0	2	Coffee Shop	0.11	2	Chinese Restaurant	0.09
3	Other Great Outdoors	0.0	3	Rental Car Location	0.11	3	Grocery Store	0.06
----Rockaway Beach----			----Somerville----					
	venue	freq		venue	freq			
0	Beach	0.11	0	Park	1.0			
1	Deli / Bodega	0.06	1	Yoga Studio	0.0			
2	BBQ Joint	0.04	2	Nail Salon	0.0			
3	Latin American Restaurant	0.04	3	Optical Shop	0.0			

DISCUSSION

Exploratory analysis shows that Astoria, Sunnyside Gardens, Jackson Heights, Woodside are the most vibrant neighborhoods in Queens due to the number of venues in each of this neighborhood. This apparently mean visitors and tourist are attracted to these neighborhoods. These neighborhoods among other things have high number of Museums, restaurants, ice cream shops, gyms and gourmet.

However, some neighborhood shows very low number of venues for anyone intending to visit Queens, but this analysis has shown it's an opportunity to create business for such neighborhood.

Examples of these neighborhood are; somerville, Brookville, Baywaters, Malba to name a few.

The population of these neighborhoods are on the rise and certainly and new venues like Butchery, Dessert and ice cream venues, Credit centers, Chinese restaurants, middle east restaurants and gift shops to serve the growing population of these neighborhood.

CONCLUSION

This analysis will be very important for individuals who intend to relocate or visit Queens for either business or pleasure but would like to identify the right neighborhood in the borough of. In addition, tourists and real estate agents can benefit from this analysis. The results can be further improved on if neighborhood attribute information is available for all the data instance of type of business, cost of doing business data from other years is also analyzed and average cost of houses in each of the neighborhood is put into consideration as part of the analysis.

