Best Neighborhoods of NYC with Maximum Venues

By:
Ashkan Nejati

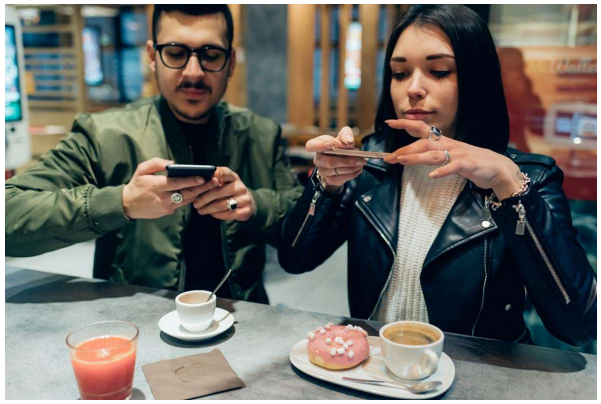
November 2019

Contents

- Introduction
- Data Acquisition and Cleaning
- Exploratory Data Analysis
- Predictive Modeling
- Results & Discussion

Introduction

- Problem Background:
  - ✓ *New York City is the most populous city in US*
  - ✓ *Tourism is a vital industry for New York City*
- Problem Description:
  - ✓ *Making the NYC the first city destination for the tourists*
  - ✓ *Invite the most famous Instagram Influencers to New York*
  - ✓ *Finding the location with maximum venues in minimum distances*



Data Acquisition and Cleaning

- Data Sources:
  - ✓ Using link https://geo.nyu.edu/catalog/nyu_2451_34572 for coordinates
  - ✓ Foursquare API to explore neighborhoods in New York City
- Data Cleaning:
  - ✓ Data scraped from multiple sources were combined into tables

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585
5	Bronx	Kingsbridge	40.881687	-73.902818
6	Manhattan	Marble Hill	40.876551	-73.910660
7	Bronx	Woodlawn	40.898273	-73.867315
8	Bronx	Norwood	40.877224	-73.879391
9	Bronx	Williamsbridge	40.881039	-73.857446

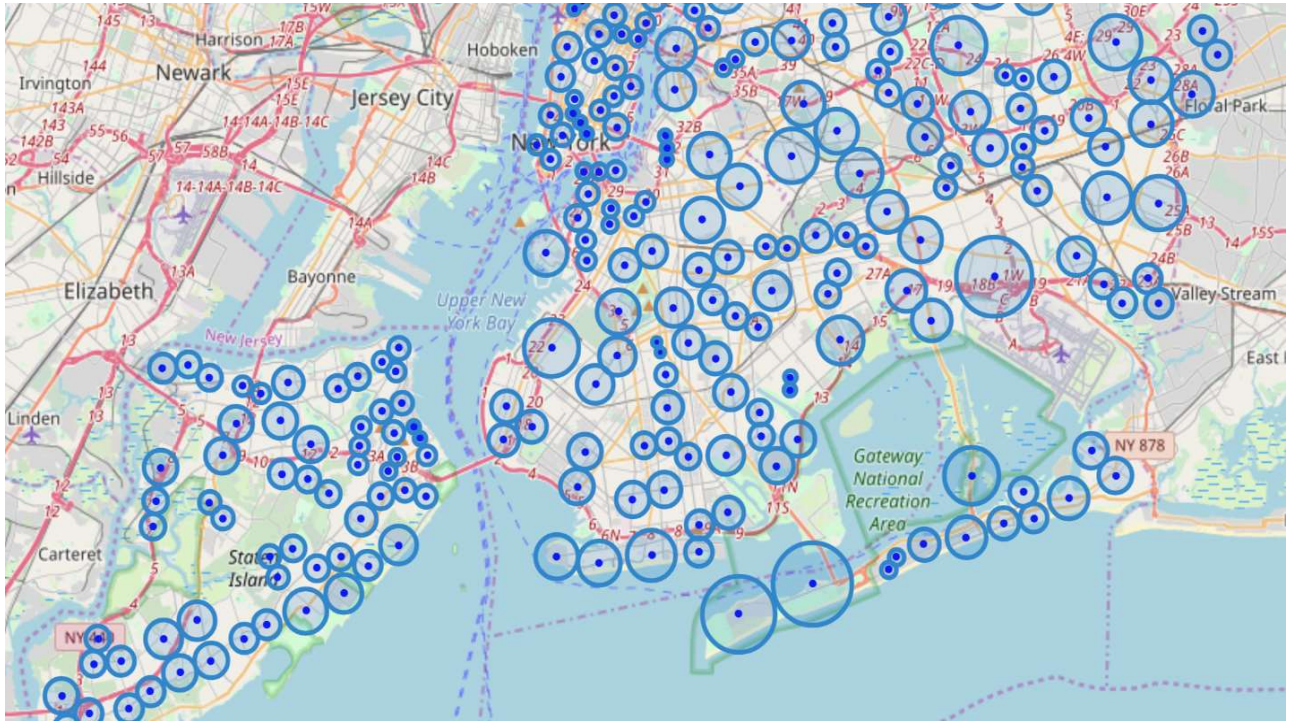
Data Acquisition and Cleaning

- Data Cleaning:
- ✓ Using foursquare API for finding different kind of venues
- ✓ Cleaning the json file and structure it into a pandas data frame

	name	categories	lat	lng
0	Fish & Ting	Caribbean Restaurant	40.885539	-73.829151
1	Cozy Cottage Restaurant	Diner	40.886332	-73.827616
2	Mario's Pizza	Pizza Place	40.888628	-73.831260
3	Dyre Fish Market	Seafood Restaurant	40.889318	-73.831453
4	Taco Bell	Fast Food Restaurant	40.883029	-73.824901
5	Dyre Deli Grocery	Deli / Bodega	40.888235	-73.831282
6	HomeGoods	Furniture / Home Store	40.890814	-73.820849
7	Smashburger	Burger Joint	40.890172	-73.820584
8	St. Paul's Church National Historic Site	Historic Site	40.893482	-73.825328
9	Dunkin'	Donut Shop	40.885384	-73.828099

Exploratory Data Analysis

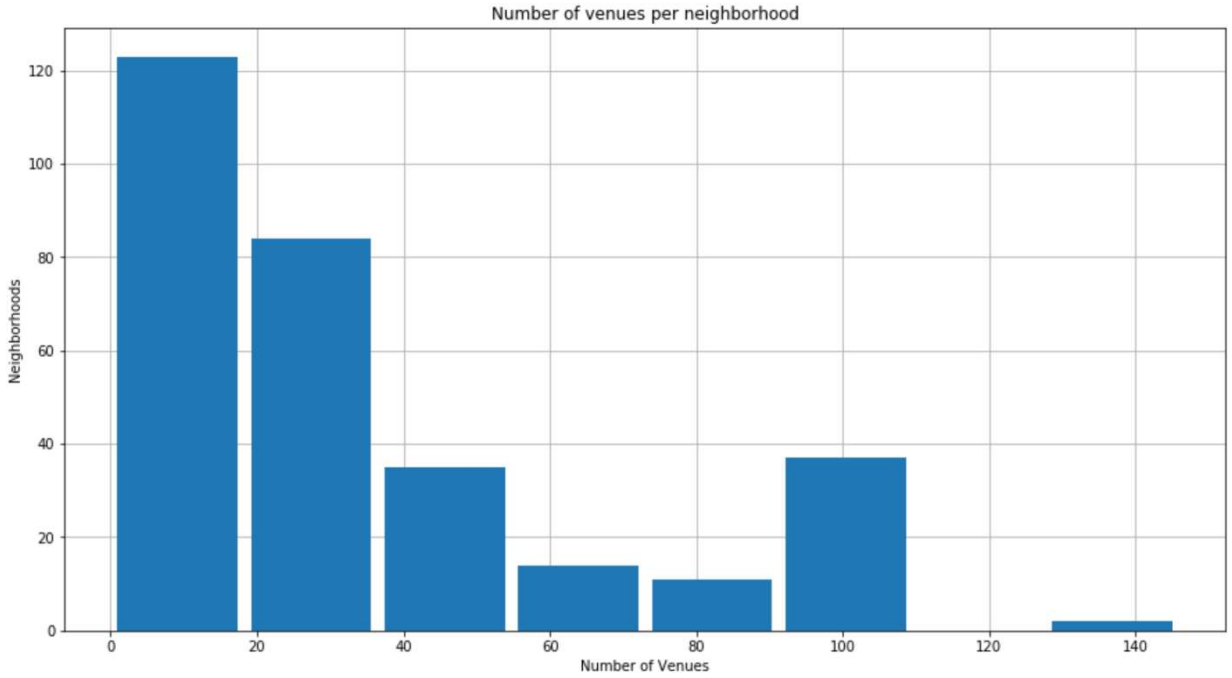
- ➢ Calculation of distances:
- ✓ Radius for each neighborhood results in a better venues search
- ✓ Find the closest points for each neighborhood
- ✓ Distance column is added to the main data frame



Exploratory Data Analysis

- Analyze venues data:
- ✓ The minimum amount of venues is 0, and the maximum is 146
- ✓ 50% of the neighborhoods presents 22 or less venues

	No. of Venues
count	306.000000
mean	34.725490
std	32.207877
min	0.000000
25%	12.000000
50%	22.000000
75%	46.000000
max	146.000000

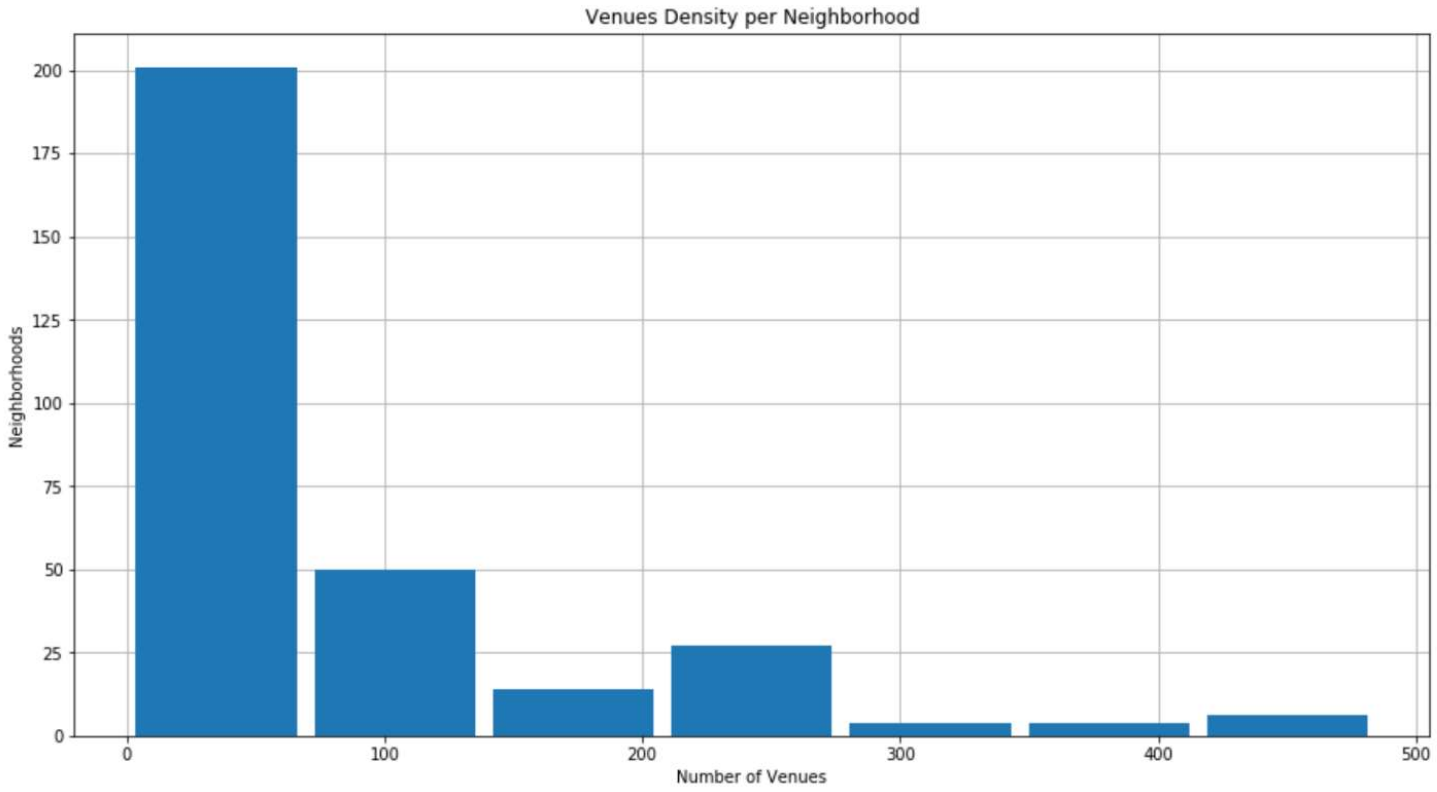


Exploratory Data Analysis

- Analyze venues data:

✓ Represent the venues per neighborhood in terms of density

✓ 50% of the neighborhoods presents a density between 0 and 45



Predictive Modeling

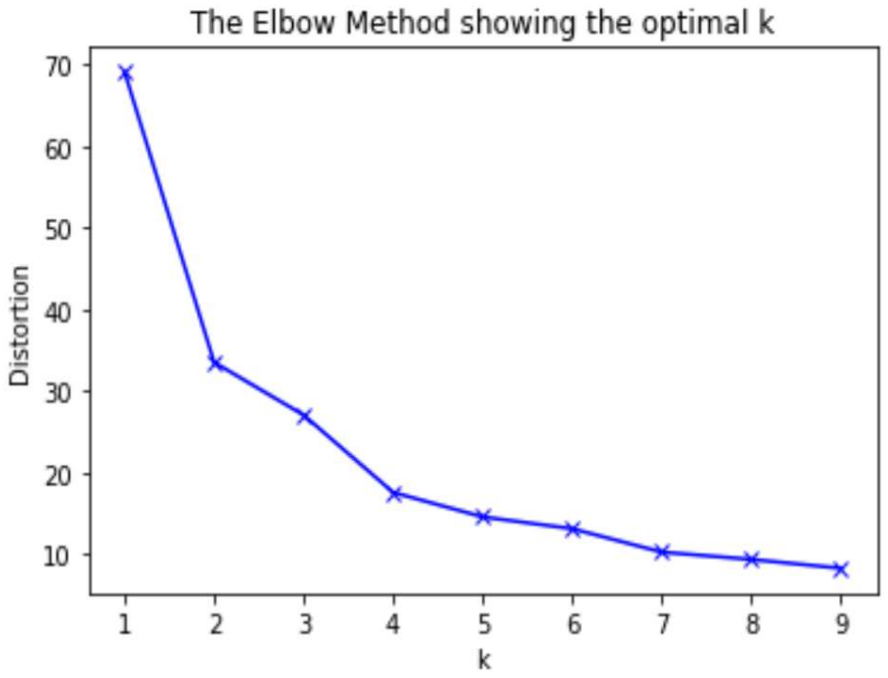
- Cluster Neighborhoods in NYC with K-means method

✓ Using the elbow method to find the optimal for k

✓ Optimal value of the number of cluster was defined as 7

✓ Applying the clustering to the neighborhood data frame

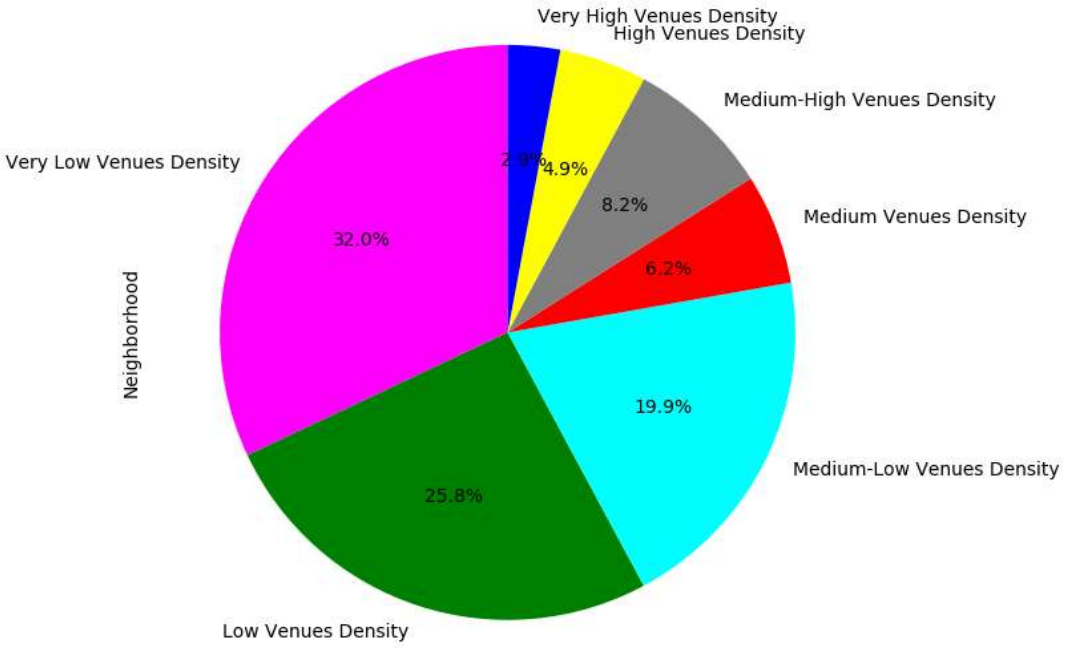
	Neighborhood	Density	Distance	Cluster
0	Allerton	90	353.0	4
1	Annadale	18	643.0	0
2	Arden Heights	5	694.0	0
3	Arlington	10	470.0	0
4	Arrochar	45	417.0	3
5	Arverne	34	515.0	3
6	Astoria	143	697.0	6
7	Astoria Heights	29	447.0	3
8	Auburndale	21	774.0	0
9	Bath Beach	70	656.0	4



Results & Discussion

- Centroids values of venues density and neighborhoods per cluster

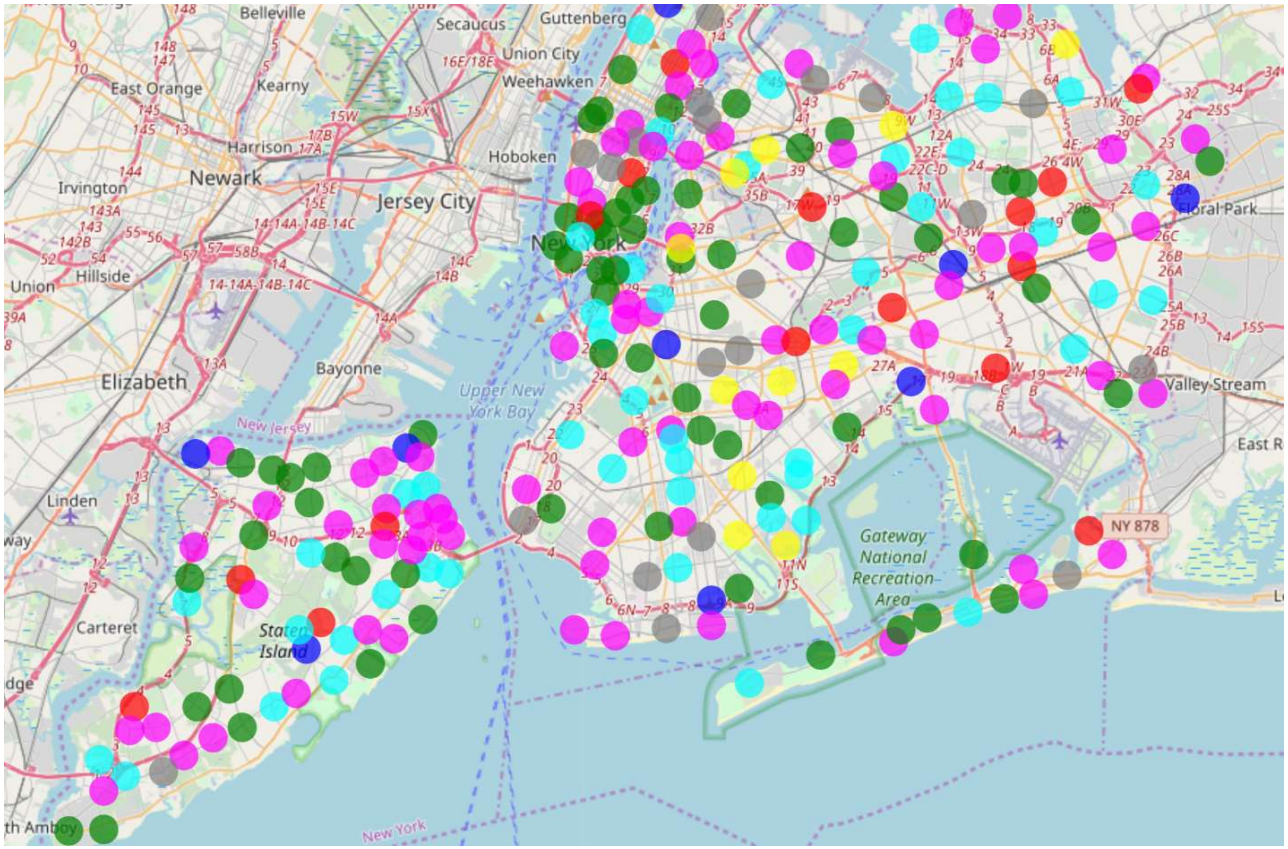
	Cluster	Centroid	Neighborhoods
0	0	14	98
1	3	38	79
2	4	74	61
3	6	129	19
4	1	213	25
5	5	278	15
6	2	442	9



Results & Discussion

- Clustered Neighborhoods:

✓ Adding the cluster column to the main data frame for neighborhoods



Results & Discussion

- Clustered Neighborhoods:
- ✓ Location of high venue density neighborhoods in the map
- ✓ Using this method for other cities to find the neighborhoods with large number of venues.

No.	Borough	Neighborhood	Latitude	Longitude	Distance	Density
0	Manhattan	Chinatown	40.71562	-73.9943	242	413
1	Manhattan	Clinton	40.7591	-73.9961	216	462
2	Manhattan	Greenwich Village	40.72693	-73.9999	266	375
3	Manhattan	Hudson Yards	40.75666	-74.0001	216	379
4	Manhattan	Little Italy	40.71932	-73.9973	212	471
5	Manhattan	Murray Hill	40.7483	-73.9783	309	478
6	Manhattan	Soho	40.72218	-74.0007	212	471
7	Brooklyn	North Side	40.71482	-73.9588	223	448
8	Brooklyn	South Side	40.71086	-73.958	206	485

Thank you for your Attention