Where to Eat? Evaluating New York's Neighborhoods by their Restaurants

Karl A. Stadnicki

1. Introduction

1.1 Background

The best neighborhood for restaurants in New York City is an often-debated topic, but few have tried using a data-driven approach to solving this problem. Restaurants in New York are a big deal and the opening of a few hot new places has the potential to increase property values in the area. Manhattan has long been considered the center of NYC food culture, but more recently several Brooklyn neighborhoods have received a lot of buzz. I attend to address this issue through a data driven quantitative approach.

1.2 Problem

Data that may contribute to determining NYC's best restaurant neighborhood will likely be found on rating websites. This data will have to be correlated with geolocation information.

1.3 Interest

Real estate agents, prospective New Yorkers and restaurateurs should be quite interested in this information. Restaurants are a major selling point for apartments in NYC. Furthermore, research shows that restaurants often pop up in clusters so that restauranteurs may profit from locating close to other successful establishments.

2. Data acquisition and cleaning

2.1 Data sources

Restaurant location data was pulled using the Four-Square Places API. The API allows users to pull data about various venues including their location and category. To complement the API datasets, I scraped TripAdvisor rating data for the top 100 restaurants in NYC.

2.2 Data cleaning

Data scraped from Trip Advisor needed to be collated with the restaurant data from the Four Square API. This necessitated extensive cleaning to match restaurant names in instances where the same restaurant was listed under two different names (for instance Paulie Gee's in one

dataset and Paulie Gee's Pizza in the other). If a highly ranked restaurant in the TripAdvisor data was not included in the Four-Square data, I dropped the restaurant and pulled another one from further down the list. Restaurants were identified by their names and identified as ranked (in the top 100) or unranked using the TripAdvisor data.

3. Methodology

3.1 Exploratory Data Analysis

I began by exploring the data in the Four-Square API. I sought to create a search query that would capture the largest possible pool of restaurants in New York attempting several different queries and assessing the size of the resulting data frame. After evaluating the restaurant data, I proceeded to identify rating data. I initially attempted to pull this data from the Four-Square API, but, I quickly exceeded daily limits on premium data pulls. At this point, I decided to utilize TripAdvisor data for the restaurant ratings choosing to focus on the top 100 restaurants as rated by TripAdvisor's algorithm (which incorporates the quantity and rating of user reviews). Collating this data with the Four-Square API data allowed me to count the number of top 100-level restaurants in each neighborhood yielding the following result:

Neighborhood	No. of Top 100 Restaurants
Greenwich Village	7
West Village	7
Soho	7
Noho	6
Midtown South	6
Flatiron	5
Midtown	5
Chelsea	4
Murray Hill	4
East Village	4
Lenox Hill	4
Turtle Bay	3
Park Slope	3
Civic Center	3
Tribeca	2
Clinton	2
North Side	2
Carnegie Hill	2
Upper West Side	2
Upper East Side	2
Sutton Place	2
South Side	2
Tudor City	1
Manhattanville	1
Chinatown	1
Financial District	1
Concord	1
Hudson Yards	1
Little Italy	1
Fulton Ferry	1
Gramercy	1

Hunters Point	1
Yorkville	1

This data suggests that Manhattan neighborhoods outperform neighborhoods in other parts of the city. Nevertheless, I found this data to be incomplete in that it only resolved part of the original question. I reasoned that the target audience for this report would be interested not only in the number of top-quality restaurants each neighborhood has, but also how many restaurants it has in general. This yielded the following result:

Neighborhood	**No. of Restaurant Venues**
Murray Hill	119
South Side	100
Financial District	100
West Village	100
Midtown	100
Lenox Hill	100
Greenwich Village	100
Midtown South	100
Noho	100
East Village	100
Soho	100
Clinton	100
North Side	100
Little Italy	100
Chinatown	100
Flatiron	100
Yorkville	88
Astoria	87
Chelsea	87
Clinton Hill	85
Downtown	84
Turtle Bay	84
Belmont	84
Tudor City	82
Sunnyside Gardens	80
Civic Center	78
Upper East Side	78
Bay Ridge	76
Brooklyn Heights	73
Carnegie Hill	71
Washington Heights	71
Flushing	69
Cobble Hill	68
Boerum Hill	68
Woodside	68
Fordham	68
Carroll Gardens	68
Sutton Place	66
Tribeca	65
Hamilton Heights	64
Bayside	61
Jackson Heights	60

Upper West Side	59
Fort Hamilton	58
Kingsbridge	57
Greenpoint	56
East Williamsburg	52
Bushwick	51
Lincoln Square	50
East Harlem	50
Gramercy	49
Lower East Side	49
Sunnyside	48
Park Slope	48
Inwood	48
Prospect Lefferts Gardens	47
Hunters Point	46
Long Island City	46
Rego Park	46
Hudson Yards	45
Fulton Ferry	45
Fort Greene	44
Central Harlem	44
College Point	44
Prospect Heights	43
Manhattanville	43
Ridgewood	43
Little Neck	42
Bath Beach	42
Manhattan Valley	42
Gowanus	41
Bedford Park	41
Morningside Heights	40
Ditmas Park	39
Dumbo	39
Concourse Village	39
Ravenswood	38
Williamsburg	36
Homecrest	35
Mill Basin	35
Prospect Park South	35
Kingsbridge Heights	35
Bensonhurst	33
Norwood	33
New Dorp	32
Kew Gardens	31
Concourse	30
Gravesend	30
Morris Park	30
West Brighton	30
Rockaway Beach	30
Jamaica Center	29
Bulls Head	29
Pelham Bay	29
Battery Park City	29
Sunset Park	29

Jamaica Hills	29
Kensington	28
Cypress Hills	28
Red Hook	28
Maspeth	28
Brighton Beach	28
High Bridge	28
Oakland Gardens	28
Ocean Hill	27
Stapleton	27
Elmhurst	27
Windsor Terrace	27
Mount Eden	27
City Line	27
Van Nest	27
Far Rockaway	26
Rosebank	26
Manhattan Terrace	25
Eltingville	25
Forest Hills	24
Howard Beach	23
Ozone Park	23
Parkchester	23
Douglaston	23
Allerton	23
Dongan Hills	23
Grant City	23
Rockaway Park	22
University Heights	22
St. George	22
Borough Park	22
Lefrak City	21
Melrose	21
North Riverdale	20
Georgetown	20
Corona	20
East New York	20
Flatbush	20
Mott Haven	20
Morrisania	20
Westchester Square	19
Forest Hills Gardens	19
East Tremont	19
Mount Hope	19
Woodhaven	19
Schuylerville	18
Vinegar Hill	18
Unionport	18
Bedford Stuyvesant	18
Ocean Parkway	18
Sheepshead Bay	18
Richmond Hill	18
Broadway Junction	18
Old Town	18

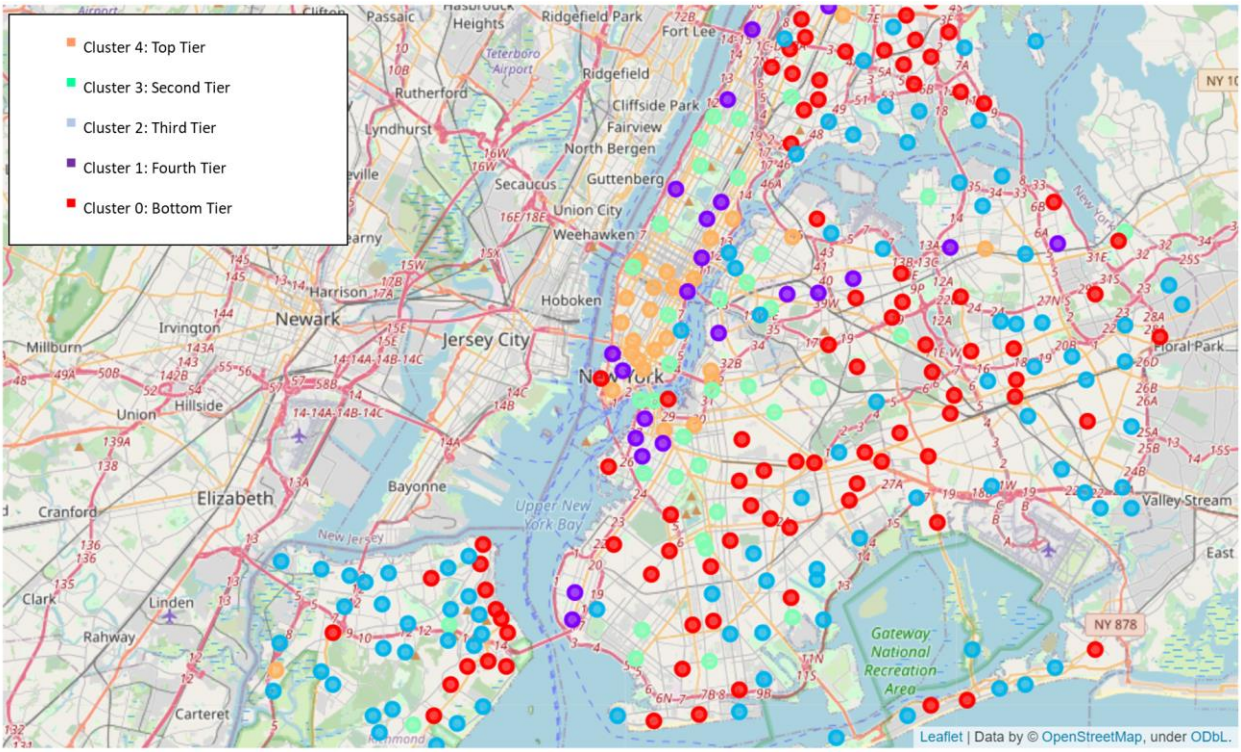
North Corona	17
Tompkinsville	17
Coney Island	17
New Lots	17
Great Kills	17
Remsen Village	16
Steinway	16
Erasmus	16
Bellerose	16
Crown Heights	16
Arrochar	16
Bay Terrace	16
Wingate	16
Weeksville	16
Woodlawn	15
Rugby	15
St. Albans	15
Hillcrest	15
Pelham Parkway	15
Queensboro Hill	15
Claremont Village	15
Edgewater Park	15
Annadale	15
Clifton	15
Middle Village	15
Marble Hill	14
Pelham Gardens	14
Grasmere	14
Eastchester	14
Shore Acres	14
Kew Gardens Hills	14
New Springville	13
Rossville	13
Olinville	13
Castleton Corners	13
Bronxdale	13
Roosevelt Island	12
Pleasant Plains	12
City Island	12
Brownsville	12
Queens Village	12
Port Richmond	12
Blissville	11
Glen Oaks	11
East Elmhurst	11
Flatlands	11
Highland Park	11
West Farms	10
Manor Heights	10
Auburndale	10
Co-op City	10
Longwood	10
Midwood	10
Elm Park	10

Woodrow	10
Charleston	10
Hunts Point	9
Baychester	9
Astoria Heights	9
New Dorp Beach	9
Soundview	9
Utopia	9
Madison	9
Belle Harbor	8
Queensbridge	8
South Ozone Park	8
Cambria Heights	8
Port Morris	8
Gerritsen Beach	8
Arverne	8
Midland Beach	8
Floral Park	8
Bellaire	8
Rochdale	7
Canarsie	7
Spuyten Duyvil	7
Morris Heights	7
Beechhurst	7
Prince's Bay	7
Castle Hill	7
Concord	7
Throgs Neck	7
Travis	7
Willowbrook	7
South Jamaica	7
Glendale	7
Stuyvesant Town	6
Richmond Valley	6
Lindenwood	6
Springfield Gardens	6
Huguenot	6
Heartland Village	6
Manhattan Beach	6
East Flatbush	6
Hollis	6
Fox Hills	6
Mariner's Harbor	6
Paerdegat Basin	5
Wakefield	5
Country Club	5
Hammels	5
Briarwood	5
Edenwald	5
Fresh Meadows	5
Rosedale	5
Starrett City	5
Pomonok	5
Greenridge	5

Randall Manor	4
Bergen Beach	4
Edgemere	4
Arlington	4
Richmond Town	4
Tottenville	4
Laurelton	4
Marine Park	4
Broad Channel	4
Sandy Ground	4
New Brighton	4
Williamsbridge	3
Lighthouse Hill	3
Dyker Heights	3
Roxbury	3
Grymes Hill	3
Holliswood	3
Silver Lake	2
Arden Heights	2
Brookville	2
South Beach	2
Egbertville	2
Clason Point	2
Westerleigh	2
Emerson Hill	2
Butler Manor	1
Sea Gate	1
Jamaica Estates	1
Graniteville	1
Malba	1
Riverdale	1
Whitestone	1
Bloomfield	1
Howland Hook	1
Park Hill	1

3.2 K-Means

Next, I conducted k-means to sort the neighborhoods into clusters. This allowed me to sort neighborhoods into tiers based on restaurant quantity and quality. This also made it easier to visualize the data. I sorted the neighborhoods into 5 clusters with cluster 4 having the best opportunities for culinary adventures and 0 having the worst. I then mapped these using folium.



4. Results

The following neighborhoods are the neighborhoods with the best food scene in Manhattan:

Neighborhood	Borough
Greenwich Village	Manhattan
West Village	Manhattan
Soho	Manhattan
Noho	Manhattan
Midtown South	Manhattan
Flatiron	Manhattan
Midtown	Manhattan
Chelsea	Manhattan
Chelsea	Staten Island
Murray Hill	Manhattan
Murray Hill	Queens
East Village	Manhattan
Lenox Hill	Manhattan
Turtle Bay	Manhattan
Clinton	Manhattan
North Side	Brooklyn
South Side	Brooklyn

Chinatown	Manhattan
Financial District	Manhattan
Little Italy	Manhattan
Yorkville	Manhattan
Astoria	Queens
Clinton Hill	Brooklyn
Downtown	Brooklyn
Belmont	Bronx

Overall, Manhattan has the best restaurants in terms of both quality and quantity in New York. Top neighborhoods in the restaurant scene tend to be centered around the lower part of Manhattan. Brooklyn is strong as well relative to other boroughs.

The strongest three neighborhoods are **Greenwich Village, West Village and Soho**.

5. Discussion

The analysis shows that the NYC food scene is large and vibrant. The strongest region for restaurants is Manhattan with a high concentration of restaurants including some of the best restaurants in the city. The search for the top neighborhood resulted in a three-way tie between Greenwich Village, the West Village, and Soho. A way to address this issue may be through applying additional criteria or expanding the TripAdvisor dataset to the top 250 restaurants.

6. Conclusion

In this study, I conducted a data-driven analysis of NYC's restaurant scene. I determined the number of restaurants in each neighborhood and the number of those restaurants that were ranked in the top 100 restaurants in NYC by TripAdvisor. I then built a classification model that groups each neighborhood into tiers based on the strength of its restaurant scenes.