

NYC's restaurants analysis

NYCDSA Web Scraping project

Leonardo Quintero
Apr 2019



Is there a relation between restaurant's prices and positive reviews broken down by cuisine?

Potential interest target

- People looking for good restaurants in town

Purpose of the Project

- Make it easy to decide which restaurants to visit and which ones to avoid in NYC.

Questions

- Does a high price guarantee quality of food?
- Are restaurants in higher end neighborhoods mean higher quality or better experiences?
- Is there a predominance of cuisine type by neighborhood?
- What neighborhoods are worth visiting in order to have a good culinary experience?



Web Scraping

Tripadvisor.com with Scrapy

Obtaining the URL from the first page “next” button and creating a for loop to iterate over all subsequent pages

avg	cuisine	name	price	rank	reviews
0 4.5	Asian,Restaurants	Taiyaki NYC	\$	1 of 10,534 Restaurants in New York City	300
1 4.5	American,Steakhouse,Latin,South American,Glute...	K Rico	-	1 of 530 Quick Bites in New York City	2,562
2 4.5	Asian,Korean,Gluten Free Options,Restaurants,D...	Jungsik		2 of 530 Quick Bites in New York City	247
3 4.5	Italian,Pizza,Vegetarian Friendly,Vegan Option...	B Side Pizzeria	-	2 of 10,534 Restaurants in New York City	340
4 4.5	Indian,Vegetarian Friendly,Vegan Options,Glute...	Indian Accent		3 of 10,534 Restaurants in New York City	441

The screenshot shows the TripAdvisor search interface for New York City restaurants. At the top, there's a green header bar with the TripAdvisor logo, a 'New York City' dropdown, and navigation links for 'Trips', 'Inbox', and a search bar. Below the header, the URL shows 'United States > New York (NY) > New York City > New York City Restaurants'. The main title is 'Restaurants in New York City'. On the left, there's a sidebar with a map, search filters ('Your Selections' including 'Restaurants' and 'Quick Bites'), and categories like 'Establishment Type' (checkboxes for Restaurants, Quick Bites, Dessert), 'Reservations' (checkbox for Online Reservations), and 'Cuisines & Dishes'. The main content area displays several sections of restaurant results, each with a thumbnail image, name, price range, and review count. For example, the 'Moderately Priced' section includes 'Piccola Cucina Osteria' (3.5 stars, 1,847 reviews), 'Piccola Cucina' (3.5 stars, 630 reviews), 'SottoCasa Pizzeria' (3.5 stars, 299 reviews), and 'Song E Napule' (3.5 stars, 239 reviews).

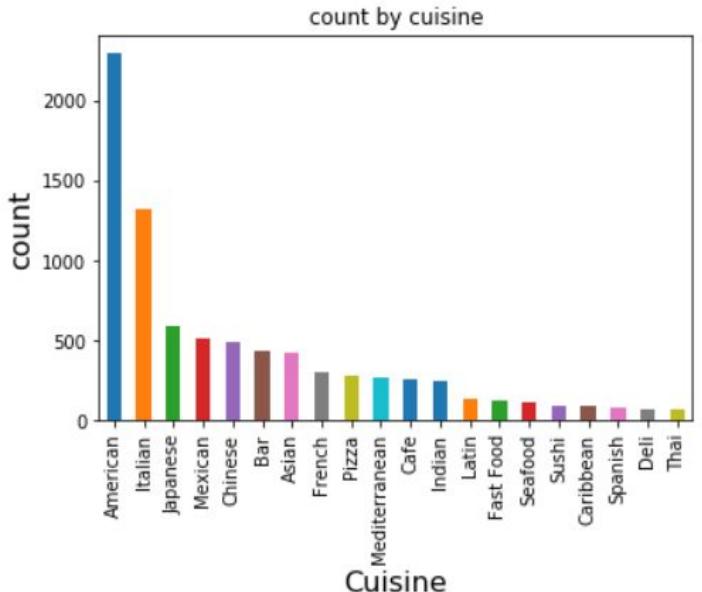


At 30 restaurants per page, 12446 restaurants were scraped including informations as shown below.

	name	cuisine	neighborhood
count	12446	9198	12446
unique	10897	88	50
top	Subway	American	Manhattan
freq	150	2298	8897

	name	cuisine	price	rank	neighborhood	reviews	avg
0	Piccola Cucina Osteria	Italian	2.0	1	Manhattan	2012.0	4.5
1	Club A Steakhouse	American	3.0	2	Manhattan	3501.0	4.5
2	SottoCasa Pizzeria	Pizza	2.0	3	Manhattan	327.0	5.0
3	Boucherie Union Square	French	3.0	4	Manhattan	641.0	5.0
4	Piccola Cucina	Italian	2.0	5	Manhattan	704.0	4.5

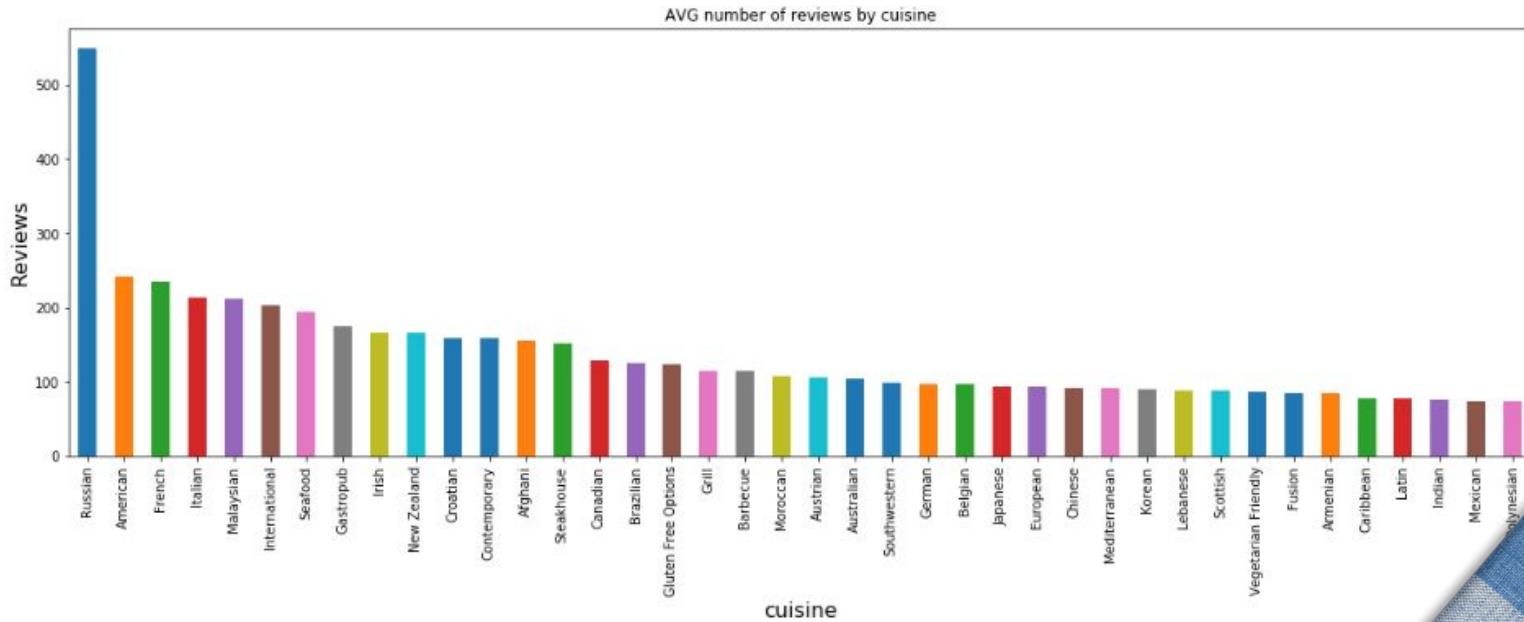
The most popular type of cuisine in NY seems to be American, followed by Italian, Japanese and Mexican.



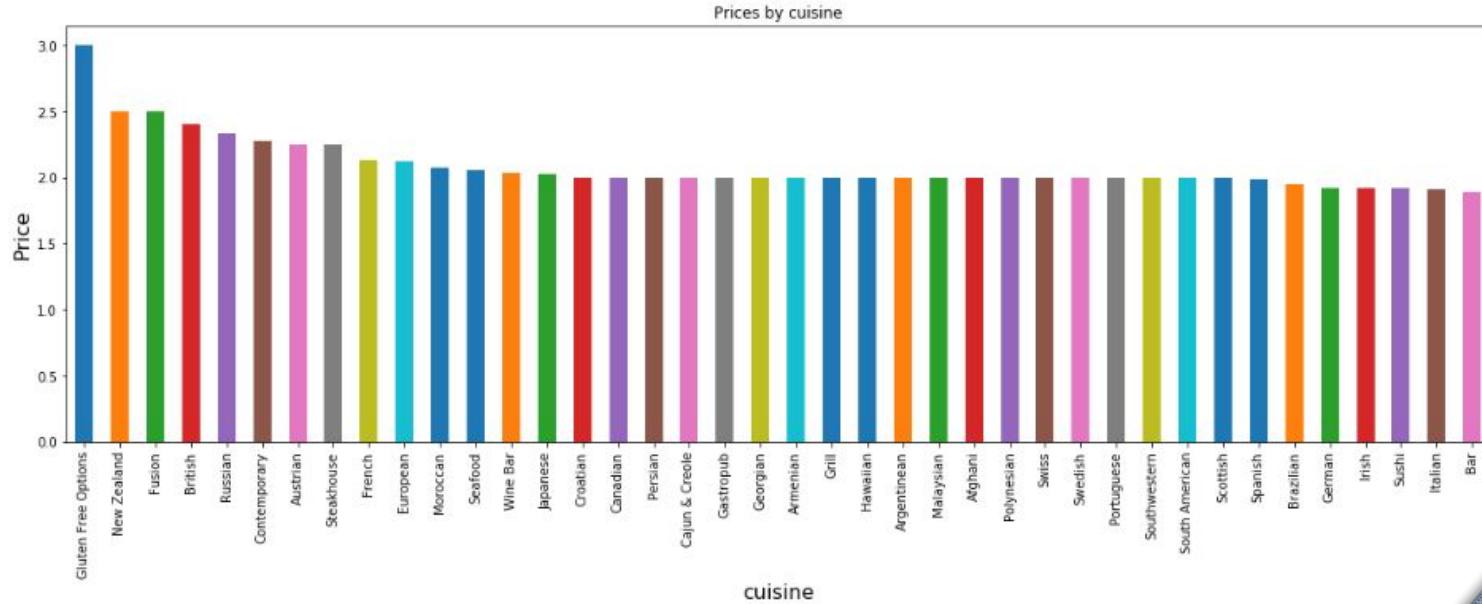
This could be cautiously used as a parameter for ethnicity segmentation in the city of NY.



But the most reviewed cuisine is Russian



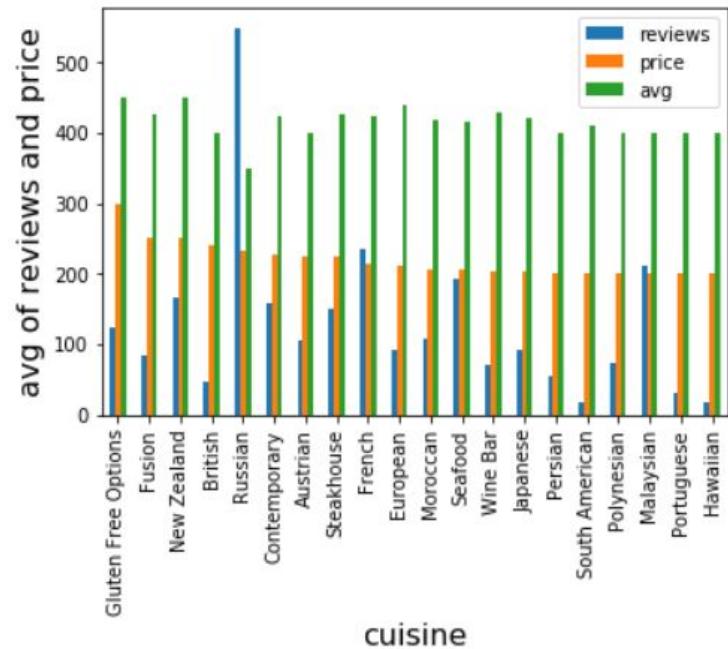
What are the most expensive types of cuisine in NY?



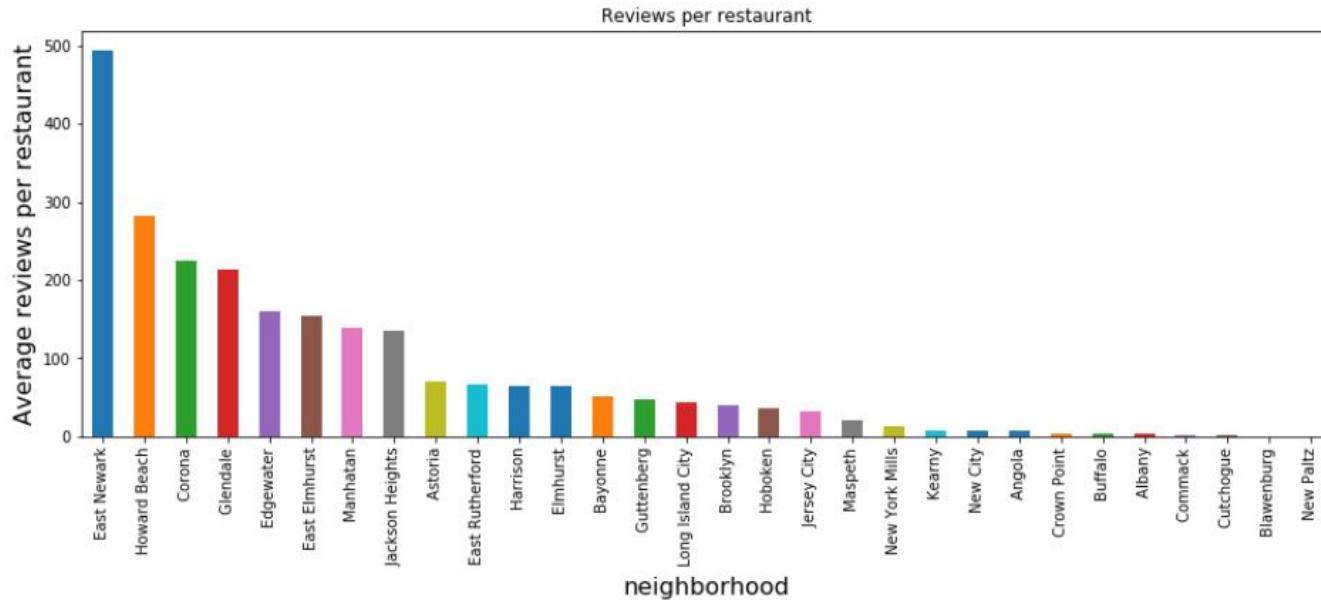
But does more expensive mean better?



Gluten free options seem to be more expensive options, less likely to be rated but with higher average rating

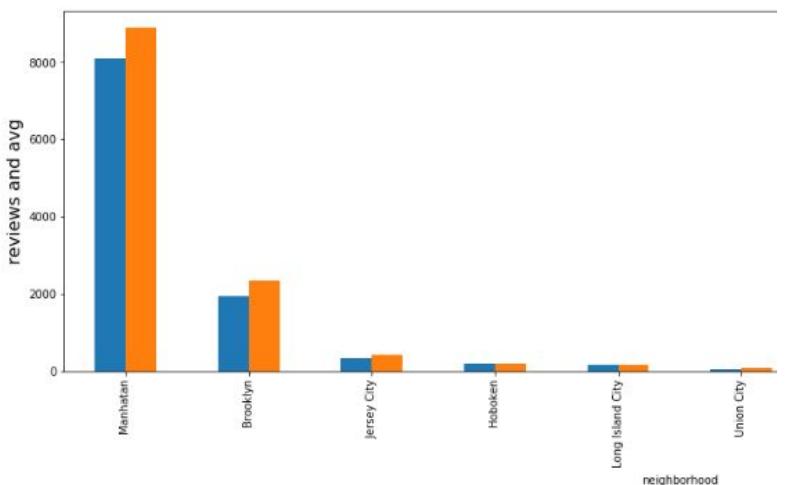


Analysing data by Neighborhood



People seem to be more likely to review restaurants in the surrounding areas, in average, with more reviews per restaurant.

With 1.122.396 reviews, Manhattan seems to be the area with the highest concentration of restaurants and reviews. Also seems to have the highest ratings.

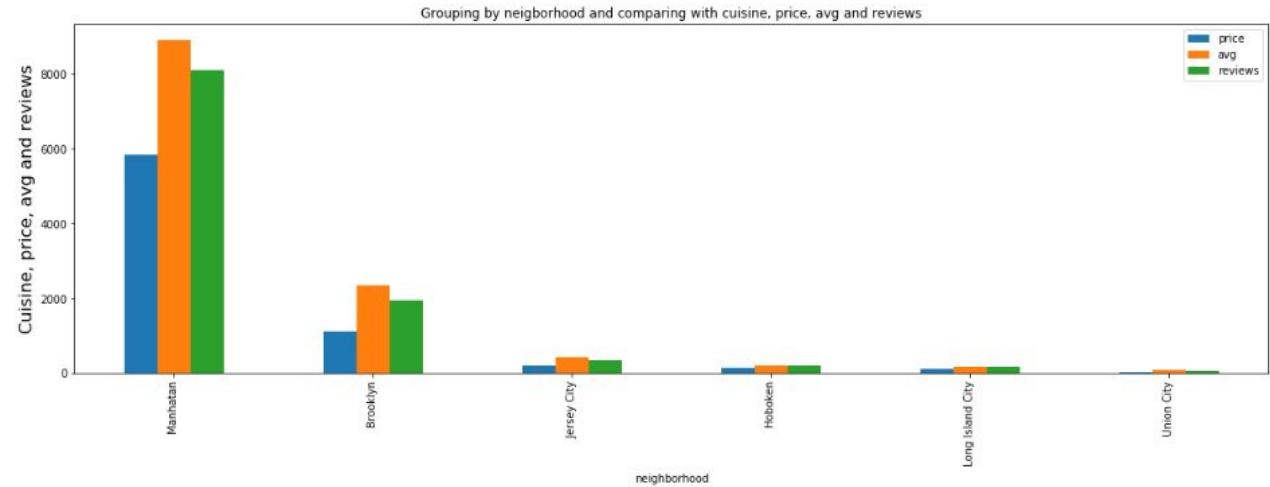


neighborhood	reviews	avg
Manhattan	1122396.0	3654250.0
Brooklyn	76810.0	992000.0
Jersey City	10810.0	168550.0
Hoboken	6735.0	82700.0
Long Island City	6632.0	68650.0
Astoria	5654.0	35300.0
Union City	655.0	33850.0
Sunnyside	1070.0	26150.0
Secaucus	3097.0	11700.0
North Bergen	1385.0	8700.0



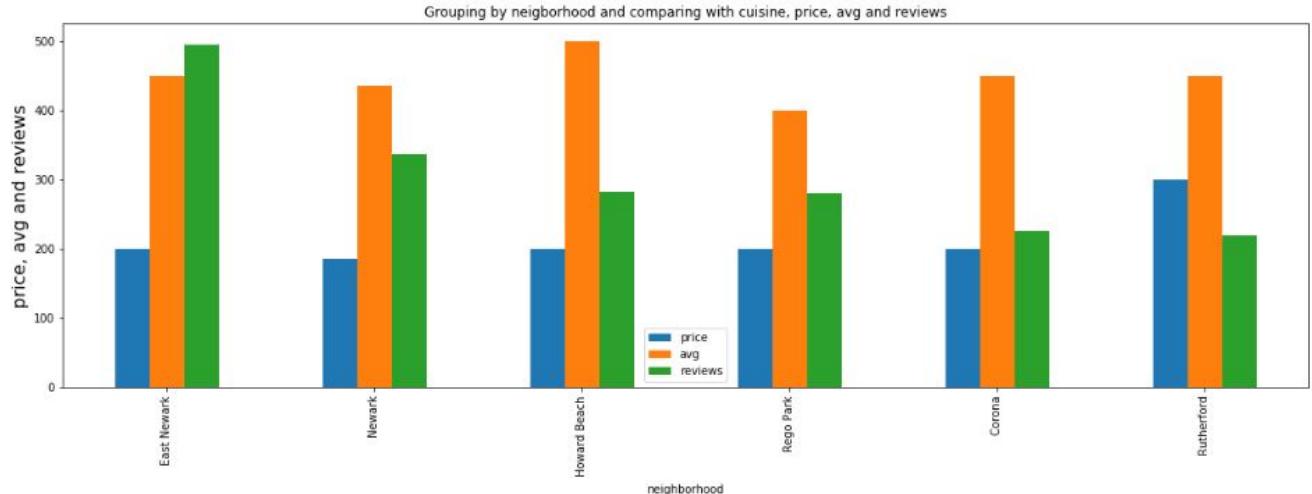


Manhattan seems to have a lower correlation between prices and average rates compared to the rest of neighborhoods





However, in average, the highest concentration of reviews, low prices and best ratings are away from manhattan.



Conclusion

- 12,446 restaurants scraped from NYC
- Scrapped data from trip advisor's NYC area, including surrounding areas.
- There is a chance to select a restaurant or cuisine by cuisine
- Vegan or gluten free options are an important part of the market, probably due to being specific types of food.

- Higher prices do not guarantee better quality of food.
- There are better chances of finding a better culinary experience depending on the neighborhood you decide to dine in.

