

Capstone Project: The Battle of Neighborhoods

Peer-graded assignment

Tobias Zimmer

Introduction - Peruvian Cuisine in New York City

Peruvian cuisine is as diverse as its environment – an immense Pacific coast crashing against the mirage of the desert, leading up to some of the tallest peaks in the Andes mountains, and then falling back down into the dense, lush Amazon. Over centuries, the original inhabitants of Peru and their vast civilizations created an array of dishes from this tremendous biodiversity and the cuisine has continued to evolve with the myriad of cultures who have come to call the country home – Spanish, Japanese, Chinese, African, and others.

Consider the famous Peruvian dish *Lomo Saltado*, for example. Lomo saltado is a popular, traditional Peruvian dish, a stir fry that typically combines marinated strips of sirloin (or other beef steak) with onions, tomatoes, french fries, and other ingredients; and is typically served with rice. The dish originated as part of the chifa tradition, the Chinese cuisine of Peru, though its popularity has made it part of the mainstream culture.

With all that it has to offer, it is no wonder that the Peruvian cuisine has successfully established itself in a city like New York. However, it is still a small branch of the culinary world. Therefore, we will have a look at Peruvian restaurants in New York city and explore where the best Pisco and Ceviche can be found.

Data

We will use the following data sources:

1. The **New York City data** (https://cocl.us/new_york_dataset) including boroughs and neighborhoods. We will use this dataset to get a list of all boroughs and neighborhoods in New York City and to get their respective geographical coordinates (latitude, longitude)
2. The **Foursquare API** to obtain venues as well as venue details. First, we will use the Places API to find all venues around given coordinates, e.g. the location of a neighborhood. Second, we will download venue details like the ratings or likes of a venue for further analysis.

Five questions we want to answer with our analysis:

1. How many Peruvian restaurants exist in New York City?
2. Which borough and which neighborhood of New York City has the highest number of Peruvian restaurants?
3. Which is the best Peruvian restaurant in New York City?
4. Which borough and which neighborhood has the best Peruvian restaurants?
5. Which borough has the highest density of top-rated Peruvian restaurants?

Methodology:

1. We start with defining the functions to request the required data.
2. We then collect the list of all boroughs and neighborhoods in New York City.
3. We identify the Peruvian restaurants in each neighborhood using the Foursquare API.
4. We then collect venue information (ratings) for all identified restaurants.
5. We combine geographical data (step 1) with venue data (step 3) to find and map the best areas for Peruvian cuisine.

1) Import packages and define functions for API requests

We define three functions:

- ... to get the coordinates of a given location using geopy:
- ... to download the boroughs and neighborhoods data from the cocl.us dataset:
- ... to get venue information for given coordinates from FourSquare:

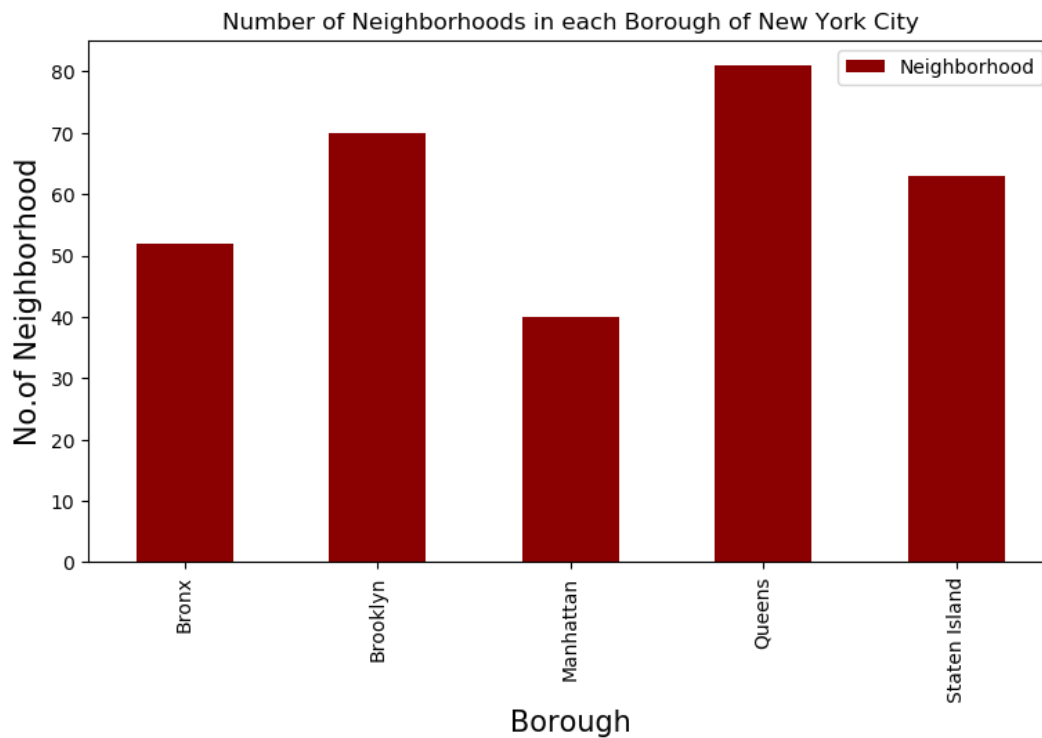
2) Get information about boroughs and neighborhoods in New York City

Before we look for restaurants, we first collect data about the boroughs and neighborhoods in New York City.

	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

We can see that there are 306 neighborhoods in New York City.

How are these 306 neighborhoods distributed among the boroughs? We group the data and plot the distribution:



We can see in this diagram that Queens has the highest number of neighborhoods.

3) Get information about Peruvian Restaurants in New York City

As we have the list of all neighborhoods, we can now count the Peruvian restaurants in each neighborhood:

	Borough	Neighborhood	ID	Name
0	Bronx	Mott Haven	4bb28791715eef3b9f5e85bb	Pio Pio
1	Bronx	Port Morris	4bb28791715eef3b9f5e85bb	Pio Pio
2	Brooklyn	Bushwick	4c117b66416620a12136d9e4	Inca Chicken
3	Brooklyn	Gowanus	4db8c0fcfa8c978590c0f6f1	Surfish
4	Brooklyn	Bath Beach	4e3c94bc483b04e17a8b15a6	La granja

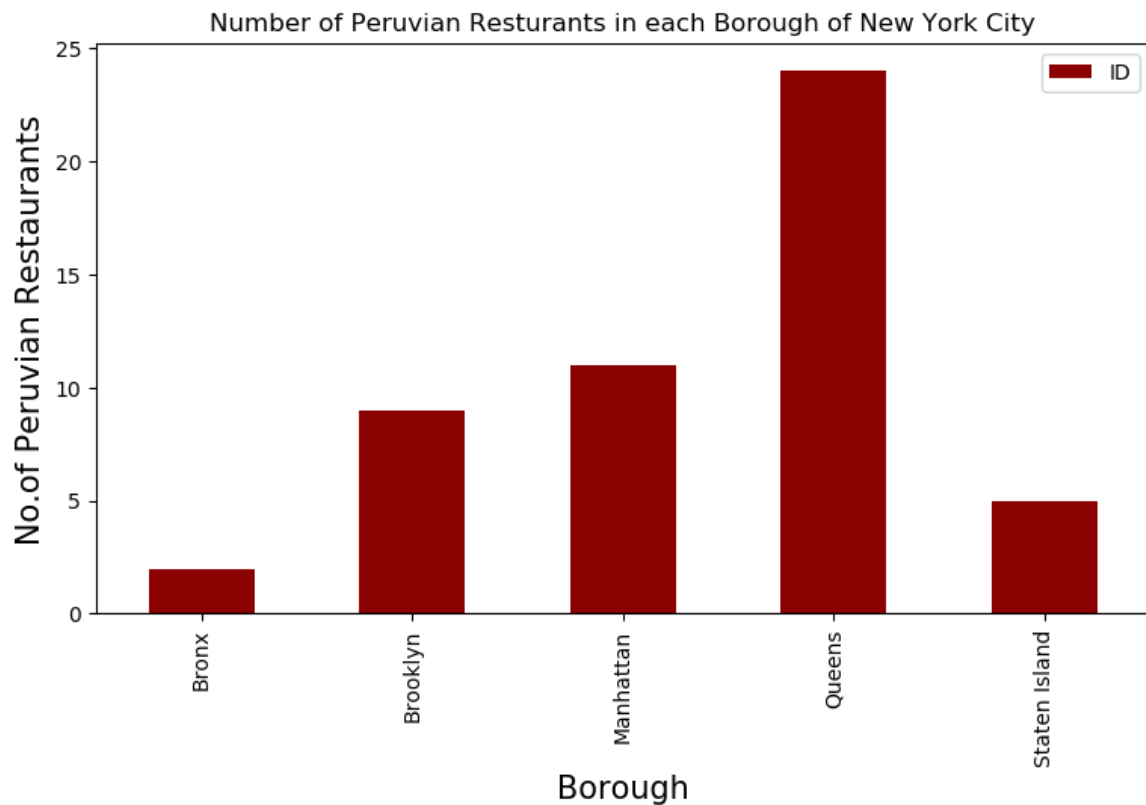
When requesting data via the Foursquare API, sometimes a KeyError occurs. As a result, only a part of the full dataset is obtained. Therefore, the results of a successful download are saved to a CSV file so that they are available even if the API request fails.

	Borough	Neighborhood	ID	Name
46	Queens	Blissville	51ec2df8498ee2a4dc8ee843	Don Pollo II
47	Brooklyn	Vinegar Hill	53d1a481498e1fac7c37f675	Pollo d'Oro
48	Brooklyn	Homecrest	4c12cbd7b7b9c928ef7ca837	Coney Island Taste
49	Manhattan	Hudson Yards	5babb7988194fc002c109112	Chirp
50	Manhattan	Hudson Yards	4b1b1f52f964a52074f823e3	Pio Pio

Question 1: How many Peruvian restaurants exist in New York City?

We can see from the dataframe that there are 51 Peruvian restaurants in New York City.

How are the restaurants distributed among the boroughs?

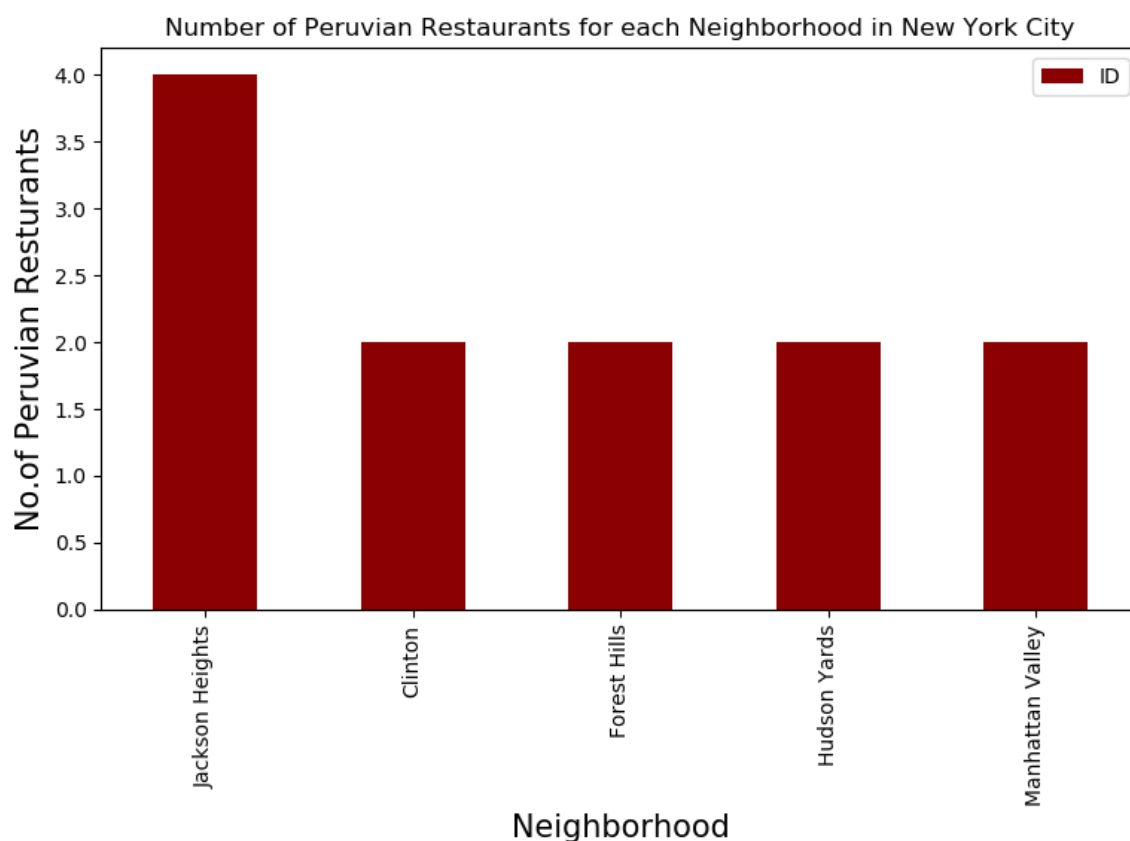


```
peru_rest_ny.groupby('Borough')['ID'].count()
Borough
Bronx          2
Brooklyn       9
Manhattan     11
Queens        24
Staten Island  5
Name: ID, dtype: int64
```

Question 2a: Which borough of New York City has the highest number of Peruvian restaurants?

We can see that the borough with the highest number of Peruvian restaurants is Queens with 24 restaurants in total.

However, which are the neighborhoods with the highest number of Peruvian restaurants?



Question 2b: Which neighborhood of New York City has the highest number of Peruvian restaurants?

We can see that there is one neighborhood, Jackson Heights, which has four Peruvian restaurants.

The four restaurants in Jackson Heights are shown below:

	Borough	Neighborhood	ID	Name
16	Queens	Jackson Heights	4b9ece55f964a520590337e3	Urubamba
17	Queens	Jackson Heights	4b8344cff964a520f0ff30e3	Pio Pio
18	Queens	Jackson Heights	4f9771e4e5e8dedda9c72330	Pio Pio to Go
19	Queens	Jackson Heights	557870be498e72ac68e380c0	Juanita's Café

4) Analyze the ratings of all restaurants

We will now analyze the ratings of all 51 Peruvian restaurants. To do so, we have to download the venue details via the FourSquare API:

	Borough	Neighborhood	ID	Name	Likes	Rating	Tips
0	Bronx	Mott Haven	4bb28791715eef3b9f5e85bb	Pio Pio	67.0	8.8	23.0
1	Bronx	Port Morris	4bb28791715eef3b9f5e85bb	Pio Pio	67.0	8.8	23.0
2	Brooklyn	Bushwick	4c117b66416620a12136d9e4	Inca Chicken	34.0	8.5	15.0
3	Brooklyn	Gowanus	4db8c0fcfa8c978590c0f6f1	Surfish	134.0	8.2	58.0
4	Brooklyn	Bath Beach	4e3c94bc483b04e17a8b15a6	La granja	6.0	6.9	6.0

When requesting data via the Foursquare API, sometimes a KeyError occurs. As a result, only a part of the full dataset is obtained. Therefore, the results of a successful download are saved to a CSV file so that they are available even if the API request fails.

We can now look for the best Peruvian restaurant. We will consider two criteria: Ratings and Likes.

First, we determine the restaurant with the best ratings:

```
peru_rest_stats_ny.iloc[peru_rest_stats_ny['Rating'].idxmax()]
Borough          Manhattan
Neighborhood      Greenwich Village
ID               5b95359481a0ea002cdce0ce
Name              Llama-San
Likes              80
Rating             9
Tips              14
Name: 12, dtype: object
```

Second, we determine the restaurant with the highest number of likes:

```
peru_rest_stats_ny.iloc[peru_rest_stats_ny['Likes'].idxmax()]
Borough          Manhattan
Neighborhood      Clinton
ID               4b1b1f52f964a52074f823e3
Name              Pio Pio
Likes             1122
Rating            8.9
Tips             329
Name: 10, dtype: object
```

Question 3: Which is the best Peruvian restaurant in New York City?

The restaurant with the highest rating is Llama-San in Greenwich Village, Manhattan. The restaurant with the highest number of likes is Pio Pio which is also located in Manhattan (Clinton).

5) Identify and plot the best areas for Peruvian cuisine in New York City

We will now combine the results from step 2) and step 3) to find the best boroughs and neighborhoods with respect to Peruvian cuisine.

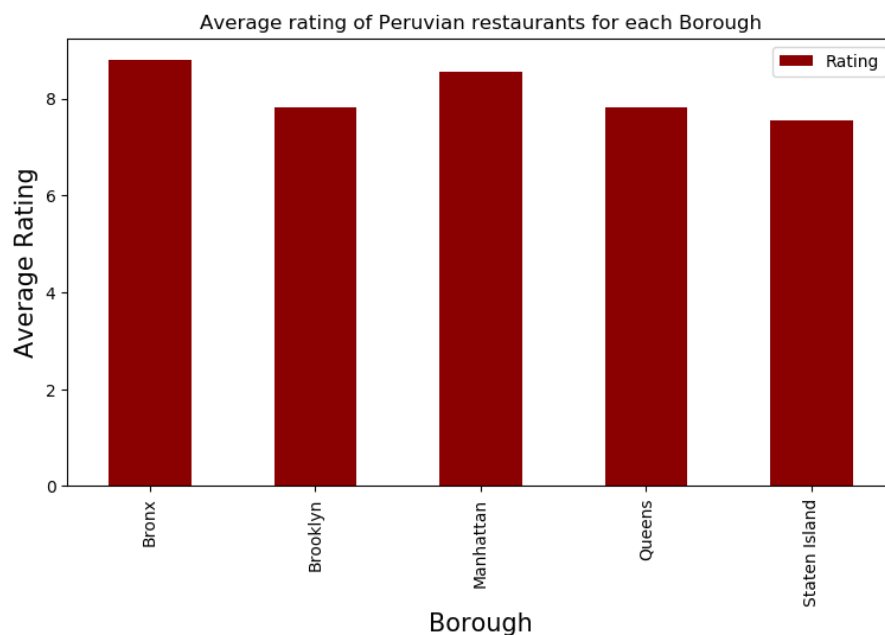
	Neighborhood	Average Rating
39	West Village	9.0
17	Greenwich Village	9.0
20	Hudson Yards	8.8
27	Mott Haven	8.8
5	Clinton	8.8
31	Port Morris	8.8
23	Little Neck	8.5
21	Hunters Point	8.5
11	Elmhurst	8.5
8	Douglaston	8.5

Question 4a: Which neighborhood has the best Peruvian restaurants?

We can see that West Village and Greenwich Village are excellent neighborhoods with regards to Peruvian cuisine as both have an average rating of 9.0.

Question 4b: Which borough has the best Peruvian restaurants?

We can see that although Queens has the most Peruvian restaurants and also the best ones, the Bronx has the highest average rating.



As we are looking for very good restaurants, we will focus now on neighborhoods with an average rating of 8.5 or higher.

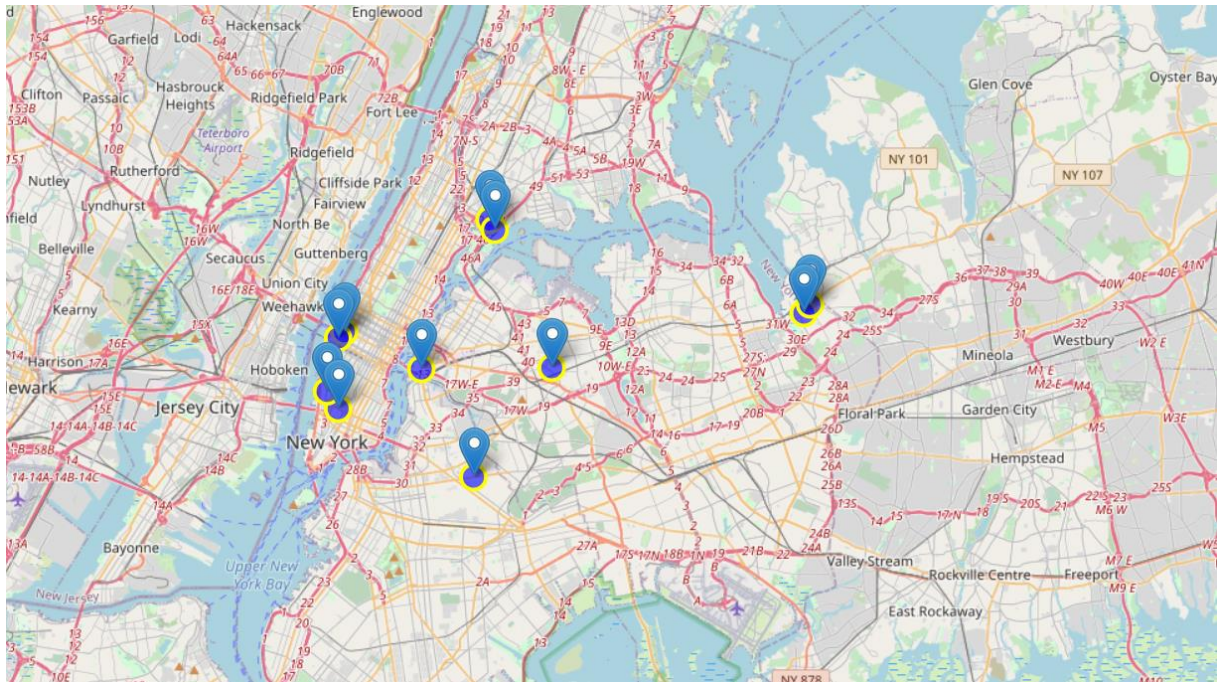
	Borough_x	Neighborhood	Latitude_x	Longitude_x	Average Rating	Label	Borough_y	Latitude_y	Longitude_y
0	Brooklyn	Bushwick	40.698116	-73.925258	8.5	Bushwick, Brooklyn(8.5)	Brooklyn	40.698116	-73.925258
1	Manhattan	Clinton	40.759101	-73.996119	8.8	Clinton, Manhattan(8.8)	Manhattan	40.759101	-73.996119
2	Queens	Douglaston	40.766846	-73.742498	8.5	Douglaston, Queens(8.5)	Queens	40.766846	-73.742498
3	Queens	Elmhurst	40.744049	-73.881656	8.5	Elmhurst, Queens(8.5)	Queens	40.744049	-73.881656
4	Manhattan	Greenwich Village	40.726933	-73.999914	9.0	Greenwich Village, Manhattan(9.0)	Manhattan	40.726933	-73.999914
5	Manhattan	Hudson Yards	40.756658	-74.000111	8.8	Hudson Yards, Manhattan(8.8)	Manhattan	40.756658	-74.000111
6	Queens	Hunters Point	40.743414	-73.953868	8.5	Hunters Point, Queens(8.5)	Queens	40.743414	-73.953868
7	Queens	Little Neck	40.770826	-73.738898	8.5	Little Neck, Queens(8.5)	Queens	40.770826	-73.738898
8	Bronx	Mott Haven	40.806239	-73.916100	8.8	Mott Haven, Bronx(8.8)	Bronx	40.806239	-73.916100
9	Bronx	Port Morris	40.801664	-73.913221	8.8	Port Morris, Bronx(8.8)	Bronx	40.801664	-73.913221
10	Manhattan	West Village	40.734434	-74.006180	9.0	West Village, Manhattan(9.0)	Manhattan	40.734434	-74.006180

To plot these neighborhoods in a map, we add the longitude and latitude from the New York data set above.

	Borough	Neighborhood	Latitude	Longitude	Average Rating
0	Brooklyn	Bushwick	40.698116	-73.925258	8.5
1	Manhattan	Clinton	40.759101	-73.996119	8.8
2	Queens	Douglaston	40.766846	-73.742498	8.5
3	Queens	Elmhurst	40.744049	-73.881656	8.5
4	Manhattan	Greenwich Village	40.726933	-73.999914	9.0
5	Manhattan	Hudson Yards	40.756658	-74.000111	8.8
6	Queens	Hunters Point	40.743414	-73.953868	8.5
7	Queens	Little Neck	40.770826	-73.738898	8.5

	Borough	Neighborhood	Latitude	Longitude	Average Rating
8	Bronx	Mott Haven	40.806239	-73.916100	8.8
9	Bronx	Port Morris	40.801664	-73.913221	8.8
10	Manhattan	West Village	40.734434	-74.006180	9.0

Finally, we plot the identified neighborhoods using Folium:



Question 5: Which borough has the highest density of top-rated Peruvian restaurants?

We can see in the map, that four highly rated restaurants can be found in Manhattan.

Summary:

We can now summarize the answers we have found to our questions:

Question 1: How many Peruvian restaurants exist in New York City?

In June 2020, there are 51 Peruvian restaurants in New York City.

Question 2: Which borough and which neighborhood of New York City has the highest number of Peruvian restaurants?

The borough with the highest number of Peruvian restaurants is Queens with 24 restaurants in total. The neighborhood with the highest number of Peruvian restaurants is Jackson Heights which is located in Queens.

Question 3: Which is the best Peruvian restaurant in New York City?

The restaurant with the highest rating is Llama-San in Greenwich Village, Manhattan. The restaurant with the highest number of likes is Pio Pio which is also located in Manhattan (Clinton).

Question 4: Which borough and which neighborhood has the best Peruvian restaurants?

The neighborhoods with the highest rating are West Village and Greenwich Village. Both are in Manhattan and have an average rating of 9. But although Queens has the most Peruvian restaurants and although the best two neighborhoods are in Manhattan, the Bronx has the highest average rating.

Question 5: Which borough has the highest density of top-rated Peruvian restaurants?

Four highly rated restaurants are located relatively close to another in Manhattan so that this borough has the highest density of top Peruvian restaurants.