Applied Data Science Capstone - The Battle of Neighborhoods

1. Introduction

São Paulo is a municipality in the Southeast Region of Brazil. The metropolis is an alpha global city (as listed by the GaWC) and the most populous city in Brazil, the Americas, the Western Hemisphere and the Southern Hemisphere. Additionally, São Paulo is the largest Portuguese-speaking city in the world. The municipality is also the world's 4th largest city proper by population. The city is the capital of the surrounding state of São Paulo, the most populous and wealthiest state in Brazil. It exerts strong international influences in commerce, finance, arts and entertainment. The name of the city honors the Apostle, Saint Paul of Tarsus. The city's metropolitan area, the Greater São Paulo, ranks as the most populous in Brazil and the 12th most populous on Earth.

São Paulo is considered the most multicultural city in Brazil. Since 1870 to 2010, approximately 2.3 million immigrants arrived in the state, from all parts of the world. The Italian community is one of the strongest, with a presence throughout the city. Of the 9 million inhabitants of São Paulo, 50% (4.5 million people) have full or partial Italian ancestry. São Paulo has more descendants of Italians than any Italian city (the largest city of Italy is Rome, with 2.5 million inhabitants).

The cuisine of São Paulo is a tourist attraction. The city has 62 cuisines across 12,000 restaurants. During the 10th International Congress of Gastronomy, Hospitality and Tourism (Cihat) conducted in 1997, the city received the title of "World Gastronomy Capital" from a commission formed by 43 nations' representatives.

As part of this project, we will explore restaurants from the main regions of São Paulo, focusing on Italian cuisine. The outcome of this project will be useful for those interested on knowing more about the gastronomic environment in the city.

We will provide answers to the following questions:

- What localities have the largest number of rated restaurants in the main regions of São Paulo ?

- What localities have the smallest number of rated restaurants in the main regions of São Paulo ?
- What localities have the best rated restaurants in the main regions of São Paulo ?
- What localities have the worst rated restaurants in the main regions of São Paulo ?
- What localities have the largest number of rated Italian restaurants in the main regions of São Paulo ?
- What localities have the best rated Italian restaurants in the main regions of São Paulo ?
- What localities have more nearby venues with Italian restaurants ?

2.   Data

For this project, we will use the following data:

- Data from restaurants situated at the main regions of São Paulo, including Locality, Cuisines, Restaurant Names, Rating, Latitude, Longitude, etc.:

> - Data source: [Zomato kaggle dataset](#). This dataset contains the required information to explore various localities of São Paulo.

- Data from nearby places in each locality of the main regions of São Paulo:

> - Data source: Foursquare API. By using this API we will get all the venues in each neighborhood.
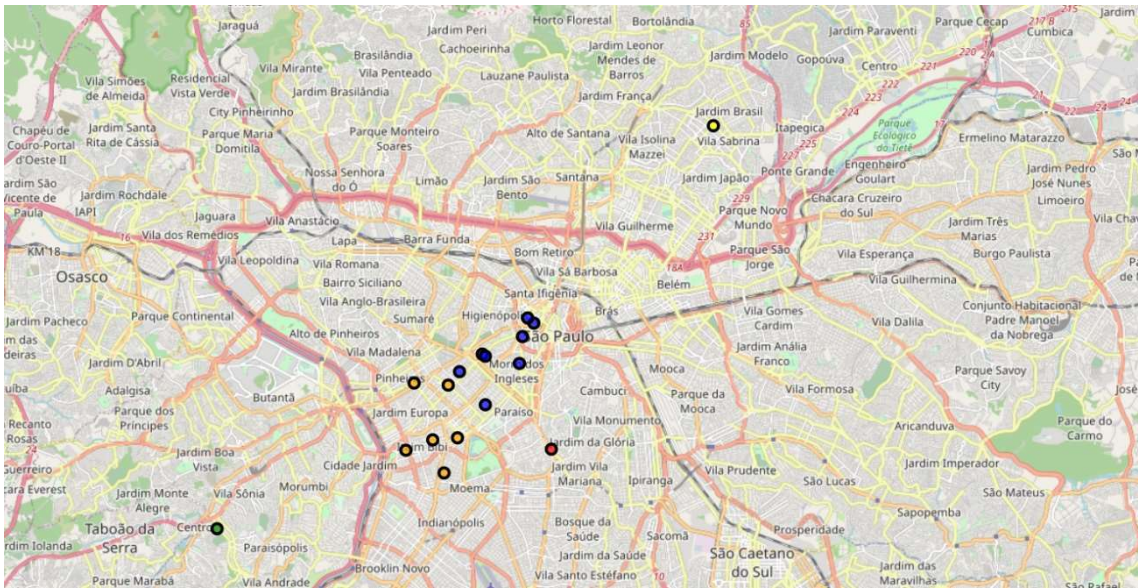
3.   Methodology

- Collect restaurants data from Zomato Kaggle dataset.

- Clean up data.

- Visualize data using Folium Python library.

- Generate plots to compare each locality with regards to restaurants.

- Group data based on locality.

- Use Foursquare API to find all venues for each neighborhood.

- Filter out all nearby venues by locality.

- Use aggregate rating from each restaurant to find the best places.

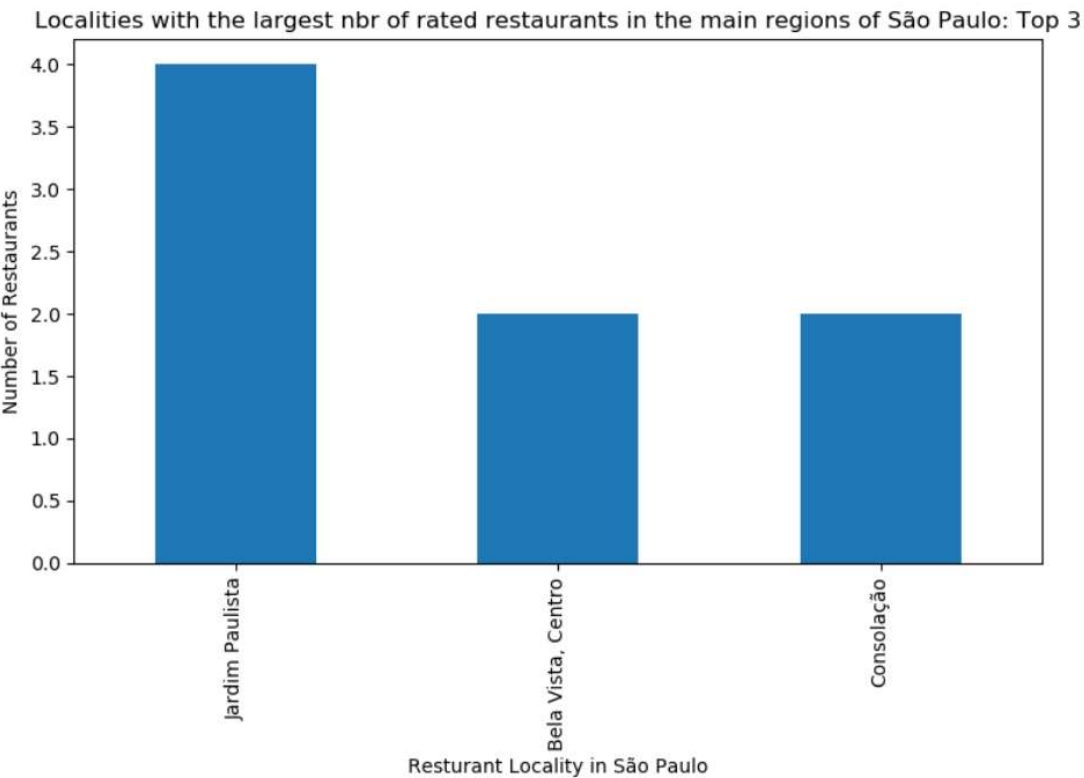- Run k-means to cluster the localities into 5 clusters.

4. Results

Restaurants by clusters

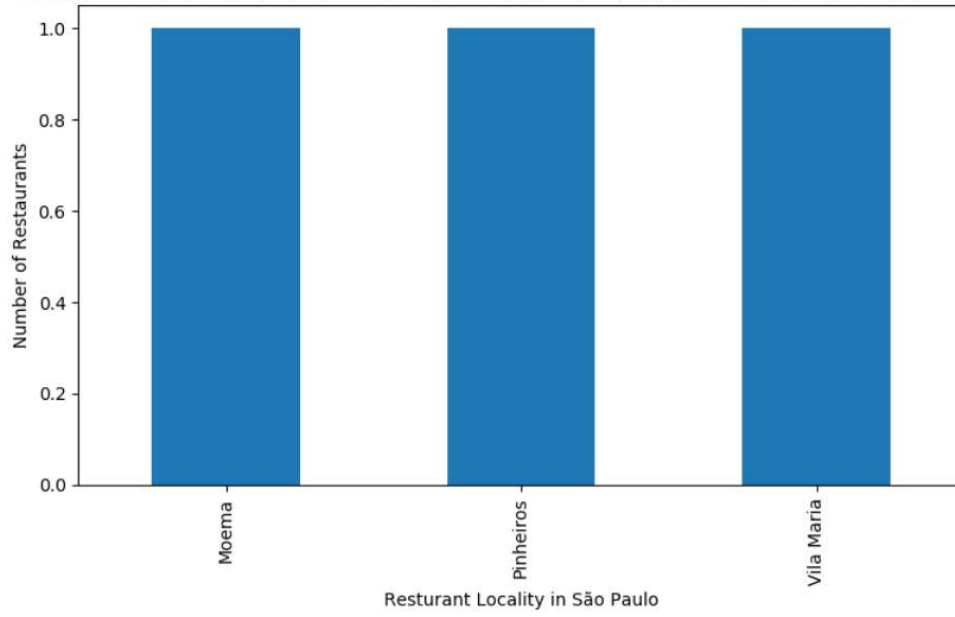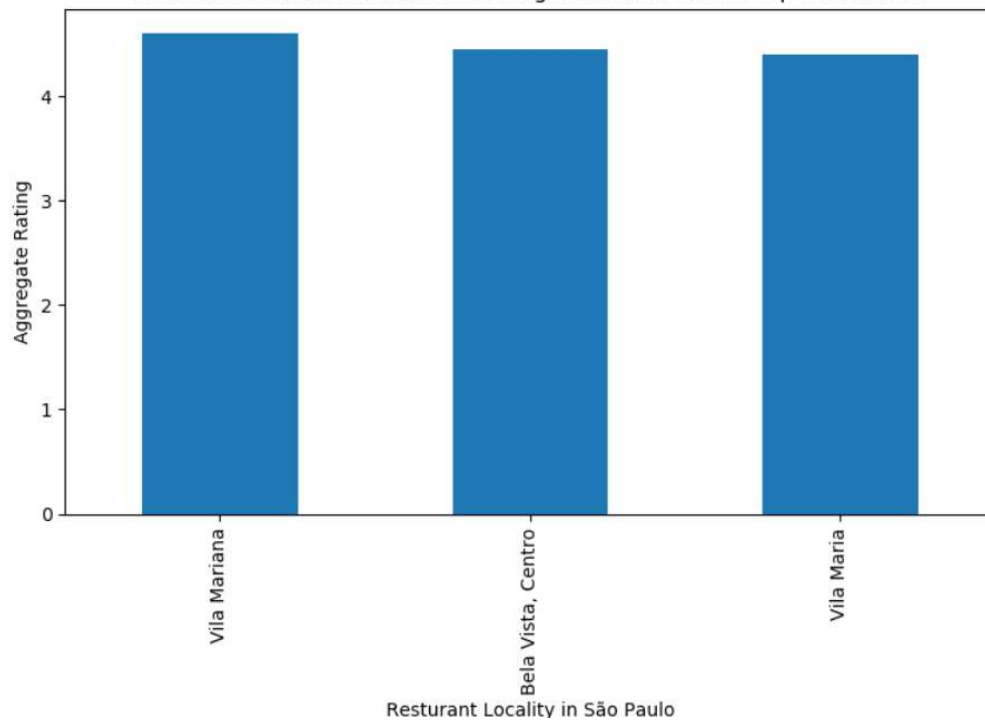|	Restaurant Name	Locality	Longitude	Latitude	Cuisines	Aggregate rating	Rating text	Votes	Cluster
0	Cantina Famiglia Mancini	Bela Vista, Centro	-46.645167	-23.550667	Italian, Pizza	4.5	Excellent	49	2
1	Templo da Carne - Marcos Bassi	Bela Vista, Centro	-46.646333	-23.559000	Steak, BBQ	4.4	Very Good	17	2
2	Gopala Hari	Consolação	-46.658667	-23.556167	Indian	3.1	Average	5	2
3	Jiquitaia	Consolação	-46.657523	-23.556710	Brazilian	4.1	Very Good	15	2
4	Skye - Hotel Unique	Jardim Paulista	-46.666851	-23.581688	Beverages, International	4.8	Excellent	59	4
5	Les 3 Brasseurs	Itaim Bibi	-46.675110	-23.582135	French, Brazilian, Beverages	4.6	Excellent	30	4
6	Red Steak & Burger	Itaim Bibi	-46.683888	-23.585324	Brazilian, Grill	3.9	Good	5	4
7	Paris 6 Classique	Jardim Paulista	-46.666041	-23.561568	French	3.4	Average	73	2
8	Kawa Sushi	Jardim Paulista	-46.657418	-23.571638	Sushi, Japanese	3.5	Good	9	2
9	A Figueira Rubaiyat	Jardim Paulista	-46.669833	-23.565500	BBQ, Grill, Brazilian	4.3	Very Good	39	4
10	Kinoshita	Moema	-46.671333	-23.592333	Sushi	3.9	Good	12	4
11	Meats	Pinheiros	-46.681333	-23.564833	Gourmet Fast Food, Burger	4.3	Very Good	68	4
12	Paribar	Republica	-46.641594	-23.546642	Brazilian, Italian	4.3	Very Good	46	2
13	Terraço Italia	Republica	-46.643425	-23.545163	Italian	4.4	Very Good	37	2
14	Esquina Mocotó	Vila Maria	-46.581672	-23.486535	Brazilian, North Eastern	4.4	Very Good	22	3
15	Veloso	Vila Mariana	-46.635667	-23.585167	Brazilian, Bar Food, Beverages	4.6	Excellent	58	0
16	Sainte Marie Gastronomia	Vila Sônia	-46.746958	-23.609207	Lebanese, Arabian	4.1	Very Good	11	1

Map with the restaurants by clusters

Characteristics of the localities where the restaurants are situated

Localities with the largest nbr of rated restaurants in the main regions of São Paulo: Top 3
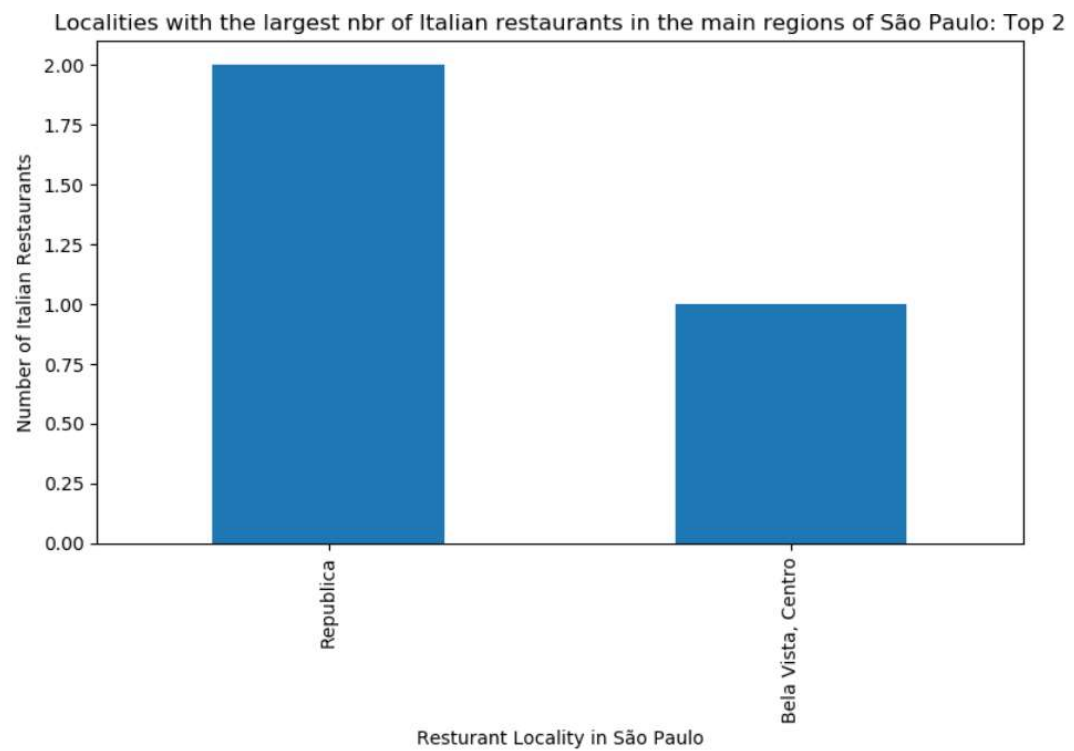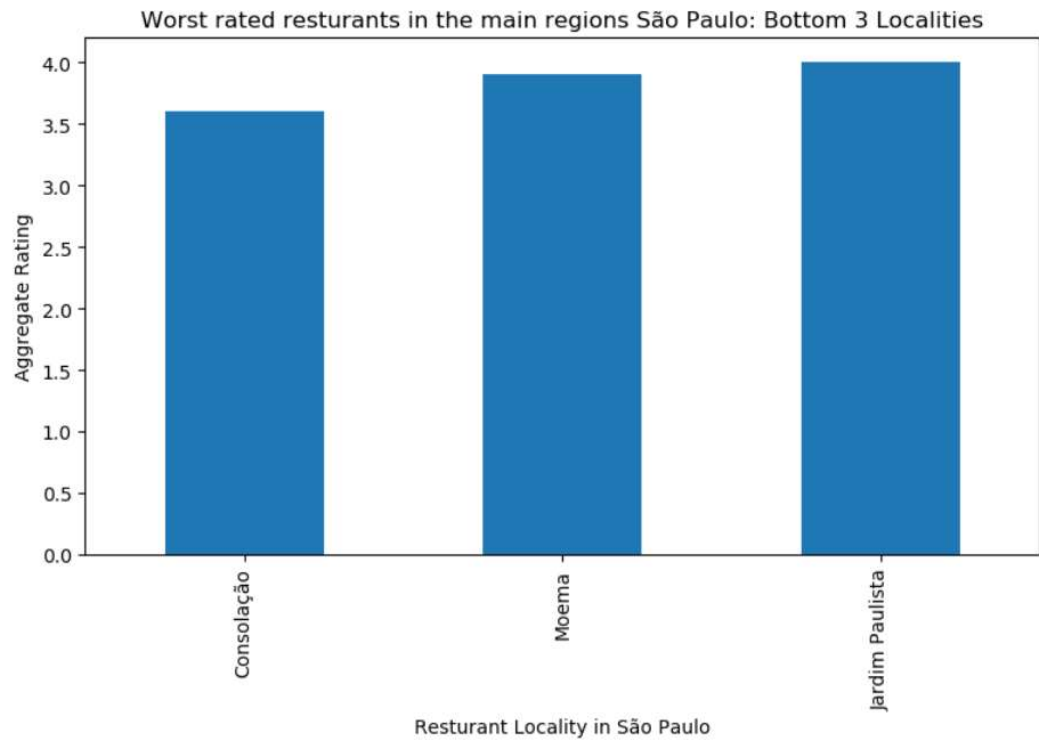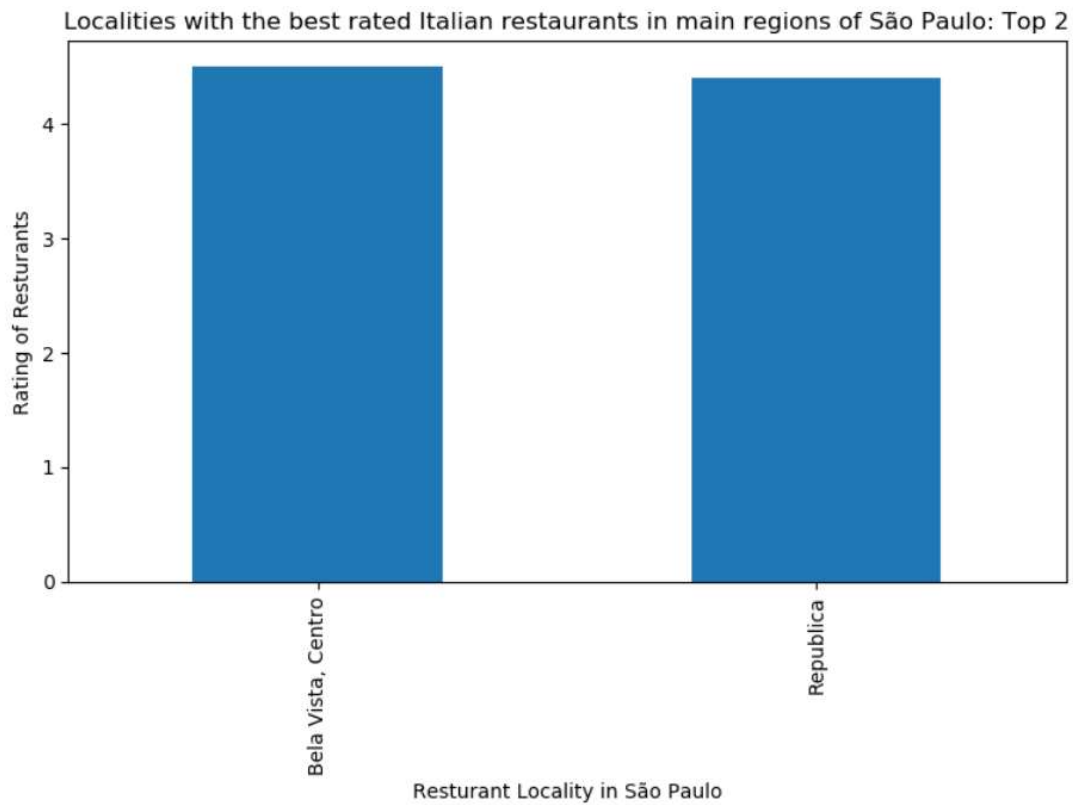
Localities with the smallest nbr of rated restaurants in the main regions of São Paulo: Bottom 3

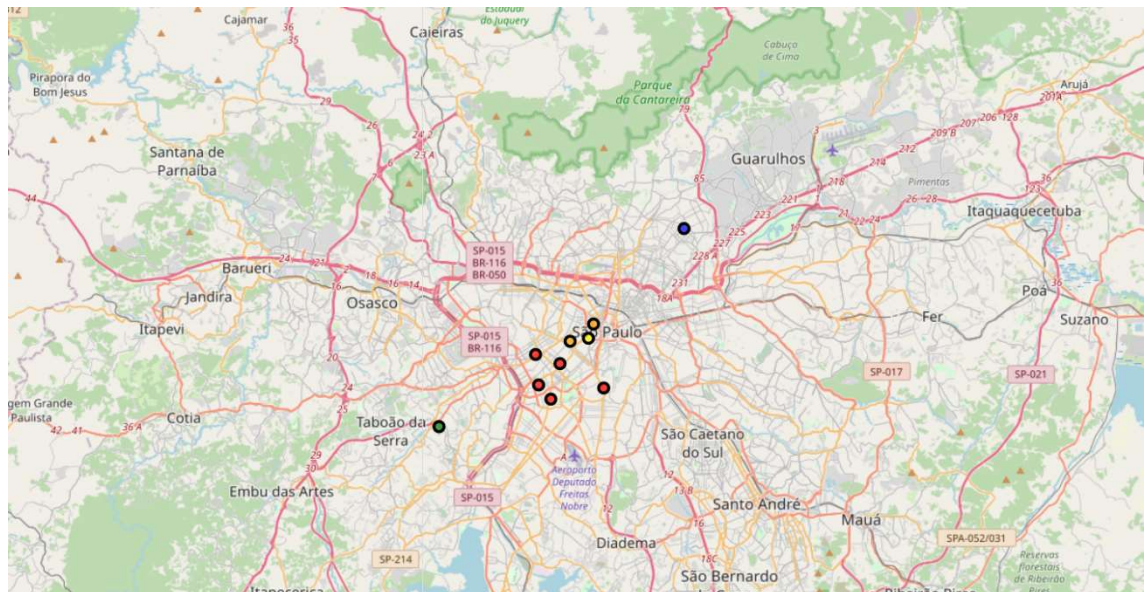Best rated resturants in the main regions of São Paulo: Top 3 Localities

Worst rated resturants in the main regions São Paulo: Bottom 3 Localities

Aggregate Rating

Consolação | Moema | Jardim Paulista

Resturant Locality in São Paulo

Localities with the largest nbr of Italian restaurants in the main regions of São Paulo: Top 2

Number of Italian Restaurants

Republica | Bela Vista, Centro

Resturant Locality in São Paulo

Localities with the best rated Italian restaurants in main regions of São Paulo: Top 2

Top 10 nearby venues of each locality (with cluster labels)

	Locality	Nbr_of_Restaurants	Cuisines	Agg_Rating	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Bela Vista, Centro	2	Italian, Pizza, Steak, BBQ	4.45	3	Italian Restaurant	Bakery	Farmers Market	Theater	Hotel	Brazilian Restaurant	Café	Snack Place	Flower Shop	Pharmacy
1	Consolação	2	Indian, Brazilian	3.60	4	Brazilian Restaurant	Bar	Burger Joint	Hotel	Gym / Fitness Center	Coffee Shop	Café	Bakery	Ice Cream Shop	Lounge
2	Itaim Bibi	2	French, Brazilian, Beverages, Brazilian, Grill	4.25	0	Italian Restaurant	Japanese Restaurant	Restaurant	Bar	Brazilian Restaurant	Hotel	French Restaurant	Ice Cream Shop	Gym / Fitness Center	Pizza Place
3	Jardim Paulista	4	Beverages, International, French, Sushi, Japan...	4.00	0	Italian Restaurant	Art Gallery	Brazilian Restaurant	Spanish Restaurant	Ice Cream Shop	Dessert Shop	Gym / Fitness Center	Restaurant	Bakery	Café
4	Moema	1	Sushi	3.90	0	Bakery	Italian Restaurant	Restaurant	Brazilian Restaurant	Café	Pizza Place	Coffee Shop	Women's Store	Hotel	Argentinian Restaurant

5	Pinheiros	1	Gourmet Fast Food, Burger	4.30	0	Italian Restaurant	Clothing Store	Burger Joint	Ice Cream Shop	Pizza Place	Restaurant	Café	Brazilian Restaurant	Bakery	Coffee Shop
6	Republica	2	Brazilian, Italian, Italian	4.35	4	Brazilian Restaurant	Coffee Shop	Bar	Dessert Shop	Café	Tea Room	Theater	Record Shop	Ice Cream Shop	Vegetarian / Vegan Restaurant
7	Vila Maria	1	Brazilian, North Eastern	4.40	2	Brazilian Restaurant	Pizza Place	Market	Northeastern Brazilian Restaurant	Asian Restaurant	Martial Arts School	Department Store	Event Space	Grocery Store	Ice Cream Shop
8	Vila Mariana	1	Brazilian, Bar Food, Beverages	4.60	0	Pizza Place	Ice Cream Shop	Spa	Café	Dance Studio	Bakery	Restaurant	General Entertainment	Vegetarian / Vegan Restaurant	Farmers Market
9	Vila Sônia	1	Lebanese, Arabian	4.10	1	Gym / Fitness Center	Construction & Landscaping	Big Box Store	Mexican Restaurant	Café	Bus Station	Middle Eastern Restaurant	Martial Arts School	Churrascaria	Brazilian Restaurant

Map with the localities by clusters

Cluster 1

	Locality	Agg_Rating	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
2	Itaim Bibi	4.25	0	Italian Restaurant	Japanese Restaurant	Restaurant	Bar	Brazilian Restaurant	Hotel	French Restaurant	Ice Cream Shop	Gym / Fitness Center	Pizza Place
3	Jardim Paulista	4.00	0	Italian Restaurant	Art Gallery	Brazilian Restaurant	Spanish Restaurant	Ice Cream Shop	Dessert Shop	Gym / Fitness Center	Restaurant	Bakery	Café
4	Moema	3.90	0	Bakery	Italian Restaurant	Restaurant	Brazilian Restaurant	Café	Pizza Place	Coffee Shop	Women's Store	Hotel	Argentinian Restaurant
5	Pinheiros	4.30	0	Italian Restaurant	Clothing Store	Burger Joint	Ice Cream Shop	Pizza Place	Restaurant	Café	Brazilian Restaurant	Bakery	Coffee Shop
8	Vila Mariana	4.60	0	Pizza Place	Ice Cream Shop	Spa	Café	Dance Studio	Bakery	Restaurant	General Entertainment	Vegetarian / Vegan Restaurant	Farmers Market

Cluster 2

	Locality	Agg_Rating	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
9	Vila Sônia	4.1	1	Gym / Fitness Center	Construction & Landscaping	Big Box Store	Mexican Restaurant	Café	Bus Station	Middle Eastern Restaurant	Martial Arts School	Churrascaria	Brazilian Restaurant

Cluster 3

	Locality	Agg_Rating	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
7	Vila Maria	4.4	2	Brazilian Restaurant	Pizza Place	Market	Northeastern Brazilian Restaurant	Asian Restaurant	Martial Arts School	Department Store	Event Space	Grocery Store	Ice Cream Shop

Cluster 4

	Locality	Agg_Rating	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Bela Vista, Centro	4.45	3	Italian Restaurant	Bakery	Farmers Market	Theater	Hotel	Brazilian Restaurant	Café	Snack Place	Flower Shop	Pharmacy

Cluster 5

	Locality	Agg_Rating	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	Consolação	3.60	4	Brazilian Restaurant	Bar	Burger Joint	Hotel	Gym / Fitness Center	Coffee Shop	Café	Bakery	Ice Cream Shop	Lounge
6	Republica	4.35	4	Brazilian Restaurant	Coffee Shop	Bar	Dessert Shop	Café	Tea Room	Theater	Record Shop	Ice Cream Shop	Vegetarian / Vegan Restaurant

5. Discussion

Based on the results obtained from Zomato dataset, we can notice that 'Republica' and 'Bela Vista' are the most recommended localities if you want to explore rated Italian restaurants in the main regions on São Paulo. On the other hand, if we consider the nearby venues given by the Foursquare API, 'Bela Vista', 'Itaim Bibi', 'Jardim Paulista', 'Moema' e 'Pinheiros' are the most recommended localities to find Italian restaurants.

6. Conclusion

Considering the main regions of São Paulo:

- Jardim Paulista, Bela Vista and Consolação are the localities with the largest nbr of rated restaurants.
- Moema, Pinheiros and Vila Maria are the localities with the smallest nbr of rated restaurants.
- Vila Mariana, Bela Vista and Vila Maria are the localities with the best rated restaurants.
- Consolação, Moema and Jardim Paulista are the localities with the worst rated restaurants.

- Republica and Bela Vista are the localities with the largest nbr of rated Italian restaurants.
- Bela Vista and Republica are the localities with the best rated Italian restaurants.

- Cluster 1 (Bela Vista, Itaim Bibi, Jardim Paulista, Moema, Pinheiros): it is most recommended for Italian Restaurant.
- Cluster 2 (Vila Maria): it is most recommended for Market, Pizza and Brazilian Restaurant.
- Cluster 3 (Consolação, Republica): it is most recommended for Brazilian Restaurant, Bar and Coffee Shop.
- Cluster 4 (Vila Mariana): it is most recommended for Ice Cream, Pizza and Bakery.
- Cluster 5 (Vila Sonia): it is most recommended for Pet Store, Gym and Construction & Landscaping.