The Battle of Neighborhoods:
identifying São Paulo's cuisine opportunities

1. Introduction

São Paulo is the capital city of the state bearing the same name. It is also the most populous city in Brazil[1], having the largest GDP in the southern hemisphere[2].

Since its inception, the city of São Paulo has received a lot of immigrants from a number of different countries. This multicultural city boasts statistics such as having more Italian descendants than any given city in Italy[3]. During the immigration process, those immigrants clustered together forming neighborhoods. One notorious example is the *Liberdade* neighborhood, containing mainly immigrants from Japan, China and Korea. *Liberdade* is characteristic for various Asian restaurants, which attracts revenues from the immigrants in the area and tourists in search for different gastronomic experiences. This multicultural aspect of São Paulo got the city to be known for its cultural diversity, including gastronomic.

1.1. Problem Description

São Paulo is a rich city containing neighborhoods born from different nationalities. One of the interesting aspects of different nationalities is their cuisine. In this sense, São Paulo seems a place full of opportunities in gastronomy.

Thus, this work targets neighborhoods absent of restaurants from the nationalities of their immigrants, e.g., a German neighborhood which does not contain a restaurant that serves typical German food.

The restaurant opportunities envisioned by this work would attract customers from two demographics: locals and tourists. The immigrant locals living in such neighborhood are likely to appreciate restaurants serving food from their country of origin. In addition, as the restaurant will be located in a neighborhood known to be of immigrant origins correspondent to that gastronomy, the tourists can be attracted by a different and authentic food experience.

1.2. Target audience

Companies or individuals interested in opening restaurants in the city of São Paulo (Brazil).

1.3. Objective

The objective of this report is to identify local opportunities to open restaurants in specific neighborhoods in the city of São Paulo (Brazil), based on the immigrants of such neighborhoods.

2. Data

This work focuses on the city of São Paulo in Brazil. The data contains information about postal code of neighborhoods, geolocation by postal code, immigration by neighborhood, and restaurants of different cuisines by neighborhood.

2.1.Postal Code of Neighborhoods

This dataset is used to obtain the geolocation of each neighborhood of São Paulo associated with its name. The data is available online on: http://cep.la/baixar. Figure 1 presents an example of the dataset.

	Postal Code	Neighborhood
0	01001000	Sé
1	01002020	Centro
2	01017000	Brás
3	01035100	República
4	01101000	Luz

Figure 1 – Example of the postal code of neighborhoods dataset.

2.2.Geolocation by Postal Code

The geopy API (https://geopy.readthedocs.io/en/stable/#) is used to associate each postal code with their respective latitude and longitude. This is necessary to associate each neighborhood with their nearby restaurants. An example is provided in Figure 2.

	Latitude	Longitude
Postal Code		
01001-000	-23.550562	-46.633653
01002-020	-23.547212	-46.637019
01017-000	-23.549744	-46.631144
01035-100	-23.540726	-46.643465
01101-000	-23.522819	-46.630261

Figure 2 – Example of the geolocation by postal code obtained through the geopy API.

2.3.Immigration by Neighborhood

This dataset gives the association of each neighborhood to its immigration origins. Figure 3 shows an example of the dataset, which can be found on: https://pt.wikipedia.org/wiki/Lista_dos_bairros_paulistanos_por_imigra%C3%A7%C3%A3o (in portuguese).

	Origin	Neighborhood
0	Argentina	Bom Retiro
1	Argentina	Brás
2	Argentina	Cerqueira César
3	Argentina	Jardim América
4	Argentina	Jardim Paulista

Figure 3 – Example of the immigration by neighborhood dataset.

2.4. Restaurants of Different Cuisines by Neighborhood

The Foursquare API (https://developer.foursquare.com/)was used to obtain the restaurants near a given geolocation, which can be associated with each neighborhood. An example is provided in Figure 4.

	Neighborhood	Venue	Venue Category
0	Barra Funda	U Mnicha	Restaurant
1	Barra Funda	Restaurace Damika	Restaurant
2	Bela Vista	Famiglia Mancini	Italian Restaurant
3	Bela Vista	Walter Mancini Ristorante	Italian Restaurant
4	Bela Vista	La Penisola	Italian Restaurant

Figure 4 – Example of restaurants of different cuisines by neighborhood obtained through the Foursquare API.

3. Methodology

First, the postal codes for each neighborhood were obtained from the dataset 2.1. However, several postal codes exist for the same neighborhood. Therefore, it was chosen to only keep the first occurrence for each neighborhood.

With the postal codes in hand, the next step was to find the geolocation for each neighborhood. The geopy API was not able to find the geolocation for some neighborhoods. In total, the geopy was unable to find geolocation data for 959 of the original 1907 neighborhoods. Figure 5 presents the location of such neighborhoods in the map of the city.

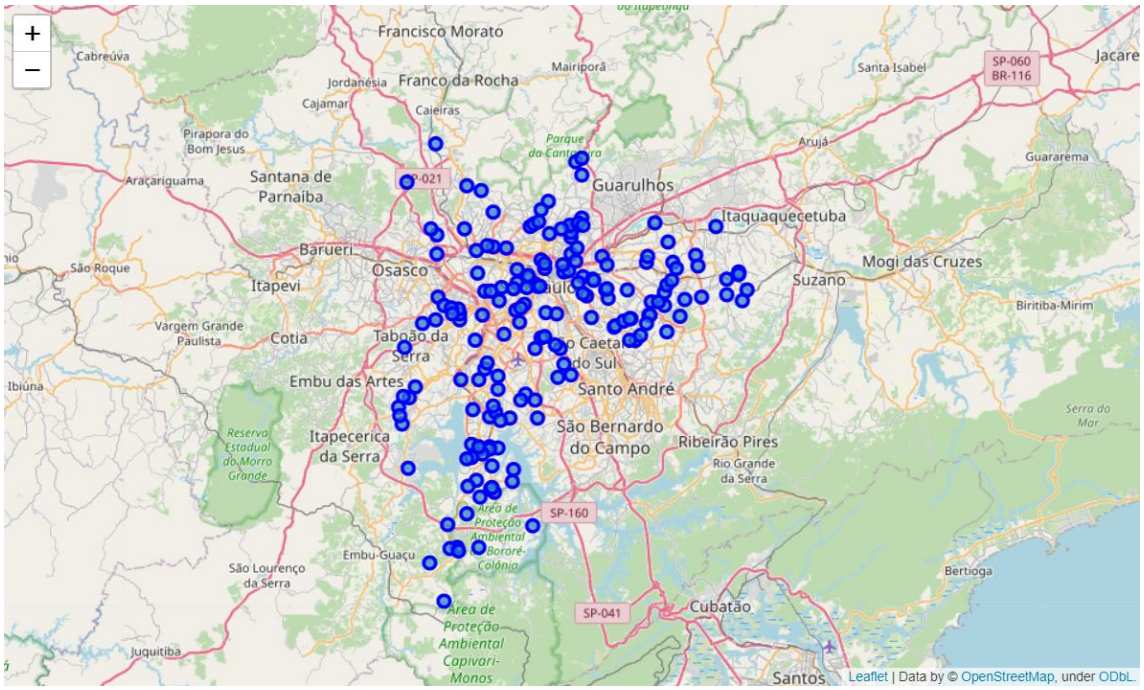

Figure 5 – Visualization of the neighborhoods of São Paulo superimposed on the map of the city.

The next step was to scrape the immigration data from the wiki page (2.3), translate it to English and convert some data to countries. For example, the page reports immigrants from the Austro-Hungarian Empire, which subsequently dissolved into several countries. Note that the Jewish immigrants were labelled as being from Israel to facilitate the matching with the

restaurants. Figure 6 shows a word cloud of the migrants and immigrants in the neighborhoods of the city.

Figure 6 – Word cloud of the migrants and immigrants in the neighborhoods of São Paulo. Note that Northeast refers to the northeastern part of Brazil.

Then, the geolocation is merged with the immigration data for the neighborhoods. This process leads to a considerable reduction in the data size due to the wiki page not containing many neighborhoods, as can be seen in Figure 7

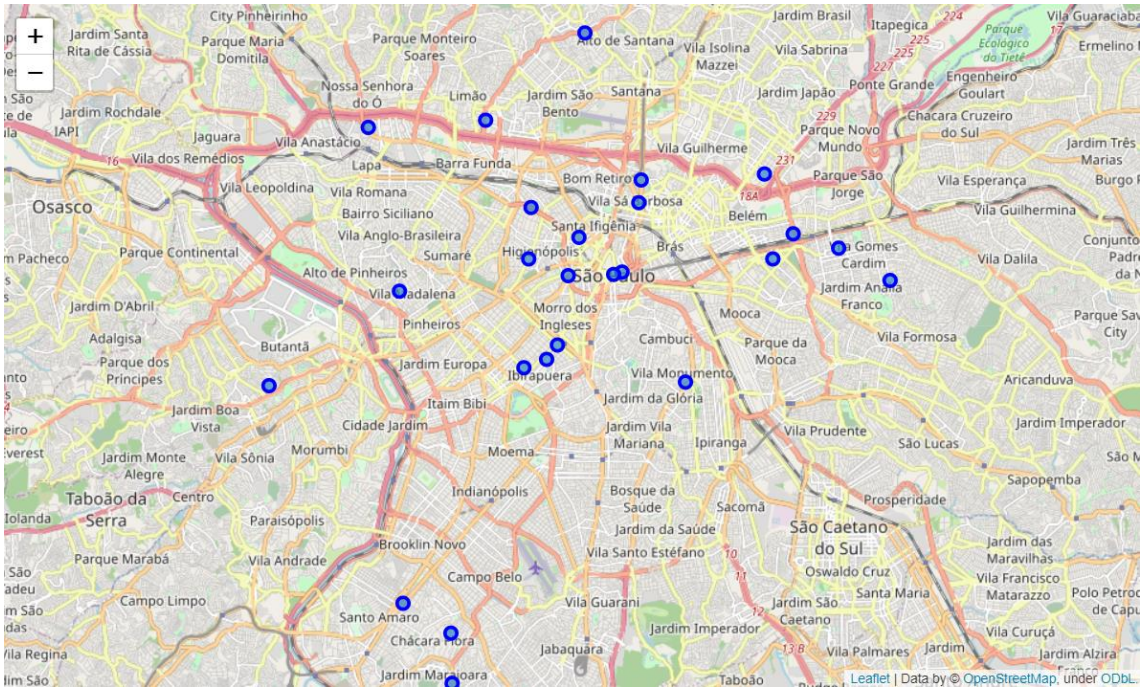

Figure 7 – Map of the neighborhoods containing geolocation and immigration data.

Originally, there were 1027 neighborhoods with geolocation data and 139 with immigration data. However, there were only 60 with both geolocation and immigration data.

Next, the Foursquare API was used to obtain the venues near each neighborhood, limited by 200 venues per neighborhood and a radius of 1000 meters from the informed location. Then the retrieved venues were filtered to get only the restaurants. Within those data, the restaurant categories which could not be attributed to a migrant or immigrant origin were dropped (see Table 1). In addition, some restaurant types were changed according to similar cuisines as Table 2 reports.

Table 1 – List of restaurant categories dropped from the dataset.

Brazilian Restaurant
Restaurant
Vegetarian / Vegan Restaurant
Mineiro Restaurant
Fast Food Restaurant
Gluten-free Restaurant
Comfort Food Restaurant
Southeastern Brazilian Restaurant
Seafood Restaurant
Hawaiian Restaurant
Fondue Restaurant

Table 2 – Restaurant name change according to cuisine similarity.

Name retrieved from Foursquare API	**New name according to similar cuisine**
Baiano Restaurant	Northeastern Brazilian Restaurant
Sushi Restaurant	Japanese Restaurant
Empanada Restaurant	Latin American Restaurant
Kosher Restaurant	Israeli Restaurant
Japanese Curry Restaurant	Japanese Restaurant
Dumpling Restaurant	Asian Restaurant
Tapas Restaurant	Spanish Restaurant
Bunsik Restaurant	Korean Restaurant
Ramen Restaurant	Asian Restaurant
Paella Restaurant	Spanish Restaurant
Halal Restaurant	Middle Eastern Restaurant
Doner Restaurant	Middle Eastern Restaurant
Kebab Restaurant	Middle Eastern Restaurant
Falafel Restaurant	Middle Eastern Restaurant
Southern / Soul Food Restaurant	American Restaurant
Donburi Restaurant	Japanese Restaurant
Bossam/Jokbal Restaurant	Korean Restaurant
Soba Restaurant	Asian Restaurant
South American Restaurant	Latin American Restaurant

This led to the reduction of the number of unique restaurant types from 52 to 25. After this data cleanup, the restaurants were ordered in terms of popularity and the top 10 were selected.

In order to compare the immigrant data with the restaurant present in each neighborhood, each nationality was given a list of restaurants that they are likely to appreciate, as Table 3 shows.

Table 3 – List of nationalities and the restaurants each are likely to appreciate.

Nationality	Restaurants likely to appreciate
United States	American Restaurant
Argentina	Latin American Restaurant, Argentinian Restaurant
Germany	German Restaurant
Bolivia	Latin American Restaurant, Bolivian Restaurant
Israel	Middle Eastern Restaurant, Israeli Restaurant
Armenia	Eastern European Restaurant, Armenian Restaurant
Italy	Mediterranean Restaurant, Italian Restaurant
Belgium	Belgian Restaurant
Spain	Mediterranean Restaurant, Spanish Restaurant
Brazil-Northeast	Latin American Restaurant, Northeastern Brazilian Restaurant
Syria	Middle Eastern Restaurant, Mediterranean Restaurant, Syrian Restaurant
Croatia	Croatian Restaurant
Peru	Latin American Restaurant, Peruvian Restaurant
Russia	Asian Restaurant, Eastern European Restaurant, Russian Restaurant
Ukraine	Asian Restaurant, Eastern European Restaurant, Ukrainian Restaurant
China	Asian Restaurant, Chinese Restaurant
Egypt	Middle Eastern Restaurant, Mediterranean Restaurant, Egyptian Restaurant
Portugal	Mediterranean Restaurant, Portuguese Restaurant
Norway	Norwegian Restaurant
France	Mediterranean Restaurant, French Restaurant
Lebanon	Middle Eastern Restaurant, Mediterranean Restaurant, Lebanese Restaurant
Nigeria	Nigerian Restaurant
Japan	Asian Restaurant, Japanese Restaurant
Haiti	Latin American Restaurant, Haitian Restaurant
Scotland	Scottish Restaurant
Colombia	Latin American Restaurant, Colombian Restaurant
England	English Restaurant
Senegal	Senegalese restaurant
Lithuania	Lithuanian Restaurant
Ecuador	Latin American Restaurant, Ecuadorian Restaurant
Paraguay	Latin American Restaurant, Paraguayan Restaurant
Sweden	Swedish Restaurant
Venezuela	Latin American Restaurant, Venezuelan Restaurant
Greece	Mediterranean Restaurant, Greek Restaurant
Angola	Angolan Restaurant
South Korea	Asian Restaurant, Korean Restaurant

Finally, a list was created containing the restaurants that are likely to be appreciated in that neighborhood but that are not contemplated in the top 10 of such neighborhood. Note that some cuisines are broader than a single nationality, thus the broad cuisines are considered contemplated if a restaurant of its subgroup is present in the neighborhood. For example, it a given neighborhood is likely to appreciate Mediterranean restaurants and the neighborhood contains Italian restaurants, then such Mediterranean cuisine needs are already satisfied. Table 4 is the full table of correspondences.

Table 4 – Correspondence between broad-cuisine restaurants and their subgroups.

Broad-cuisine restaurant	Restaurants within that broad classification
Latin American Restaurant	Argentinian Restaurant, Northeastern Brazilian Restaurant, Peruvian Restaurant, Colombian Restaurant
Middle Eastern Restaurant	Israeli Restaurant
Mediterranean Restaurant	Italian Restaurant, Spanish Restaurant, Portuguese Restaurant, French Restaurant, Greek Restaurant
Asian Restaurant	Chinese Restaurant, Japanese Restaurant, Korean Restaurant

Note that these broad cuisines group more restaurants than the ones listed in Table 4. However, Table 4 only lists the ones that were contained in the Foursquare search for the analyzed neighborhoods.

4. Results

Following the methodology presented in section 3, one is able to get a table of restaurants which are likely to be appreciated, but are not contemplated in each neighborhood, such as Table 5.

Table 5 – List of the neighborhoods of São Paulo with the restaurants likely to be appreciated but not contemplated in that neighborhood.

Neighborhood	Restaurants not contemplated
Bela Vista	Japanese Restaurant, Portuguese Restaurant
Bom Retiro	Syrian Restaurant, Ecuadorian Restaurant, Korean Restaurant, Bolivian Restaurant, Italian Restaurant, Lebanese Restaurant, Venezuelan Restaurant
Brás	Ecuadorian Restaurant, Angolan Restaurant, Colombian Restaurant
Campo Grande	German Restaurant, English Restaurant, Belgian Restaurant
Canindé	Bolivian Restaurant
Chácara Flora	Norwegian Restaurant, Portuguese Restaurant, English Restaurant, Spanish Restaurant
Chácara Monte Alegre	Belgian Restaurant
Higienópolis	Israeli Restaurant, Egyptian Restaurant
Imirim	-
Jardim América	Belgian Restaurant
Jardim Bonfiglioli	Norwegian Restaurant
Jardim Bélgica	Belgian Restaurant
Lapa	Syrian Restaurant, Portuguese Restaurant, Lebanese Restaurant, English Restaurant, Spanish Restaurant
Limão	Northeastern Brazilian Restaurant
Paraíso	Syrian Restaurant, Lebanese Restaurant
Quarta Parada	Croatian Restaurant
República	Senegalese restaurant, Nigerian Restaurant
Santa Cecília	-
Santo Amaro	Norwegian Restaurant, Swedish Restaurant, Portuguese Restaurant
Saúde	German Restaurant
Sé	-
Tatuapé	Syrian Restaurant, Lebanese Restaurant
Vila Alpina	-
Vila Bela	Eastern European Restaurant, Ukrainian Restaurant
Vila Carrão	Portuguese Restaurant
Vila Gomes Cardim	-
Vila Madalena	Spanish Restaurant, French Restaurant
Vila Maria	Bolivian Restaurant
Vila Mariana	Lithuanian Restaurant, Syrian Restaurant, German Restaurant, Lebanese Restaurant
Vila Mazzei	-
Vila Monumento	Japanese Restaurant, Portuguese Restaurant

Note that the restaurants not contemplated assume that the restaurant is not among the top 10 most common restaurants in that neighborhood.

Now, one can organize the restaurants that most appear in the not contemplated group for our neighborhoods, as shown in Figure 8.

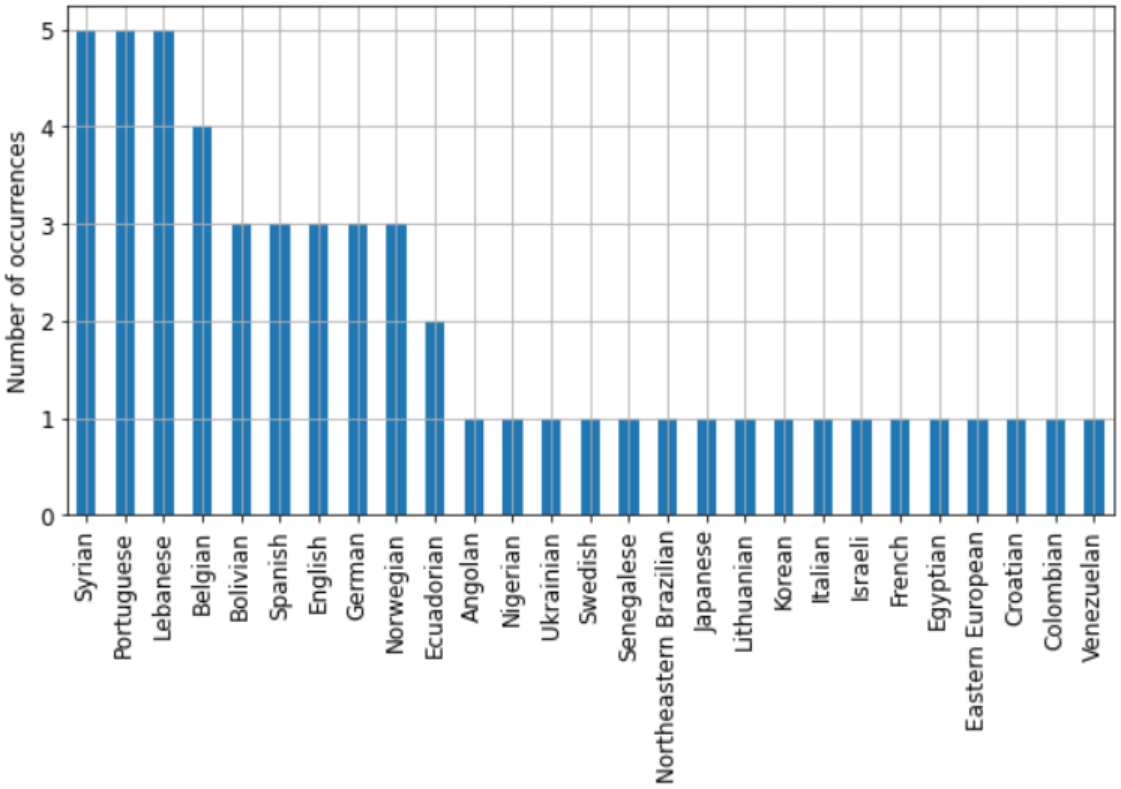

Figure 8 – Number of neighborhoods which each restaurant kind appear as not contemplated.

Therefore, São Paulo appears to have more opportunities for Syrian, Portuguese and Lebanese restaurants. However, one should note that some cuisines taken into account in this analysis may be too exotic, e.g., Belgian or Norwegian. These cuisines may not be good candidates for new restaurants as they might not attract as many tourists.

5. Discussions

With the above analysis, it is possible to make to following recommendations:

- The neighborhoods with the most variety of opportunities for restaurants of different cuisines are Bom Retiro, Chácara Flora, Lapa, and Vila Mariana;
- São Paulo presents great opportunities for new restaurants on Syrian, Portuguese, and Lebanese cuisines;
- Other notable opportunities come from Bolivian, Spanish, English, and German cuisines;
- Those opportunities are located in the neighborhoods listed by Table 5.

Note that these recommendations are based on some assumptions of the analysis, namely:

- Only one postal code for each neighborhood was used, leading to some neighborhoods being excluded from the analysis when geopy was unable to find the geolocations for them;
- The radius of opportunity for each neighborhood was considered as 1000 meters from the postal code location of each one;
- The restaurants likely to be appreciated by each nationality follow the list in Table 3;
- The recommended opportunities are based on the absence of a restaurant likely to be appreciated in the top 10 most common restaurants in each neighborhood.

6. Conclusions

This report lists a number of opportunities for new restaurants to be opened in the city of São Paulo in Brazil. These recommendations are neighborhood-centered, based on their immigration origins and aim to attract both locals and tourists looking for authentic experiences.

Based on the assumptions of the analysis, this report could underline 25 neighborhoods with different opportunities for new restaurants in 27 different cuisines, albeit some of them may be too exotic for a business oportunity.

References

[1] 'List of cities in Brazil by population' (2020) *Wikipedia*. Available at https://en.wikipedia.org/wiki/List_of_cities_in_Brazil_by_population (Accessed: 3 May 2020).

[2] 'List of cities by GDP' (2020) *Wikipedia*. Available at https://en.wikipedia.org/wiki/List_of_cities_by_GDP (Accessed: 3 May 2020).

[3] 'Sobre São Paulo (Cidade)' (2011) *Encontra São Paulo*. Available at https://www.encontrasaopaulo.com.br/sobre-sao-paulo.php (Accessed: 3 May 2020). *In Portuguese*.