Rent Apartament in São Paulo

Fábio Henrique Frazão Mendes

January 14, 2020

1. Introduction

1.1. Scenario and Background

I live in São Paulo in Brazil and work at the Secretariat of Finance and Planning and to get around in São Paulo I use the subway. I am currently analyzing whether I change my neighborhood or stay in my apartment. It should be emphasized that an important factor that should not be ignored is location. After all, the place where people live has a major influence on quality of life and any choice in this regard must be made by respecting a set of criteria that must be defined before the search for a new home begins.

1.2. Problem to be resolved

When searching for a property you need to consider a number of important aspects that include both negotiating and choosing the property itself, including the best neighborhood for your needs. So a good place to live is critical for you to live well. A safe place, close to work and neighborhood facilities like supermarket, hospital, pharmacy and restaurant and near the best places in town. The challenge to solve is being able to find an apartment for rent that combines the factors presented.

1.3. Interested Audience

This project will be relevant for a person considering moving to Sao Paulo and also for new residents, as the approach and methodologies used here are applicable in all cases. The use of FourSquare data and mapping techniques combined with data analysis will help resolve the key issues raised. Finally, this project is a good case study for developing data science skills.

2. Data

Based on definition of our problem, factors that will influence our decission to rent a apartment in Sao Paulo are:

- ✓ Safe place;
- ✓ Distance from work;
- ✓ Near the subway;
- ✓ Rent price;
- ✓ Neighborhood Facilities.

Following data sources will be needed to extract/generate the required information:

- 7 safe neighborhoods were chosen to live in search on the site survey https://meulugar.quintoandar.com.br/bairros-mais-seguros-de-sao-paulo/. The neighborhoods are Paraíso, Itaim Bibi, Vila Nova Conceição, Barra Funda, Higienopolis, Mooca and Pinheiros. It is necessary to mention who I live in Paraíso.
- The Google maps was used to find the latitude and longitude of places, distance from work and distance from the subway.
- The neighborhood facilities like example restaurant, hospital, pharmacy were obtained Foursquare API.
- The price of the apartment for rent was found on the website www.quintoandar.com.br

3. Methodology

This section represents the main component of the report where the data is gathered, prepared for analysis. The tools described are used here and the Notebook cells indicates the execution of steps.

3.1 The analysis and the stratey

The strategy is based on mapping the above described data in section 2.0, in order to facilitate the choice of at least two candidate places for rent. The choice is made based on the demands imposed : location near a subway, rental price, distance from woork, safe place and neighboring places.

In first step we selected 7 locations that are safe, close to the subway, not far from work and within the market price average. So we created a dataframe with this information with the help of google maps and the site www.quintoandar.com.br.

The second step was collected the required data: seven locations and facility type (category) of every neighborhood.

The third step was to divide these rental places into three clustering(using k-means clustering). After that, the cluster with the largest places to rent was chosen. The next step was to discard places that are farthest from the workplace and the subway.

This visual approach and maps with popups labels allow quick identification of location, price and feature, thus making the selection very easy.

4. Analysis

4.1 **Create the Dataframe with 7 locations**

Let's analyze the data to infer the best neighborhood to live in Sao Paulo. To do this, we need to create a dataframe with the seven locations to rent, then list the locations obtained by fourSquare for the purpose of creating 3 Clusters. After that, we will analyze the dataframe information and maps to find the best place to live.

	Neighborhood	Latitude	Longitude	Distance Work	Distance Metro	M2 price rent
0	Paraiso	-23.574218	-46.639347	2580	310	46
1	Itaim Bibi	-23.584194	-46.681096	6430	1000	67
2	Vila Nova Conceicao	-23.596479	-46.668131	6400	1600	68
3	Barra Funda	-23.527149	-46.669275	4680	250	33
4	Higienopolis	-23.544800	-46.656316	2660	650	34
5	Mooca	-23.549403	-46.594401	3650	1340	26
6	Pinheiros	-23.562716	-46.690196	6240	614	51

Figure 1. Dataframe with 7 locations.

4.2 FourSquare API

Get locations by neighborhood from FourSquare to create a dataframe with the 10 common points. After that prepare the data to find the clusters.

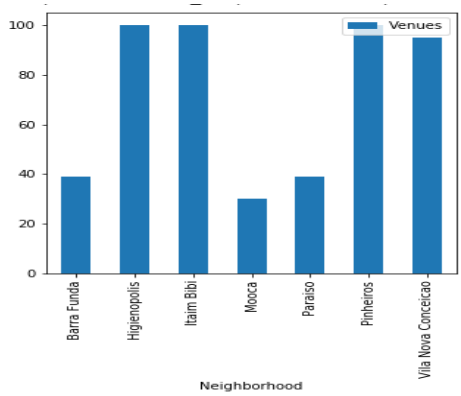

Figure 2. Amount obtained per neighborhood.

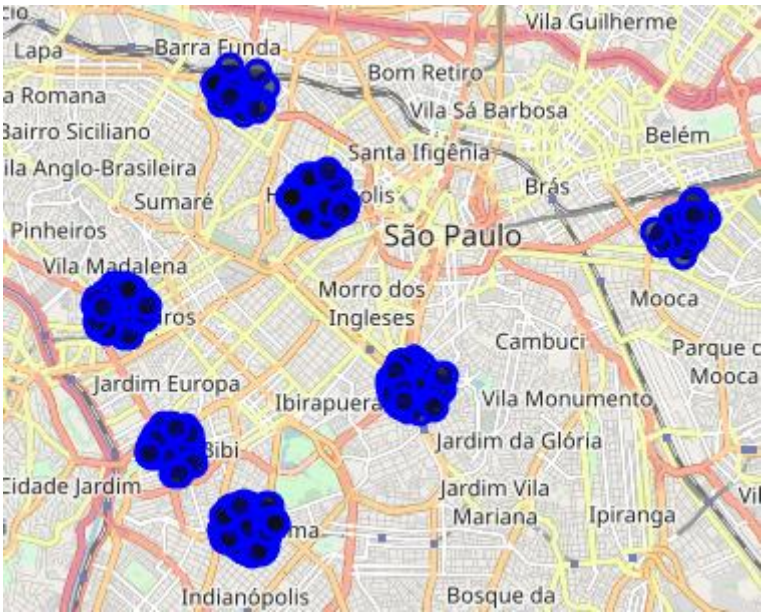

Figure 3. Show in the São Paulo Maps amount obtained.

	Neighborhood	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	Barra Funda	Brazilian Restaurant	Music Venue	Restaurant	Residential Building (Apartment / Condo)	Bookstore	Boutique	Farmers Market	Salon / Barbershop	Fair	Café
1	Higienopolis	Italian Restaurant	Bakery	Café	Coffee Shop	Pizza Place	Middle Eastern Restaurant	Jewelry Store	Bar	Pastry Shop	Food & Drink Shop
2	Itaim Bibi	Italian Restaurant	Restaurant	Bar	Japanese Restaurant	Brazilian Restaurant	Burger Joint	Hotel	Coffee Shop	Steakhouse	French Restaurant
3	Mooca	Bar	Pizza Place	Steakhouse	Pet Store	Bakery	Gym / Fitness Center	Farmers Market	Burger Joint	Plaza	Fried Chicken Joint
4	Paraiso	Coffee Shop	Hostel	Middle Eastern Restaurant	Tennis Court	Spa	Hotel	Restaurant	Gym	Nightclub	Music Venue

Figure 4. Dataframe with 10 comon Venues.

4.3 Cluster

Create 3 cluster using k-means for the 7 places.

```
# Cluster 0
saopaulo_merged.Neighborhood[saopaulo_merged['Cluster Labels'] == 0]
```

```
5    Mooca
Name: Neighborhood, dtype: object
```

```
# Cluster 1
saopaulo_merged.Neighborhood[saopaulo_merged['Cluster Labels'] == 1]
```

```
0                Paraiso
1             Itaim Bibi
2    Vila Nova Conceicao
4            Higienopolis
6               Pinheiros
Name: Neighborhood, dtype: object
```

```
# Cluster 2
saopaulo_merged.Neighborhood[saopaulo_merged['Cluster Labels'] == 2]
```

```
3    Barra Funda
Name: Neighborhood, dtype: object
```

Figure 5. List 3 Clusters.

After examining various cluster data, I concluded that cluster 1 has the best locations in Sao Paulo, providing guidance on where to look for the future apartment.

	Neighborhood	Latitude	Longitude	Distance Work	Distance Metro	M2 price rent	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
0	Paraiso	-23.574218	-46.639347	2580	310	46	1	Coffee Shop	Hostel	Middle Eastern Restaurant	Tennis Court	Spa	Hotel
1	Itaim Bibi	-23.584194	-46.681096	6430	1000	67	1	Italian Restaurant	Restaurant	Bar	Japanese Restaurant	Brazilian Restaurant	Burger Joint
2	Vila Nova Conceicao	-23.596479	-46.668131	6400	1600	68	1	Furniture / Home Store	Gym / Fitness Center	Italian Restaurant	Bakery	Supermarket	Pharmacy
4	Higienopolis	-23.544800	-46.656316	2660	650	34	1	Italian Restaurant	Bakery	Café	Coffee Shop	Pizza Place	Middle Eastern Restaurant
6	Pinheiros	-23.562716	-46.690196	6240	614	51	1	Bar	Dessert Shop	Farmers Market	Burger Joint	Art Gallery	Pet Store

Figure 6. List Cluster 1.

It will be discarded the locations: Itaim Bibi, Vila Nova Conceicao and Pinheiros due to the price and distance from work and subway. We will work with two locations Paraíso and Higienópolis.

5. Results and Discussion

Seven places were chosen to live in São Paulo. We did a cluster analysis to find similar places. After that, we found that only Paraíso and Higienópolis met the following variables: safe place, distance from work, near the subway, close price and neighborhood facilities.

Which one to choose? Firstly, you should visit these places at varying times and pay attention to the movement of the street, whether it is a more deserted place or there is constant movement of people. Remember that good infrastructure and adequate lighting help to reduce the risk of robbery and exposure to violence.

This work demonstrates that using machine learning techniques will help you choose where to rent, but before renting you should know the place to see the neighborhood yourself to see if it meets your needs.

6. Conclusion

The paper presents the use of python language, use of Foursquare API to explore Sao Paulo with machine learning techniques to find a place to rent according to the variables: safe place, distance from work, near the subway, close price and neighborhood facilities. Also, it should be clear that using artificial intelligence will help in choosing, but choosing a property to rent is not an easy task that depends on your lifestyle.