DETERMINING POSSIBLE LOCATIONS FOR A NEW ITALIAN RESTAURANT IN PORTO ALEGRE, BRAZIL

BRUNO PEIXOTO AGUILAR

JANUARY 29, 2021

INTRODUCTION AND BUSINESS PROBLEM

Porto Alegre is the largest city in the south of Brazil. Porto Alegre is home to more than 1.5 million people. In addition to its population, there are more than 3 million people who live in its metropolitan area. Economically, Porto Alegre is also a very important city of Brazil as its GDP is the eighth highest among all cities of Brazil.

Determining the location of a business is a crucial decision that business owners must make.

The present project will attempt to answer this question supported by data.

INTRODUCTION AND BUSINESS PROBLEM

Basic criteria to determine the ideal neighbourhood for the new restaurant:

1. The neighbourhood must have a significant number of restaurants
2. The ideal neighbourhood must have a minimum population with financial conditions to be a customer of the restaurant
3. The focus will be to determine a neighbourhood with no Italian restaurants, if possible.

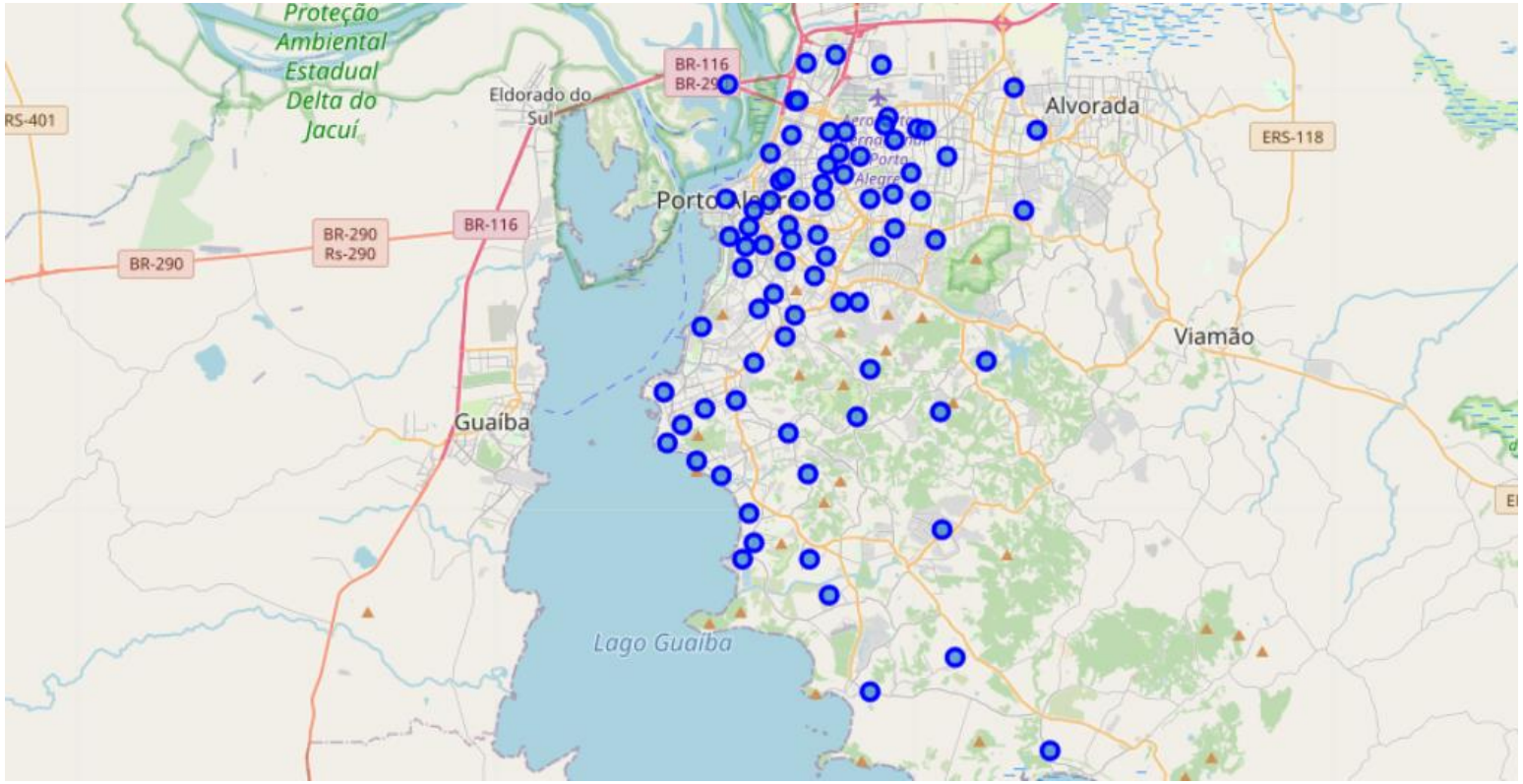
DATA ACQUISITION AND PREPROCESSING

The required data and their sources are listed below:

1. Wikipedia list of Porto Alegre boroughs, their population and average income (79 rows)
2. Coordinates for each boroughs of Porto Alegre retrieved geocoder.
3. Foursquare data on restaurants locations (620 rows)

DATA ANALYSIS

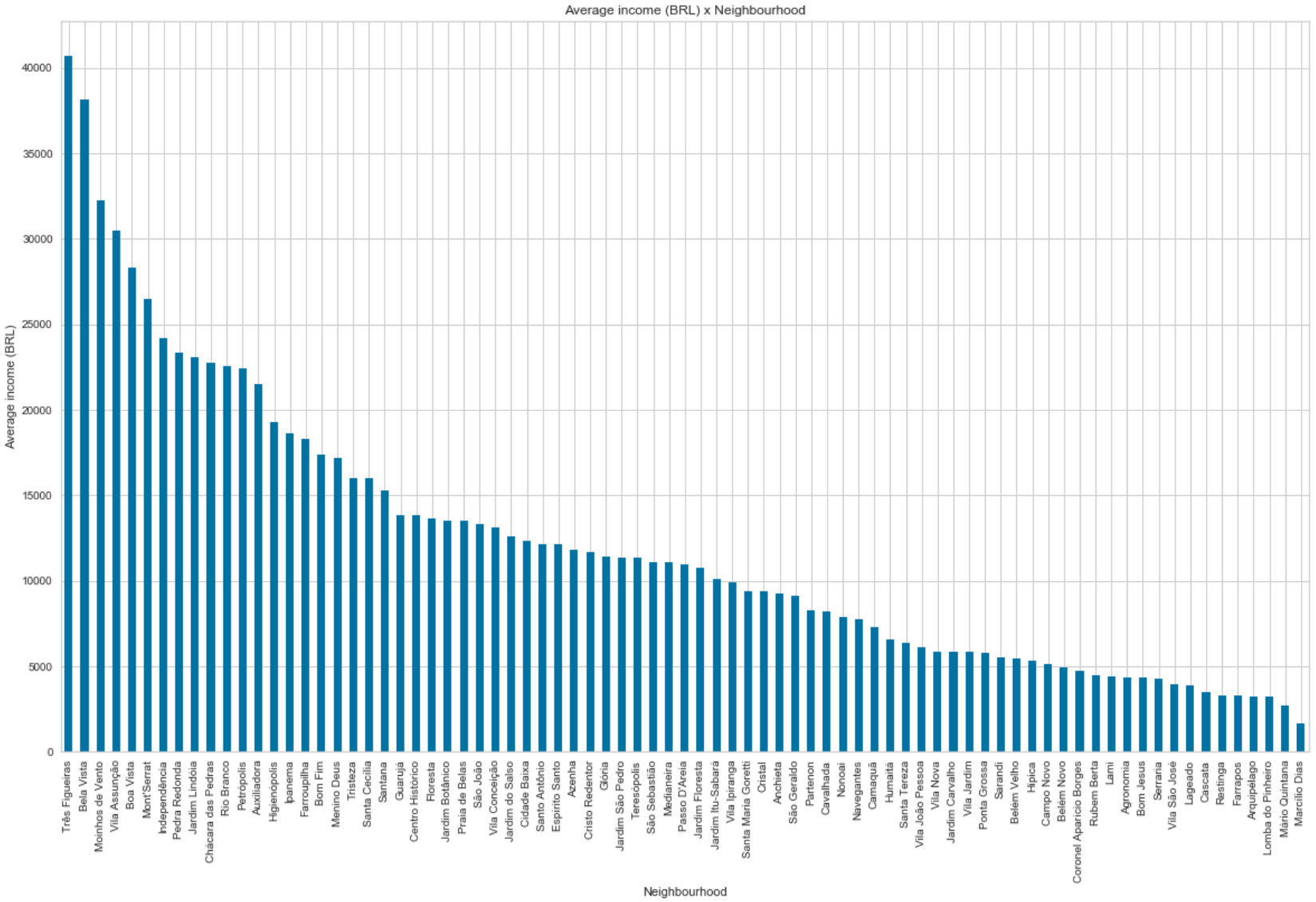
The map of Porto Alegre with the neighbourhoods was initially generated to assure that coordinates were precise enough for the analysis.



DATA ANALYSIS

The bar chart shows the average income for each neighbourhood.

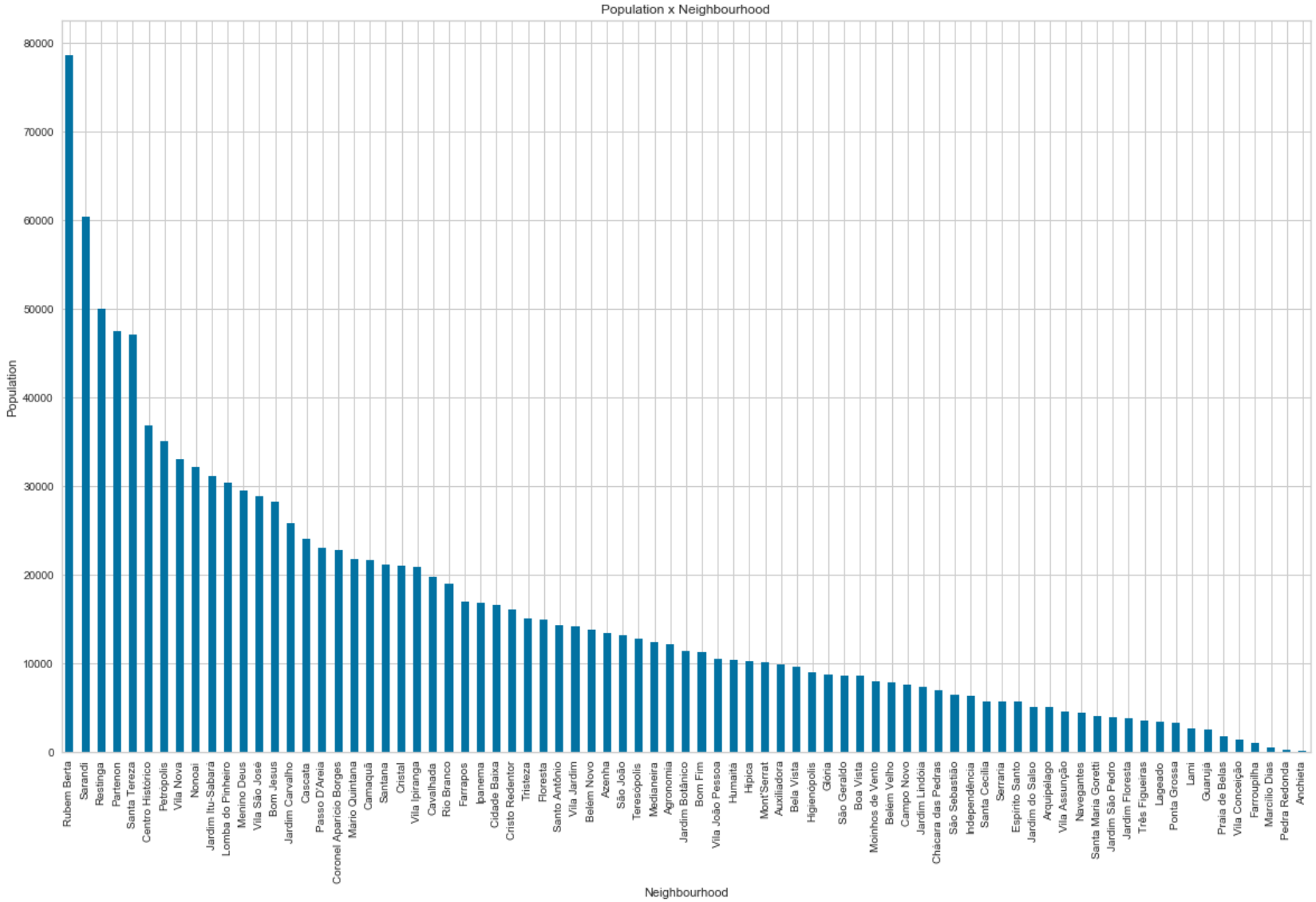
Neighbourhoods with average income of less than BRL 10000,00 were excluded from the analysis.



DATA ANALYSIS

The bar chart shows the population for each neighbourhood.

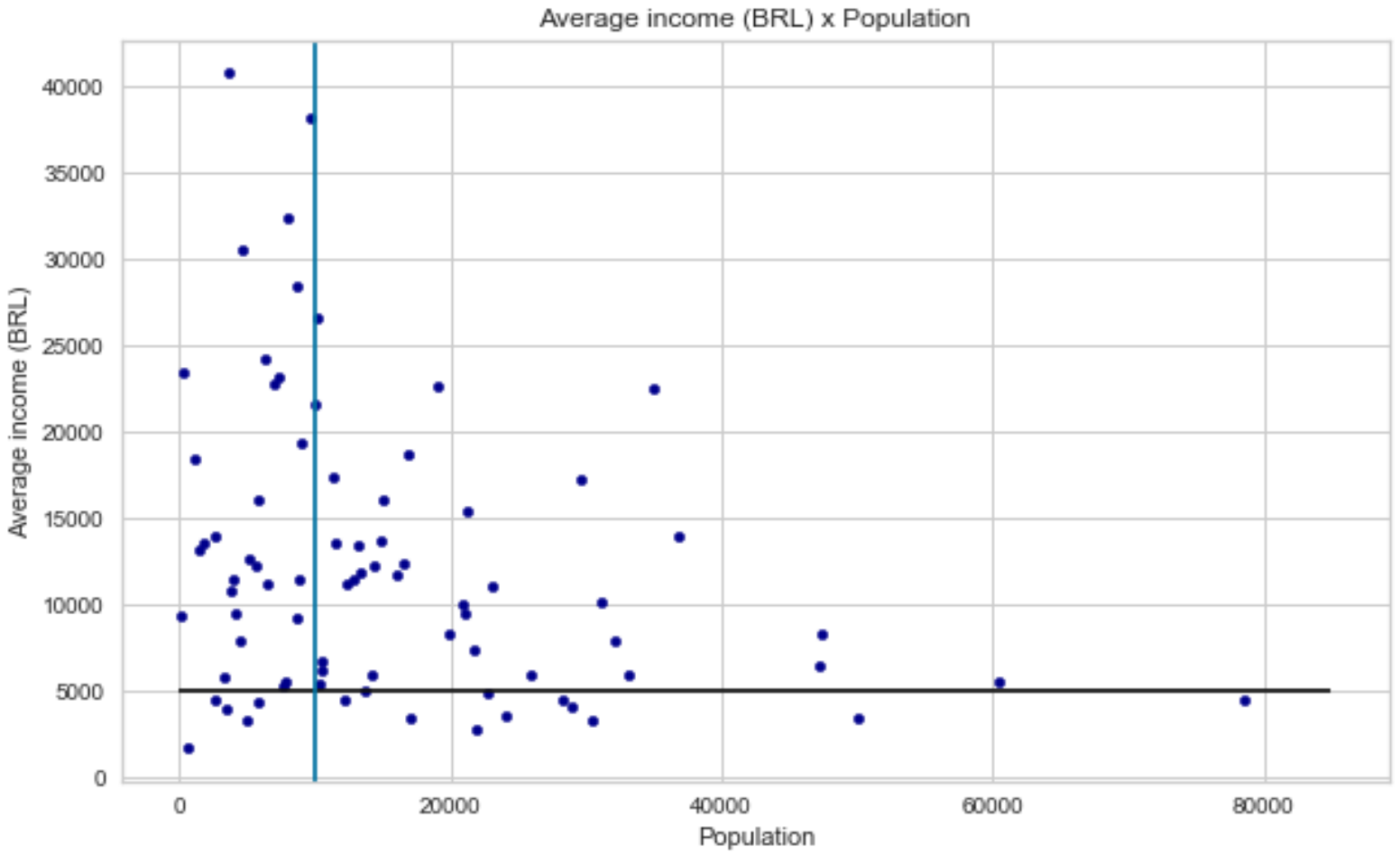
Neighbourhoods with population less than 5,000 were excluded from the analysis.



DATA ANALYSIS

To understand if there was any relationship between the variables, scatter plot was used.

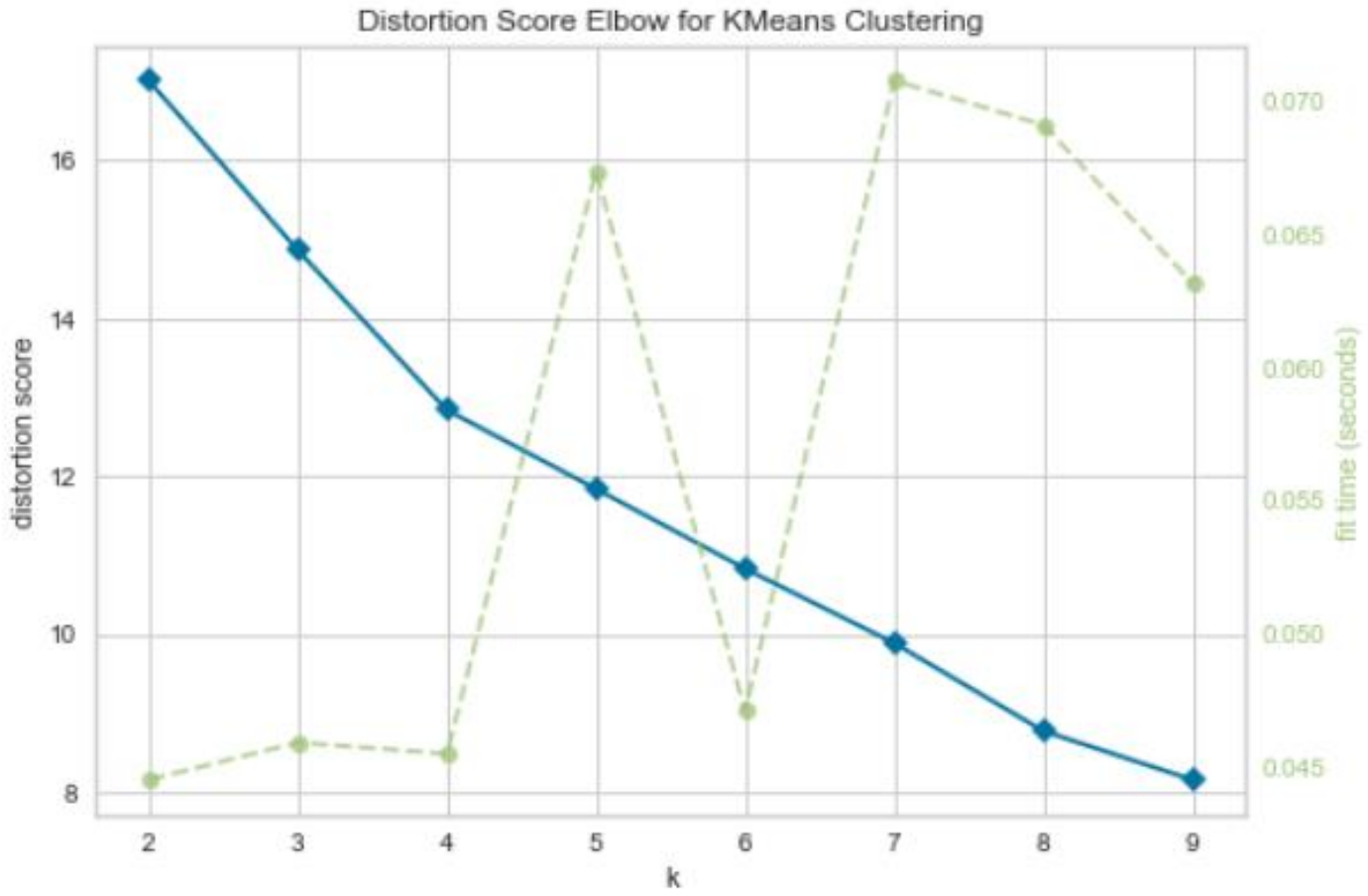
The scatter plot allowed the visualization of the neighbourhoods of interest.



DATA ANALYSIS

Elbow analysis to determine the k for k-means clustering did not show a clear elbow. However, the most significant elbow is on k=4.

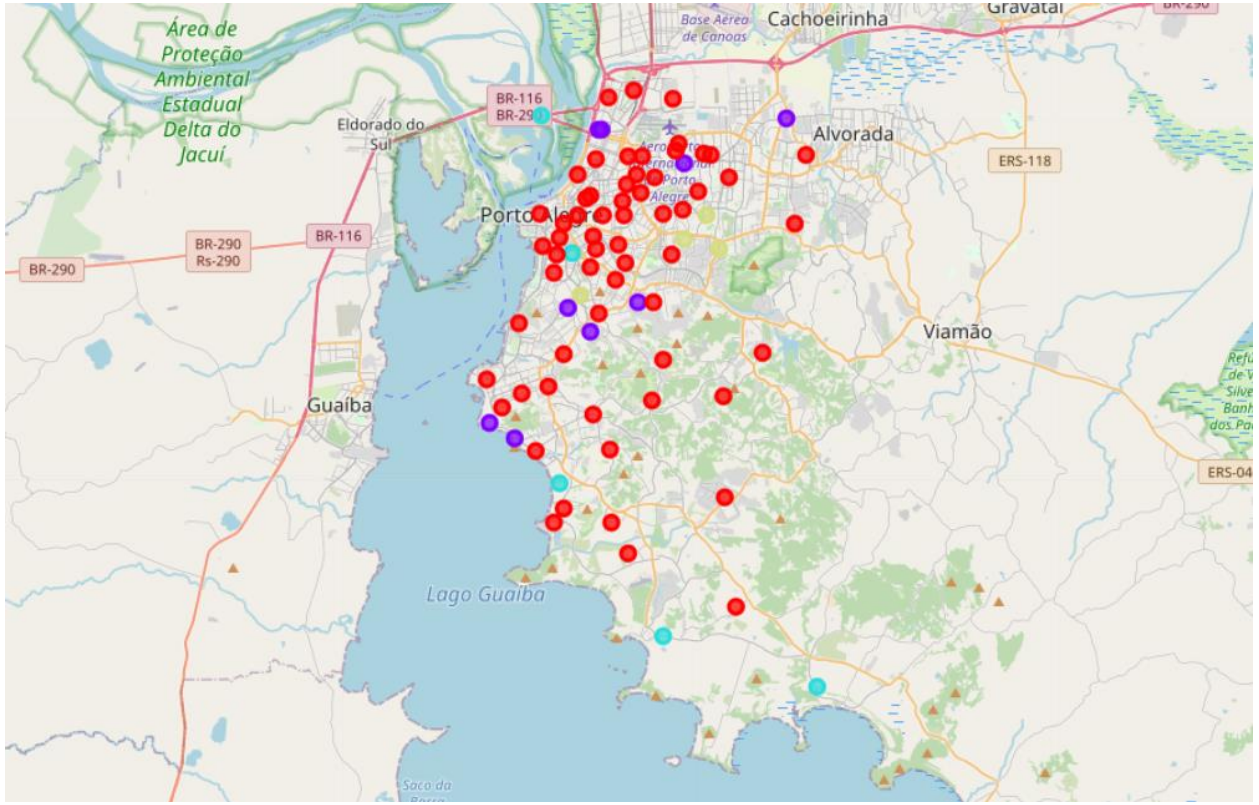
Therefore, there will be 4 clusters.



DATA ANALYSIS

K-Means clustering did not show significant differences between cluster 1, 2 and 3.

Cluster 4 and 5 were excluded from the list of possible locations for the characteristics of the venues and the lack of data, respectively.



DATA ANALYSIS

As there are many neighbourhoods with no Italian restaurants, neighbourhoods with at least one Italian restaurant were excluded from the analysis.

Neighbourhood	Venue Category
Auxiliadora	2
Boa Vista	1
Bom Fim	2
Centro Histórico	1
Chácara das Pedras	2
Cidade Baixa	2
Espírito Santo	1
Humaitá	1
Independência	1
Jardim Botânico	2
Jardim Lindóia	1
Moinhos de Vento	2
Rio Branco	5
Santana	2
São Geraldo	2
São Sebastião	2
Tristeza	1

DATA ANALYSIS

The final list of possible neighbourhoods to open a new Italian restaurant in Porto Alegre is comprised of 16 neighbourhoods.

	Neighbourhood	Population	Average income (BRL)	Latitude	Longitude
0	Azenha	13449	11,803.00	-30.05	-51.22
1	Bela Vista	9621	38,148.00	-30.03	-51.19
2	Cristo Redentor	16103	11,671.00	-30.01	-51.16
3	Floresta	14941	13,629.00	-30.02	-51.21
4	Glória	8809	11,407.00	-30.07	-51.20
5	Higienópolis	9096	19,283.00	-30.02	-51.18
6	Ipanema	16877	18,634.00	-30.13	-51.23
7	Jardim do Salso	5143	12,584.00	-30.05	-51.17
8	Menino Deus	29577	17,160.00	-30.06	-51.22
9	Mont'Serrat	10236	26,477.00	-30.03	-51.19
10	Passo D'Areia	23083	10,956.00	-30.02	-51.18
11	Petrópolis	35069	22,407.00	-30.05	-51.19
12	Santa Cecília	5800	15,983.00	-30.04	-51.21
13	Santo Antônio	14392	12,133.00	-30.05	-51.21
14	São João	13238	13,354.00	-30.01	-51.19
15	Teresópolis	12844	11,341.00	-30.08	-51.21

RESULTS AND DISCUSSION

The analysis showed that restaurants are pulverized in all Porto Alegre. In other words, there is no specific area in which restaurants are extremely concentrated.

Moreover, there are some areas in which Foursquare did not provide any data on their restaurants

After collecting all data required for this project, the first criteria to narrow down the possible neighbourhoods for establishing a new restaurant were the population and the average income in each neighbourhood. This process excluded many locations, reducing the candidate neighbourhoods to 33.

RESULTS AND DISCUSSION

K-Means clustering provided some more information to exclude more neighbourhoods based on the characteristic of the restaurants in those clusters.

It allowed only two clusters to be excluded. The first excluded cluster was characterized by the pubs and bars. These neighbourhoods were excluded because the areas of interest must be characterized by restaurants specifically. The other cluster excluded from the analysis was characterized by not having any data on its restaurants.

Therefore, the last step was to exclude all other neighbourhoods that already have an Italian restaurant. The result of this process was a list of 16 neighbourhoods with no Italian restaurant and with people who have financial conditions to be a customer.

CONCLUSIONS

The main objective of this project was to determine the location of a new Italian restaurant in the city of Porto Alegre, Brazil.

By using Foursquare data to determine the distribution of restaurants in the city and using socio-economical data from the neighbourhoods, it was possible to reduce the area where the restaurant could be established.

The process is replicable to other cities depending on the available data.

Lastly, to determine the final location for the restaurant, other factors would have to be taken into consideration including other characteristics of the neighbourhoods and the strategy to be implemented for this new restaurant.