Opening a restaurant in Santiago City.

Part 1 - A description of the problem and a discussion of the background.

Introduction:

A. Background

Santiago, also known as Santiago de Chile, is the capital and largest city of Chile as well as one of the largest cities in the Americas. It is the center of Chile's most densely populated region, the Santiago Metropolitan Region, whose total population is 7 million, of which more than 6 million live in the city's continuous urban area. The city is entirely located in the country's central valley.

Santiago is the cultural, political and financial center of Chile and is home to the regional headquarters of many multinational corporations.

Santiago generates 45% of the country's GDP. Some international institutions, such as ECLAC (Economic Commission for Latin America and the Caribbean), have their offices in Santiago. The strong economy and low government debt is attracting migrants from Europe and the United States. in addition, it is one of the most profitable and stable economies in the region.

Santiago's steady economic growth over the past few decades has transformed it into a modern metropolis. The city is now home to a growing theater and restaurant scene, extensive suburban development, dozens of shopping centers, and a rising skyline, including the tallest building in Latin America, the "Gran Torre Santiago". It includes several major universities, and has developed a modern transportation infrastructure, including a free flow toll-based, partly underground urban freeway system and the Metro de Santiago, South America's most extensive subway system.

B. Problem Description

First of all I would like to comment that I have been lucky to have traveled a lot and one of my best experience was about the traditional foods of Eastern Europe especially the food of Hungary.

Santiago City of Chile offers many business opportunities due to its characteristics, but there are a large number of competitors and the complexity and infrastructure of the city brings greater challenges in case of opening a business.

C. Objective

Downtown Santiago is considered the core of the city and it is an excellent spot for opening a Hungarian restaurant.

The question is to clarify and locate what is the neighborhood with the fewest competitors in downtown Santiago.

D. Target Audience

The target audience is all interested in clarifying the number of competitors in case of opening a Hungarian restaurant in the downtown Santiago.

Part 2 - A description of the data and how it will be used to solve the problem.

Description of the data

The data consists of the main neighborhoods in Santiago that will be processed and disaggregated using Foursquare's location data to locate the main places and businesses, followed by performing statistical analysis and visualizations to locate the neighborhood with the fewest competitors.

Firstly, to obtain the names, latitude and longitude of the main neighborhoods of the city of Santiago, it will be necessary to scrape the following wikipedia page.
https://es.wikipedia.org/wiki/Anexo:Comunas_de_Chile

Secondly, since not all the required neighborhoods are on the web page from above, it will also be necessary to get neighborhoods from the page below that allow extracting a KML file that will be transformed into an Excel file using Google Earth.
https://www.google.com/maps/d/viewer?ie=UTF8&oe=UTF8&msa=0&mid=1ACWDQ2UW_t_jLBHcV49D4CuIWLE&ll=-33.49700571058003%2C-70.633807&z=11

Lastly, to obtain exactly the requested names of the neighborhoods in downtown Santiago, it will be necessary to scrape the following wikipedia page.
https://es.wikipedia.org/wiki/Anexo:Barrios_de_Santiago_de_Chile

Once the data is obtained and processed from the first three steps, Foursquare location data will be performed to get the competitors and revise other aspects, in addition. The necessary statistical analysis will be performed to clarify the problem.

Part.3 Data Acquisition - Processing - Visualization.

E. Methodology

Data Acquisition, Processing and Visualization.

The overall task is done by an iterative procedure to process and form the necessary charts by means of scraping and forming dataframes which would permit to form the corresponding visualizations.

The visualization would be done mainly by map plots and bar plots to locate the neighborhood with fewest competitors.

Some of the data are in Spanish, thus it is necessary to translate them except the names of the neighborhoods.

The first web scraping would be done by using BeautifulSoup library and it corresponds to the main neighborhoods of each capital city per region in Chile that are conveniently arranged together with is

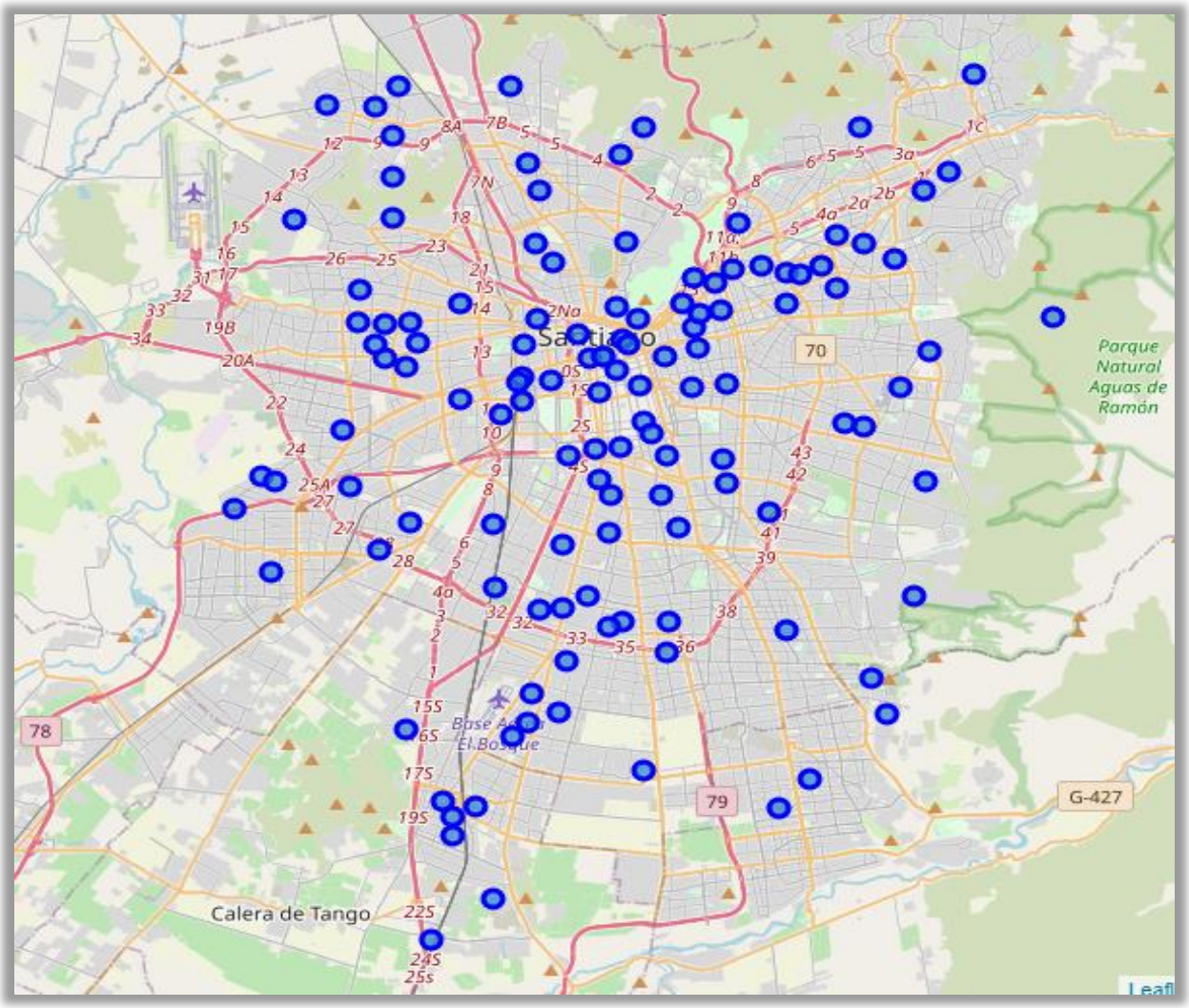
longitude and latitude and other data. This is a wikipedia's page that permits to get the information by means of a wikitable sortable table.

This information would be clean and process with the aim of obtaining a dataframe to display only the neighborhoods of Santiago City.

Because not all the required neighborhoods are in the first web scraping. It would be necessary a second web scraping.

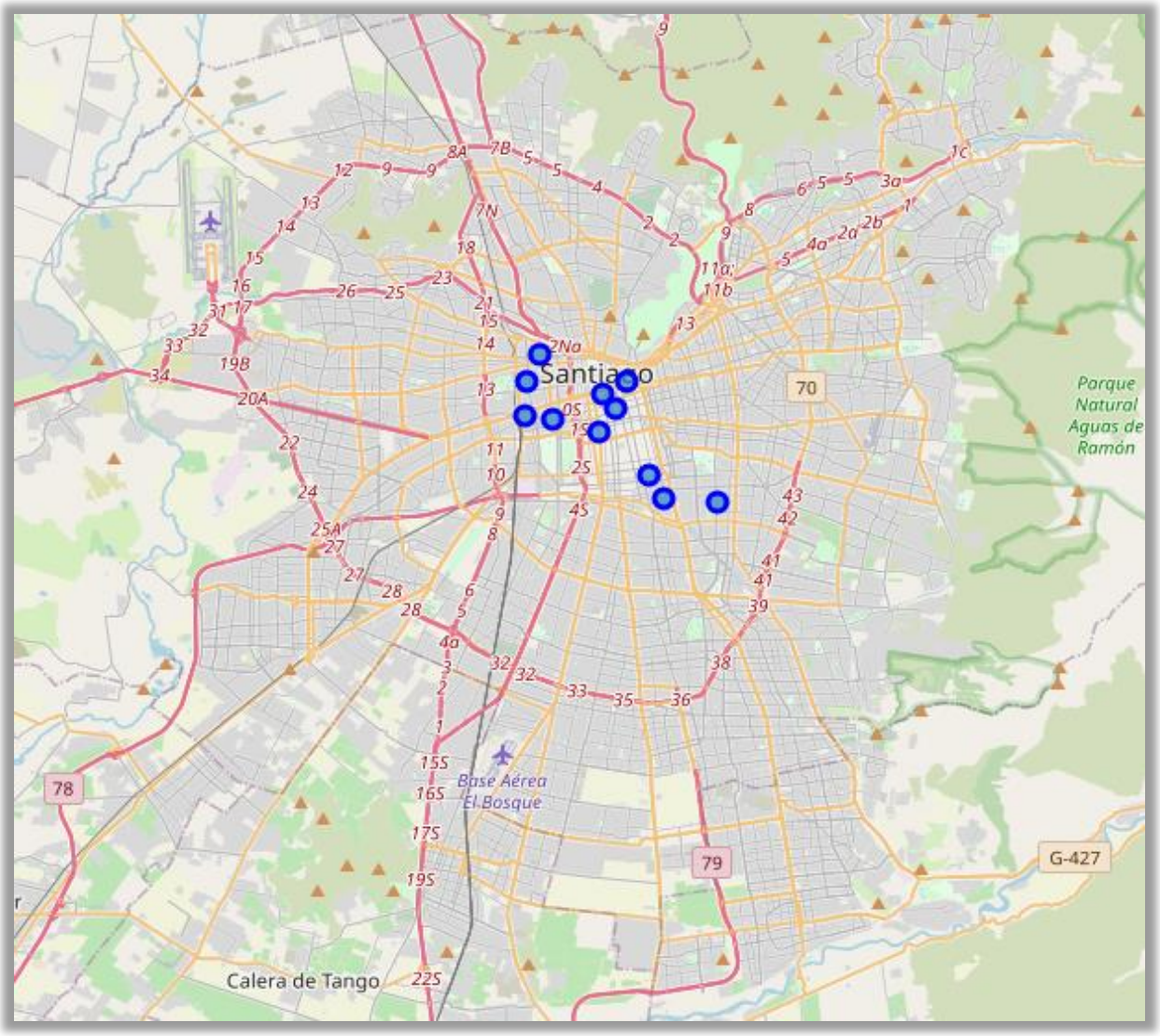
This second action would obtain the data from a web page that allows the data to be downloaded in a KML file, therefore, through Google Earth, the data is obtained through a cvs file that permits to form a second data frame.

As a third action the first two dataframes would be merged to get the total neighborhoods and visualize the main neighborhoods in Santiago City.



The last web scraping is be done by using BeautifulSoup library in a wikipedia page that has the main neighborhoods from Downtown Santiago. Therefore allowing us to do the same procedure as the first web scraping because it has a wikitable sortable table as well.

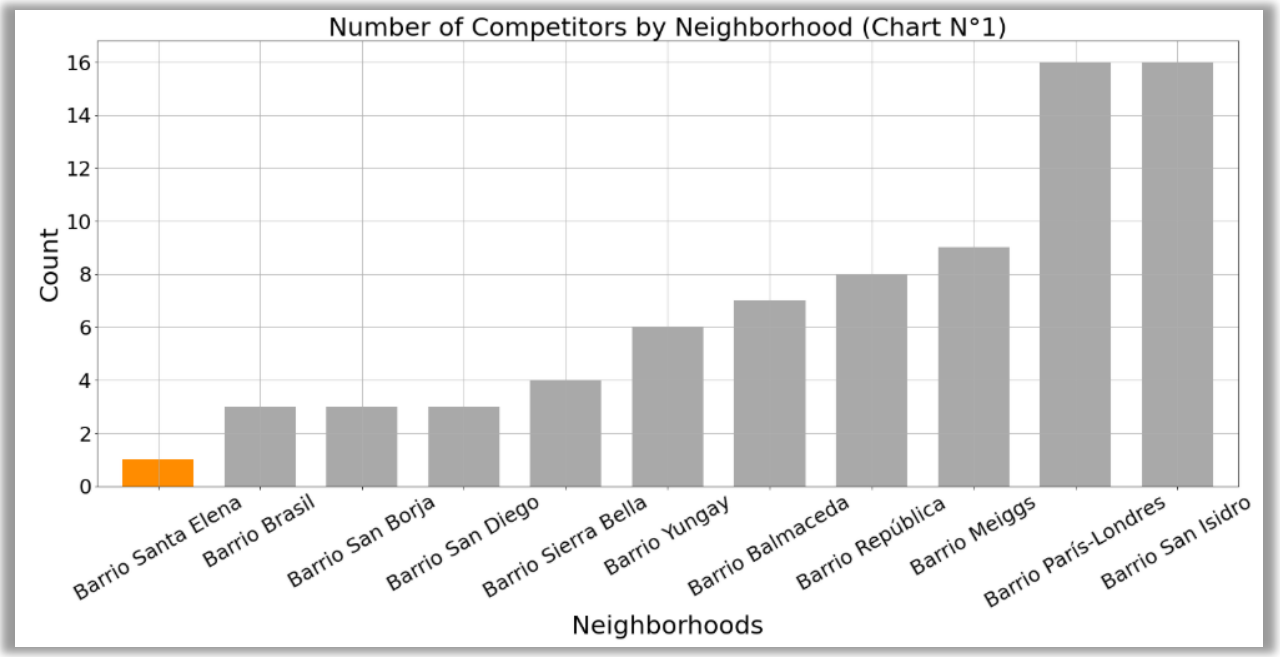
Now we have the dataframes to obtain the latitude and longitude from the Downtown's neighborhoods by matching the resulting dataframes and dropping the unnecessary columns to form a new dataframe to visualize the neighborhoods within Downtown.



The following objective would be to create a dataframe to see the top 10 Common venues per neighborhood in Downtown by means of using the Foursquare location data to explore the Neighborhoods and put that into a pandas dataframe.

	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	Barrio Balmaceda	Peruvian Restaurant	Pharmacy	South American Restaurant	Pizza Place	Mobile Phone Shop	Chinese Restaurant	Farmers Market	Martial Arts Dojo	Park	Skate Park
1	Barrio Brasil	Peruvian Restaurant	Liquor Store	Soccer Stadium	Restaurant	Seafood Restaurant	Discount Store	Dive Bar	Donut Shop	Electronics Store	Event Space
2	Barrio Meiggs	Shopping Mall	Pharmacy	Fried Chicken Joint	Sandwich Place	Restaurant	Café	Business Service	Asian Restaurant	Fast Food Restaurant	Nightclub
3	Barrio París-Londres	Coffee Shop	Sandwich Place	Pizza Place	Bookstore	Plaza	Restaurant	Café	Theater	Chinese Restaurant	Hotel
4	Barrio República	Pizza Place	Chinese Restaurant	Pub	Hot Dog Joint	Burrito Place	Bus Station	Restaurant	Lounge	Food	Martial Arts Dojo
5	Barrio San Borja	Hotel	Art Gallery	Art Museum	Gift Shop	Coffee Shop	Park	Yoga Studio	Mountain	Restaurant	Deli / Bodega
6	Barrio San Diego	Diner	Restaurant	Men's Store	Bus Station	Bus Line	Café	Food & Drink Shop	General Entertainment	Flea Market	Bistro
7	Barrio San Isidro	Pizza Place	Sushi Restaurant	Restaurant	Peruvian Restaurant	Latin American Restaurant	Fried Chicken Joint	Indian Restaurant	Gym / Fitness Center	Playground	Pharmacy
8	Barrio Santa Elena	Peruvian Restaurant	Farmers Market	Convenience Store	Sandwich Place	Park	Shoe Store	Bakery	Food	Flea Market	Fast Food Restaurant
9	Barrio Sierra Bella	BBQ Joint	American Restaurant	South American Restaurant	Health Food Store	Italian Restaurant	Paper / Office Supplies Store	Pet Store	Plaza	Chinese Restaurant	Hot Dog Joint
10	Barrio Yungay	Bar	Peruvian Restaurant	Museum	Restaurant	Latin American Restaurant	Performing Arts Venue	Historic Site	Metro Station	Event Space	Coffee Shop

Once the dataframe is done the direct competitors would be extracting to create a new dataframe of competitors with the aim of plotting a bar chart and visualize the number of them per neighborhood.

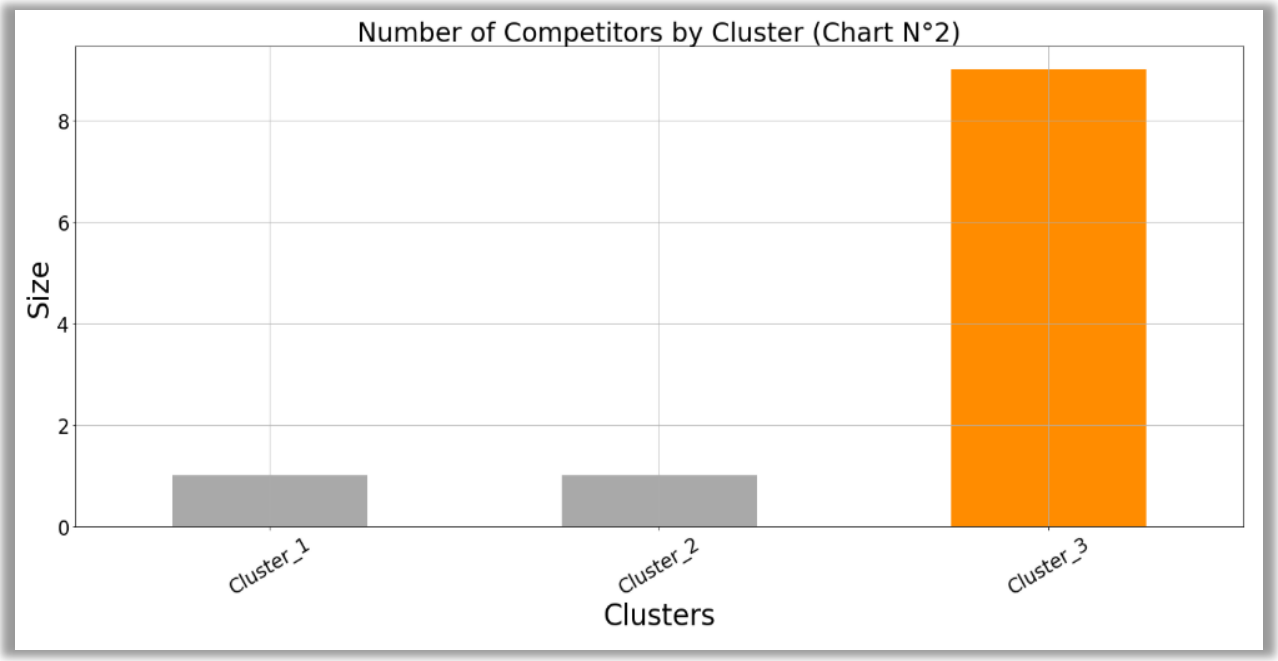


The previous tasks showed that the neighborhood "Barrio Santa Elena" is the place with fewest competitors.

Now a different approach by using "K-Means Clustering" would be taken by clustering the neighborhoods by similarities and visualizing the numbers of competitors per cluster.

The clustering and subsequent visualization showed that the neighborhood "Barrio santa Elena" the place with fewest competitors and it is also standing out according to similarities.

The neighborhood "Barrio Santa Elena“ belongs to the cluster_1



Part 4. Results - Discussion - Conclusion

F. Results

The results reflected that is possible to answer the initial question whether if it is possible to start a Hungarian restaurant or not according to the number of competitors within Downtown Santiago if the neighborhood with fewest competitors is chosen.

G. Discussion

On the one hand, one point to highlight is that the task perfectly clarified the initial question, leaving no doubts on the matter, but on the other hand, raised questions about whether it is better to perform new questions and searches or not.

As a recommendation for those who start operating a Hungarian restaurant or any type of restaurant within Downtown Santiago, look for the place with the fewest competitors nearby is a good idea but it is necessary to think over the problem because this question don't answer all the perspectives to clarify the best spot to place it.

H. Conclusion

As a consequence of the results, it is possible to answer the question that is to clarify and locate what is the neighborhood with the fewest competitors in downtown Santiago.

The answers are the neighborhood called "Barrio Santa Elena" and the location is shown in the “Visualization N°3”. But due to the amount of information collected during the task is not enough, it is possible to conclude that it is necessary to perform new questions and analyzes to make sure where place is the best to open a Hungarian restaurant.

I. Reference

1. Communes and Neighborhoods of Chile.
https://es.wikipedia.org/wiki/Anexo:Comunas_de_Chile

2. Main Neighborhoods of Santiago City.
https://www.google.com/maps/d/viewer?ie=UTF8&oe=UTF8&msa=0&mid=1ACWDQ2UW_t_jLBHcV49D4CuIWLE&ll=-33.49700571058003%2C-70.633807&z=11

3. Main Neighborhoods of Downtown Santiago.
https://es.wikipedia.org/wiki/Anexo:Barrios_de_Santiago_de_Chile