Analysis of the Venues per District in the city of Santiago, Chile

Juan Pablo Dib

May 3, 2019

1. Introduction

1.1 Business Problem Description

Santiago de Chile is the largest city in Santiago, concentrating over 7 million people in a country which total population is 17.5 million, comprising more than 40% of Chile's total population in one city. This city is divided into 34 counties/districts located in an area of 837.89 squared kilometers(source)(s2). This creates huge differences in density for each district translating into a critical decision for a new business on where to focus their efforts to start with. As an entrepreneur the idea of creating a new business in this city could cause big headaches as to where should it be located and what should be the target market. This report will be guided by the following questions:

```
1. What is the best district in the city of Santiago, Chile for a
new business?
2. What should be the focus of this new business to thrive in that
county?
3. Can that business expand to other districts?
```

We will leverage the results of the Chilean National Institute of Statistics regarding population by district, correlating it to their geographical placement and the venue density, and type in each district via the Foursquare API.

1.2 Data Description

The following information sources will be considered for this project.

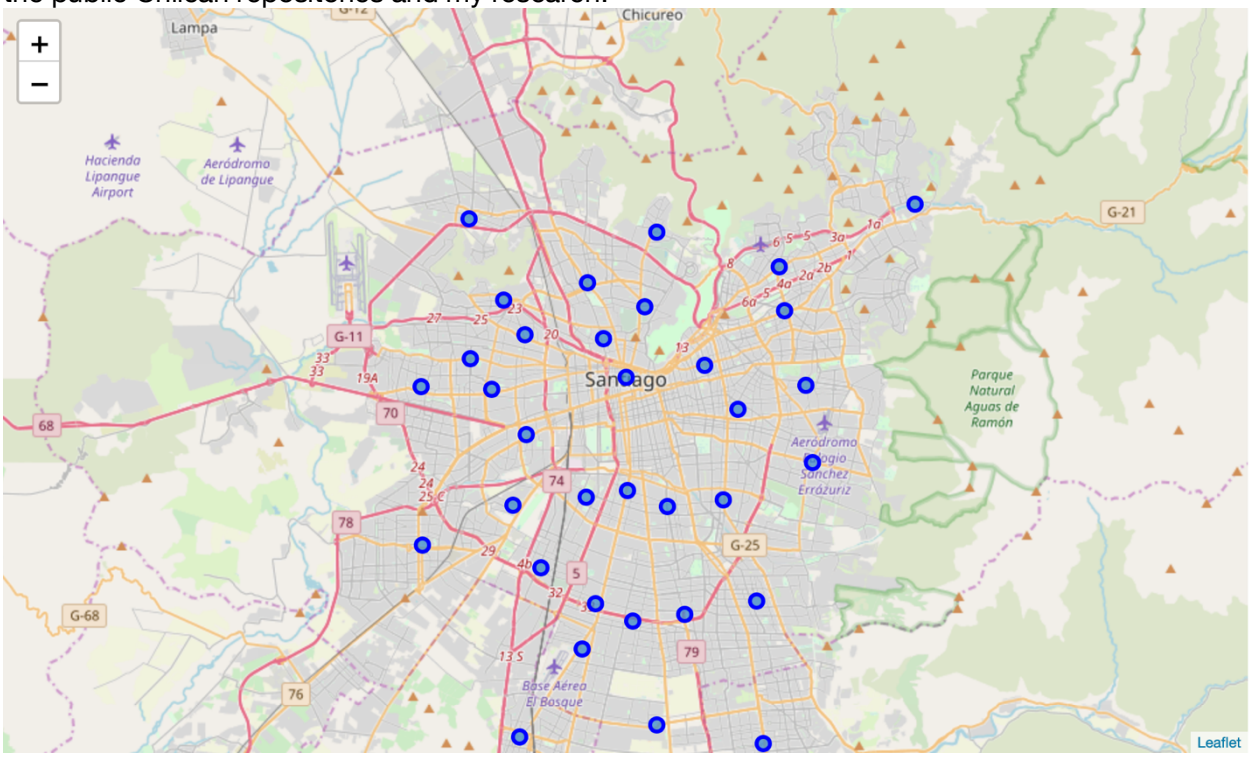
- Venue Information -- Forsquare API to retrieve every venue located in the city of Santiago.
- Demographics -- Information is being leveraged from the INE (National Institute of Statistics) regarding the population of all the districts.
- Geographical information -- I researched for the location of all the districts in Santiago.

2. Methodology

As explained in point 1.2 I merged the demographical and geografical information onto one pandas table. This includes the columns *Comunas(district), Latitude, Longitude, Hombres(men), Mujeres(women), Total*. The last three columns refer to the population of Men, Women and total of each District.

	Comuna	Latitude	Longitude	Hombres	Mujeres	Total
0	San Ramon	-33.542156	-70.647184	40873	42027	82900
1	Renca	-33.403697	-70.713522	72681	74470	147151
2	Pedro Aguirre Cerda	-33.488703	-70.670953	49513	51661	101174
3	La Granja	-33.538966	-70.619992	57025	59546	116571
4	El Bosque	-33.553833	-70.673269	79372	83133	162505

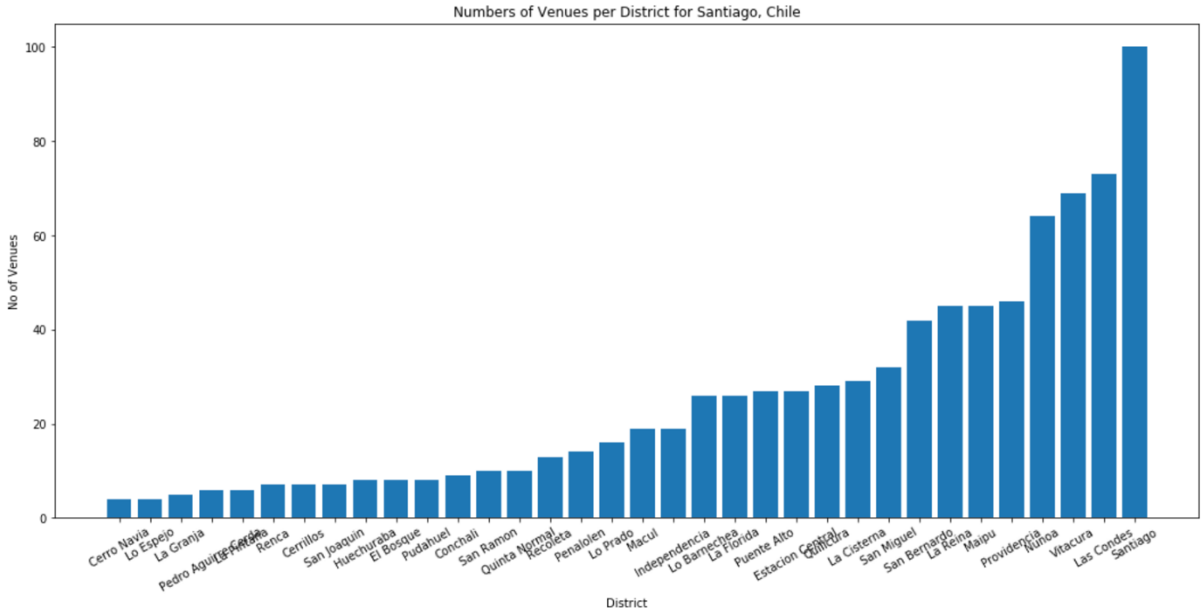
Utilizing the Folium library, I could map the city on a Zoom of 11. This was also completed by adding markers with each of the districts coordinates, for which I processed the databases obtained from the public Chilean repositories and my research.



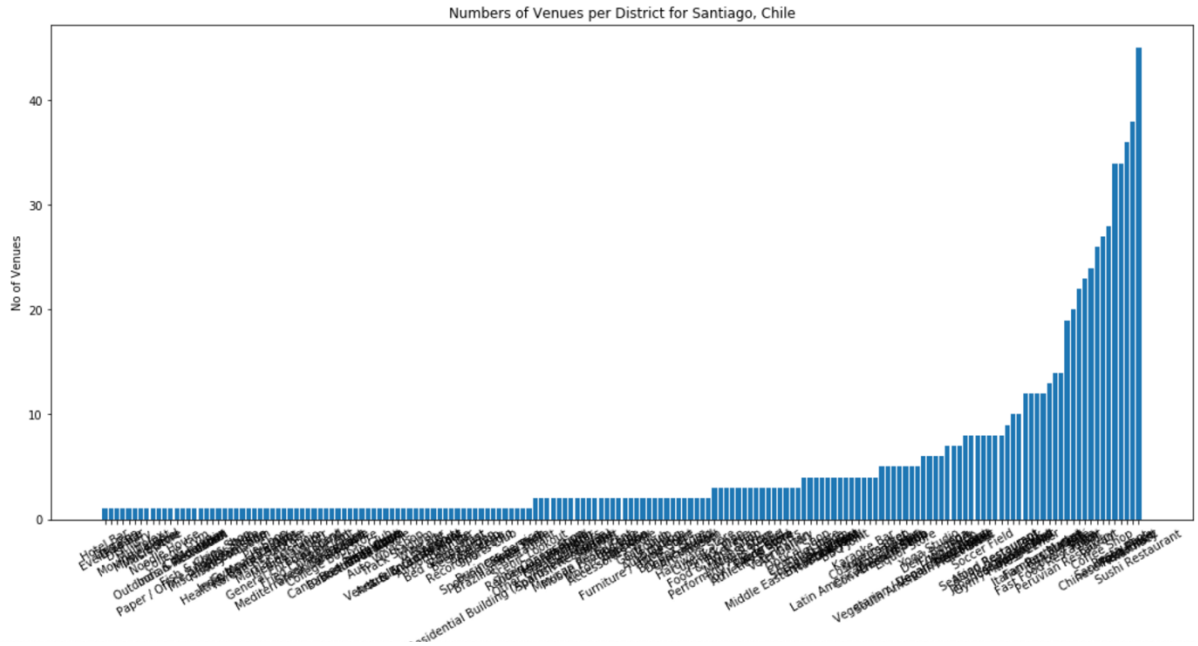
The Foursquare API was utilized to obtain all the venues close to the districts center, limits were stablished as 750 meter radius for each one, with a venue limit of 100 restricted by the API itself. The following mini table shows an example of the results.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	San Ramon	-33.542156	-70.647184	Parque La Bandera	-33.545825	-70.645319	Park
1	San Ramon	-33.542156	-70.647184	Feria Libre Fernandez Albano	-33.539116	-70.649408	Farmers Market
2	San Ramon	-33.542156	-70.647184	Señor del Bajón	-33.542524	-70.644486	Sandwich Place
3	San Ramon	-33.542156	-70.647184	Metro San Ramón	-33.539978	-70.645304	Metro Station
4	San Ramon	-33.542156	-70.647184	El Señor de los Bajones 2 - San Ramon	-33.541119	-70.643669	Food Truck

We can see from the list above that the venues top places by number of venues were Santiago, Las Condes, Vitacura, Nunoa, Providencia. Results that are shown here rely on the location of the district center, regarding that, results could be affected by this and the API imposed venue limit of 100 per query. There were 174 unique venues divided in all the city.



As we can see in the chart above, this is the distribution of venues per district. Where there is a big gap between different districts, with more than 20 times the amount of venues. Taking into account the following chart, we can see the distribution of venues by their type, identifying which of them are the favorites.



This is reviewed in the following chart which identifies the top 5 venues.

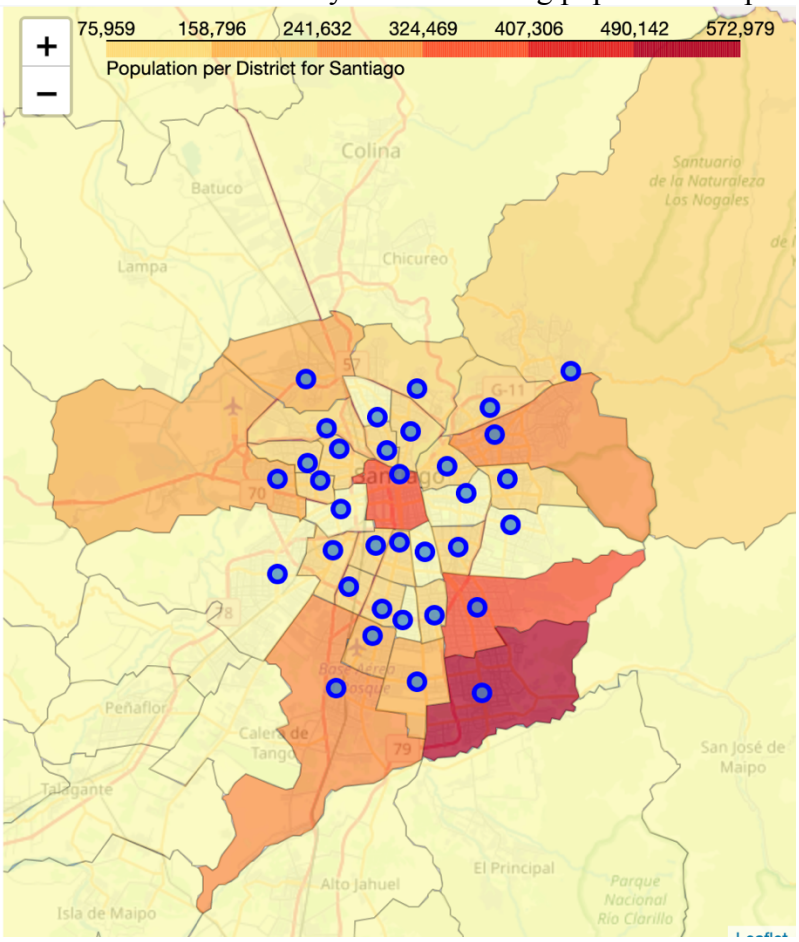
	Venue Category	Neighborhood
143	Sandwich Place	34
140	Restaurant	34
131	Pizza Place	36
129	Pharmacy	38
161	Sushi Restaurant	45

3. Results

Now we will view all the results of this experiment. Right below we can view some districts and their top 10 types of venues.

	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	Cerrillos	Rental Car Location	Airport	Art Gallery	Chinese Restaurant	Music Venue	Restaurant	Bus Station	Yoga Studio	Fast Food Restaurant	Farmers Market
1	Cerro Navia	Pizza Place	Japanese Restaurant	Cantonese Restaurant	Asian Restaurant	Yoga Studio	Diner	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Event Space
2	Conchali	Park	Japanese Restaurant	Cupcake Shop	Pub	Bus Station	Sandwich Place	Asian Restaurant	Pharmacy	Gym / Fitness Center	Department Store
3	El Bosque	Pharmacy	Department Store	Bar	Flea Market	Sushi Restaurant	Shopping Mall	Ice Cream Shop	Park	Distillery	Fish & Chips Shop
4	Estacion Central	Bakery	Burger Joint	Asian Restaurant	Pharmacy	Food	Farmers Market	Metro Station	Bus Station	Pool	Plaza

So further on we can analyze the following population map of all Santiago.



The biggest opportunity lies on districts like Cerro Navia, La Pintana, Pudahuel, Lo Espejo and La Granja.

	Comuna	Latitude	Longitude	Hombres	Mujeres	Total	Venue	
16	Cerro Navia	-33.429087	-70.730442	65438	67184	132622	4	33155.
10	La Pintana	-33.586674	-70.634802	87044	90291	177335	6	29555.
7	Pudahuel	-33.440915	-70.755812	112412	117881	230293	8	28786.
25	Lo Espejo	-33.518828	-70.694159	49146	49658	98804	4	24701.
18	La Granja	-33.538966	-70.619992	57025	59546	116571	5	23314.

4. Discussion

Santiago is the biggest city in Santiago with a population of over 7 million people. This is quite a big number for a country this long. The districts studied have provided us further information on how it is distributed by venues, population, the favorite venues and the ‘hotest’ districts.

So, we can see from the table above that the best opportunity does not rely on the most populated areas or the densest by amount of venues, but by the amount of people per venue per district. In this list the top 5 would be Cerro Navia, La Pintana, Pudahuel, Lo Espejo and La Granja. Further on we have discussed what types of venues are the favorites for people that live in Santiago.

In future studies there could be done more statistics, basing the location on areas instead of specific points, taking into consideration also the density of the population for the districts, improving this model with valuable information as it is based on basic hypothesis.

5. Conclusion

As a result, there could be real potential in the questions asked before which are:

1. What is the best district in the city of Santiago, Chile for a new business?
    - Based on the Population against venues analysis, the best district would be Cerro Navia. Taking into account that there is a great opportunity as the market has a gap on the amount of venues it have, enhancing the potential success for it.
2. What should be the focus of this new business to thrive in that county?
    - The business should focus on a Sushi House, Pharmacy. I discarded the Pizza Place as it is the most common venue for that district, leaving small opportunity for another business to thrive in that economy.
3. Can that business expand to other districts?
    - Yes, there is a great opportunity in La Pintana, Pudahuel, Lo Espejo and La Granja.

The process of gathering and polishing this data may help and expand the chances of success for future ventures. Not only on this city, but this same process can be expanded to different areas and markets.

Regards,

Juan Pablo Dib