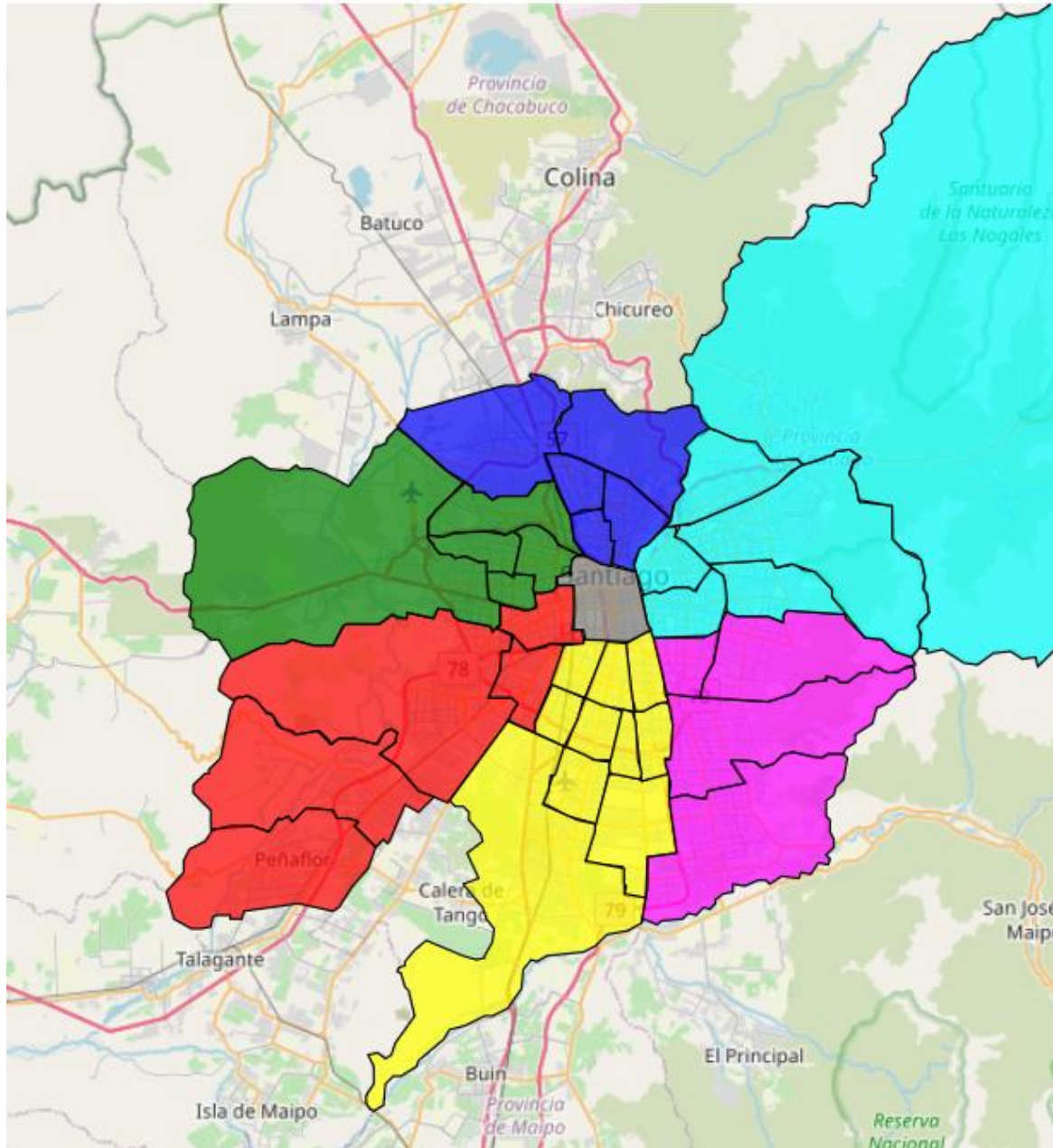




Coffee Shops in SCL

Final Project for applied data
science capstone course



Introduction

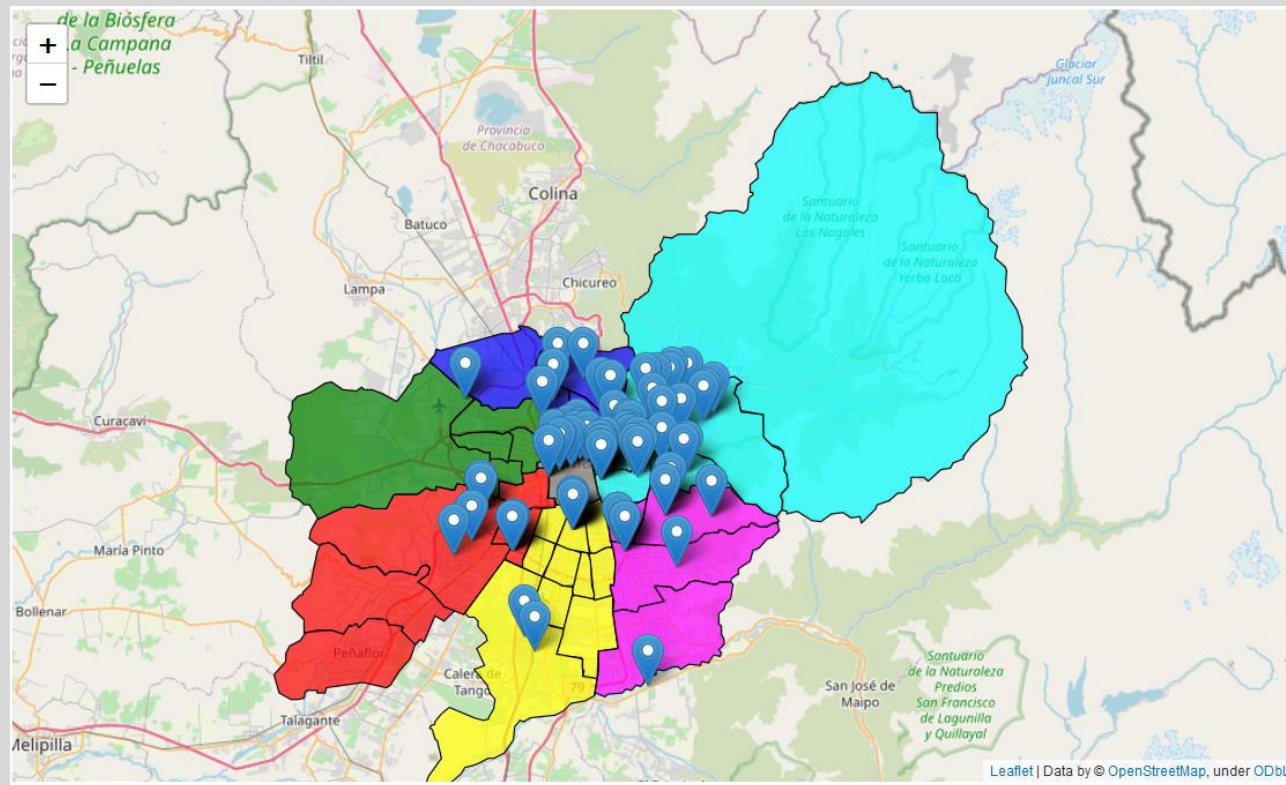
- The goal of this Project is to analyze and propose a location for a new Coffee Shop in the city of Santiago de Chile (SCL).

Data

Data will be extracted from 3 different sources:

- Wikipedia for the city characteristics: Sectores (áreas), Comunas (neighborhoods) and Población (population).
 - https://es.wikipedia.org/wiki/Anexo:Comunas_de_Santiago_de_Chile
- Geojson data from a github repository
 - <https://github.com/jlhonora/geo>
- Coffee Shops data from Foursquare (name, address, foursquare's id and rating).

Location of Coffee Stores in SCL



Data processing

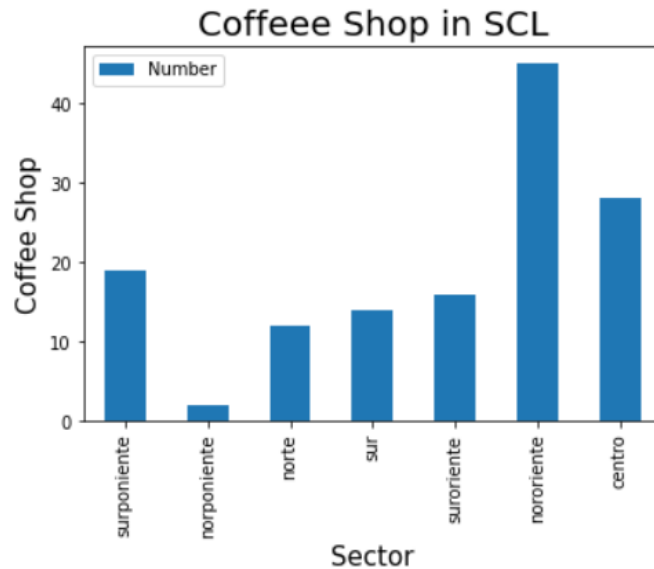
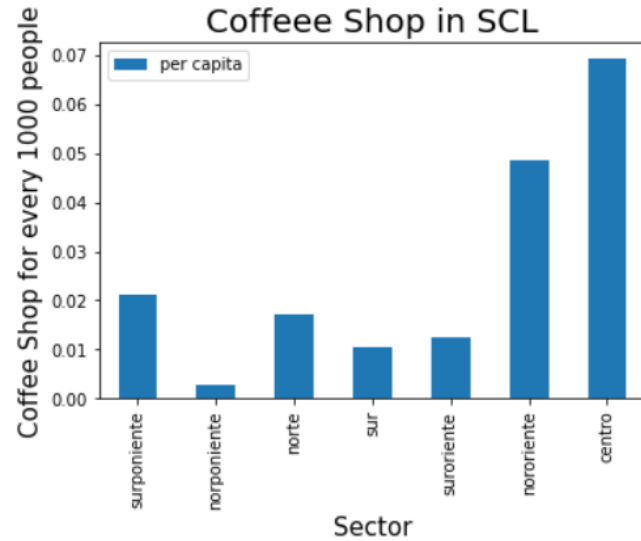
For the data processing next steps are done:

- A map of the city is shown on a choropleth map, with each Sector (área) on different color.
- Data is obtained from Foursquare. The category of the store must include the words 'Coffee', 'Café' or 'Cafetería'. If not, the store is deleted from the dataframe.
- If the store doesn't have address, latitude and longitude, address is left as nan, and is assigned the Comuna's geographical location.

Criteria

The criteria of decision for the location, is that the Sector (area) must have a high number of Coffee Shops (So the zone already is attracting people for Coffe and Food stores) but not the highest density (to avoid a high competetition). To decide the Comuna, it will be the one with the best rating

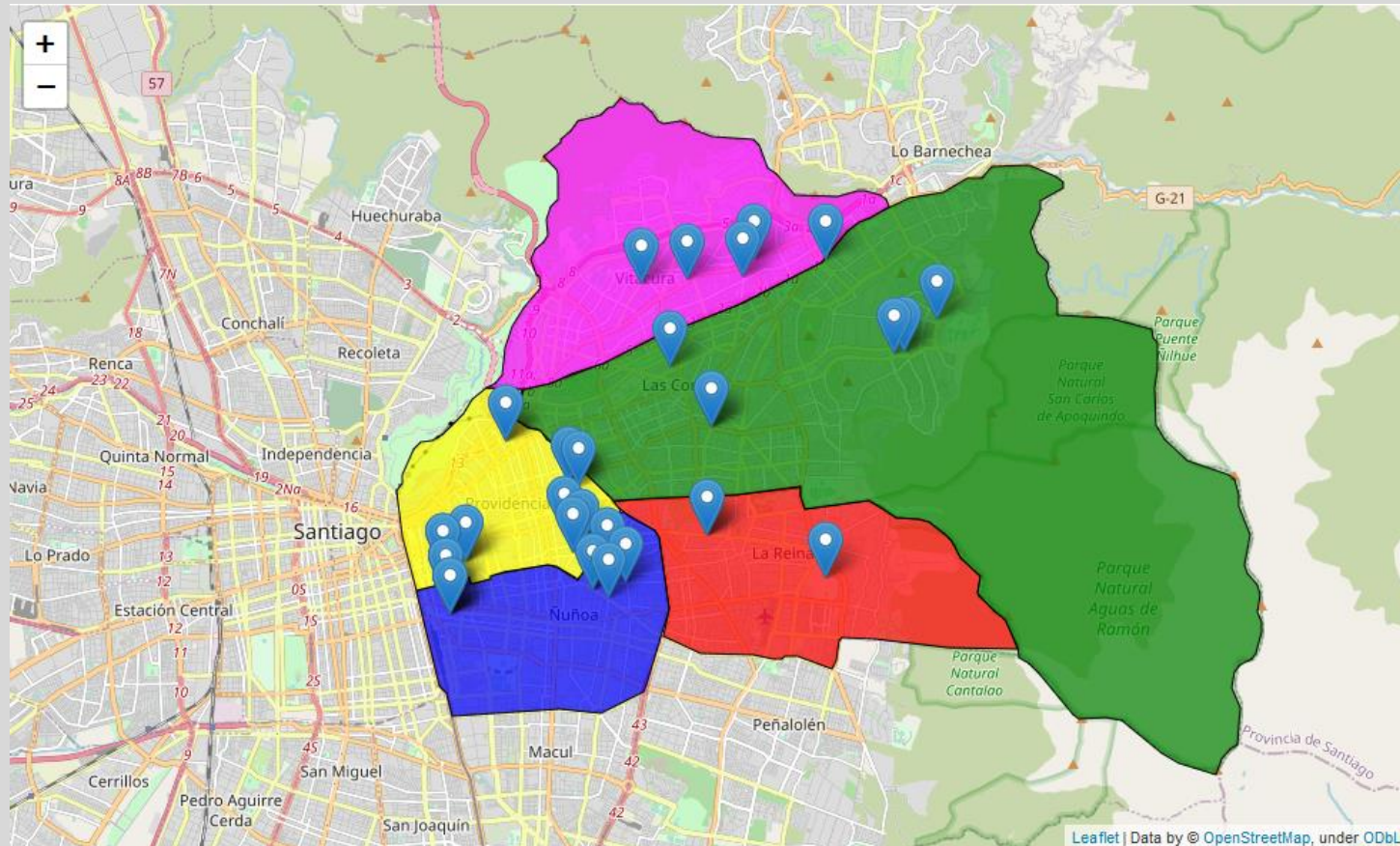
- Sector:
 - Highest density is ruled out.
 - Of the other sectors, the one with the highest number of stores Will be selected.
- Comuna:
 - Highest density is ruled out.
 - Of the remaining comunas, the one with the best average rating is selected.

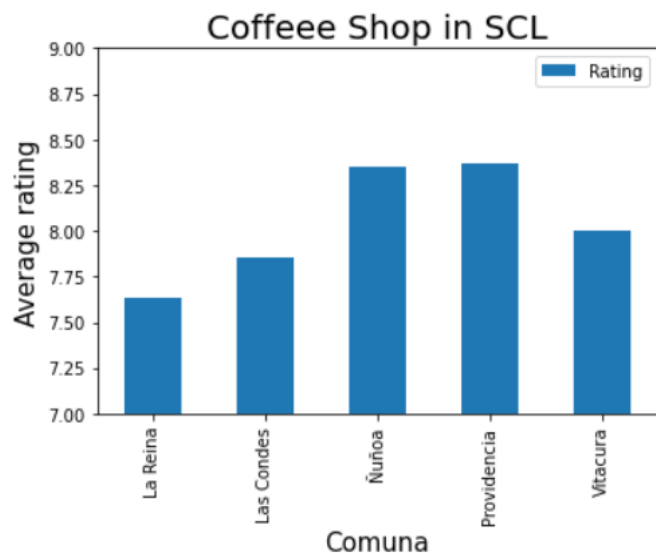
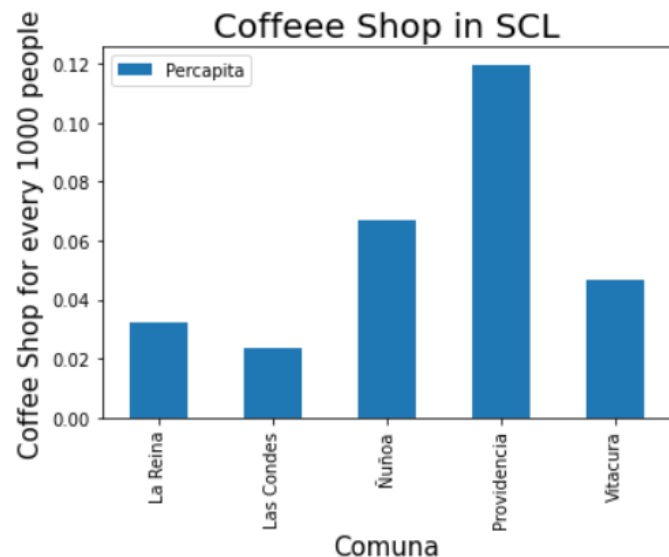


Analysis of Sectors

- It's shown on left that the central area of the city is much more dense (centro), followed by the northeastern (nororiente) area and then by southwestern (surponiente) area.
- There are more stores in the northeastern area (nororiente), followed by central (centro) and southwestern (suroriente) areas.
- Based on the criterion chosen, the Sector selected is the northeastern (nororiente).

Northeastern Coffee Shops in SCL





Analysis of Comunas

- It's shown on left that Providencia has the highest density, followed by Ñuñoa and Vitacura.
- The best average ratings are in Providencia, Ñuñoa and Vitacura.
- Based on the criterion chosen, the neighborhood proposed for a new Coffee Shop is in Ñuñoa.

Discussions

- The last image shows that there is a high concentration of Coffee Shops in the limit between Providencia and Ñuñoa, creating a concentration of stores in the northern zone of the Comuna. This might be due to a special condition of the zone like a commercial area or office buildings that are attractive for Coffee and food business
- It must be noticed that the number of Coffee Shops might be higher, because it will depend if the place is registered in the Foursquare database and if it fits with the criteria imposed in the data analysis.
- Another point to notice in the last image is that six stores can be counted, this low number can lead to a misrepresentation, so in further work data must be checked against other databases.
- A conclusion is achieved within the scope of this work, but for a solid advice for a location, other aspects must be taken in consideration.

Conclusion

- With the analysis made, the selected neighborhood for a new Coffe Shop in Santiago de Chile is the Comuna of Ñuñoa.
- This Comuna presents another Coffee shops that already have a good rating average but without the density of Providencia.
- The second alternative (that also match the criteria imposed) is the Comuna of Vitacura.

