SANTIAGO DE CHILE'S FOOD VENUES CLUSTERIZATION

Camilo Fuentes Moenne

January 2020

Introduction and Problem

- Very recent social crisis in Chile affecting a large amount of little businesses.
- There has never been a categorization of how the different communes behave in terms of food venues preferences. This investigation will propose a method to group and assign preferences to the group of communes.

Goal:

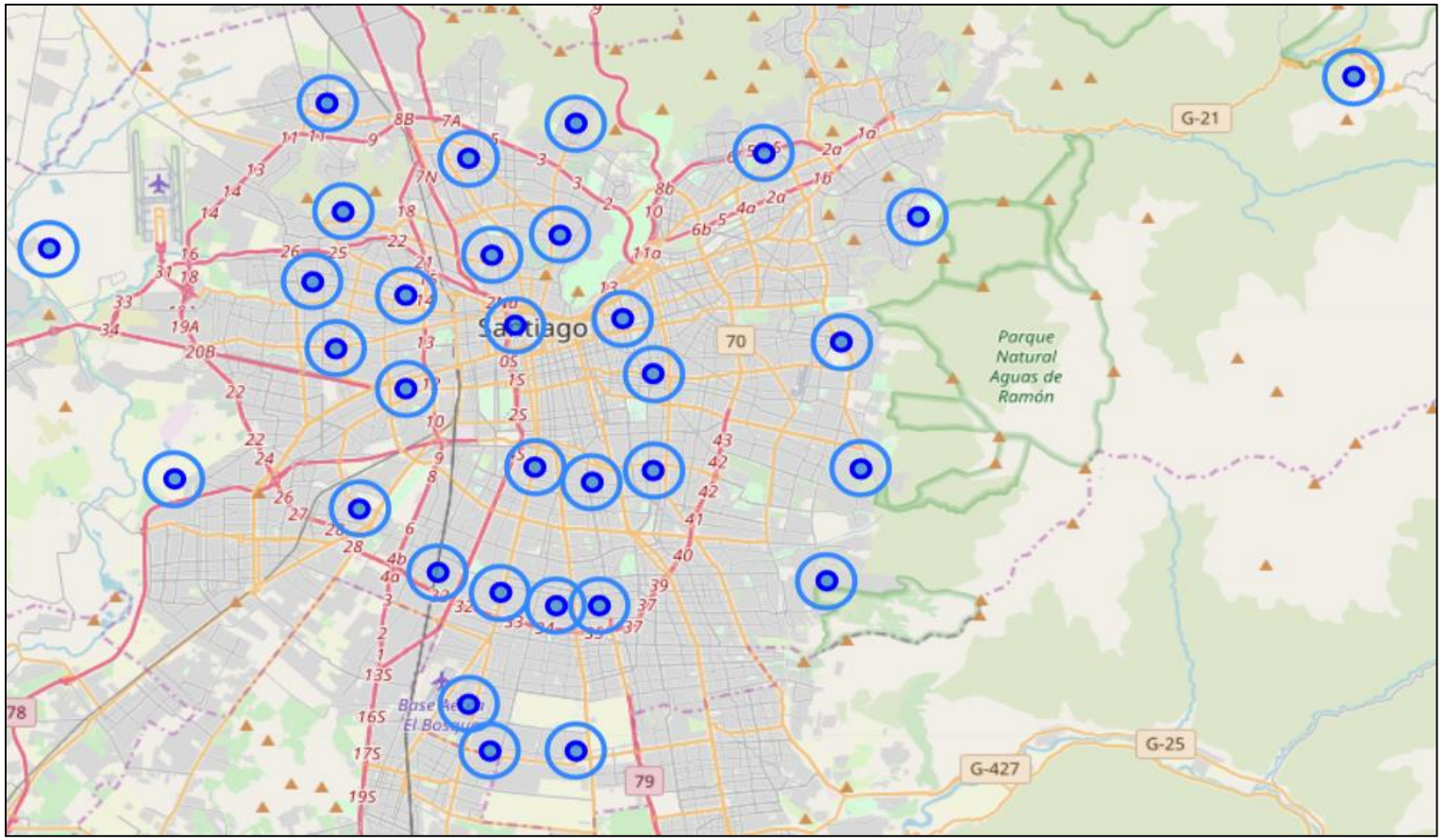
The results of this investigation can be used to suggest possible expansion or relocation for food venues based on the popularity of their category in a group of communes.

Data Acquisition

- Foursquare for the venues information

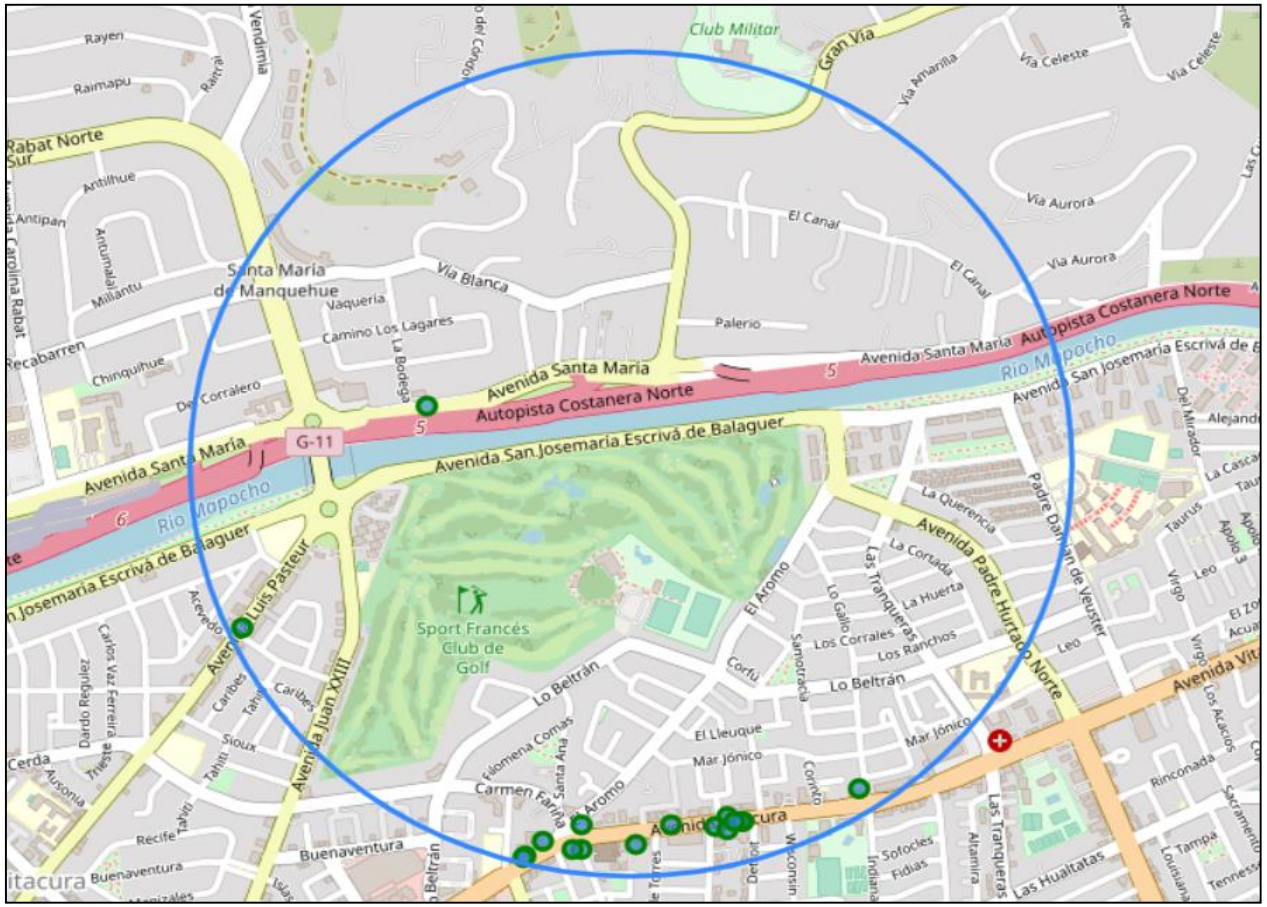
- Public data for commune geospatial location.

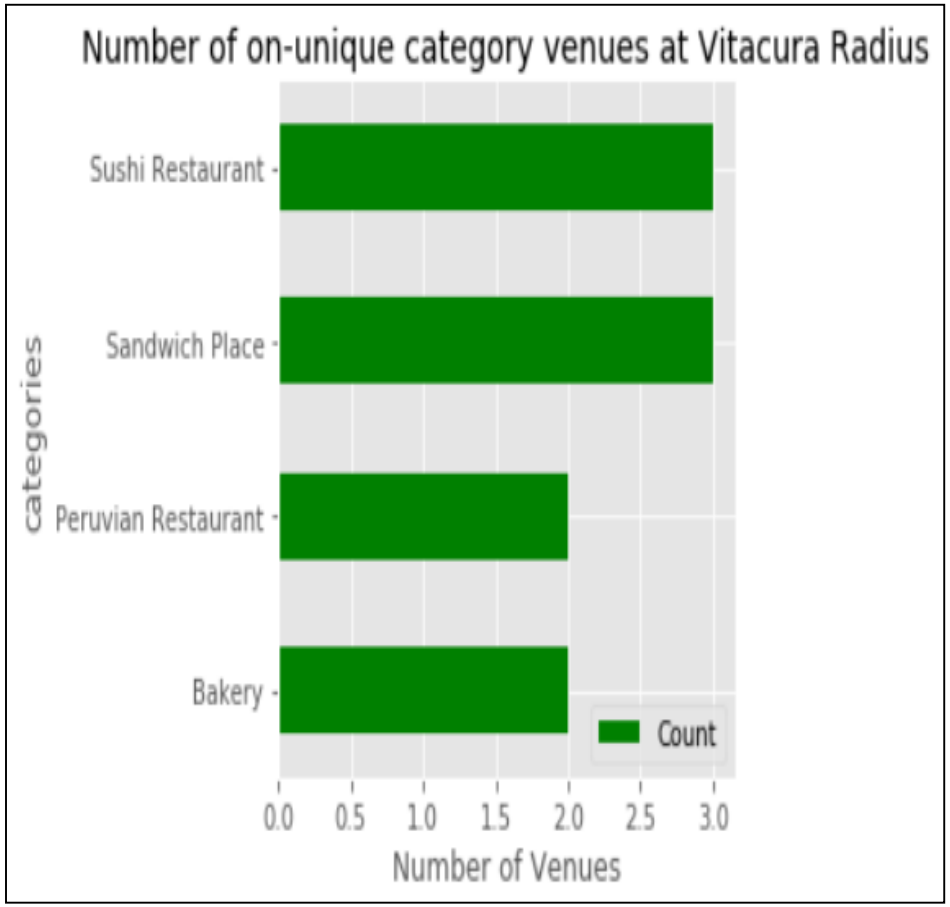
Defining initial radiuses and heuristic:



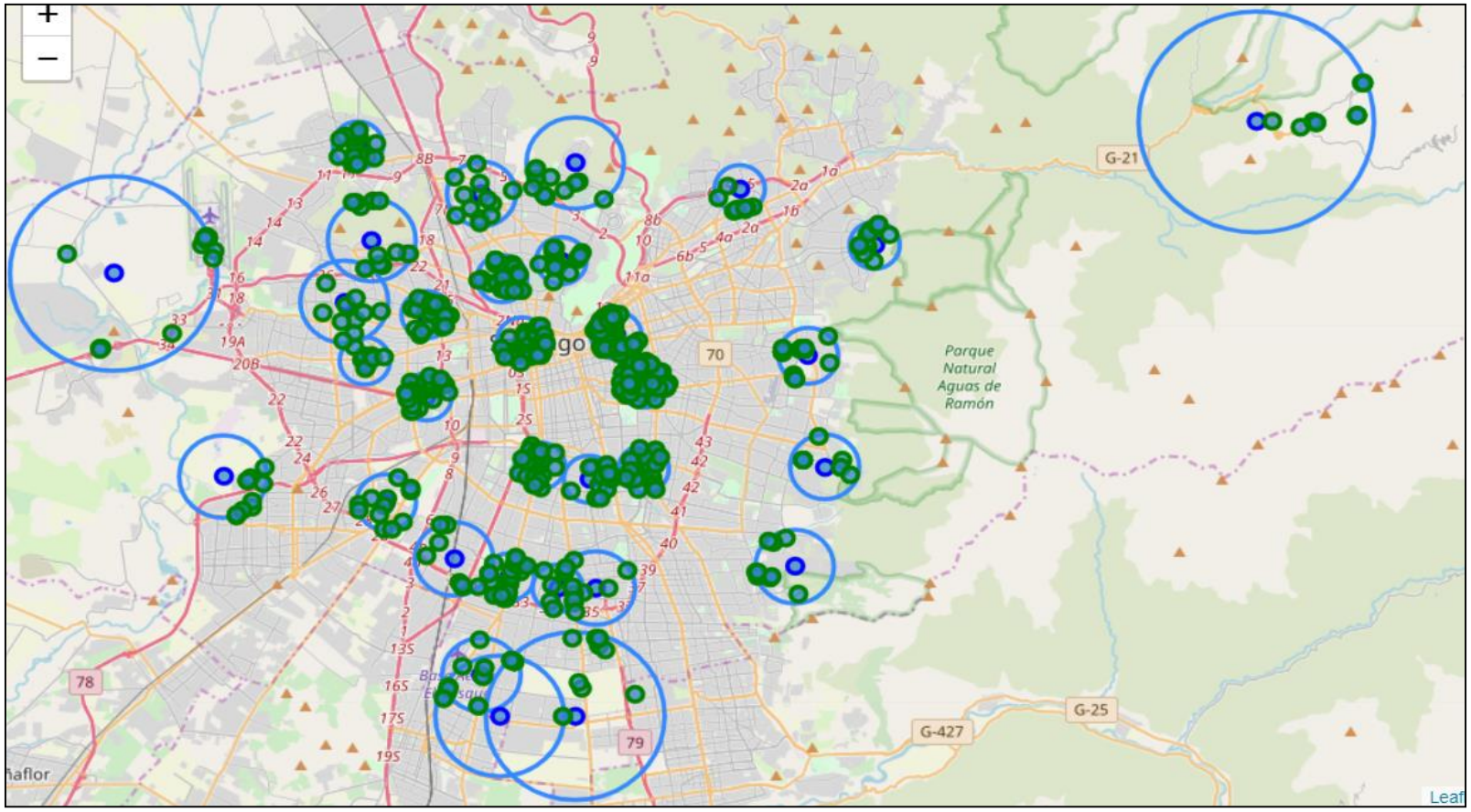
- 1 km radius
- Increses until reaches 10 venues per commune.

Test Sample: Vitacura

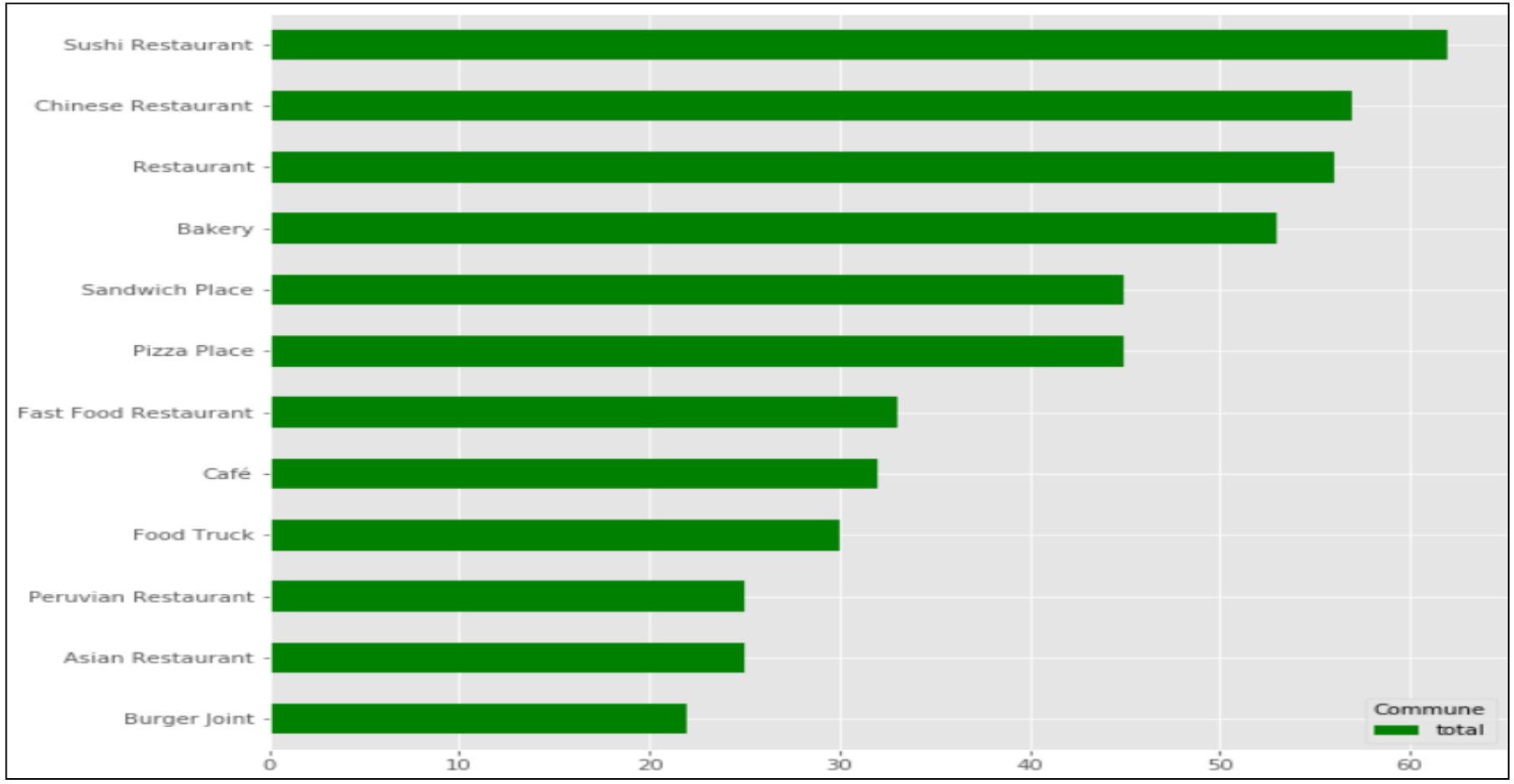




Complete Sample Overview with final radiuses:



Complete Sample Overview

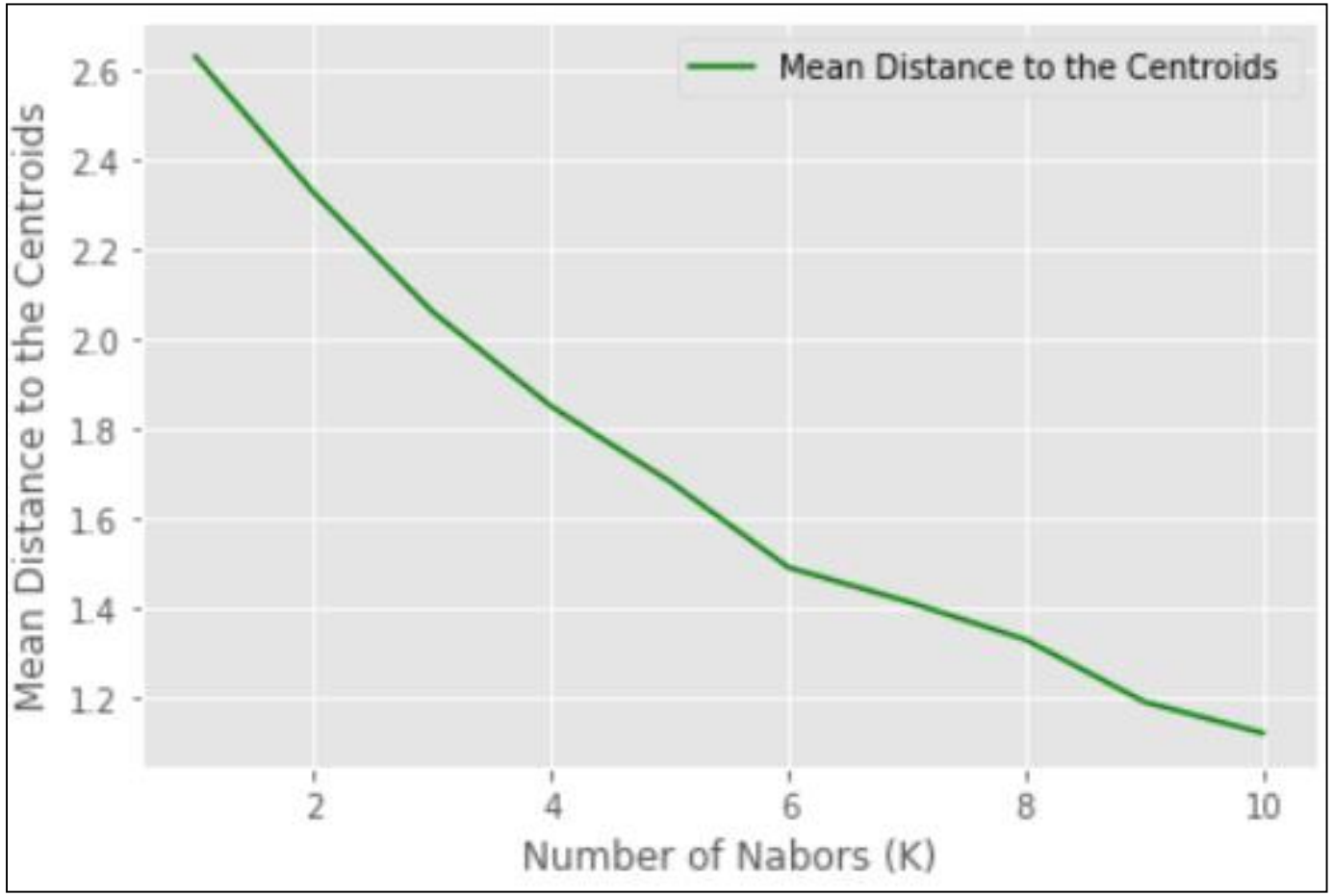


Clusterization using K-Means

- Use the top 3 most popular venues categories to cluster the communes in Santiago.

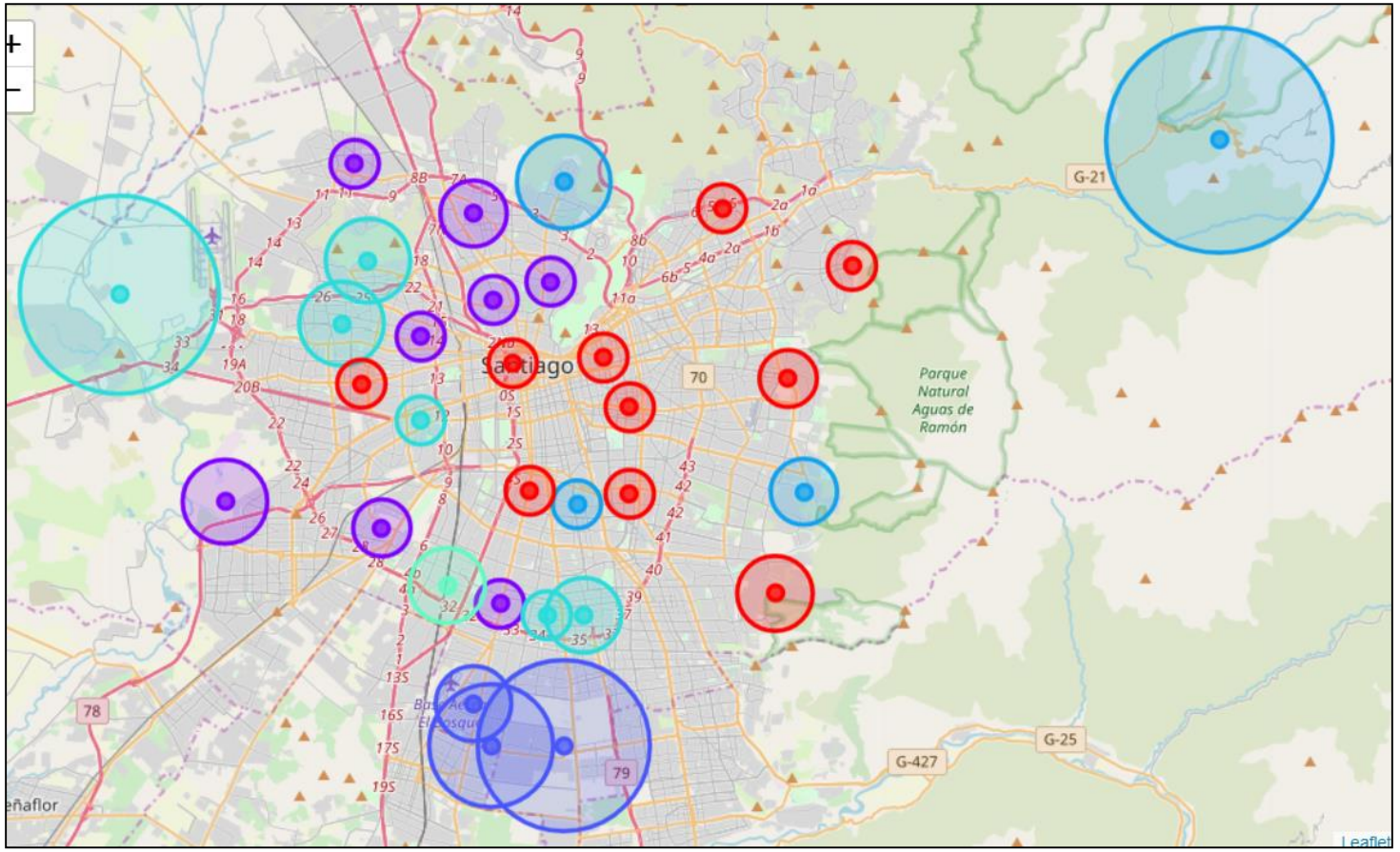
- Use elbow method to evaluate the optimal number for k.

K-Means for clusteirng



- Elbow method shows the optimal k is 6.

Clusterization Results



Cluster	Name	Color
1	Sushi and Sandwiches	
2	Chinese and Fast Food Restaurants	
3	Food Truck	
4	Food Truck and Snacks	
5	Bakery and Pizza	
6	Not Defined	

Conclusions

- Good initial approach but several areas uncovered.
- Limitations due to Foursquare’s standard user.
- Areas succesfully clustered based on initial approach.

Suggestions

- Try other geospatial points such as neighborhoods or more defined comunnal shapes.
- Investigate other types of venues.
- Investigate the influence of category in the survival of a business after placement.