## **Buenos Aires city Subway Station Analysis using Foursquare.com**

## **Study for determining a new Store Shop location**

Polarizaria

Services

Ser

Coursera Capstone Project Pablo Pedrazas



- •Subway network concentrates a high number of persons in the nearby
- Store owners try to locate their business close to the Subway
   Stations to capitalize this movement.
- •Foursquere.com let us explore a greographical area returning places of interest



•Analyzing Foursquere data with other data sets available from of traffic, population, etc. from City government will determine best locations for a new Coffe Shop





## From the Government of the city of Buenos Aires, official information:

- Neighborhood information, with geographical location
- Population for each Neighborhood from same source, but based on 2010 official data
- Subway stations information, including geographical location and line
- Premetro stations information (additional stations complimenting the subway network)
- Traffic information of the subway stations

### From Foursquare.com

- Exploration of all Venues near subways stations
- Exploration of specific categories of Coffee Shops near subway stations



#### **Buenos Aires Subway Stations over Neighborhood population representation**

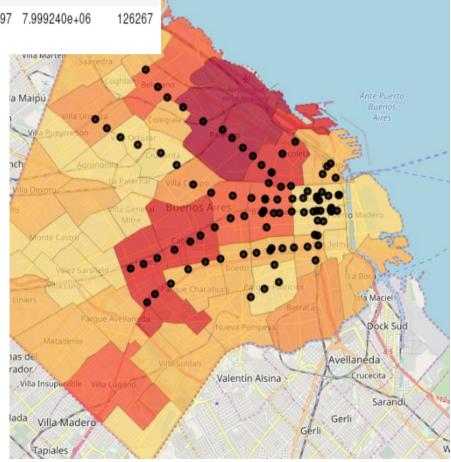
	WKT	Borough	comuna	perimetro	area	Population
0	POLYGON ((-58.4771156675186 -34.5951149914833,	AGRONOMIA	15	6556.167772	2.122169e+06	13912
1	POLYGON ((-58.4128700313089 -34.6141162515854,	ALMAGRO	5	8537.901368	4.050752e+06	131699
2	POLYGON ((-58.4119188098038 -34.5980030767748,	BALVANERA	3	8375.821811	4.342280e+06	138926
3	POLYGON ((-58.3703353711449 -34.6329258371189,	BARRACAS	4	12789.791771	7.953453e+06	89452
4	POLYGON ((-58.4505669109009 -34.5356104340406,	BELGRANO	13	20609.775397	7.999240e+06	126267

→ Data adquisition, cleaning, and ellaboration

→ Visual representation



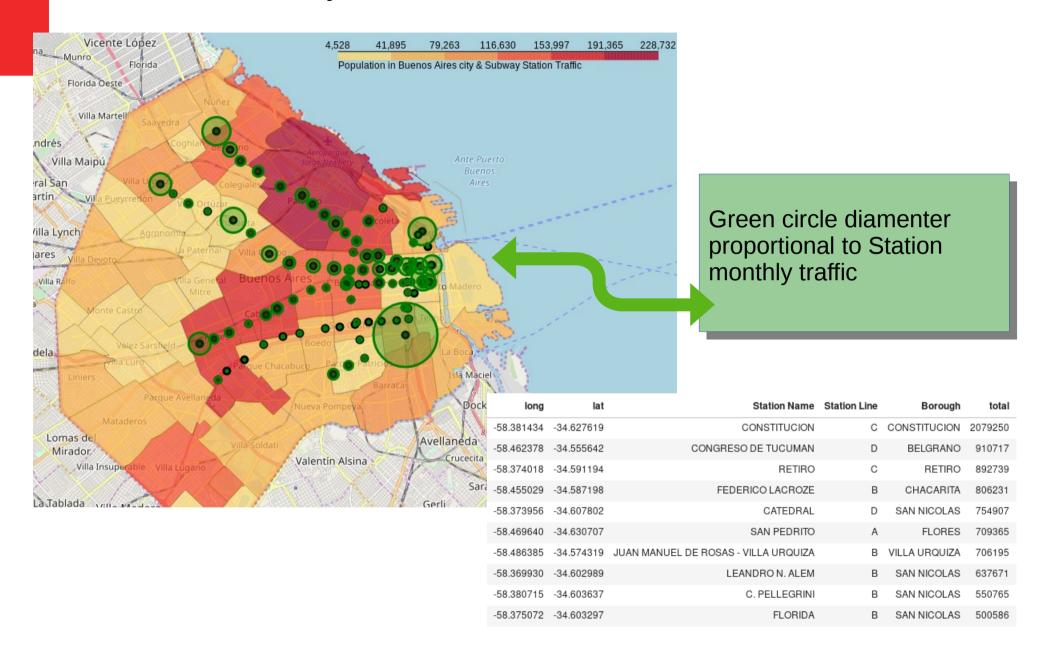
	long	lat	Station Name	Station Line	total	totalN
0	-58.380574	-34.604245	9 DE JULIO	D	229045	0.110158
1	-58.436429	-34.618280	ACOYTE	Α	361379	0.173803
2	-58.407161	-34.591628	AGÜERO	D	280248	0.134783
3	-58.401208	-34.609834	ALBERTI	Α	119855	0.057643
4	-58.420962	-34.603165	ALMAGRO - MEDRANO	В	426167	0.204962



116,630

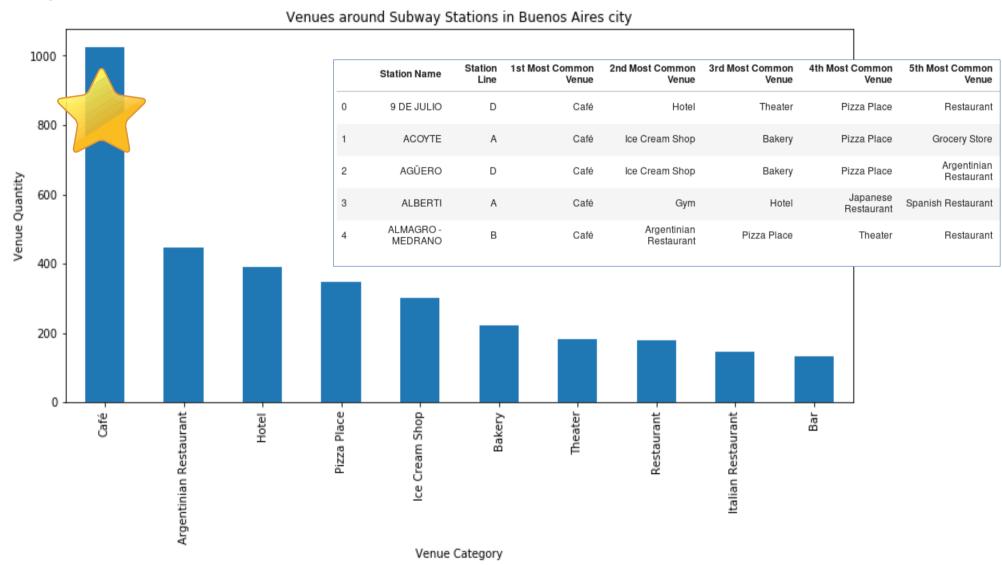
Population in Buenos Aires city

#### **Buenos Aires Subway Stations with Traffic data**



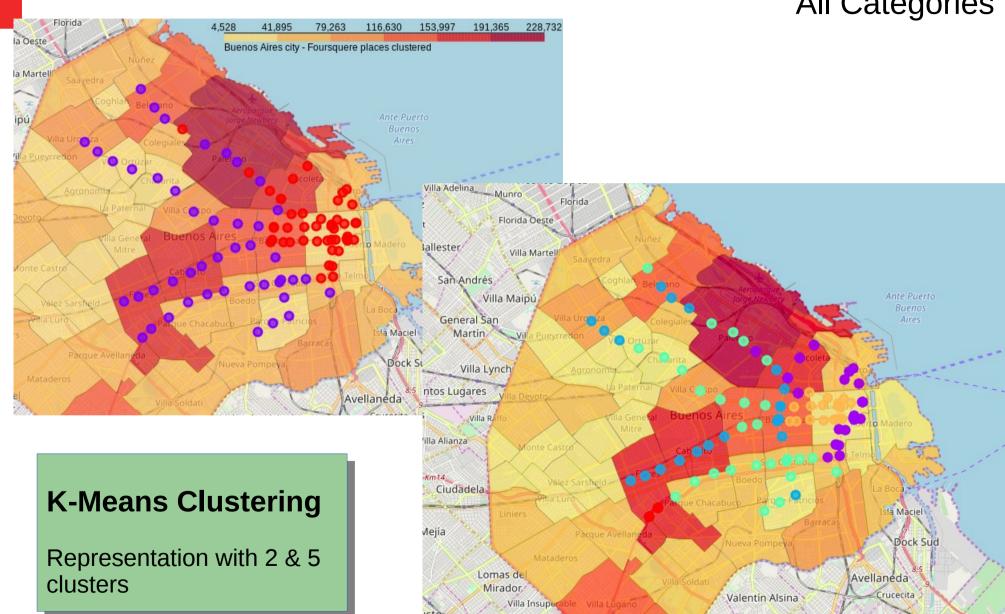
### Foursquere top categories from nearby Subway Stations





## **Clustering Foursquere findings**

All Categories



## **New exploration for only Coffee Shops**

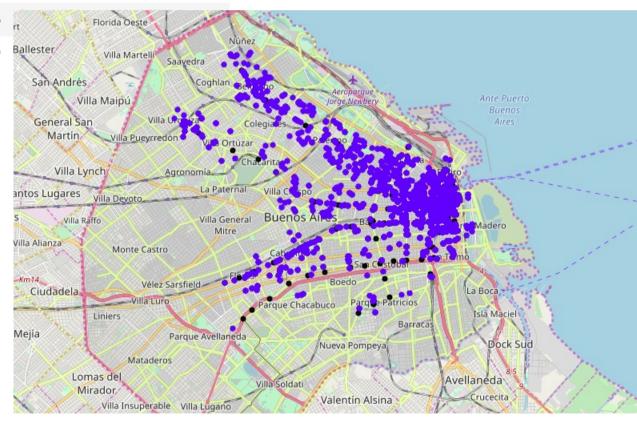


100

totalN Cafes



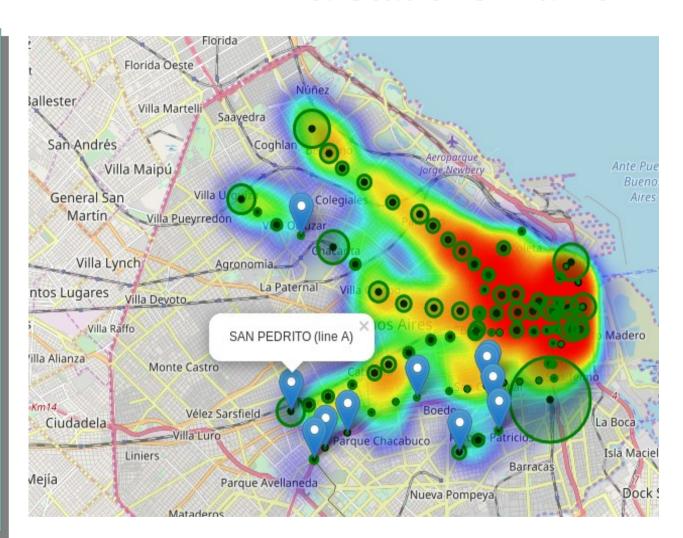
totalN	total	Borough	Station Line	Station Name	lat	long
0.110158	229045	SAN NICOLAS	D	9 DE JULIO	-34.604245	-58.380574
0.173803	361379	CABALLITO	А	ACOYTE	-34.618280	-58.436429
0.134783	280248	RECOLETA	D	AGÜERO	-34.591628	-58.407161
este	Florida O	A rt Florida		ALBERTI	-34.609834	-58.401208
artelli	Villa Ma	Ballester	В	ALMAGRO - MEDRANO	-34.603165	-58.420962



## Heatmap of Coffee Shops & Stations Traffic

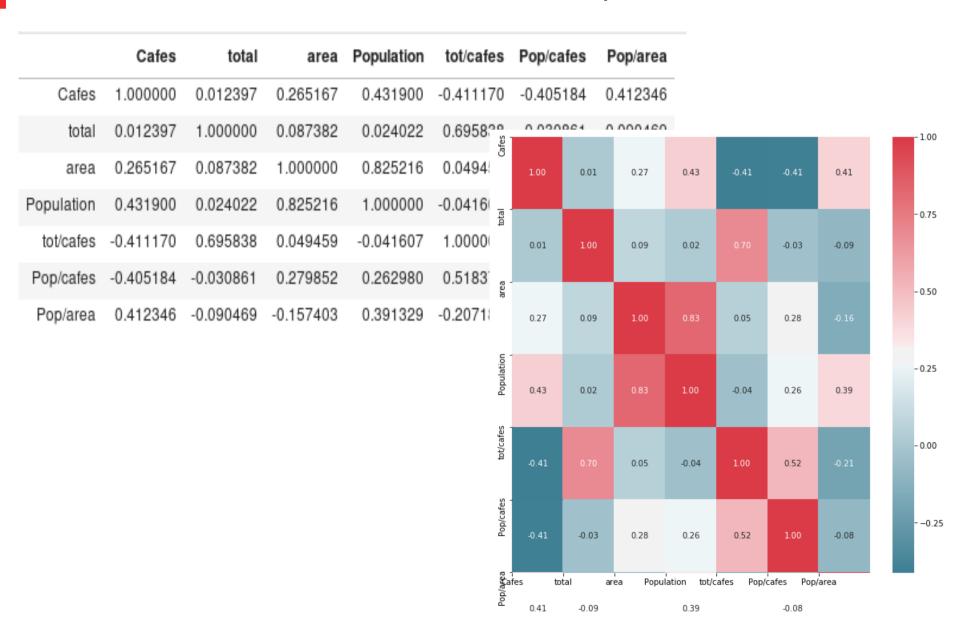
### **Heatmap**

- Green circles indicates traffic proportionally with ratio
- Blue markers indicates stations with less "Cafes" or coffee shops, targets for final decission of new shop location



### **Correlation matrix**

Cafes vs Population, Area, Traffic, etc.



# Final Decission After applying feature population as predictor



### **Conclusion and Observations**

- Serveral data sets avaiable were studied applying data science techniques
- Subway network data, together with neighborhoods, area and population around stations were studied.
- Criteria was based on population as predictor over Stations with less stores found
- Graphical representations assisted in the decision to made recommending to explore the following locations:
  - 1. San Pedrito (Line A)
  - 2. Varela (Line E)
  - 3. Plaza de los Virreyes Eva Peron (Line E)
- Considering the distribution of the subway lines is not evenly spread about all neighborhoods, and that some areas of them might not be suitable for the use of the subway, the analytic approach might not be exact
- There are many other variables not covered in this study that should follow this initial recommendation to compliment and therefore, confirm or reject the result provided as: Store rent value, properties availability, state, investment, etc.