




# Predicting São Paulo's financial districts location using coffee shops concentration through clustering techniques

IBM Professional Certificate Capstone  
São Paulo, April of 2019

Rafael Yonezawa de Mello  
[LinkedIn profile](#)



A decorative network diagram in the top-left corner, featuring a complex web of interconnected nodes and lines. The nodes are represented by circles of varying sizes, some with concentric rings, and the lines are thin and grey. The diagram is partially cut off by the top and left edges of the slide.

1.

# Introduction

Business fundamentals



## Business fundamentals

- ◎ Culturally, brazilians usually tend not to have coffee on-the-go
- ◎ Coffee shops are usualy associated with business meetings and once in a while purchases
- ◎ Coffee Shops tend to be expensive



“

*Can we predict the location of the biggest business/financial districts of São Paulo just staring at coffee shops?*

A decorative network diagram in the top-left corner, featuring a complex web of interconnected nodes and lines. The nodes are represented by circles of varying sizes, some with concentric rings, and the lines are thin and grey. The diagram is partially cut off by the left edge of the slide.

# 2.

## Data sources

Gathering the data we need

# Foursquare's API

Using Foursquare's API we are able to gather all Coffee Shops data necessary



## Ok, but first, Coffee.

Looking out for “Coffee”, “Café” (Coffe in portuguese) and “Cafeteria” (Coffee Shop) in Foursquare’s API, I’ve found all venues I need

```
Café
Coffee
Cafeteria
We've just made 672 queries through Foursquare's API
```

```
In [5]: Coffee_Shops.head()
```

Out[5]:

	categories	hasPerk	id	location.cc	location.city	location.country	location.distance	location.formattedAddress	loc
0	[[{'id': '4bf58dd8d48988d116941735', 'name': 'B...'}]]	False	52be1cd1498efd3ee4c8c521	BR	NaN	Brasil	1495	[Brasil]	
1	[[{'id': '4bf58dd8d48988d1e0931735', 'name': 'C...'}]]	False	4fccf703e4b09949e4fc7ce4	BR	NaN	Brasil	1298	[Brasil]	
2	[[{'id': '4f4532974b9074f6e4fb0104', 'name': 'D...'}]]	False	51f57da8498eeefdf8be716	BR	Guarulhos	Brasil	1307	[Guarulhos, SP, Brasil]	
0	[[{'id': '4bf58dd8d48988d1e0931735', 'name': 'C...'}]]	False	5972817d364d972852d1236b	BR	Guarulhos	Brasil	988	[Avenida Mulungú, 262, Guarulhos, SP, 07151-38...]	
1	[[{'id': '4bf58dd8d48988d1e0931735', 'name': 'C...'}]]	False	4fccf703e4b09949e4fc7ce4	BR	NaN	Brasil	1420	[Brasil]	



3.

# Methodology

Unsupervised learning and Coffee Shops



## Unsupervised learning

- ◎ In order to identify agglommerations of Coffee Shops, I used clustering models, which means we're going to solve this question with unsupervised learning
- ◎ The only features I used to cluster venues are their latitude and longitude

```
In [9]: Cluster_Dataframe.head()
```

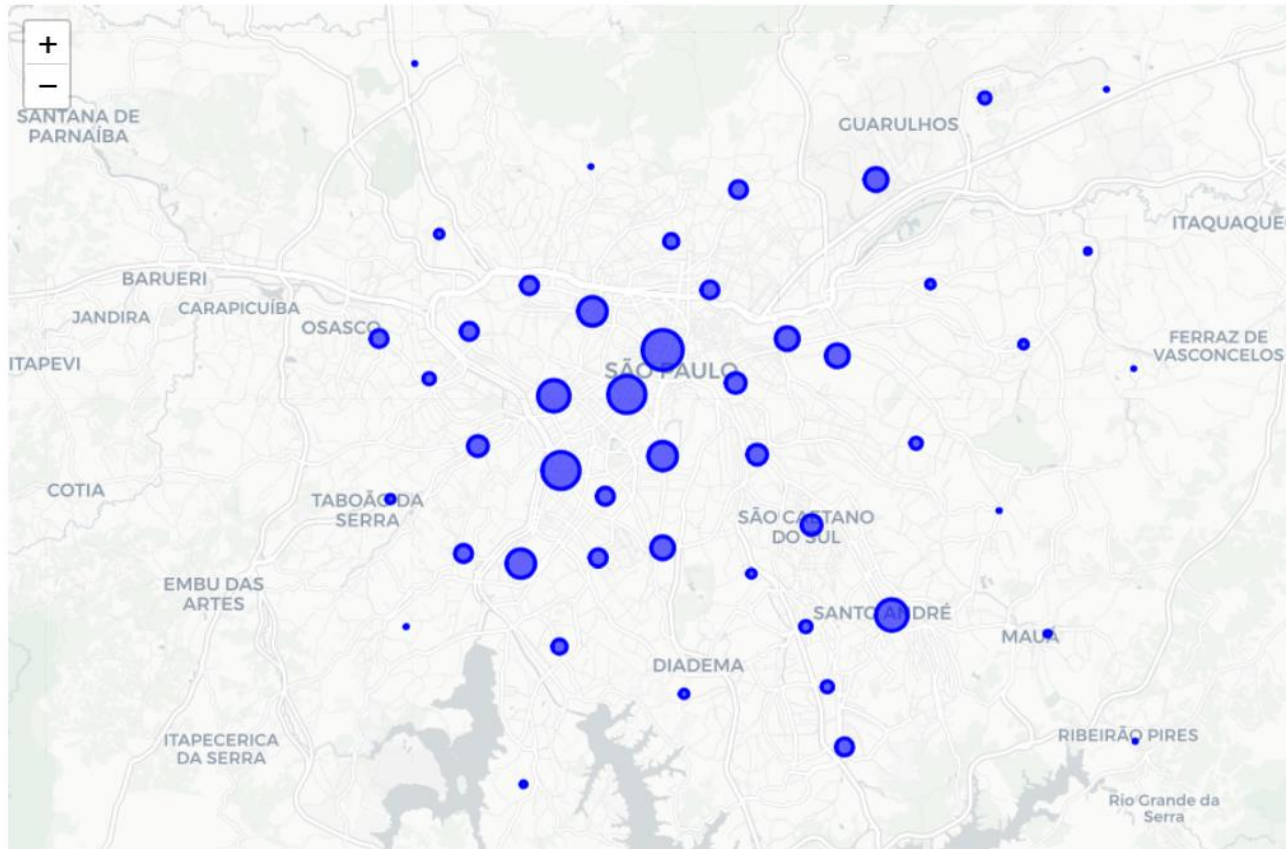
```
Out[9]:
```

	Venues	Latitude_center	Longitude_center
Cluster			
0	32	-23.692481	-46.626891
1	104	-23.547329	-46.600895
2	142	-23.596515	-46.686092
3	110	-23.658285	-46.527461
4	71	-23.583679	-46.727140

# Coffee Shops Clusters

In [19]: clusters\_map

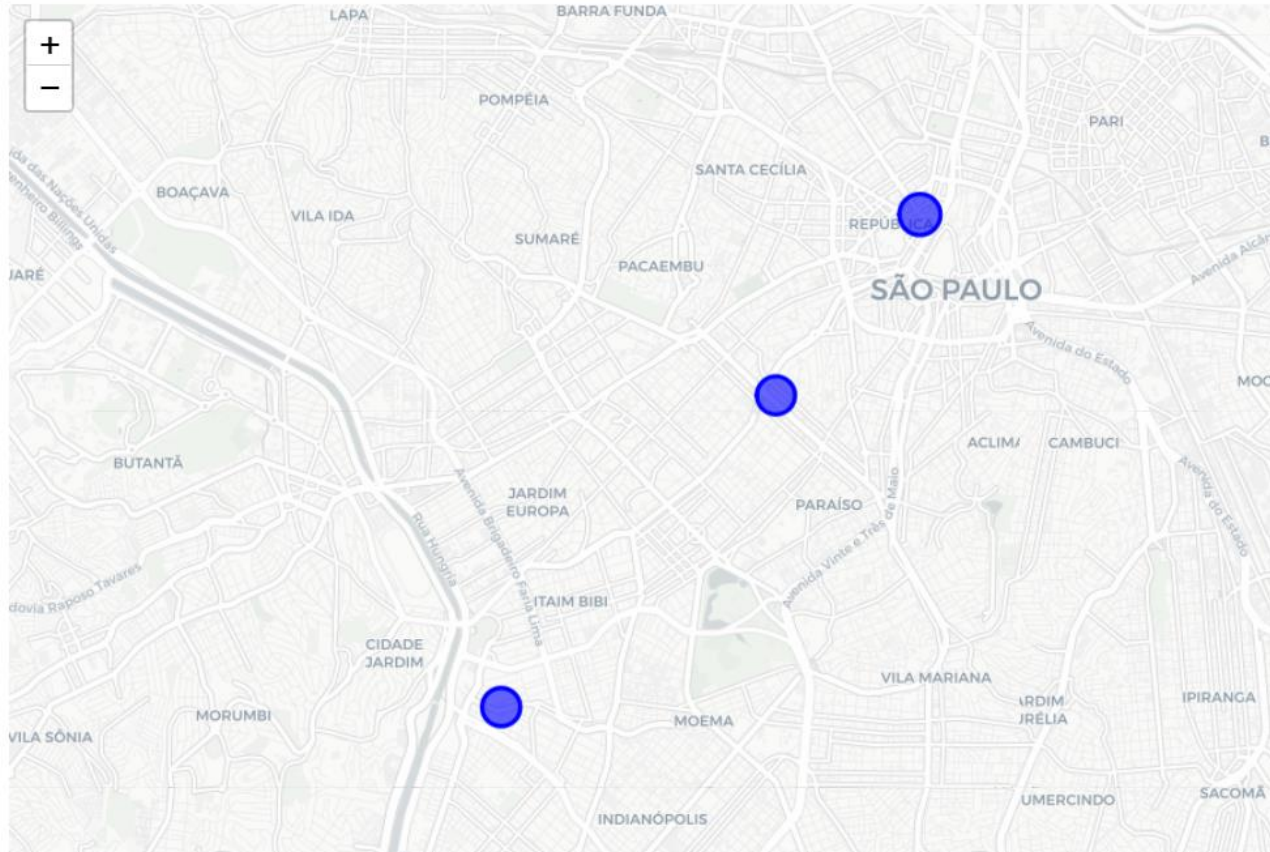
Out[19]:



## Top 3 Clusters in number of venues

In [22]: top\_clusters\_map

Out[22]:





4.


# Results

Are the top 3 clusters near financial districts?



## São Paulo's main financial districts

According to Wikipedia, there are 7 financial districts in São Paulo. The biggest 3 are:

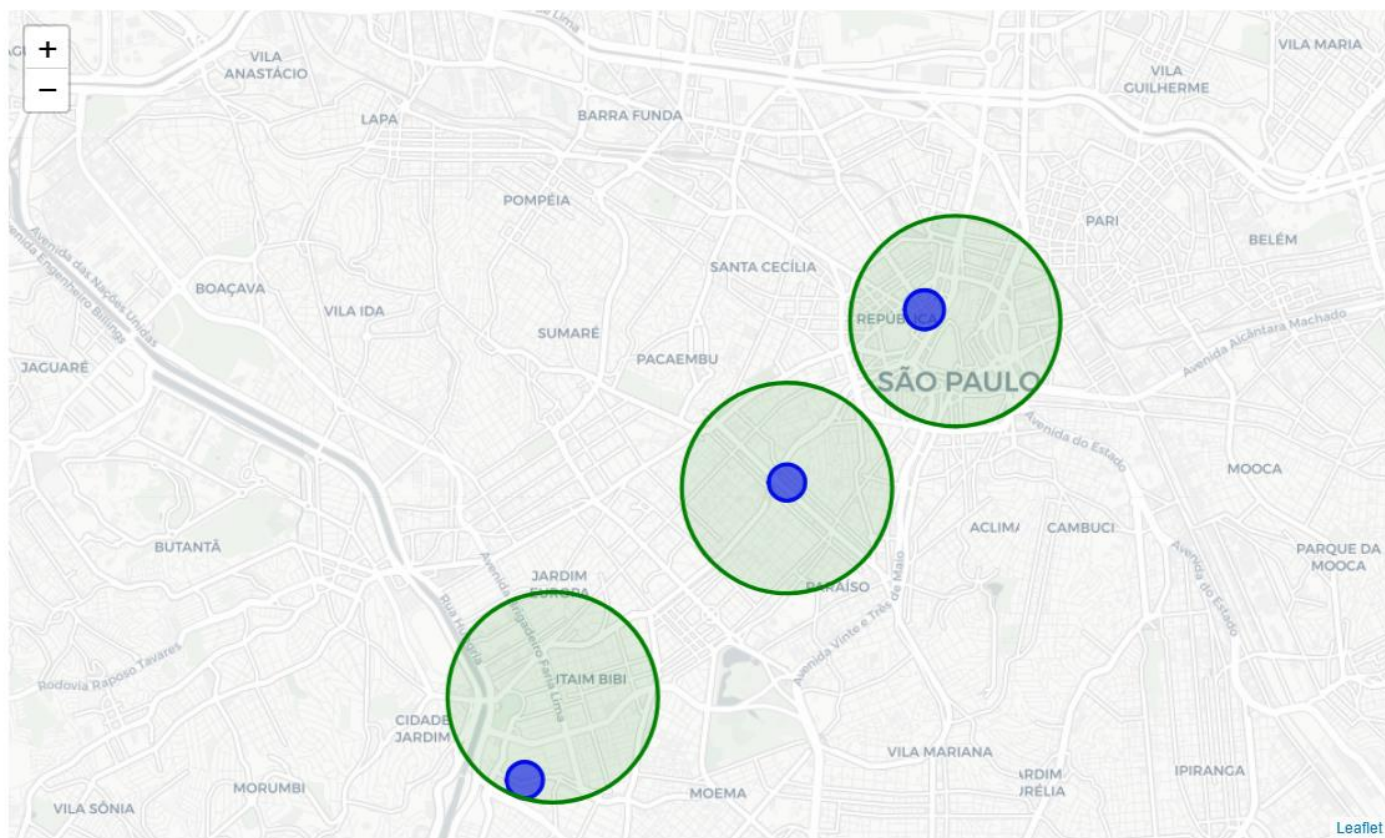
- ◎ **Avenida Paulista:** mainly known as one of the biggest tourist attractions in the city for its cultural value, it holds head offices of some banks and multinational companies. Located at [-23.5629, -46.6544].
  - ◎ **Historical Midtown:** although this part of the city seems like it's been forgotten over time, it's still the district that holds the head office of the Brazilian's stock exchange (B3), several companies and hotels. It's located at [-23.5442, -46.6339].
  - ◎ **Avenida Brigadeiro Faria Lima:** modern and luxurious, Faria Lima is by far the district with the highest economic development over the last years in São Paulo. It's the favorite spot for companies and hotels that are looking for new opportunities in the city. Located at [-23.5863, -46.6831].
- 

Plotting them with our clusters...



# Clusters and Financial Districts

Out[23]:



A decorative network diagram in the top-left corner, featuring a complex web of interconnected nodes and lines. The nodes are represented by circles of varying sizes, some with concentric rings, and the lines are thin and grey. The diagram is partially cut off by the left edge of the slide.

# 5. **Discussions and Conclusions**



## Satisfactory results

As expected, the 3 biggest agglomerations of coffee shops in São Paulo are near the biggest financial districts of the city. That support our thesis that those are good indicators of economic growth as their appearance are often associated with business development and not just populational density.

In conclusion, we can see that the results are satisfactory and the answer to our main question is yes.





# Thanks!

## Any questions?

You can find me at:

@rafael.yonezawa

[rafael.yonezawa@gmail.com](mailto:rafael.yonezawa@gmail.com)

[www.linkedin.com/rafaelyonezawa](https://www.linkedin.com/rafaelyonezawa)