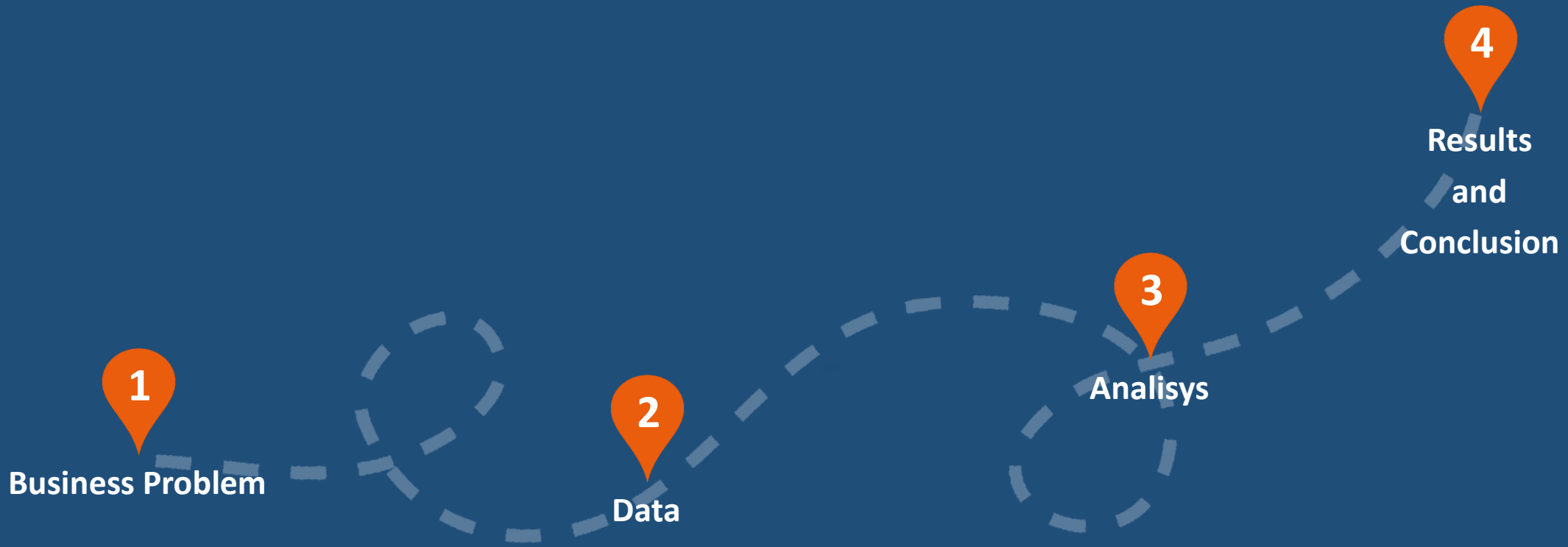
The optimal location for great coffee in Lisbon

The optimal location for great coffee in Lisbon



The optimal location for great coffee in Lisbon

CONTENTS

BUSINESS PROBLEM

The goal for this analysis is to find a good location for a ‘Café’

- **Why a ‘Café’?**

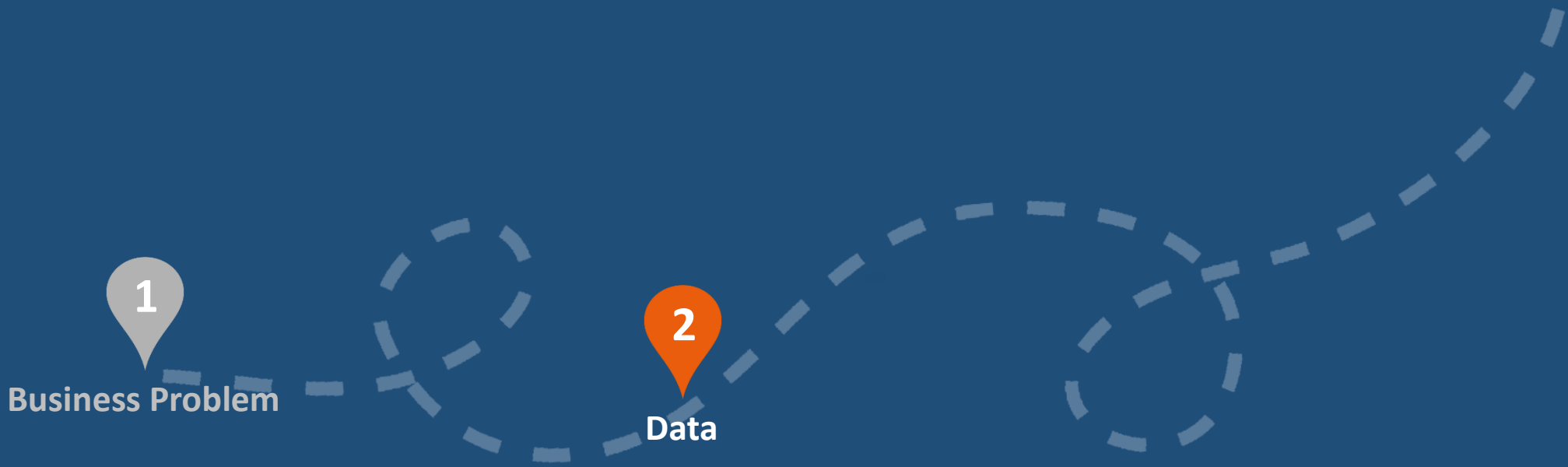
Because coffee plays a very big role in portuguese life and everyone drinks it on a daily basis;

- **Why Lisbon?**

Because it is Portugal’s capital and a city that as seen a big growth in food related businesses in the last 5 years so it could be a good investment;

The optimal location for great coffee in Lisbon

CONTENTS



DATA

Data collected for the analysis

- **Location Data**

Collected for portuguese government agencies in geogrpahic format;

- **Population Data**

Collected from Lisbons municiplality;

- **Venues Data**

Collected from the *Foursquare API*;

DATA

Data Frames used

- **Location Data**

	Neighbourhood	Latitude	Longitude
0	Ajuda	38.712174	-9.198653
1	Alcantara	38.709594	-9.183390
2	Alvalade	38.753884	-9.146498
3	Areeiro	38.741376	-9.133527
4	Arroios	38.727653	-9.137942

- **Population Data**

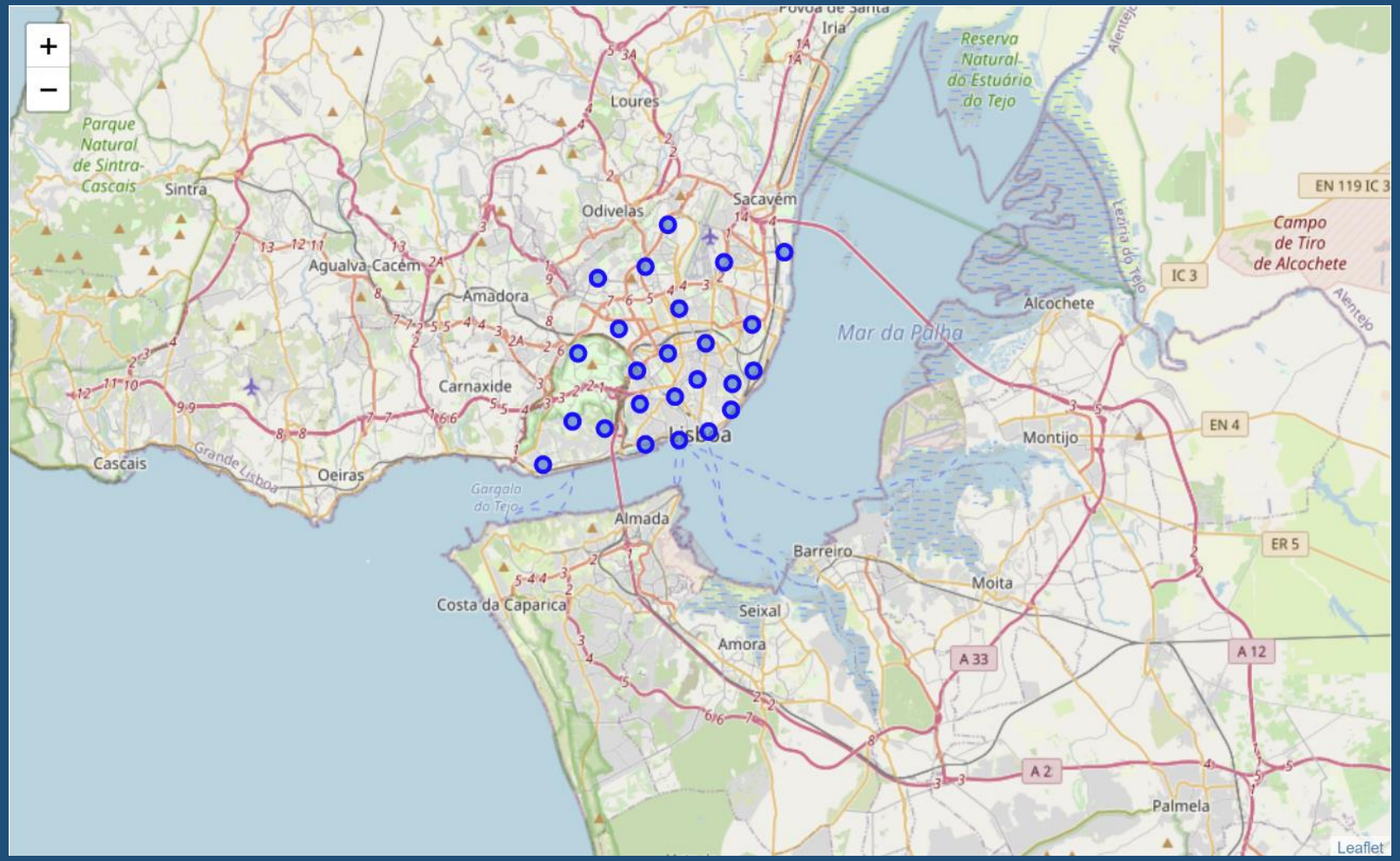
	Neighbourhood	Population	Area	Pop_Dens(pop/km2)
0	Ajuda	17958	2.85	6301.052632
1	Alcantara	14443	4.42	3267.647059
2	Alvalade	34217	5.33	6419.699812
3	Areeiro	21035	1.76	11951.704545
4	Arroios	33210	2.15	15446.511628

The optimal location for great coffee in Lisbon

CONTENTS

ANALYSIS

Map the Data Frames



ANALYSIS

Using *Foursquare API:*

- **List the venues** in each neighbourhood;
- Check the **number of venues** in each neighbourhood;

Venue Category	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude
Portuguese Restaurant	206	206	206	206	206	206
Café	110	110	110	110	110	110
Restaurant	108	108	108	108	108	108
Hotel	87	87	87	87	87	87
Bakery	65	65	65	65	65	65
Coffee Shop	61	61	61	61	61	61
Plaza	48	48	48	48	48	48
Italian Restaurant	40	40	40	40	40	40
Park	40	40	40	40	40	40
Seafood Restaurant	36	36	36	36	36	36

ANALYSIS

Using *Foursquare API:*

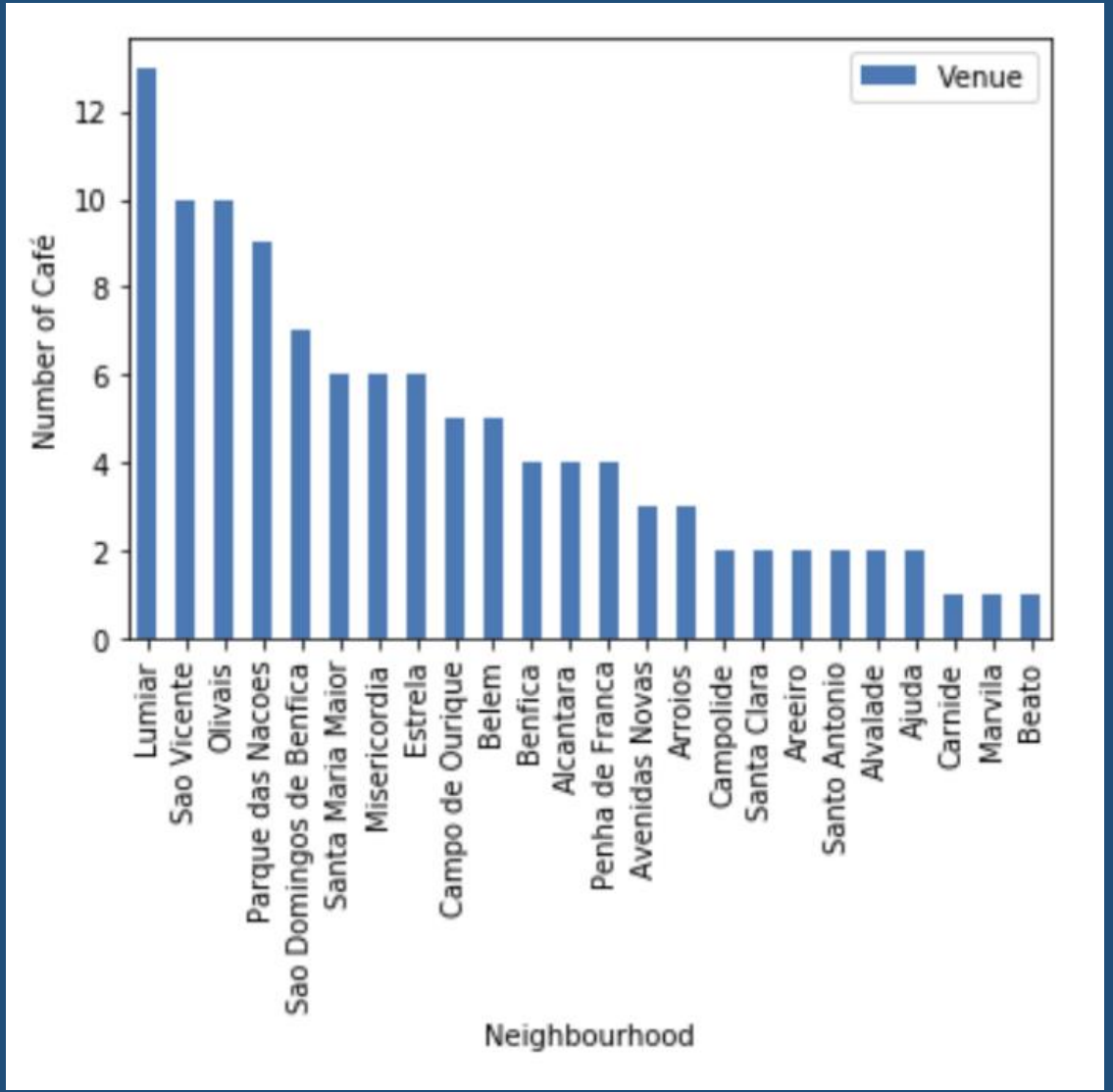
- Find the **Top 10 categories** in each neighbourhood;

	Neighbourhood	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	Ajuda	Restaurant	Portuguese Restaurant	Park	Coffee Shop	Soccer Stadium	BBQ Joint	Gym	Café	Supermarket	Soccer Field
1	Alcantara	Restaurant	Portuguese Restaurant	Café	Mediterranean Restaurant	Seafood Restaurant	BBQ Joint	Coffee Shop	Plaza	Museum	Bakery
2	Alvalade	Portuguese Restaurant	Restaurant	Bar	Coffee Shop	Bakery	Burger Joint	Indian Restaurant	Italian Restaurant	Ice Cream Shop	Bookstore
3	Areeiro	Portuguese Restaurant	Hotel	Italian Restaurant	Bakery	Gym / Fitness Center	Vegetarian / Vegan Restaurant	Restaurant	Sushi Restaurant	Ice Cream Shop	Indian Restaurant
4	Arroios	Hotel	Portuguese Restaurant	Plaza	Seafood Restaurant	Cocktail Bar	Breakfast Spot	Scenic Lookout	Café	Restaurant	Park
5	Avenidas Novas	Portuguese Restaurant	Hotel	Italian Restaurant	Restaurant	Vegetarian / Vegan Restaurant	Bakery	Burger Joint	Japanese Restaurant	Steakhouse	Coffee Shop
6	Beato	Restaurant	Brewery	Theater	Coffee Shop	Plaza	Paper / Office Supplies Store	Supermarket	BBQ Joint	Climbing Gym	Music Venue
7	Belem	Portuguese Restaurant	Café	Monument / Landmark	Ice Cream Shop	Garden	Sushi Restaurant	Bakery	Restaurant	Art Museum	Hotel
8	Benfica	Portuguese Restaurant	Restaurant	Café	Bakery	Seafood Restaurant	Park	Pizza Place	BBQ Joint	Burger Joint	Grocery Store
9	Campo de Ourique	Portuguese Restaurant	Hotel	Coffee Shop	Italian Restaurant	Café	Bakery	Restaurant	Garden	Ice Cream Shop	Indian Restaurant
10	Campolide	Hotel	Portuguese Restaurant	Restaurant	Park	Gym	Italian Restaurant	Seafood Restaurant	Garden	Japanese Restaurant	Art Museum

ANALYSIS

Using *Foursquare API:*

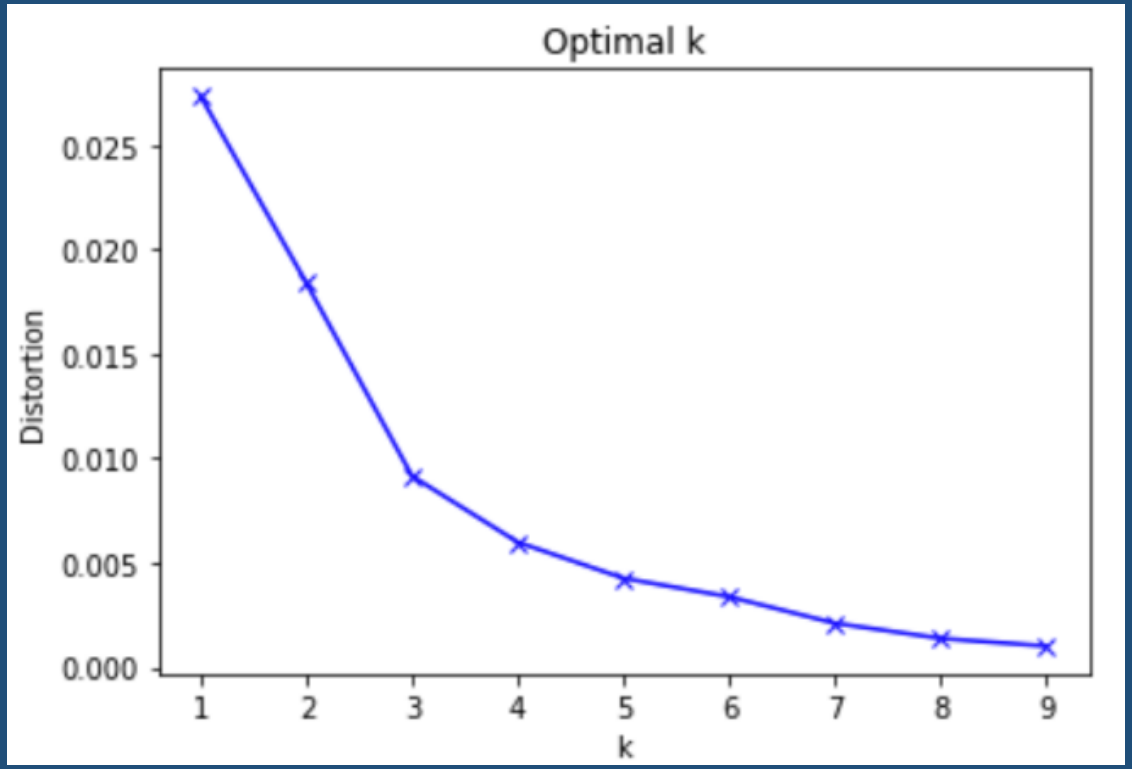
- Check the **number of 'Cafés'** in each neighbourhood;



ANALYSIS

Using the ***K-Means*** algorithm to cluster the 'Cafés' in groups:

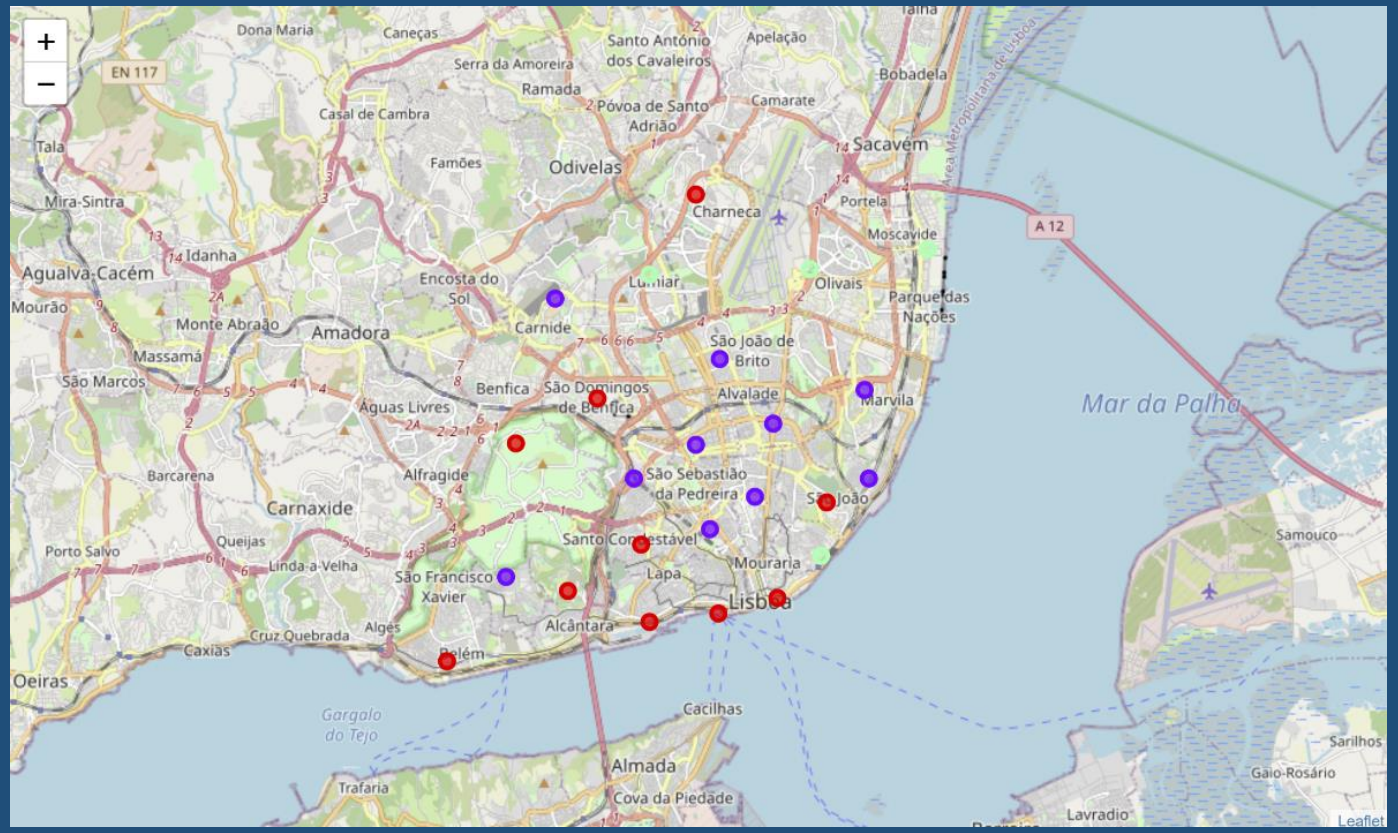
- Find the optimal **k** using the Elbow Method;



ANALYSIS

Using the ***K-Means*** algorithm to cluster the ‘Cafés’ in groups:

- Create the 3 clusters and map the results;



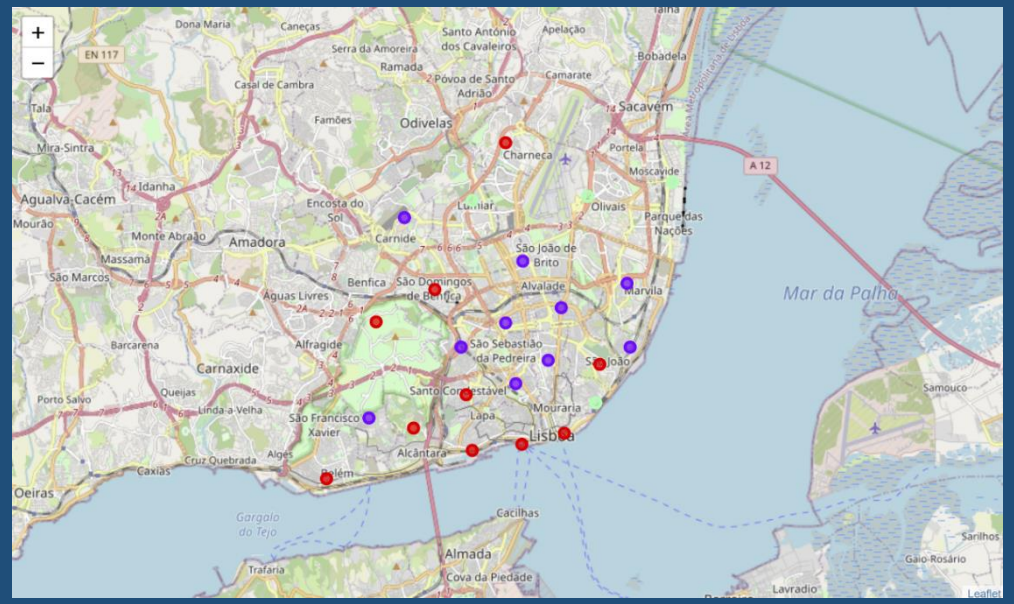
ANALYSIS

The map shows the three clusters that can be described as:

- Cluster 0 (in red),
  neighbourhoods with **few** ‘Cafés’

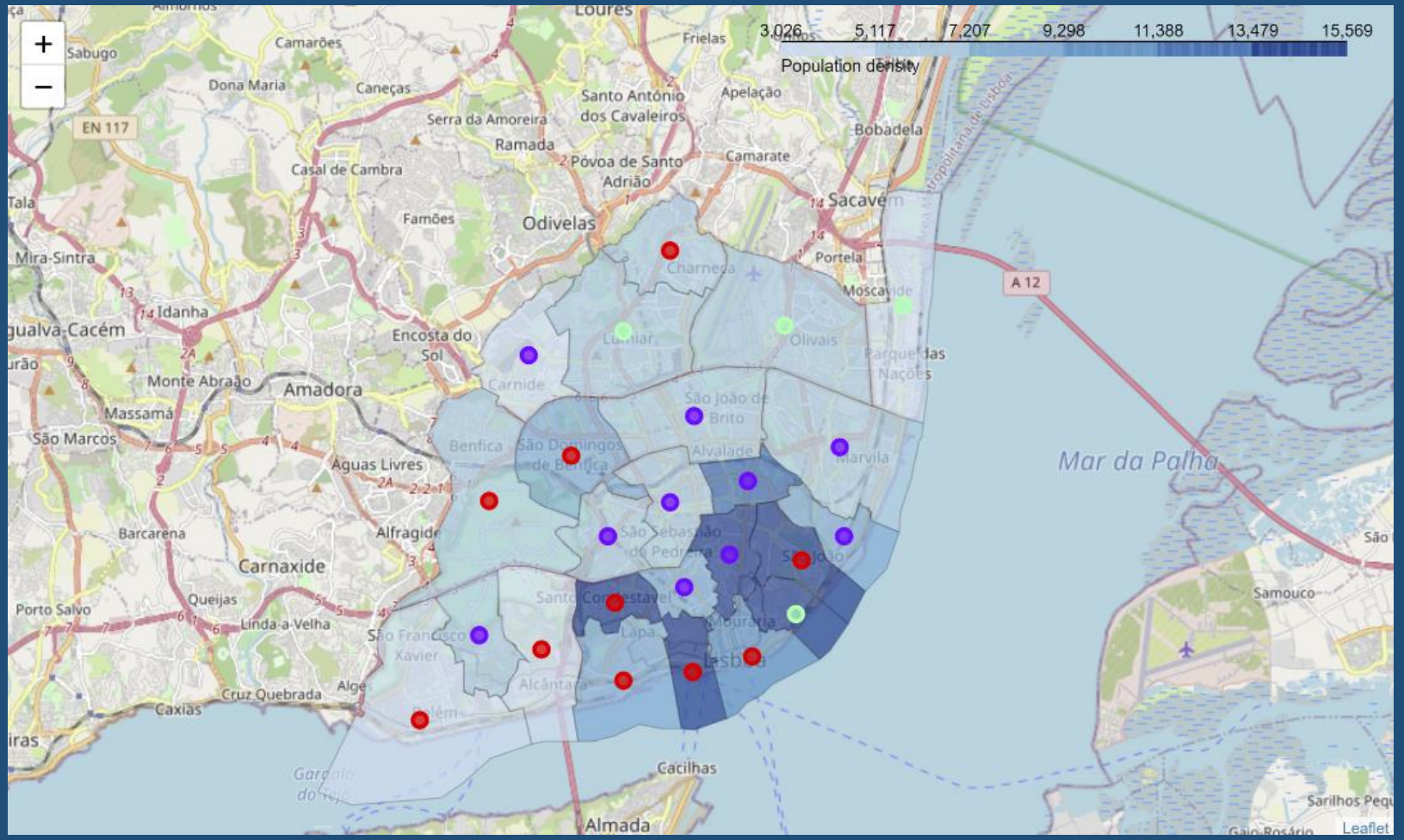
- Cluster 1 (in violet)
  neighbourhoods with an **average** number of ‘Cafés’

- Cluster 2 (in green).
  neighbourhoods with the **highest** number of ‘Cafés’



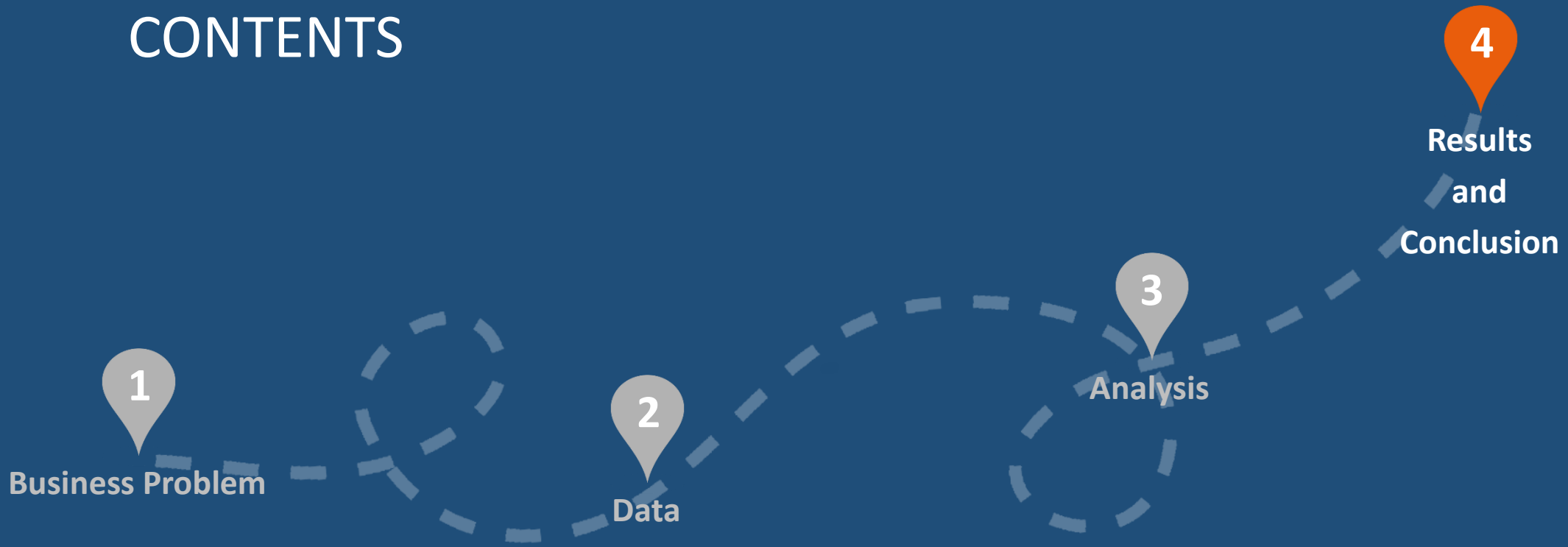
ANALYSIS

Merging the population data frame with the cluster data frame and map the results:



The optimal location for great coffee in Lisbon

CONTENTS



RESULTS

- ‘Cafés’ are among the **top 3 venue**s found in almost every neighbourhood in Lisbon though rarely the in the first place;

- The neighbourhoods with the **highest population** density are in **Lisbon's city centre** (high density = more clients) which could suggest a higher rent price;

- Although more densely populated the city centre does not have the neighbourhoods with the **highest number of ‘Cafés’**;

- There are a large number of neighbourhoods with few ‘Cafés’ (Alvalade, Areeiro, Santo Antonio, Carnide, Marvila) that could be a **good location for a new ‘Café’**.

CONCLUSION

- The results say that **Campo de Ourique** and **Misericórdia** neighbourhoods could be the most likely place to open a ‘Café’;
  Why?
  - Both neighbourhoods are in high densely populated areas and have a low number of ‘Cafés’;

- Another relevant data that could be added to this analysis:
  - data from renting prices for commercial spaces (average rent price/m2)
  - mobility data regarding the existence of public transportation and parking spaces useful to see if the business could attract clients from outside the neighbourhood;

The optimal location for great coffee in Lisbon

Applied Data Science Capstone by IBM/Coursera

Presentation created by Cláudia Morgado

12 September 2020