

# BIGGEST PORTUGUESE COMPANIES DATA ANALYSIS



[www.linkedin.com/in/monique-souza-pt](https://www.linkedin.com/in/monique-souza-pt)



[www.linkedin.com/in/joanavale1996](https://www.linkedin.com/in/joanavale1996)

## Datasets Collection



## Data

Financial

Social Networks

Counties

## Mining

Sabi

Facebook and  
Instagram

Pordata

# Web Scraping Process

Social networks from  
each district

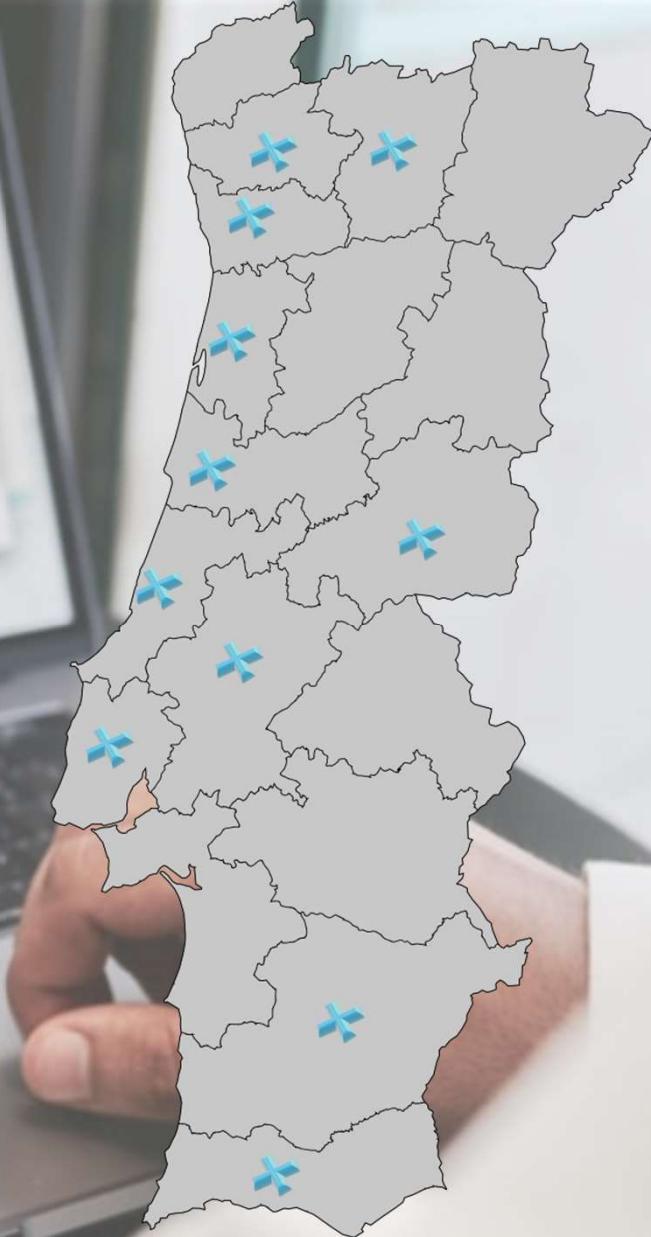
Likes on Facebook



Number of photos posted on  
Facebook

Facebook Followers

#Cityname on Instagram



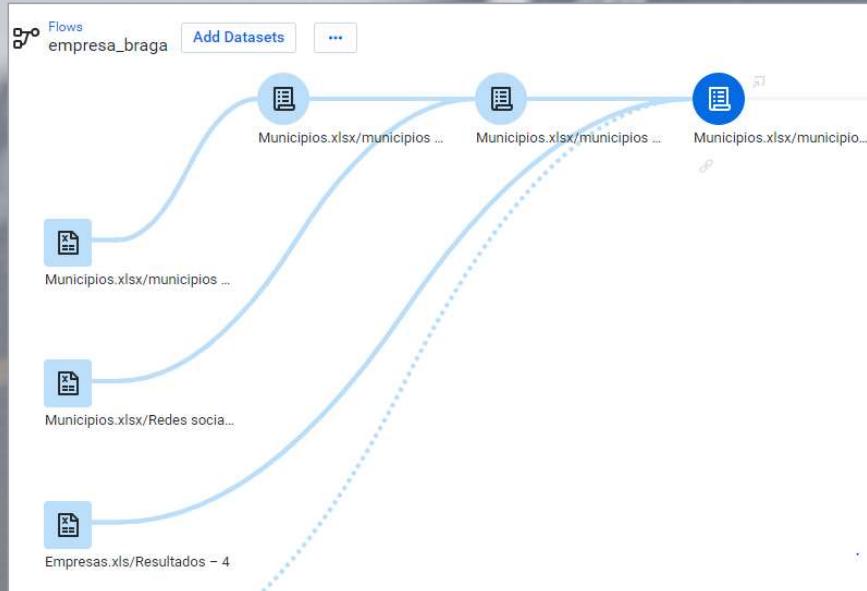
## Cleaning and Processing Data procedure:

Trifecta Software

For data exploration and preparation we made 2 Inner Joins to transform 3 CSV's in one and we also removed:

- ✓ All blanks and nulls
- ✓ All symbols
- ✓ All duplicates





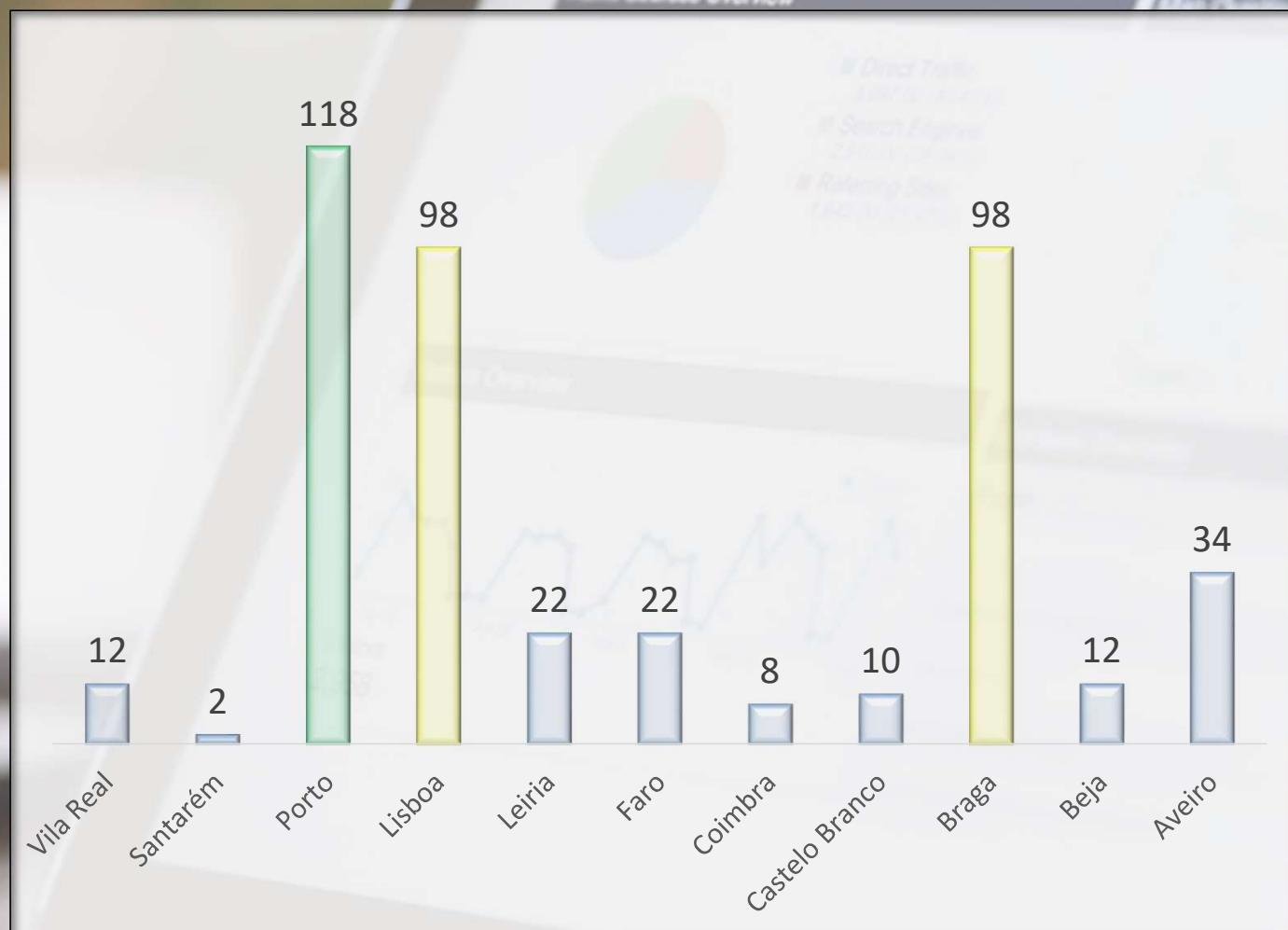
In the end we got only one CSV file

Then we used Python at VS Code editor,  
plus

Pandas      Library

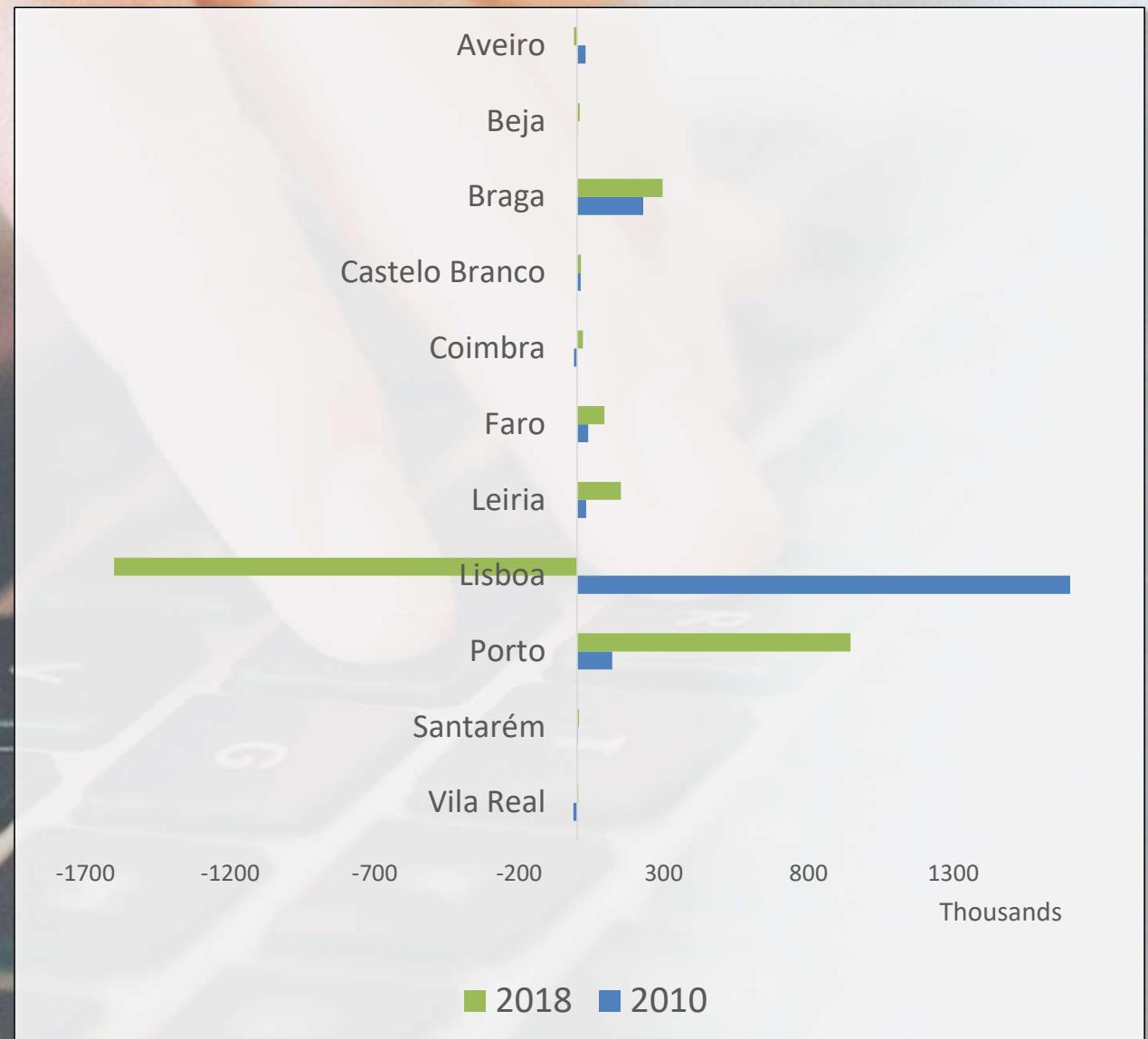
## 1<sup>th</sup> Question

Which district has the highest number of companies?



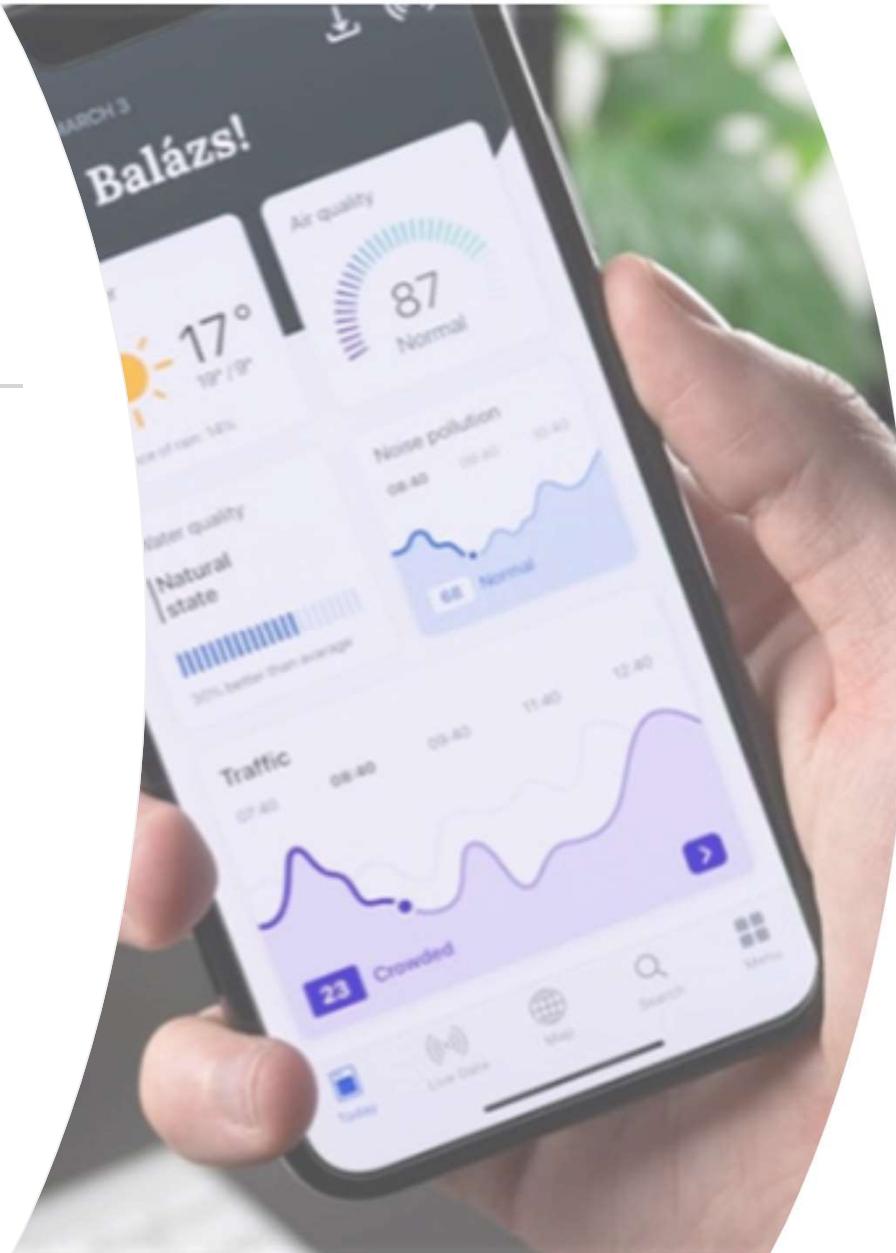
## 2<sup>th</sup> Question

Which district has the highest turnover?



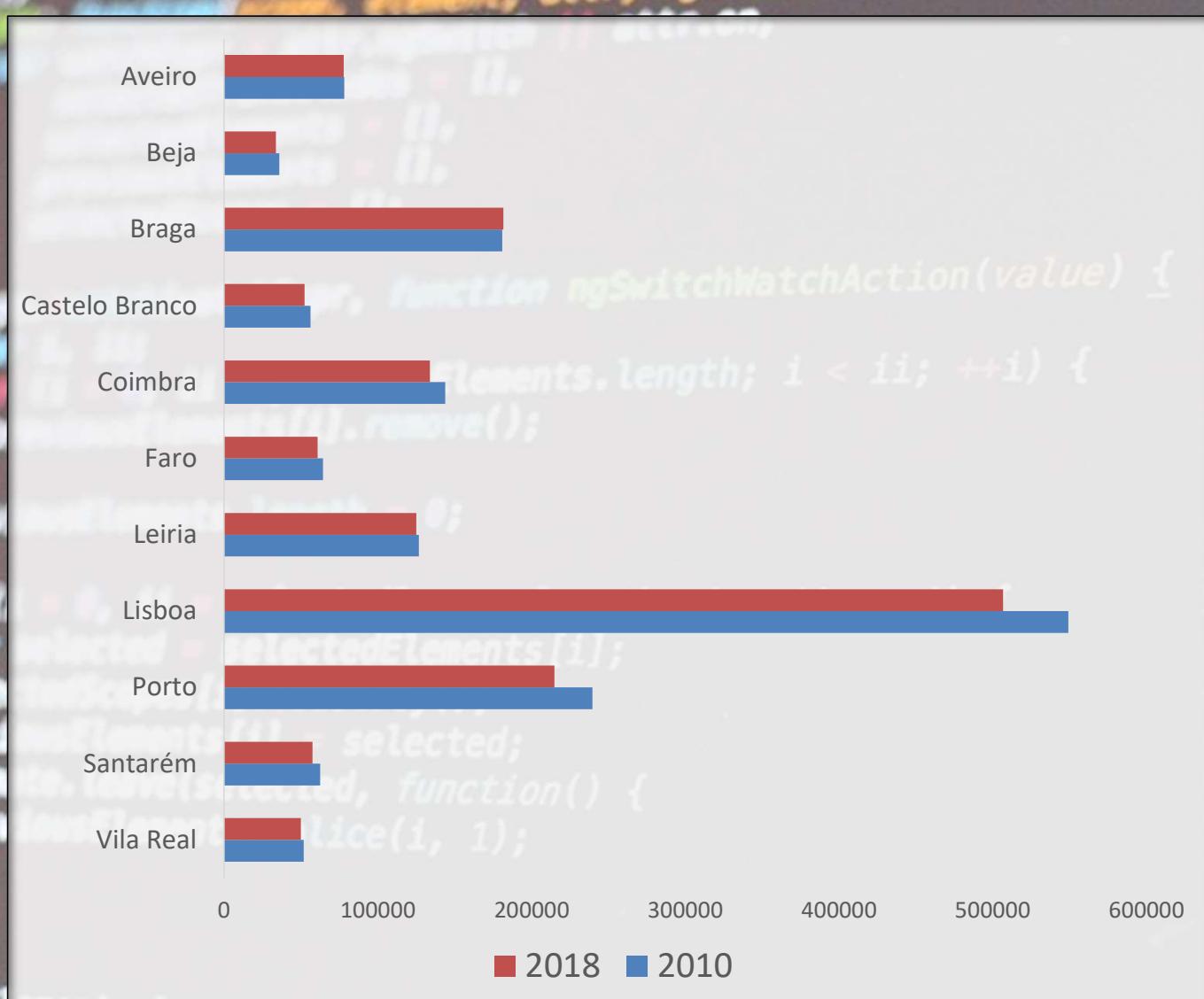
## *3<sup>th</sup> Question*

Which is the economic sustainability of each district?



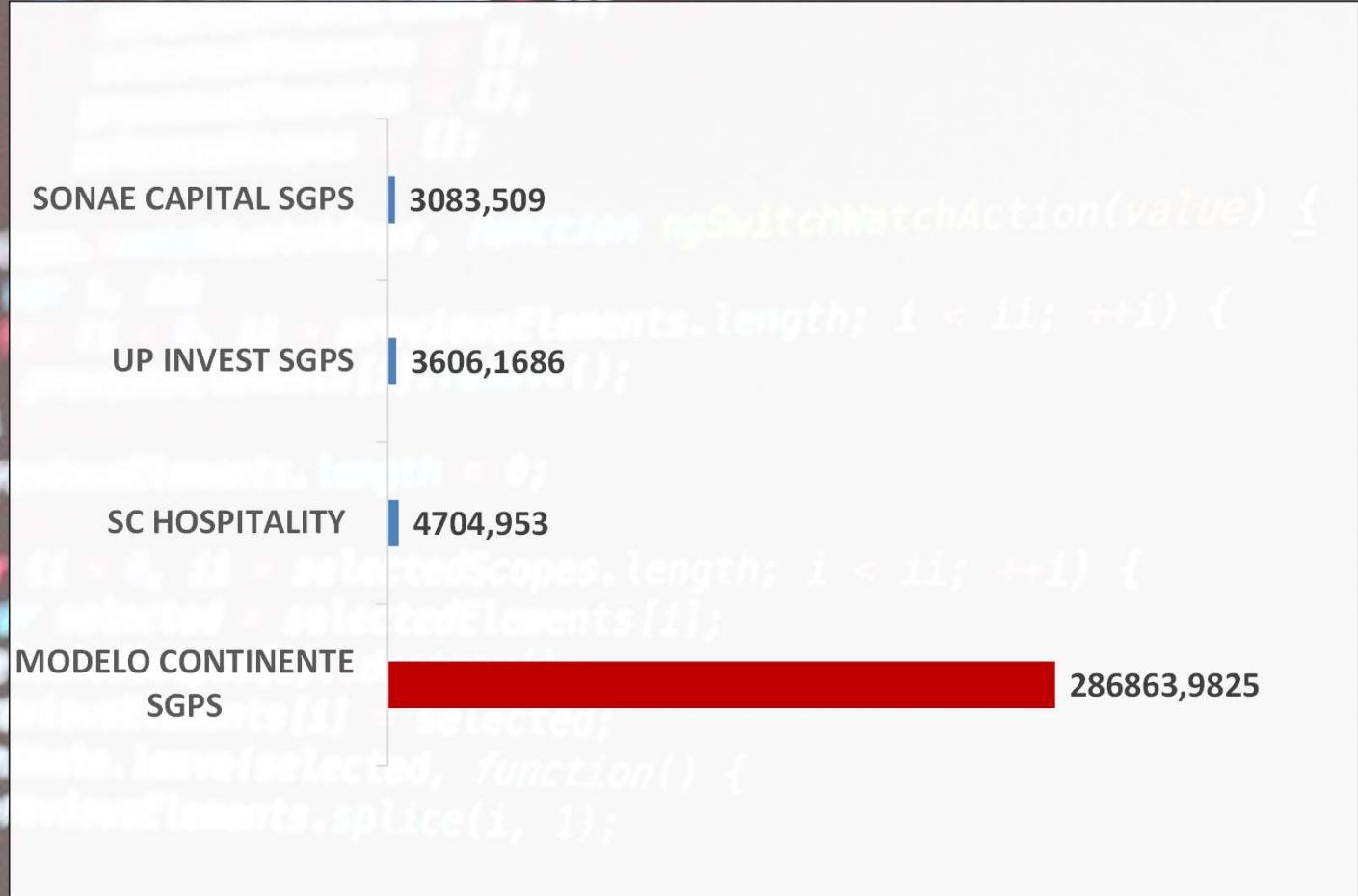
## 4<sup>th</sup> Question

Which district has the highest population density in 2010 and 2018?



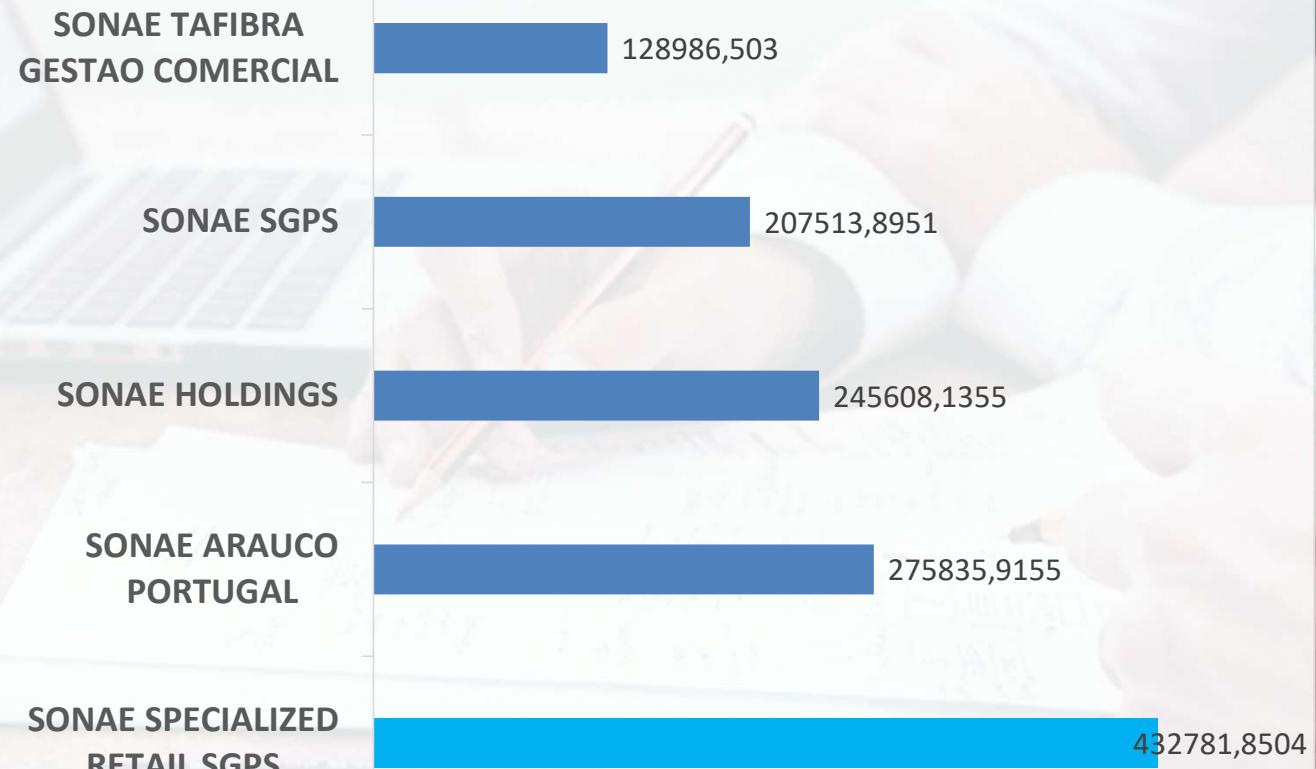
## 5<sup>th</sup> Question

Does any of the companies present a loss?



## 6<sup>th</sup> Question

What is the most profitable company?



## 7<sup>th</sup> Question

What is the top 5 companies with better and worst Operational Result (OR) average, in 2010 and 2018?

2010

2018

Company Name	District	OR Average
A ASINDUSTRIA DE MOVEIS	Aveiro	-0,08
ACTOS DE PRIMAVERA	Braga	319,5875
ADIVINHA PRIMAVERA	Faro	11,1685
VERTICE DA PRIMAVERA	Vila Real	-497,6396
VOZES DA PRIMAVERA	Braga	-15,5519

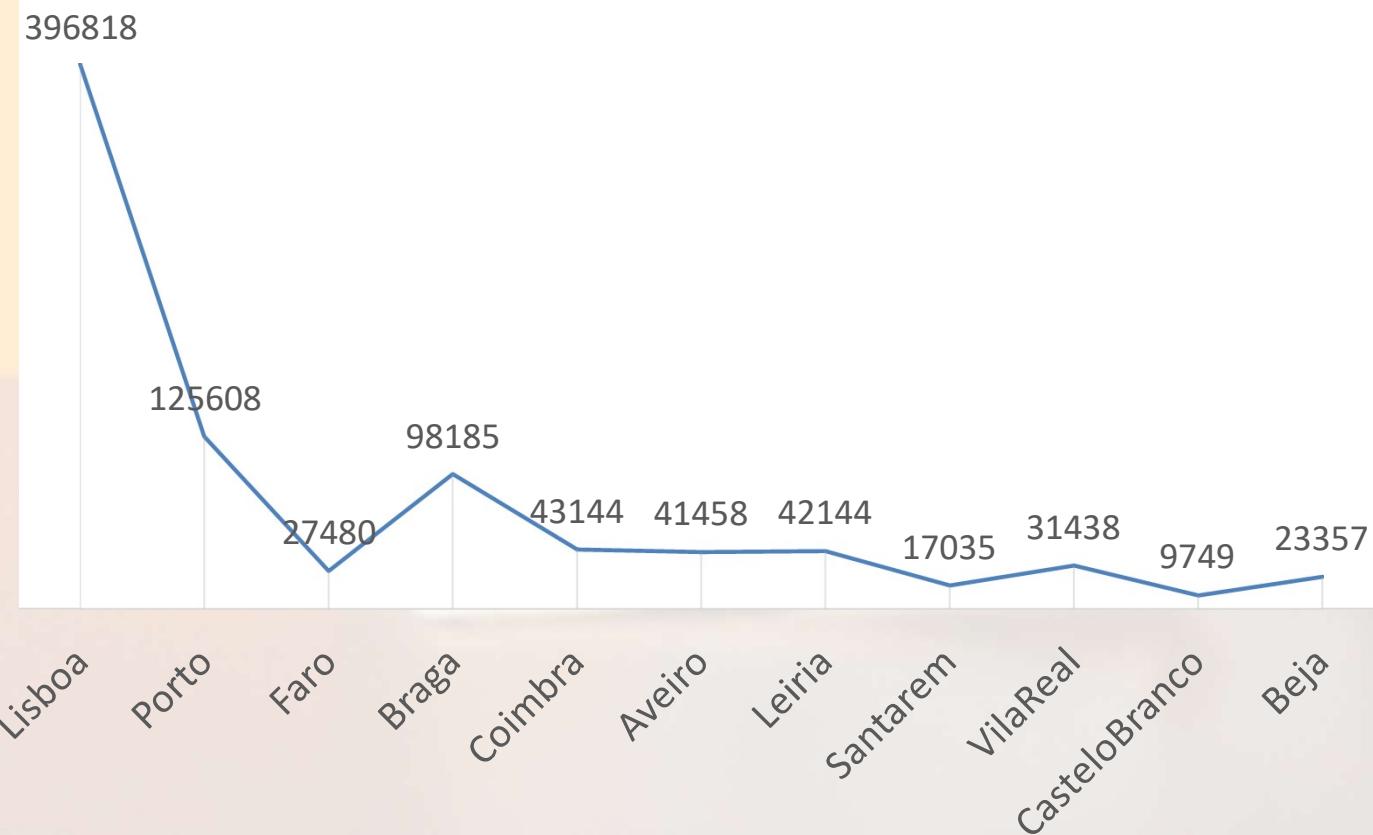
Company Name	District	OR Average
VERTICE DA PRIMAVERA	Vila Real	-497,6396
VOZES DA PRIMAVERA	Braga	-15,5519
A ASINDUSTRIA DE MOVEIS	Aveiro	-0,08
VARANDAS DA PRIMAVERA EMPREENDIMENTOS	Faro	0
ACTOS DE PRIMAVERA	Braga	319,5875

## 8<sup>th</sup> Question

Which district has more likes on Facebook?



Likes on Facebook, in January 2020



## 9<sup>th</sup> Question

Which district has more followers on Facebook?



Facebook followers, in January 2020



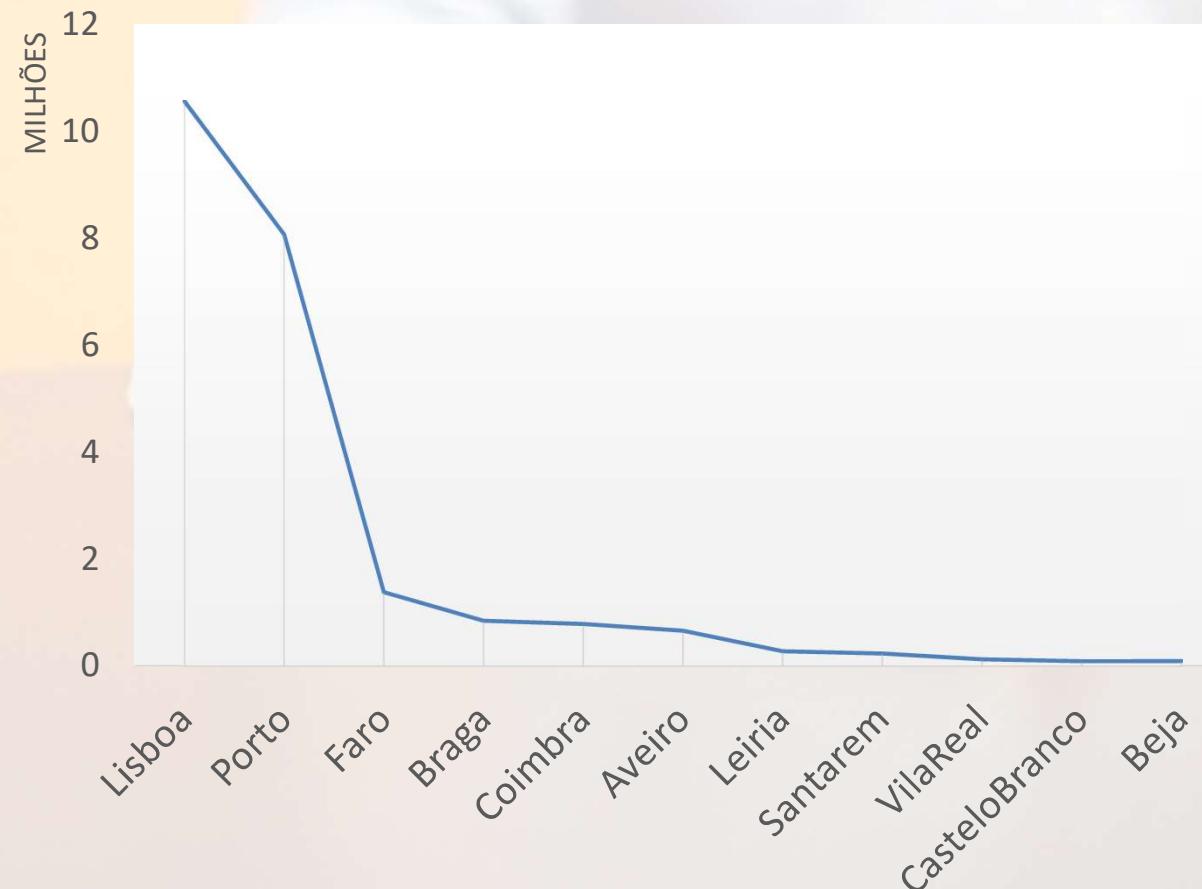
## 10<sup>th</sup> Question

What is the city most mentioned on Instagram?

Instagram



#CityName on Instagram, in January 2020



## 11<sup>th</sup> Question

What is the top 4 companies with better and worst Operating Result (OR) per employee, in 2010 and 2018?



2010

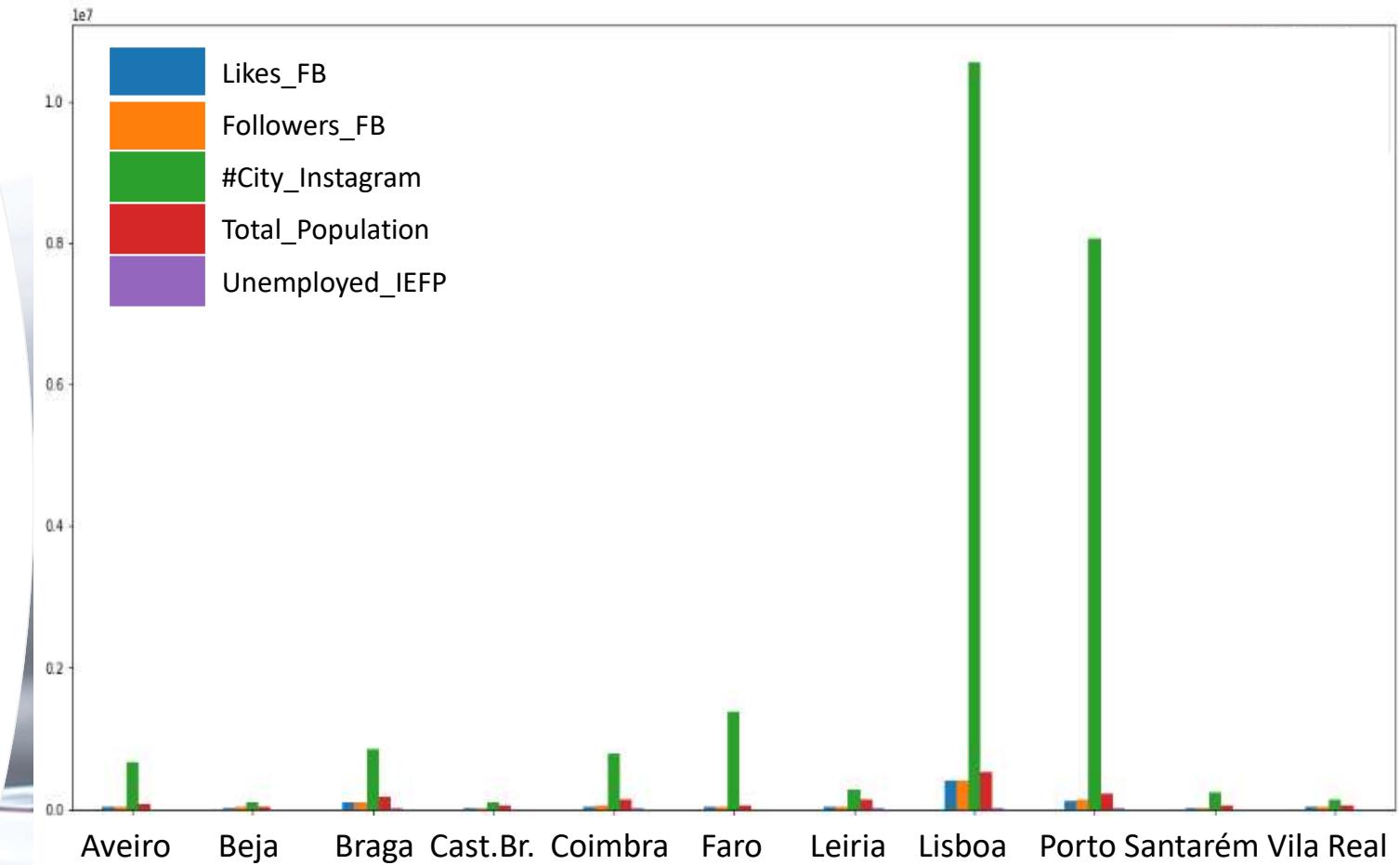
2018

Name	OR/Employee
SONAE SPECIALIZED RETAIL SGPS	417682.0112
MODELO CONTINENTE SGPS	105534.8065
UP INVEST SGPS	-3677.5041
SC HOSPITALITY	-4793.5290

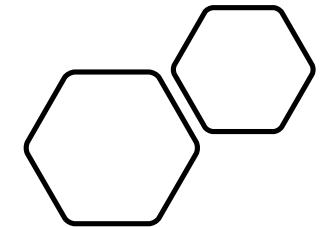
Name	OR/Employee
SONAE SPECIALIZED RETAIL SGPS	417682.0112
MODELO CONTINENTE SGPS	105534.8065
OPTIMUS SGPS	49325.5376
DST SGPS	-2124.0354
UP INVEST SGPS	-3677.5041
SC HOSPITALITY	-4793.5290

## 12<sup>th</sup> Question

Make a graph showing the performance of each district on social networks along with it's total population and the total number of unemployed.



# Questions?



Thank's for your  
attention!