

# A new business in Tarragona?

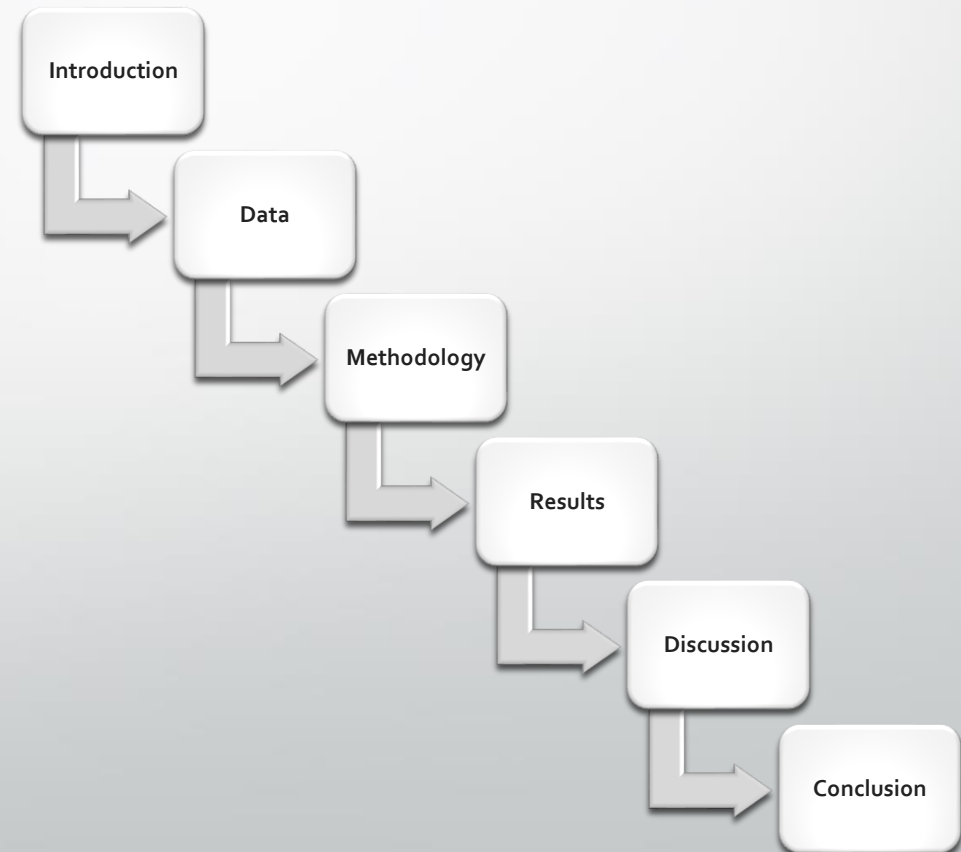
Damián Morales Sánchez

Tarragona (Spain)

August 5th, 2019

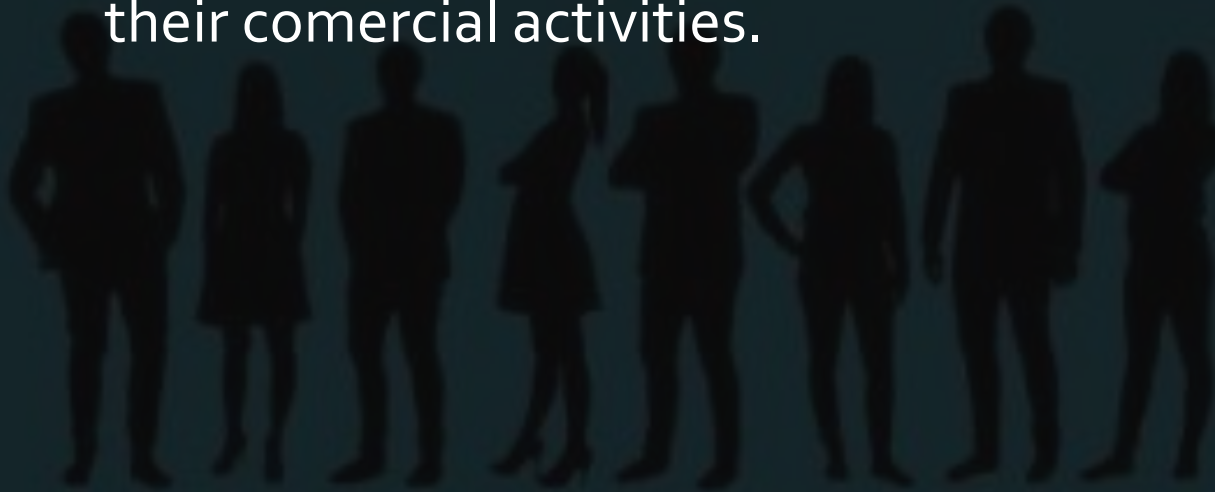
# We will talk about the...

1. Introduction
2. Data collection
3. Methodology
4. Results
5. Discussion
6. Conclusion



# Introduction

- An investment company wants to invest in Tarragona
- For this purpose, we analyze six neighbourhoods belonging to Tarragona to extract computationally their commercial activities.



# The neighbourhoods

An abstract graphic design featuring a dark grey background. On the right side, there are several thick, parallel lines in bright blue and dark grey. These lines form a corner-like structure, with one line extending diagonally upwards and another extending diagonally downwards, meeting at a point. The lines have a slight 3D effect, with the blue line appearing slightly above the grey line.

An aerial photograph of a crowded beach. The left side of the image is covered by a dark, semi-transparent overlay. A geometric graphic consisting of several parallel diagonal lines in blue and grey divides the image. The right side shows a sandy beach with many people, bordered by a dense line of green pine trees. In the background, a coastal town with buildings and a hill is visible under a clear sky.

# La Pineda



# Sant Pere i Sant Pau



Im  
Ba



Salou





An aerial photograph of a large, bustling outdoor market. The market is filled with numerous stalls, many of which are covered with black or white tarps. People are seen walking through the aisles between the stalls. In the foreground, there are several white vans and trucks. The market is set in an open area with trees visible in the background. A large blue diagonal stripe runs across the image from the top right towards the bottom left. The word "Bonavista" is written in white text on the left side of the image.

# Bonavista



The background image shows a wide-angle view of a football stadium. The pitch is green with white markings. In the background, there are several tall, multi-story apartment buildings. The foreground shows a concrete wall and some trees. A large, stylized graphic element consisting of several overlapping diagonal bands in shades of blue and grey runs from the top right towards the bottom left, partially obscuring the stadium view.

# Torreforta



# Vila-Seca



# Data collection

We collect the data from two different sources: Google Maps and Foursquare API.



Google Maps provide us the coordinates of each neighbourhood and the Foursquare API allows us to extract the 10th most important venues from each one.

# Methodology

- We create a CSV file with the latitude and longitude of each neighbourhood.

	A	B	C	D	E
1	ciudad, barrio, latitud, longitud				
2	Tarragona,	Sant Pere i Sant Pau,	41.137686,	1.152032	
3	Tarragona,	Fonteta,	41.118550,	1.216849	
4	Tarragona,	Salou,	41.080084,	1.132025	
5	Tarragona,	La Pineda,	41.079023,	1.182240	
6	Tarragona,	Vila Seca,	41.102044,	1.146068	
7	Tarragona,	Bonavista,	41.122614,	1.192402	
8					
9					



# Methodology

- Then, we create a map from the coordinates collected.



venues\_sorted.head()

# Methodology

- From Foursquare API we extract the 10th more important venues of each 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
Farmers Market	Tapas Restaurant	Bar	Spanish Restaurant	Gym	Cocktail Bar	Diner	Diner	Department Store
Ice Cream Shop	Café	Hotel	Restaurant	Gastropub	Diner	Pizza Place	Tapas Restaurant	Cocktail Bar
Hotel	Mediterranean Restaurant	Spanish Restaurant	Café	Restaurant	Pizza Place	Donut Shop	Tapas Restaurant	Train Station
Bakery	Spanish Restaurant	Soccer Field	Brewery	Train Station	Cocktail Bar	Donut Shop	Diner	Department Store
Supermarket	Stadium	Department Store	Restaurant	Hotel	Train Station	Café	Diner	Concert Hall



# Methodology

- After that, we apply a Machine Learning algorithm to cluster the neighbourhoods with a similar commercial activity. Concretely, we apply a K-Means Clusters Algorithm and we drew five clusters.

Cluster 2

```
In [126]: tarragona_merged.loc[tarragona_merged['Cluster labels'] == 2, tarragona_merged.columns[[1] + list(range(5, tarragona_merged.shape[1]))]]
```

Out[126]:

	barrio	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
2	Salou	Hotel	Mediterranean Restaurant	Spanish Restaurant	Café	Restaurant	Pizza Place	Donut Shop	Tapas Restaurant	Bar	Grocery Store
3	La Pineda	Ice Cream Shop	Café	Hotel	Restaurant	Gastropub	Diner	Fish & Chips Shop	Tapas Restaurant	Cocktail Bar	Italian Restaurant

Cluster 5

```
In [128]: tarragona_merged.loc[tarragona_merged['Cluster labels'] == 5, tarragona_merged.columns[[1] + list(range(5, tarragona_merged.shape[1]))]]
```

Out[128]:

	barrio	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
5	Bonavista	Farmers Market	Tapas Restaurant	Bar	Spanish Restaurant	Gym	Cocktail Bar	Donut Shop	Diner	Apartment	Concert Hall



# Results



# Cluster 1

## Cluster 1

```
In [122]: tarragona_merged.loc[tarragona_merged['Cluster Labels'] == 0, tarragona_merged.columns[[1] + list(range(5, tarragona_merged.shape[1]))]]
```

Out[122]:

	barrio	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
1	Torreforta	Supermarket	Stadium	Department Store	Restaurant	Hotel	Train Station	Café	Diner	Concert Hall	Coffee Shop

# Cluster 2

## Cluster 2

```
In [123]: tarragona_merged.loc[tarragona_merged['Cluster Labels'] == 1, tarragona_merged.columns[[1] + list(range(5, tarragona_merged.shape[1]))]]
```

Out[123]:

	barrio	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	Sant Pere i Sant Pau	Bakery	Spanish Restaurant	Soccer Field	Brewery	Train Station	Cocktail Bar	Donut Shop	Diner	Department Store	Concert Hall



# Cluster 3

## Cluster 3

```
In [126]: tarragona_merged.loc[tarragona_merged['Cluster Labels'] == 2, tarragona_merged.columns[[1] + list(range(5, tarragona_merged.shape[1]))]]
```

```
Out[126]:
```

	barrio	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
2	Salou	Hotel	Mediterranean Restaurant	Spanish Restaurant	Café	Restaurant	Pizza Place	Donut Shop	Tapas Restaurant	Train Station	Grocery Store
3	La Pineda	Ice Cream Shop	Café	Hotel	Restaurant	Gastropub	Diner	Fish & Chips Shop	Tapas Restaurant	Cocktail Bar	Italian Restaurant

# Cluster 4

## Cluster 4

```
In [128]: tarragona_merged.loc[tarragona_merged['Cluster Labels'] == 3, tarragona_merged.columns[[1] + list(range(5, tarragona_merged.shape[1]))]]
```

```
Out[128]:
```

	barrio	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
5	Bonavista	Farmers Market	Tapas Restaurant	Bar	Spanish Restaurant	Gym	Cocktail Bar	Donut Shop	Diner	Department Store	Concert Hall



# Cluster 5

## Cluster 5

```
In [129]: tarragona_merged.loc[tarragona_merged['Cluster Labels'] == 4, tarragona_merged.columns[[1] + list(range(5, tarragona_merged.shape[1]))]]
```

```
Out[129]:
```

	barrio	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
4	Vila Seca	Basketball Court	Restaurant	Pool	Train Station	Café	Donut Shop	Diner	Department Store	Concert Hall	Coffee Shop

```
In [ ]:
```



## **Discussion and Conclusion**

Which is the best business in each neighbourhood?



A photograph of a shopping cart in a supermarket aisle. The cart is filled with various groceries, including bags of potatoes and other produce. A person's hands are visible on the handle of the cart. The background shows shelves stocked with various products. The image is overlaid with a blue and grey geometric design consisting of several diagonal lines.

# Torreforta





# Sant Pere i Sant Pau

The image is a collage of various cakes and pastries. On the left, there are several cakes on a shelf, including one labeled 'FUDGE BROWNIE' and another labeled 'ASSORTED BARS'. On the right, there are more cakes, including one labeled 'SHERRYBREAD COOKIES Small Large' and another labeled 'Baklava'. The cakes are decorated with various toppings like fruit, cream, and chocolate. A blue and grey geometric overlay is present in the center of the image.





# Salou & La Pineda





# Bonavista





Vila-Seca



**Thanks for the attention!**