

Lighthouse Classification

IBM Data Science Capstone

Iván Martínez Suárez

1 Introduction

2 Data

3 Lighthouse clustering

4 Final Results

5 Conclusions

Introduction (I)

Background

- Galicia: spanish region with many lighthouses along its coastline
- Government's plan to boost tourism in villages near lighthouses
- Grants for newly opened food business located less than 3 km from a lighthouse
- Grants proportional to the age of the lighthouse

Introduction (II)

Business problem

- A company is interested in creating a new catering business
- Problem: many lighthouses are already populated with restaurants and bars
- Which is the best lighthouse to place the business?
- Find the best place according to restaurant density and age of the lighthouse

Data source

Source: spanish lighthouses at Wikipedia

Faro	Imagen	Localización & Coordenadas	Provincia / Comunidad	Autoridad Portuaria	Cuerpo de agua	Año de construcción	Altura del soporte	Altura del plano focal	Rango
Adra ⁶		Adra 🌐 36.74796, -3.03098	Almería Andalucía	Almería	Mar Mediterráneo	1889	26 metros (85 pies)	49 metros (161 pies)	16 millas náuticas (30 km)
Águilas ⁷		Águilas 🌐 37.40169, -1.57802	Murcia	Cartagena	Mar Mediterráneo	1860	23 metros (75 pies)	30 metros (98 pies)	13 millas náuticas (24 km)
Ahorcados ^{8 9}		Isla de los Ahorcados 🌐 38.81451, 1.41178	Islas Baleares	Baleares	Mar Mediterráneo	1856	17 metros (56 pies)	27 metros (89 pies)	10 millas náuticas (19 km)
Albir ^{10 11}		Alfaz del Pi 🌐 38.56363, -0.05007	Alicante Comunidad Valenciana	Alicante	Mar Mediterráneo	1863	8 metros (26 pies)	112 metros (367 pies)	15 millas náuticas (28 km)

Figure: Some of the Galician lighthouses and their features before cleaning (Wikipedia).

Data cleaning and selection

Cleaned data of Galician lighthouses

	Lighthouse Name	Latitude	Longitude	Community	YearBuilt	TowerHeight	FocalHeight	Range
0	Cabo Corrubedo	42.57627	-9.09009	Galicia	1854.0	14.0	32.0	15.0
1	Cabo Finisterre	42.88236	-9.27196	Galicia	1853.0	17.0	143.0	23.0
2	Cabo Ortegal	43.77107	-7.87018	Galicia	1954.0	10.0	124.0	18.0
3	Cabo Prior	43.56761	-8.31453	Galicia	1853.0	7.0	107.0	22.0
4	Cabo Prioriño	43.45879	-8.34033	Galicia	1854.0	5.0	36.0	23.0

Figure: Some of the Galician lighthouses and their features after cleaning.

Foursquare data

Most common food venues for every lighthouse

Lighthouse Name	Bakery	Brazilian Restaurant	Breakfast Spot	Burger Joint	Cafeteria	Café	Diner	Food	Food Court	Gastropub
Cabo Corrubedo	0.0	0.0	0.0	0.0	0.00	0.0	0.0	0.0	0.0	0.00
Cabo Finisterre	0.0	0.0	0.0	0.0	0.00	0.0	0.1	0.1	0.0	0.00
Cabo Ortegal	0.0	0.0	0.0	0.0	0.00	0.0	0.5	0.0	0.0	0.00
Cabo Priorioño	0.0	0.0	0.0	0.0	0.00	0.0	0.0	0.0	0.0	0.00
Cabo Silleiro	0.0	0.0	0.0	0.0	0.25	0.0	0.0	0.0	0.0	0.00

Figure: Frequency of venues for some lighthouses (from Foursquare).

Lighthouse distribution

Map of Galician lighthouses

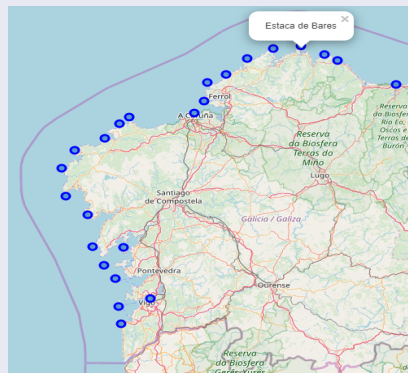


Figure: Map of Galician lighthouses.

Lighthouse clustering

Map of clusters

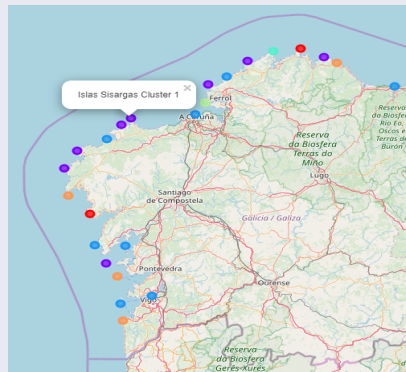


Figure: Clustering of Galician lighthouses using kMeans.

The best cluster for our purpose

Cluster 1 (in purple on the map)

- Total absence of food venues in their vicinities
- Best candidates to install a new catering business

	Lighthouse Name	Latitude	Longitude	Community	YearBuilt	TowerHeight	FocalHeight	Range	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
3	Cabo Prior	43.56761	-8.31453	Galicia	1853.0	7.0	107.0	22.0	1	None	None	None
6	Cabo Villano	43.16041	-9.21093	Galicia	1896.0	25.0	104.0	28.0	1	None	None	None
8	Isla de Sálvora	42.46586	-9.01308	Galicia	1852.0	16.0	40.0	21.0	1	None	None	None
11	Islas Sisargas	43.35991	-8.84457	Galicia	1853.0	11.0	110.0	23.0	1	None	None	None
16	Punta Candelaria	43.71083	-8.04709	Galicia	1954.0	9.0	89.0	21.0	1	None	None	None
20	Punta Nariga	43.32068	-8.91010	Galicia	1998.0	39.0	55.0	22.0	1	None	None	None
21	Punta Roncadoura	43.73576	-7.52530	Galicia	1984.0	14.0	94.0	21.0	1	None	None	None
22	Touriñán (Toriñana)	43.05336	-9.29814	Galicia	1898.0	14.0	65.0	24.0	1	None	None	None

Lighthouses sorted by age

Cluster 1

Lighthouse Name	YearBuilt
Isla de Sálvora	1852.0
Cabo Prior	1853.0
Islas Sisargas	1853.0
Cabo Villano	1896.0
Touriñán (Toriñana)	1898.0
Punta Candelaria	1954.0
Punta Roncadoira	1984.0
Punta Nariga	1998.0

Figure: List of the best Galician lighthouses to place a new food business.

Conclusions and Future Work

- Reasonable results obtained
- Be careful! Isla de Sálvora is an inhabited island
- Future improvement: number of clusters and other issues