Introduction

If you want to visit cities in Europe, what do you expect to find? Maybe cafés, italian restaurants, or museums. There are 58 capital cities in Europe, ones of them are very different from the others. The experience visiting those cities depends strongly on the places that can be found. So, which city should I visit according to my own preferences? How can I plan a highly attractive tour through Europe by selecting the right cities?

The problem

There is no tool or platform that helps and guides tourists through their trip to Europe in a way that they can fulfill their desires and expectations of places. In order to solve this problem a machine learning methodology will be developed to categorize the european capitals by their main venues.

Interest

The people that will benefit from this work are the tourists, the local citizens of european cities and enterprises that promote tourism. This work can help to show and promote cities of different countries that are not well positioned in the european tourist radar.

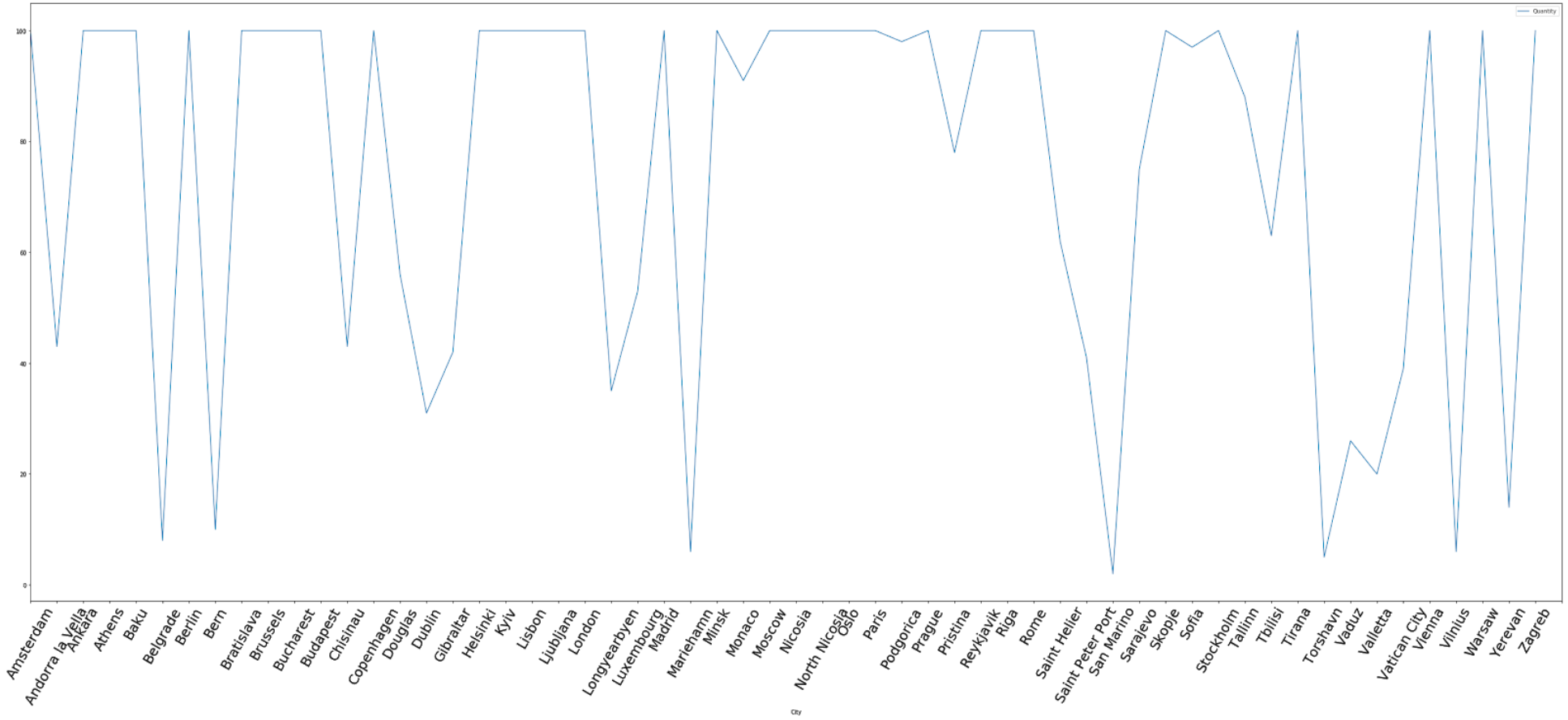
Data acquisition and processing methodology

To obtain data of the venues of every european city an API to Foursquare was done. The API returned the 100 closest venues to the very center of the city within a radius of 1000 meters. After that the data was consolidated into one single data frame with the venues categories.

After that a k-mean classification of the cities based on the category of their venues was done. The k-mean classification considered 8 clusters. After those clusters were obtained, then an exploration of them was done in order to understand the differences in the venues of different cluster cities.

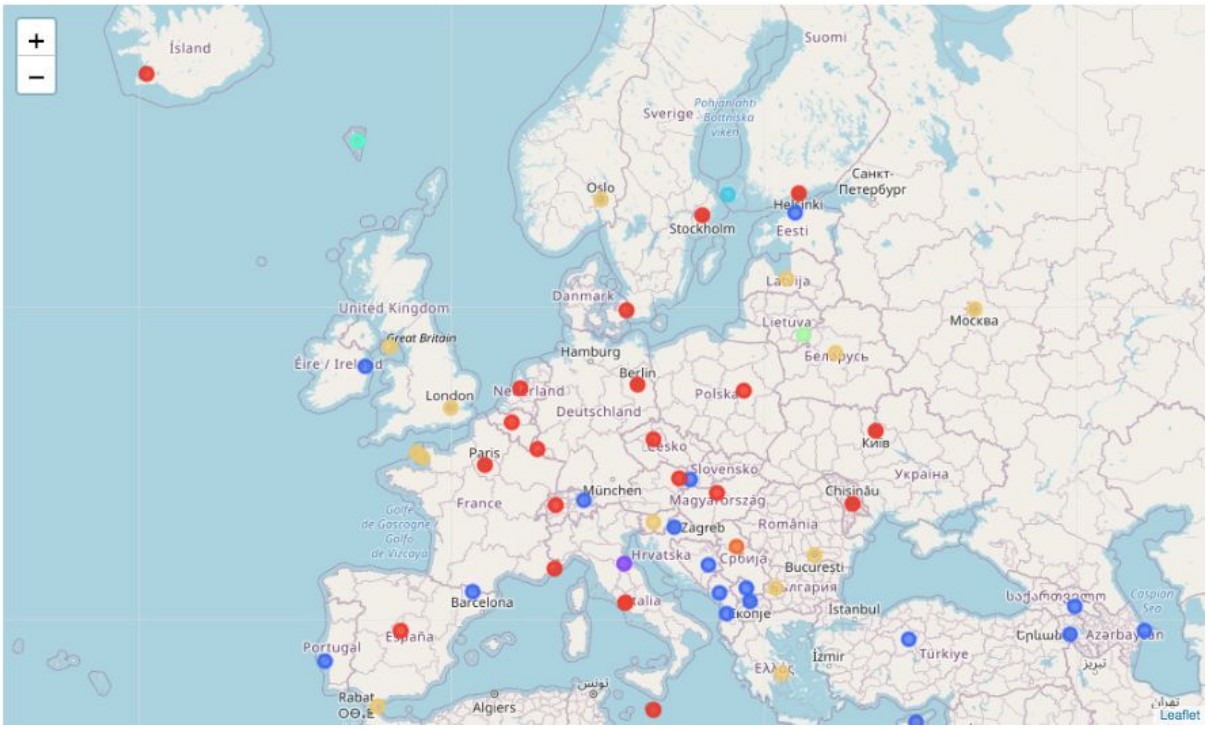
Data exploration

A first understanding of the data was exploring the quantity of venues that each capital has with a maximun of 100 venues. The following plot illustrate it.


Quantity
100
80
60
40
20
0
Amsterdam
Andorra la Vella
Ankara
Athens
Baku
Belgrade
Berlin
Bern
Bratislava
Brussels
Bucharest
Budapest
Chisinau
Copenhagen
Douglas
Dublin
Gibraltar
Helsinki
Kyiv
Lisbon
Ljubljana
London
Longyearbyen
Luxembourg
Madrid
Mariehamn
Minsk
Monaco
Moscow
Nicosia
North Nicosia
Oslo
Paris
Podgorica
Prague
Pristina
Reykjavik
Riga
Rome
Saint Helier
Saint Peter Port
San Marino
Sarajevo
Skopje
Sofia
Stockholm
Tallinn
Tbilisi
Tirana
Torshavn
Vaduz
Valletta
Vatican City
Vienna
Vilnius
Warsaw
Yerevan
Zagreb
City

K-mean results

1.- Geographical plot



In the plot you can see how the cluster capital cities are distributed through the continent.

2.- Clusters analysis

Per each cluster two tables are shown. The first one contains the cities that belong to the cluster and the second table contains the top 10 types of venues (category) of the cluster cities ordered by the most frequent places.

- **CLUSTER 0:**

City
Amsterdam
Berlin
Bern
Brussels
Budapest
Chisinau
Copenhagen
Helsinki
Kyiv
Longyearbyen
Luxembourg
Madrid
Monaco
Paris
Prague
Reykjavik
Rome
Stockholm
Valletta
Vatican City

	value
Hotel	0.059620
Italian Restaurant	0.043850
Café	0.036595
Bar	0.032923
Coffee Shop	0.032108
Restaurant	0.025819
Plaza	0.024194
Ice Cream Shop	0.017281
Sandwich Place	0.017075
Pizza Place	0.016660

- **CLUSTER 1:**

	value
cluster	1.0
Construction & Landscaping	0.5
Health Food Store	0.5

City
San Marino

- **CLUSTER 2:**

	City
1	Andorra la Vella
2	Ankara
4	Baku
8	Bratislava
15	Dublin
19	Lisbon
30	North Nicosia
33	Podgorica
35	Pristina
42	Sarajevo
43	Skopje
46	Tallinn
47	Tbilisi
48	Tirana
50	Vaduz
56	Yerevan
57	Zagreb

	value
cluster	2.000000
Café	0.103879
Restaurant	0.076326
Hotel	0.068919
Bar	0.040812
Coffee Shop	0.028821
Italian Restaurant	0.019567
Bakery	0.019234
Cocktail Bar	0.016128
Pub	0.015841
Park	0.014972

- **CLUSTER 3:**

	City
25	Mariehamn

	value
cluster	3.000000
Burrito Place	0.166667
Cosmetics Shop	0.166667
Gym	0.166667
Lake	0.166667
Paper / Office Supplies Store	0.166667
Supermarket	0.166667

- **CLUSTER 4:**

	City
49	Torshavn

	value
cluster	4.0
Boat or Ferry	0.2
Bus Station	0.2
Furniture / Home Store	0.2
Harbor / Marina	0.2
Steakhouse	0.2

- **CLUSTER 5:**

	City
54	Vilnius

	value
cluster	5.000000
Grocery Store	0.333333
Art Gallery	0.166667
Beer Garden	0.166667
Motorcycle Shop	0.166667
Supermarket	0.166667

- **CLUSTER 6:**

	City
3	Athens
10	Bucharest
14	Douglas
16	Gibraltar
20	Ljubljana
21	London
26	Minsk
28	Moscow
29	Nicosia
31	Oslo
37	Riga
39	Saint Helier
40	Saint Peter Port
44	Sofia

	value
cluster	6.000000
Coffee Shop	0.076195
Bar	0.051895
Hotel	0.048008
Pub	0.043079
Restaurant	0.042877
Café	0.039568
Cocktail Bar	0.025358
Park	0.022609
Plaza	0.021906
Italian Restaurant	0.019945

- **CLUSTER 7:**

City		value
5	Belgrade	cluster 7.000
		Flower Shop 0.250
		Restaurant 0.250
		Seafood Restaurant 0.250
		Gas Station 0.125
		Train Station 0.125

Conclusions and recommendations

If you like flower shops I suggest you visit Belgrade, if you prefer bars and coffee shops I suggest you visit the cities of cluster 6 and 0 such as Amsterdam, Athens, Bucharest, Douglas, Berlin, Brussels, etc. There are also capital cities in which it is not easy to find a café or a restaurant such as Vilnius and Torshavn.

Future directions

The next steps of this work is to include costs of tourism in European cities. This will serve to reach all the tourist expectations and their pockets capacity.