PERFECT CITY FOR OPENING A CAFÉ BAR

1. INTRODUCTION

I always wanted to open a cafe bar by myself. The problem is that I don't want that my cafe bar is on a random place in a random city. I want a cafe bar in one of the biggest European cities, and not just a regular one placed in the middle of nowhere, I want it to be placed at the city center.

So, I did a research on the list of the most populated cities in Europe based on the number of similar contents that serve coffee in the center area of the city.

I would like to present my research.

2. IMPORTING AND REDEFINING THE DATASET

First step was importing all of the libraries needed. Then I imported the dataset I wanted from wikipedia. I used a data from the wikipedia page with list of the biggest European cities („https://en.wikipedia.org/wiki/List_of_European_cities_by_population_within_city_limits"), and when I did a little arrangement it looked like this:

	Rank	City	Country	Population
0	1	Istanbul[a]	Turkey	15,519,267[2]
1	2	Moscow[b]	Russia	12615279
2	3	London	United Kingdom	9126366
3	4	Saint Petersburg	Russia	5383890
4	5	Berlin	Germany	3748148
5	6	Madrid	Spain	3223334
6	7	Kiev	Ukraine	2950800
7	8	Rome	Italy	2844750
8	9	Paris	France	2140526
9	10	Bucharest	Romania	2106144
10	11	Minsk	Belarus	1982444
11	12	Hamburg	Germany	1930996
12	13	Vienna	Austria	1899055
13	14	Warsaw	Poland	1802237
14	15	Budapest	Hungary	1768073
15	16	Barcelona	Spain	1636762
16	18	Munich	Germany	1471508
17	17	Kharkiv	Ukraine	1451132
18	19	Milan	Italy	1389834
19	20	Prague	Czech Republic	1308632
20	21	Nizhny Novgorod	Russia	1259013
21	22	Kazan	Russia	1243500
22	23	Sofia	Bulgaria	1238438
23	24	Birmingham	United Kingdom	1223000
24	25	Brussels	Belgium	1211035
25	27	Samara	Russia	1170910
26	28	Belgrade	Serbia	1166763
27	29	Ufa	Russia	1121429
28	30	Rostov-on-Don	Russia	1119875
29	31	Cologne	Germany	1085664
30	32	TekirdaÄ□	Turkey	1055412
31	33	Voronezh	Russia	1054537
32	34	Perm	Russia	1051583
33	35	Volgograd	Russia	1013533
34	36	Odessa	Ukraine	1011494

3. GETTING LOCATION DATA FOR THE CITIES

If you noted that there's no city ranked on no. 26 (not the 26th index value), it's a mistake by the Wikipedia database, I didn't just clear it away.

The second thing I did was discovering the coordinates of each city from our list. I did it using the geographical libraries from python. After writing some codes and arranging the dataframe we got this (you'll see only the first ten rows).

	Rank	City	Country	Population	latitude	longitude
0	1	Istanbul	Turkey	15519267	41.009633	28.965165
1	2	Moscow	Russia	12615279	55.479205	37.327330
2	3	London	United Kingdom	9126366	51.507322	-0.127647
3	4	Saint Petersburg	Russia	5383890	59.960674	30.158655
4	5	Berlin	Germany	3748148	52.517037	13.388860
5	6	Madrid	Spain	3223334	40.416705	-3.703582
6	7	Kiev	Ukraine	2950800	50.450034	30.524136
7	8	Rome	Italy	2844750	41.893320	12.482932
8	9	Paris	France	2140526	48.856697	2.351462
9	10	Bucharest	Romania	2106144	44.436141	26.102720

4. GETTING THE DATA FOR ALL THE CONTENTS BASED IN THE CENTRE OF EACH CITY

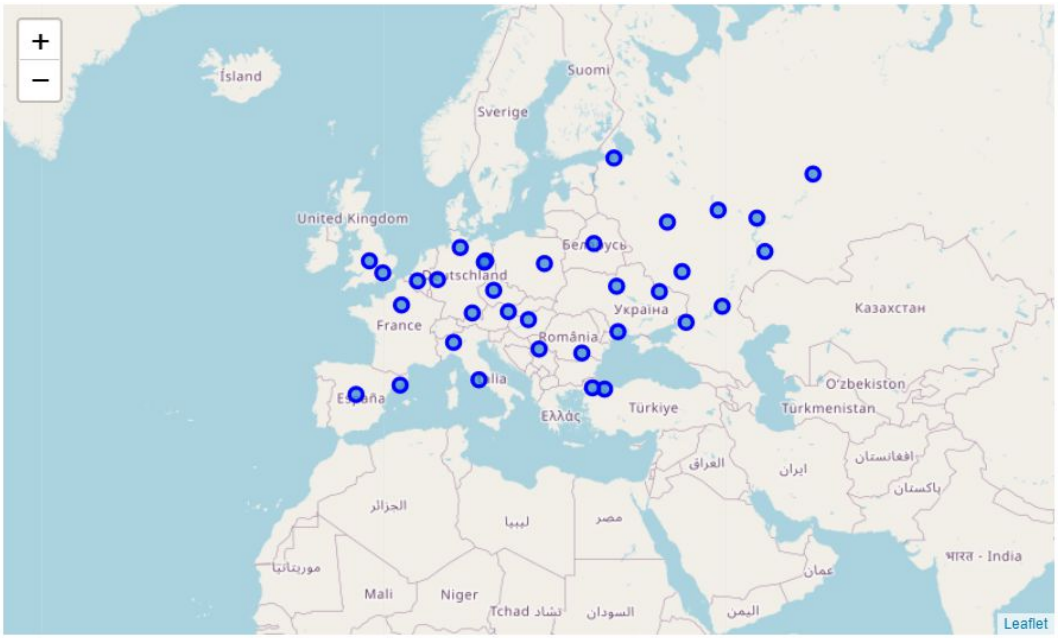
Now that we have a dataframe filled with all the things about the cities that we need for our project, the next thing is to get the best data about the contents in the city center. I did it using the Foursquare API from which I got data about all the contents that are based strictly in the center of each city from our list.

Here you can see a sample of the list containig the information about the city, the name of the content, its coordinates and a category to which it belongs.

	City	City Latitude	City Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Istanbul	41.009633	28.965165	Bayezid Camii	41.009736	28.964500	Mosque
1	Istanbul	41.009633	28.965165	Day Day Pastanesi	41.009440	28.968023	Bakery
2	Istanbul	41.009633	28.965165	Kapalıçarşı	41.009725	28.968315	Historic Site
3	Istanbul	41.009633	28.965165	Nuruosmaniye Camii	41.010151	28.970701	Mosque
4	Istanbul	41.009633	28.965165	Cevahir Bedesteni (Oldbazaar)	41.010894	28.969446	Antique Shop
...	...	...	...	...	...	...	...
3061	Odessa	46.487319	30.739278	Детская площадка у "Немо"	46.477106	30.764376	Playground
3062	Odessa	46.487319	30.739278	Wine Story - Винфорт	46.463149	30.741597	Wine Shop
3063	Odessa	46.487319	30.739278	Make My Cake Cafe	46.463345	30.754717	Dessert Shop
3064	Odessa	46.487319	30.739278	Дельфинарий Немо / Nemo Dolphinarium (Дельфіна...	46.476975	30.765269	Aquarium
3065	Odessa	46.487319	30.739278	Одесса, Пляж Аркадия, Центральная Аллея	46.473283	30.764131	Beach

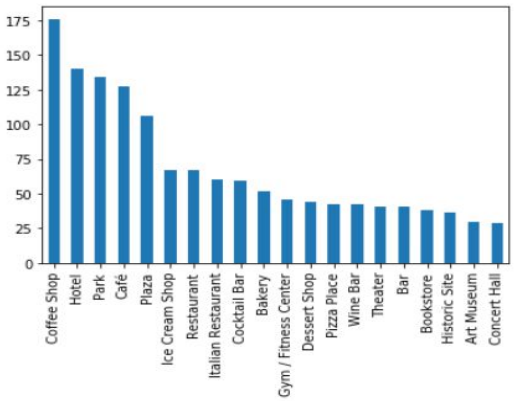
5. CREATING A MAP THAT SHOWS US LOCATIONS ON EVERY CITY FROM THE DATABASE

In the next step, I used the geographical data from the dataset and the Folium library to create a map that shows the locations of all the biggest cities.



6. SHOWING THE MOST COMMON VENUES ACROSS ALL THE CITIES

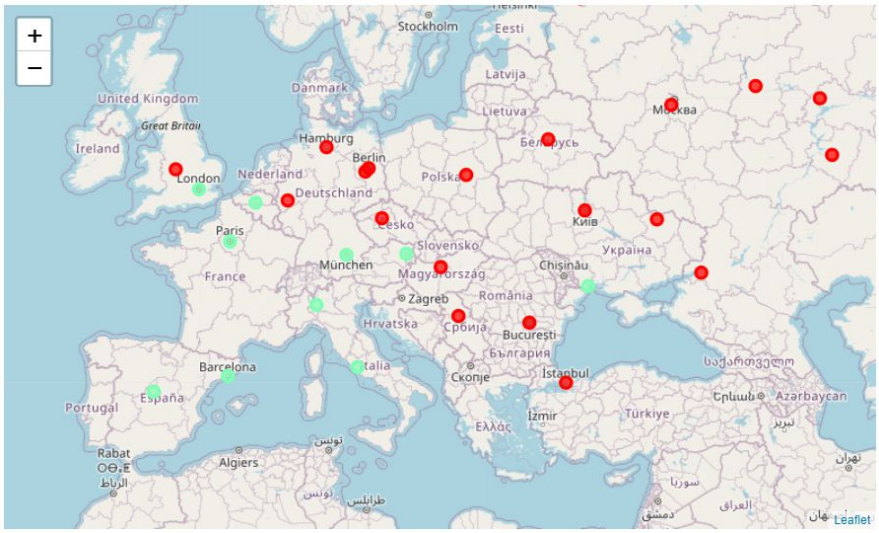
In the fourth paragraph we saw our dataframe with all the venues, its location and its category. Our project is based on the coffee shops and café bars so it is important to see how many of them are really across all the biggest cities. So I created a bar chart in which you can see that coffee shops are the most common thing that you can find in the city center. The second most common thing is hotel, third are parks, and again, in the fourth place we have café-s. So, based on the parameters from all the cities, we can see that you can easily find a café bar in most of the biggest European cities.



7. SHOW A MAP WITH CLUSTERED GROUPS WITH SIMILAR CONTENTS

I created a function that calculated the of the most common venues in every city from the list. After the program made the list for every city, I used the KMeans for clustering the list and to split our cities in two different groups. We can see the spots in our Map. You remember that points were all blue? Now they are red end turquoise, with each color representing a different group of cities.

The red color shows us cities that are full of café bars and similar contents in the city center and the turquoise one is showing the cities that have different things in offer in the center. In the eight paragraph you can see the two groups.



8. SHOWING THE DATA FROM THE TWO GROUPS (RED AND TURQUOISE) WITH DETAILED LIST OF MOST COMMON VENUE CATEGORY FOR EACH ONE

As I explained in the previous paragraph, here are the two dataframes clustered by the things they offer in the city center. You can show the city, its population, coordinates and the first five of most common venue categories you can find in the city center. The first one, red one, has got a big amount of coffee shops, café bars and similar things to offer you in the city center. The second one, turquoise has got a small amount of similar things to offer so it is obvious that we'll use it for the further examination.

RED

	City	Population	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Istanbul	15519267	41.009633	28.965165	0	Café	Historic Site	Coffee Shop	Hotel	History Museum
1	Moscow	12615279	55.479205	37.327330	0	Park	Supermarket	Café	Restaurant	Convenience Store
3	Saint Petersburg	5383890	59.960674	30.158655	0	Bakery	Salon / Barbershop	Park	Hookah Bar	Beach
4	Berlin	3748148	52.517037	13.388860	0	Coffee Shop	Park	Bookstore	Ice Cream Shop	Concert Hall
6	Kiev	2950800	50.450034	30.524136	0	Coffee Shop	Park	Tatar Restaurant	Wine Bar	Dessert Shop
9	Bucharest	2106144	44.436141	26.102720	0	Coffee Shop	Dessert Shop	Burger Joint	Park	Hotel
10	Minsk	1982444	53.902334	27.561879	0	Coffee Shop	Bar	Restaurant	Cocktail Bar	Park
11	Hamburg	1930996	53.543764	10.009913	0	Café	Park	Coffee Shop	Pizza Place	Ice Cream Shop
13	Warsaw	1802237	52.233717	21.071411	0	Café	Park	Dessert Shop	Sushi Restaurant	Cocktail Bar
14	Budapest	1768073	47.498382	19.040471	0	Coffee Shop	Bakery	Hotel	Ice Cream Shop	Park
17	Kharkiv	1451132	49.990279	36.230389	0	Coffee Shop	Gym / Fitness Center	Park	Yoga Studio	Italian Restaurant
19	Prague	1308632	50.087465	14.421254	0	Café	Park	Burger Joint	Ice Cream Shop	Cocktail Bar
20	Nizhny Novgorod	1259013	56.328571	44.003506	0	Coffee Shop	Café	Hotel	Scenic Lookout	Hookah Bar
21	Kazan	1243500	55.782355	49.124227	0	Coffee Shop	Boutique	Park	Restaurant	Hotel
23	Birmingham	1223000	52.479699	-1.902691	0	Pub	Bar	Indian Restaurant	Coffee Shop	Hotel
25	Samara	1170910	53.198627	50.113987	0	Coffee Shop	Beach	Pedestrian Plaza	Hotel	Café
26	Belgrade	1166763	44.817813	20.456897	0	Park	Coffee Shop	Ice Cream Shop	Cocktail Bar	Bar
27	Ufa	1121429	52.387061	13.122812	0	German Restaurant	Supermarket	Café	Park	Historic Site
28	Rostov-on-Don	1119875	47.221386	39.711420	0	Restaurant	Hotel	Café	Park	Coffee Shop
29	Cologne	1085664	50.938361	6.959974	0	Café	Park	Coffee Shop	Italian Restaurant	Ice Cream Shop

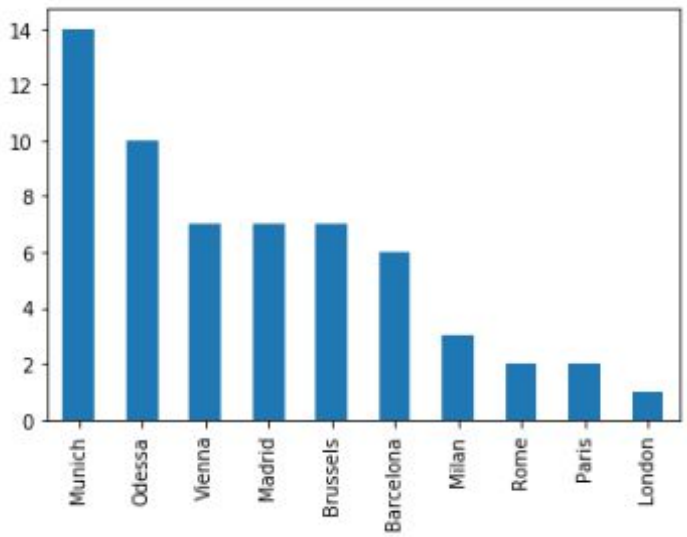
TURQUOISE

	City	Population	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
2	London	9126366	51.507322	-0.127647	2	Hotel	Cocktail Bar	Theater	Clothing Store	Hotel Bar
5	Madrid	3223334	40.416705	-3.703582	2	Plaza	Hotel	Restaurant	Tapas Restaurant	Spanish Restaurant
7	Rome	2844750	41.893320	12.482932	2	Ice Cream Shop	Pizza Place	Italian Restaurant	Plaza	Sandwich Place
8	Paris	2140526	48.856697	2.351462	2	Plaza	Wine Bar	French Restaurant	Bookstore	Italian Restaurant
12	Vienna	1899055	48.208354	16.372504	2	Hotel	Plaza	Ice Cream Shop	Austrian Restaurant	Park
15	Barcelona	1636762	41.382894	2.177432	2	Hotel	Park	Plaza	Wine Bar	Bakery
16	Munich	1471508	48.137108	11.575382	2	Café	Plaza	Ice Cream Shop	Beach	Hotel
18	Milan	1389834	45.466800	9.190500	2	Hotel	Boutique	Monument / Landmark	Plaza	Park
24	Brussels	1211035	50.846557	4.351697	2	Bar	Hotel	Concert Hall	Coffee Shop	Plaza
34	Odessa	1011494	46.487319	30.739278	2	Hotel	Coffee Shop	Plaza	Park	Café

9. WE CAN SEE THE LIST WITH THE NUMBER OF COFFEE SHOPS AND CAFÉ BARS FROM EACH CITY IN THE TARGET GROUP AS WELL AS THE BAR CHART FOR THE LIST

Finally, we got the data we wanted. Based on all of the research we did, here is the final data about the amount of contents similar to café bars in the strict center of each of the biggest cities in Europe. We can see in the table, as well as on the bar chart, that the city of Milan has got three contents similar to café bars in the city center, Paris and Rome two, and London has got only one café bar in the city center.

	City
Munich	14
Odessa	10
Vienna	7
Madrid	7
Brussels	7
Barcelona	6
Milan	3
Rome	2
Paris	2
London	1



10. CONCLUSION

We found the thing that we wanted. It's the city that's got the least contents similar to café bars in the city center – London. London is, based on the population, the third biggest city in Europe and I'm sure that the most of the 9 million people there would like to have more café bars in the center. That concludes that London is the most perfect city to start a café bar from all the biggest cities in Europe.