“Fashion Clothing Store”

Choosing the best district in Zagreb, (Croatia)

Index

1. Introduction

1. 1. ***Background***

Zagreb is the capital and the largest city of Croatia. It is located in the northwest of the country, along the Sava river, at the southern slopes of the Medvednica mountain. Zagreb lies at an elevation of approximately 122 m (400 ft) above sea level. The estimated population of the city in 2018 was 802 762, an increase of 2,8% since 2007.The population of the Zagreb urban agglomeration is about 1.2 million, approximately a quarter of the total population of Croatia. Zagreb is divided into 17 districts, that have different surroundings and things to offer. When starting a new business it would be good to know, based on type of business we are opening, which district to choose for our work.

1. 2. ***Problem***

Based on demographics data, and on machine learning models, we might be able to determine which district would be best choice for our new business. Which in this case happens to be a „Fashion store". In which most of artifacts are going to be hand made bags, wallets, dresses and similar artifacts. As with any new business we need to look for things that are important elements in order for our business not to fail. In this case main elements are going to be: demographics, foot traffic, competition, proximity and costs.

1. 3. ***Interest***

The individuals who might be interested in this, are fashion designers, and others in similar type of work who are interested in opening new venue in order to sell and showcase their own unique creations. The same mode could be applied to other types of businesses such as: traditional clothing, antique book stores, antique shops, jewelry shops and so on.

2. Data

2. 1. ***Data sources***

For this project finding all of data was hard task, since most of Croatian datasets are not available online, even so far that finding GeoJson file of Zagreb's districts turned out to be impossible. Even so, some data has been found at it was mostly at these sources:

https://en.wikipedia.org/wiki/Districts_of_Zagreb
https://www.zagreb.hr/statistika/30
https://zagreb.maps.arcgis.com/apps/MapSeries/index.html?appid=73e33727fad74298a9df9859a2331cd6

2. 2. ***Data cleaning and organization***

Most of data was collected from the internet and put into spread sheet document. Some of it like data from maps had to be literally written down by hand and added to the dataset. Among data found online most of it was useless for this project. For example weather in Zagreb, earthquakes and so on, but with material that was at hand the basic dataset was made for this project.

[3]:

	district_name	lat	lng	size_km2	pop	pop_den	host	hot	priva	unem_men	unem_wom	unem_tot	house_prices
0	Brezovica	45.715581	15.920606	127	1203	95	0	0	8	111	130	241	973
1	Črnomerec	45.832531	15.938308	24	38546	1606	2	1	128	512	514	1026	1569
2	Donja Dubrava	45.826214	16.051219	11	36363	3306	1	1	17	609	713	1322	1259
3	Donji Grad	45.808333	15.976111	3	37024	12341	15	14	752	517	547	1064	1901
4	Gornja Dubrava	45.841125	16.055661	40	61841	1546	2	0	29	904	1022	1926	1250
5	Gornji Grad–Medveščak	45.828247	15.979356	10	30962	3096	11	6	411	324	427	751	2062
6	Maksimir	45.832156	16.019858	14	48902	3493	3	3	147	604	701	1305	1740
7	Novi Zagreb – istok	45.773764	15.990944	17	59055	3474	0	2	49	821	876	1697	1253
8	Novi Zagreb – zapad	45.768653	15.952889	63	58103	922	2	7	90	799	1007	1806	1382
9	Peščenica – Žitnjak	45.793442	16.048461	35	56487	1614	1	2	61	993	1130	2123	1362

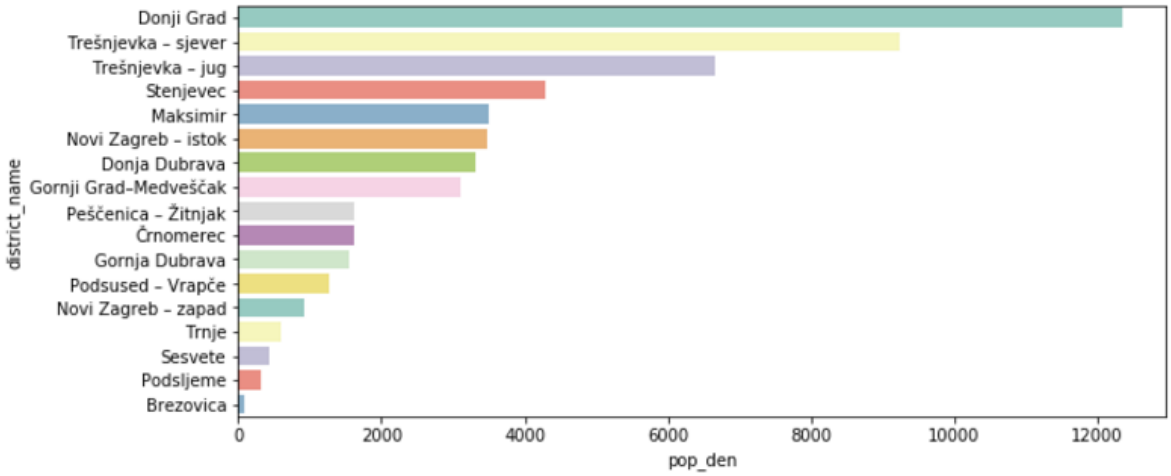
- district_name: *Name of the district*
- lat: *District's latitude*
- lng: *District longitude*
- size_km2: *District size in square kilometers*
- pop: *Number of population per district*
- pop_den: *Population density*
- host: *Number of hostels*
- hot: *Number of hotels*
- priva: *Number of private accommodation*
- unem_men: *Number of unemployed men*
- unem_women: *Number of unemployed women*
- unem_tot: Total *unemployment number*
- house_prices: House and venue buying prices. Euro per square meter.

3. Methodology

3.1 Exploratory data analysis

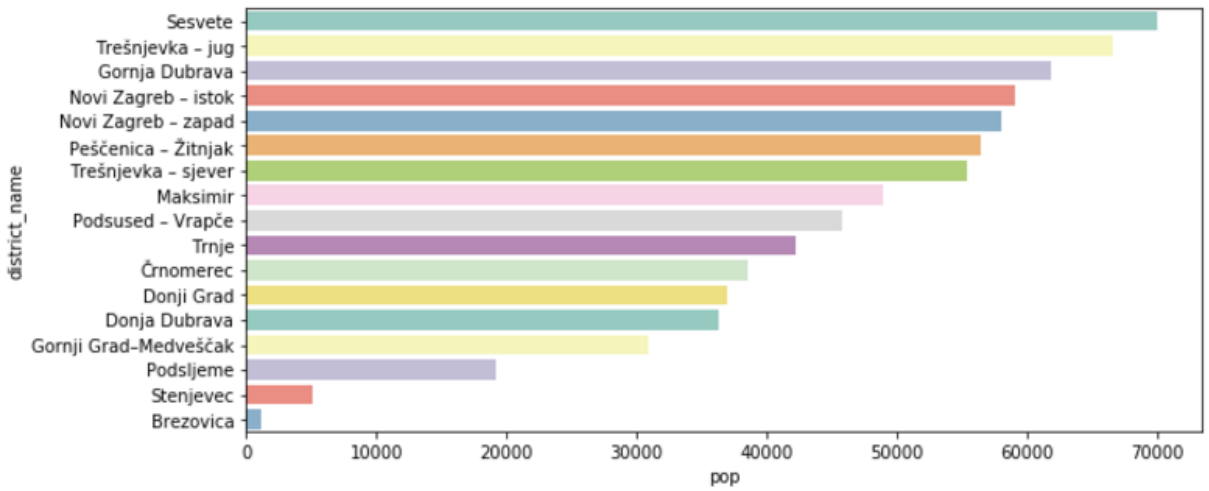
3.1.a Population and population density:

For this type of business, we are looking for places with large foot traffic, and one of the indicators of it can be, population density. The more the district is dense the bigger the concentration of people, and less of vehicle traffic is around.



Population density by the district

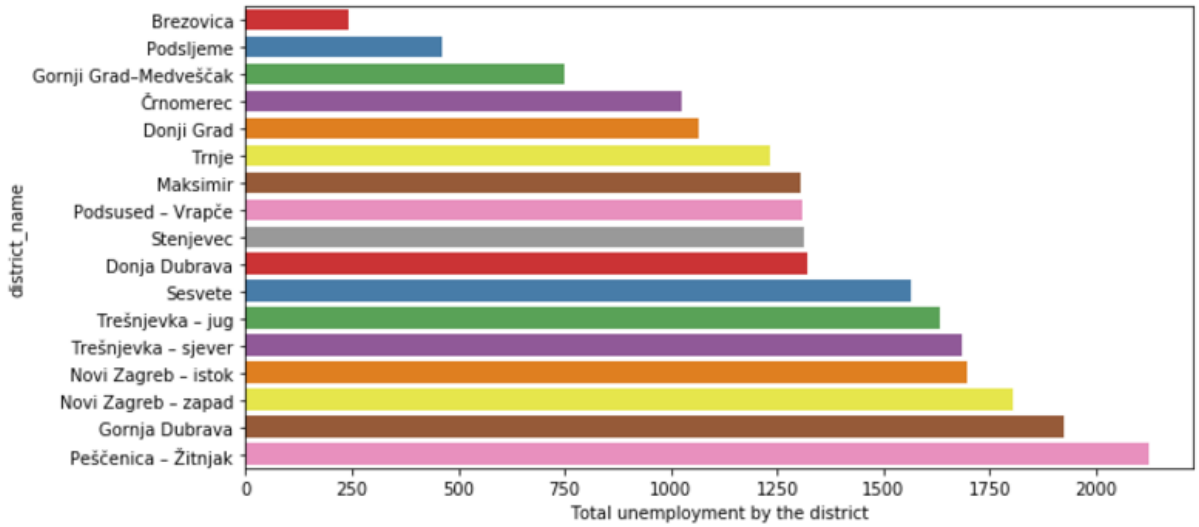
As we can see from char above, the districts with largest population density are: Donji Grad, Trešnjevka – sjever, Trešnjevka – jug, Stenjevec and Maksimir. Even tho those places don't hold largest number of population as shown below:



Population number by the district

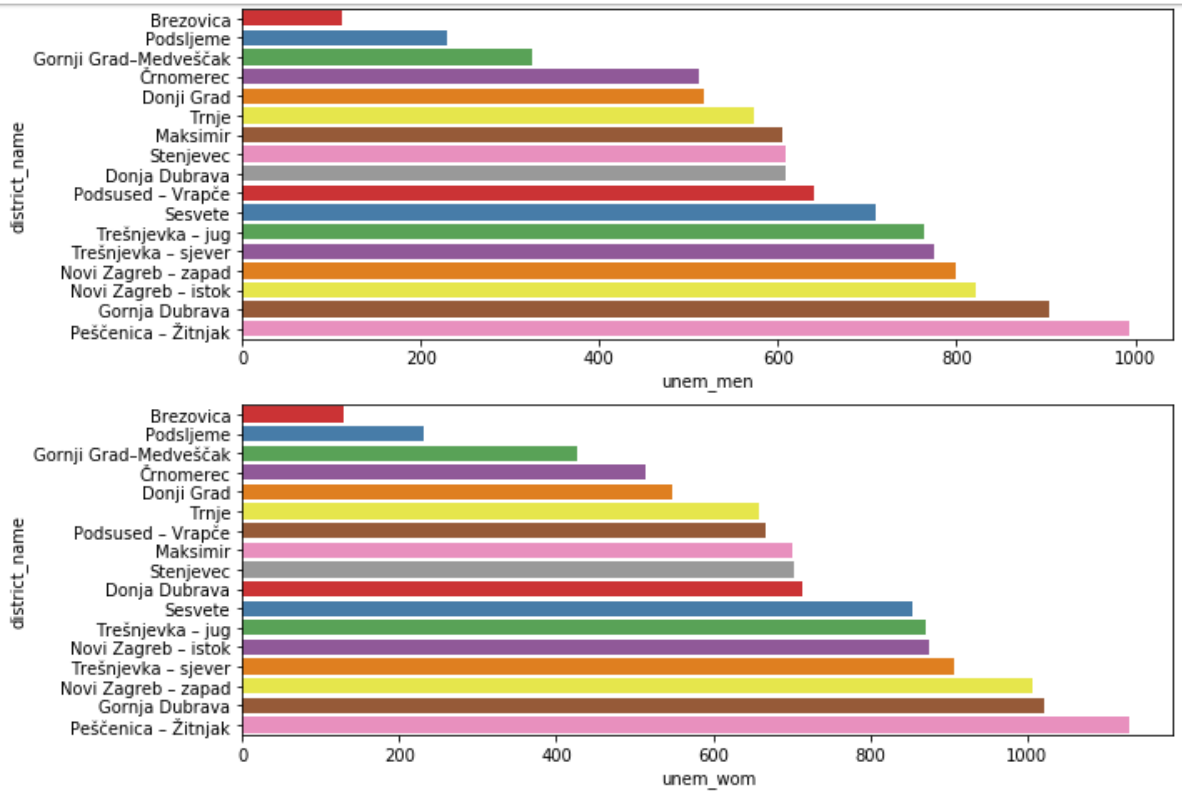
3. 1. b. Unemployment

One of the important aspects for the type of business we are planning to open is purchasing ability of surrounding population. So therefore we are looking for districts with least unemployment, and therefore higher purchasing ability.



Total unemployment by the district

Also let's look at the unemployment by the gender:



Unemployment by the gender for each district

As we can see in of top 5 districts, unemployment is lowest for women in district Donji Grad, which is good indicator and good news for the type of business we are trying to open.

3. 1. c. Using Foursquare API to explore districts

In order to explore district further Foursquare API service has been used. Where we obtained coordinates for each of the districts, the venues in them have been explore, grouped collected in datasets. One of the main indicators of diversity, and large number of people would be a number of unique venues in each district.

[16]:

District	District Latitude	District Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Donji Grad	100	100	100	100	100	100
Trnje	70	70	70	70	70	70
Trešnjevka – jug	62	62	62	62	62	62
Stenjevec	54	54	54	54	54	54
Novi Zagreb – istok	53	53	53	53	53	53
Novi Zagreb – zapad	48	48	48	48	48	48
Trešnjevka – sjever	47	47	47	47	47	47
Donja Dubrava	36	36	36	36	36	36
Gornji Grad–Medveščak	35	35	35	35	35	35
Peščenica – Žitnjak	16	16	16	16	16	16
Gornja Dubrava	15	15	15	15	15	15
Podsljeme	13	13	13	13	13	13
Podsused – Vrapče	12	12	12	12	12	12
Maksimir	11	11	11	11	11	11
Sesvete	10	10	10	10	10	10
Črnomerec	4	4	4	4	4	4

Number of unique venues by the district

As we can see in set above Donji Grad again comes up at the top of given districts, and it's a good indicator that is might be the most vibrant and diverse place in the whole city. With districts Trnje, Trešnjevka – jug and Stenjevec following it.

Further more lets look into top venues by the each district :

[143]:

	District	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	Donja Dubrava	Café	Pizza Place	Bar	Supermarket	Light Rail Station	Gym / Fitness Center	Coffee Shop	Grocery Store	Pharmacy	Market
1	Donji Grad	Plaza	Bar	Café	Restaurant	Dessert Shop	Hotel	Hostel	Theater	Italian Restaurant	Mediterranean Restaurant
2	Gornja Dubrava	Grocery Store	Pharmacy	Bar	Supermarket	Coffee Shop	Clothing Store	Bakery	Women's Store	Café	Shopping Mall
3	Gornji Grad–Medveščak	Dessert Shop	Café	Bar	Supermarket	Pizza Place	Nightclub	Lounge	Pub	Chinese Restaurant	Multiplex
4	Maksimir	Café	Lake	Supermarket	Grocery Store	Soccer Field	Restaurant	Bus Station	Farmers Market	Electronics Store	Eastern European Restaurant
5	Novi Zagreb – istok	Café	Pizza Place	Supermarket	BBQ Joint	Pet Store	Smoke Shop	Bakery	Grocery Store	Bar	Italian Restaurant
6	Novi Zagreb – zapad	Café	Grocery Store	Bakery	Clothing Store	Liquor Store	Pizza Place	Dessert Shop	Bus Station	Mobile Phone Shop	Bar
7	Peščenica – Žitnjak	Fast Food Restaurant	Hardware Store	Bar	Multiplex	Chinese Restaurant	Drugstore	Café	Bus Station	Food	Shopping Mall
8	Podsljeme	Restaurant	Light Rail Station	Café	Market	Fruit & Vegetable Store	Grocery Store	Travel & Transport	Trail	Mountain	Electronics Store
9	Podsused – Vrapče	Supermarket	Chinese Restaurant	Restaurant	Café	Soccer Field	Sports Club	Bar	BBQ Joint	Park	Pizza Place

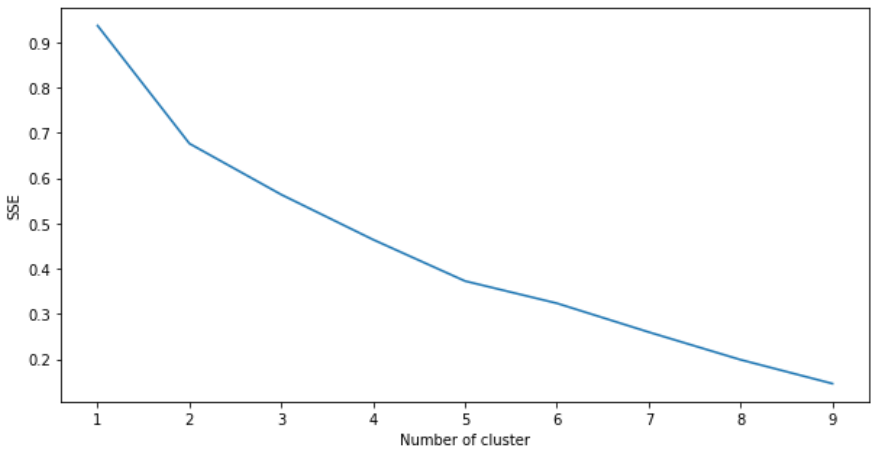
As we can see Donji Grad again comes to our interest because it has most plazas, hotels, hostels, and theaters. All places that are good proximity for the type of business we are planning to open.

1	Donji Grad	Plaza	Bar	Café	Restaurant	Dessert Shop	Hotel	Hostel	Theater	Italian Restaurant	Mediterranean Restaurant

3. 2. k – Means clustering

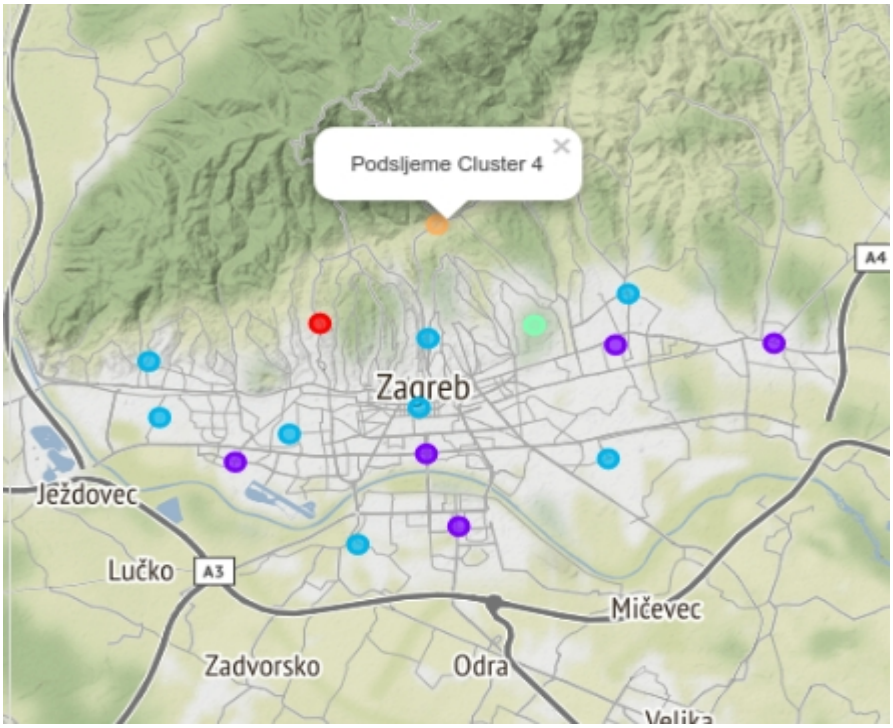
In order to find similarity between districts, an unsupervised, partition based, machine learning algorithm „k – Means clustering" has been used. It's an algorithm that is being used for segmentation based on similarity of given examples in the dataset. It works in such way that examples within a cluster will be very similar, and at the same time examples between clusters will be very different.

One of the first steps in preparation for the algorithm was, to determine how many clusters are to be used. In other words, into how many cluster districts are city districts to be divided in. The method for determining such number is called Elbow – point:



As we can see in the graph above, the best number of clusters for this assignment is k = 5 .

After k has been determined, and machine learning model has been fitted with proper dataset, the algorithm produced a group of 5 clusters of Zagreb's districts. And for the representation of clusters on the map, the Folium library has been used.



Districts of Zagreb, organized into 5 different clusters

As we can see from the map, districts in the center and towards south – west (Cluster 2) are more more similar that ones towards east (Cluster 1) . And at the north we have 3 completely unique districts (Cluster 0, Cluster 3, Cluster 4).

Further more, lets look at the names of districts within each cluster:

[64]:

	Clusters	district_name
1	0	Črnomerec
2	1	Donja Dubrava
7	1	Novi Zagreb – istok
12	1	Sesvete
14	1	Trešnjevka – jug
16	1	Trnje
3	2	Donji Grad
4	2	Gornja Dubrava
5	2	Gornji Grad–Medveščak
8	2	Novi Zagreb – zapad
9	2	Peščenica – Žitnjak
11	2	Podsused – Vrapče
13	2	Stenjevec
15	2	Trešnjevka – sjever
6	3	Maksimir
10	4	Podsljeme

We can see that districts similar to district of our interest Donji Grad are:
Gornja Dubrava, Gornji Grad – Medvedščak, Novi Zagreb – zapad, Peščenica – Žitnjak, Podsused – Vrapče and Stenjevac. These places can be used as an alternative locations to explore if for some reason our targeted district based on our previous analysis falls out of our reach.

3. 3. Venues of interest

Lets look into venues of interest in the proximity of Donji Grad. Since we want our business to be accessible to local people and tourists as well, lets look into cultural and artistic venues in the area.

3. 3. a. Plazas:

[96]:

	District	District Latitude	District Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
41	Donji Grad	45.808333	15.976111	Trg Nikole Šubića Zrinskog | Zrinjevac	45.810244	15.978109	Plaza
42	Donji Grad	45.808333	15.976111	Trg kralja Tomislava	45.806442	15.978627	Plaza
59	Donji Grad	45.808333	15.976111	Strossmayerov trg	45.808578	15.978560	Plaza
62	Donji Grad	45.808333	15.976111	Cvjetni trg	45.812251	15.974208	Plaza
69	Donji Grad	45.808333	15.976111	Trg bana Josipa Jelačića	45.813019	15.977229	Plaza
103	Donji Grad	45.808333	15.976111	Trg Europe	45.813235	15.979996	Plaza
114	Donji Grad	45.808333	15.976111	Strossmartre	45.814536	15.972880	Pedestrian Plaza
127	Donji Grad	45.808333	15.976111	Trg Republike Hrvatske	45.809648	15.969468	Plaza
138	Donji Grad	45.808333	15.976111	Rooseveltov trg	45.808075	15.968083	Plaza

Donji Grad has nine plazas in proximity which is great indicator of foot traffic and large number of tourists and other venues as well.

3. 3. b. Museums:

	District	District Latitude	District Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
107	Donji Grad	45.808333	15.976111	Muzej za umjetnost i obrt (MUO)	45.809333	15.968855	Art Museum
126	Donji Grad	45.808333	15.976111	Galerija Klovićevi dvori	45.815022	15.974649	Museum
139	Donji Grad	45.808333	15.976111	Muzej prekinutih veza | Museum of Broken Relat...	45.814900	15.973476	Museum

We can see there are two Museums and one Art Museum in our proximity.

3. 3. c. Theaters:

	District	District Latitude	District Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
54	Donji Grad	45.808333	15.976111	Zagrebačko kazalište mladih (ZKM)	45.811199	15.976042	Theater
60	Donji Grad	45.808333	15.976111	Hrvatsko narodno kazalište u Zagrebu (HNK)	45.809444	15.969960	Theater
77	Donji Grad	45.808333	15.976111	Zagrebačko kazalište lutaka (ZKL)	45.807140	15.977001	Theater
86	Donji Grad	45.808333	15.976111	Kino Europa	45.811830	15.973250	Indie Movie Theater
112	Donji Grad	45.808333	15.976111	Kazalište Kerempuh	45.812323	15.970908	Theater

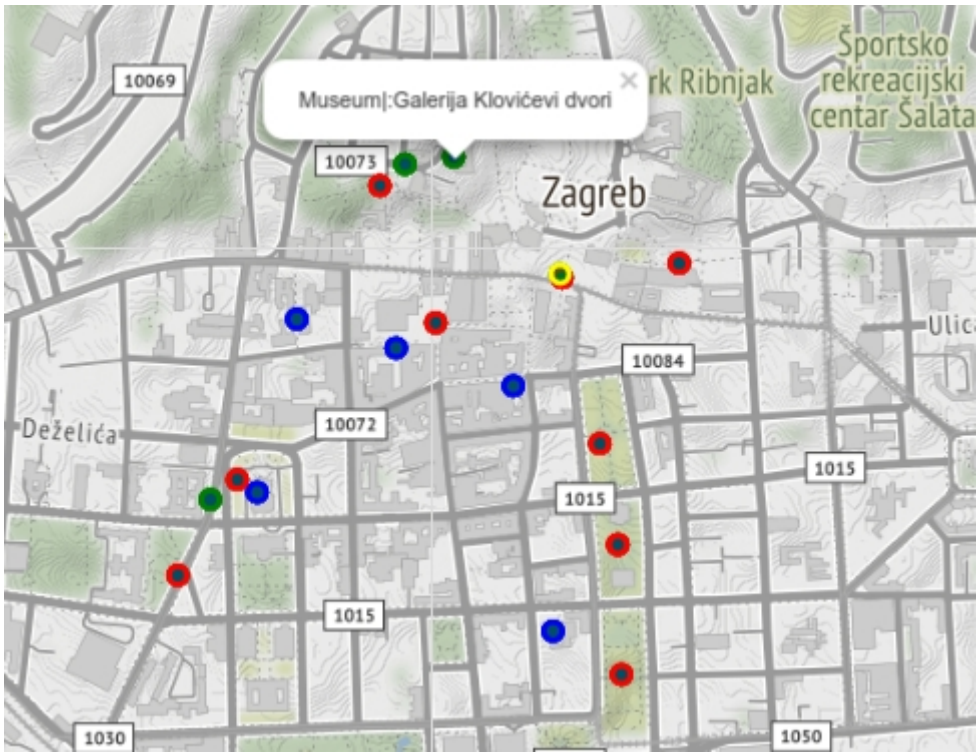
There are five theaters in our proximity, Croatian national theater included (HNK).

3. 3. d. Historic sites:

	District	District Latitude	District Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
88	Donji Grad	45.808333	15.976111	Ban Jelacic square	45.81308	15.977171	Historic Site

We can see there is one Historic Site in our proximity, which happens to be Ban Jelačić square.

After isolating venue categories of our interest, locations of such venues, together with their name, and category had been plotted on map of Zagreb, using Folium library.



O – Museums , **O** – Theaters , **O** – Historic sites, **O** – Plazas

3. 4. Accommodation

Zagreb is an important tourist center, not only in terms of passengers traveling from the rest of Europe to the Adriatic Sea, but also as a travel destination itself. Since the end of the war, it has attracted close to a million visitors annually, mainly from Austria, Germany and Italy, and in recent years many tourists from far east (South Korea, Japan, China, and last two years, from India). It has become an important tourist destination, not only in Croatia, but considering the whole region of southeastern Europe.

In 2010 more than 600,000 tourists visited the city, with a 10% increase seen in 2011. In 2012 a total of 675 707 tourists visited the city. The record number of tourists visited Zagreb in 2017. – 1.286.087, up 16% compared to the year before, which generated 2.263.758 overnight stays, up 14,8%.

From the dataset built for the project we can easily see the number of hostels, hotels and private accommodation spaces, based on each district.

	district_name	host	hot	priva
3	Donji Grad	15	14	752
5	Gornji Grad–Medveščak	11	6	411
16	Trnje	9	12	185
15	Trešnjevka – sjever	4	4	161
6	Maksimir	3	3	147

As we can see the district Donji Grad again comes up at the top, with 15 hostels, 14 hotels and 752 privately owned places, available for tourists and visitors to rent.

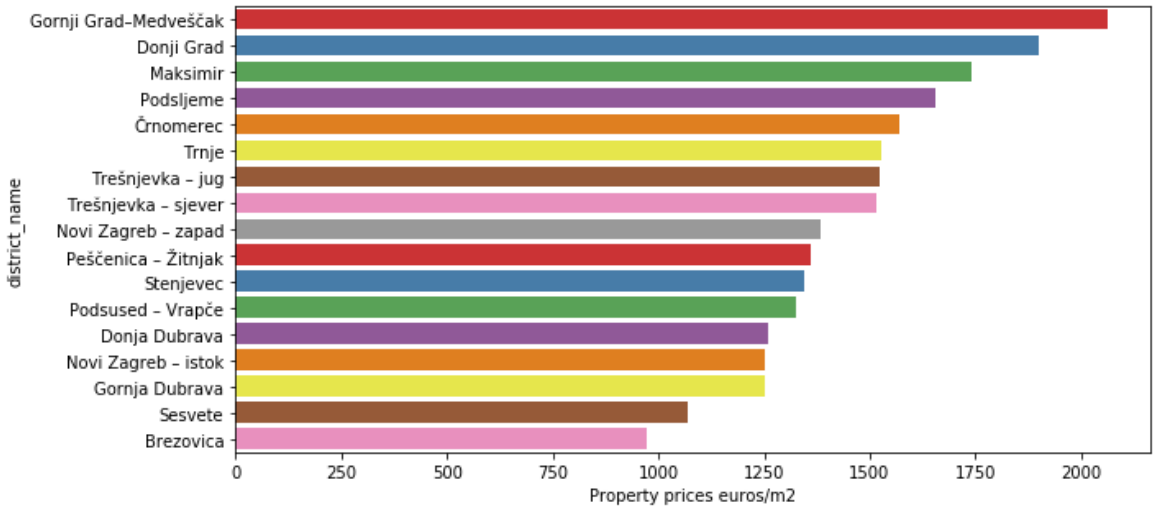
3. 5. Prices

One of the important aspects of any enterprise are property prices. Prices of properties had been collected during data gathering process, and they show value for each district. The prices shown below are prince in euros, per square meter.

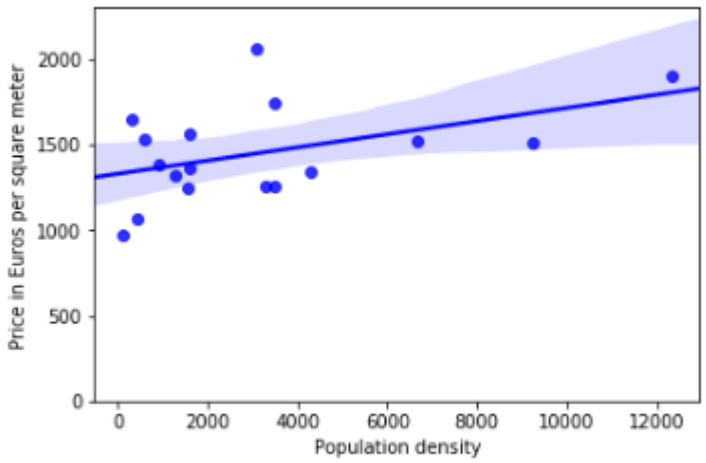
[52]:

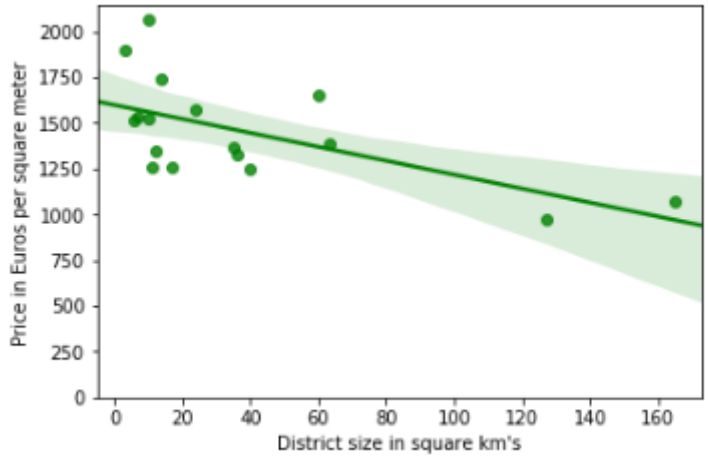
	district_name	house_prices
5	Gornji Grad–Medveščak	2062
3	Donji Grad	1901
6	Maksimir	1740
10	Podsljeme	1654
1	Črnomerec	1569
16	Trnje	1528
14	Trešnjevka – jug	1523
15	Trešnjevka – sjever	1517
8	Novi Zagreb – zapad	1382
9	Peščenica – Žitnjak	1362
13	Stenjevec	1346
11	Podsused – Vrapče	1326
2	Donja Dubrava	1259
7	Novi Zagreb – istok	1253
4	Gornja Dubrava	1250
12	Sesvete	1069
0	Brezovica	973

We can see from the set that our target district is in second place with price of 1901 euros per square meter.



Such high level of price is due to location of the district, large number of venues, and cultural landmarks. But also we can see that there is a slight positive correlation between population density of districts and price of land. At the same time, physical size of district seems to be slightly negatively correlated with the price of properties. Makes sense, since bigger the district the more room is left to spare.





4. Conclusion

During this project, we have built a dataset, analyzed it's values, found different venues withing districts, grouped them using k – Means and visualized the values on the map. From all the give data at this time, one can conclude that best district for opening „Fashion Store” happens to be Donji Grad. The whole area has features that are really desirable for given type of business...

- Highest population density: 12 341
- Decent number of population: 37 024
- Lower unemployment rate for women: 547 women unemployed
- Highest number of unique venue categories: 100
- It surrounded by district that are similar to it, therefore the business plan can go beyond the limits of the district itself.
- It is surrounded by museums, theaters, galleries and historic sites, which are great elements for attracting tourist foot traffic.
- Largest number of hostels, hotels and units of private accommodation: 15 hostels, 14 hotels and 752 privately owned places.
- Only downside is rather high rate of prices for properties, but considering other aspects and elements of the district, it could be a decent and good investment.

5. Future directions

For the future direction, a larger dataset needs to be built so more models and predictions can be made. Unfortunately, for the scope of this assignment such endeavor would take months to collect, and it wouldn't be finished in time. Also it would be advisable to do the same exploration of surrounding districts like: Gornji Grad – Medvedščak, Stenjevec and Trešnjevka – sjever. One interesting district is also Trnje, that falls into cluster number 1; and it borders with district of Donji Grad. It is as well pretty much close to all cultural places, and has the second largest number of unique venues. This area would really need further exploration.