Exploring boroughs of Oslo

Julia Vasilyeva

July 22, 2020

Table of Contents

Introduction

Background

Oslo is the capital of Norway, a global city which was ranked "Beta World City" in studies carried out by the Globalization and World Cities Study Group and Network in 2008.

Source: https://en.wikipedia.org/wiki/Oslo#/media/File:Norway_Counties_Oslo_Position.svg

Oslo is not only the capital of Norway, but also economic and governmental centre of Norway. The city is a hub of Norwegian trade, banking, industry and shipping. According to the official Norwegian job portal 27,77% of all job opportunities in the country belong to Olso:

Source: https://www.finn.no/job

This make it attractive especially for professionals aiming to build carrier and building experience. That's why relocation to Oslo is often a good option.

Oslo was ranked number one in terms of quality of life among European large cities in the European Cities of the Future 2012 report by fDi magazine. However, it’s a very expensive city as well. A survey conducted by ECA International in 2011 placed Oslo as the second most expensive city in the world for living expenses after Tokyo.

Problem

Finding a suitable place to live is not easy. Not only price and quality of housing has to be taken into account, but also neighborhood. Very popular and wide-used Norwegian resource finn.no make it easy to search on many different parameters, such as area, price, year of building, energy use and borough. It's possible to use the address and then google maps for looking into the surroundings. However, it's not enough for solving the problem of searching for the neighborhood, which suits the personal requirements.

This project aiming to investigate the boroughs and cluster it, so that it would be easier to find the borough based on personal preferences .

Data wrangling

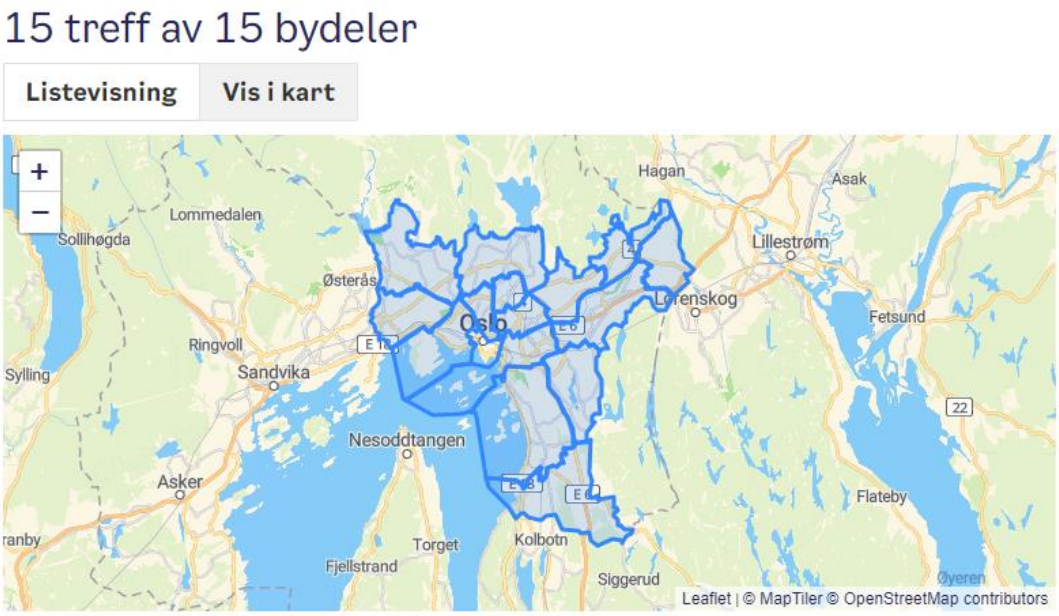
Data wrangling included pre-processing, dealing with missing or wrong data and data formatting. For solving the problem it was necessary to use several sources and approaches for collecting and gathering the data.

List of Oslo boroughs

Wikipedia page evaluates following 15 boroughs:

Boroughs	Inhabitants (2020)[26]	Area in km^2	number
Alna	49,801	13.7	12
Bjerke	33,422	7.7	9
Frogner	59,269	8.3	5
Gamle Oslo	58,671	7.5	1
Grorud	27,707	8.2	10
Grünerløkka	62,423	4.8	2
Nordre Aker	52,327	13.6	8
Nordstrand	52,459	16.9	14
Sagene	45,089	3.1	3
St. Hanshaugen	38,945	3.6	4
Stovner	33,316	8.2	11
Søndre Nordstrand	39,066	18.4	15
Ullern	34,569	9	6
Vestre Aker	50,157	16.6	7
Østensjø	50,806	12.2	13
Overall	**688,027**	**151.8**	

Official Oslo municipality website also shows 15 boroughs:



So the list of 15 boroughs was considered to use for the project.

Location Data

Location data is data describing places and venues, such as their geographical location, category, working hours, full address, and so on. For this project **Foursquare** location data was be used.

For getting the coordinates of boroughs **Geopy** geocoding web services was used. Unfortunately not all data were correct. This happened because there are several locations in Norway having this name. Ti fix the data Google Maps and Nearby functions were used to get the correct coordinates which were fixed manually directly in the resulting dataframe.

Housing prices

After a research made in the Internett several information sources with the information about property prices were found:

- Business Data Platform: statista.com
- Website listing all properties for sell in Norway: finn.no
- Official Oslo municipality data: oslo.kommune.no

The last one was evaluated as the most trusted and structured data about average prices for the different areas of Oslo. After applying filters to the data in the website following table was produced:

År	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Boligtype	Boligtype i alt	Boligtype i alt	Boligtype i alt	Boligtype i alt	Boligtype i alt	Boligtype i alt	Boligtype i alt	Boligtype i alt	Boligtype i alt	Boligtype i alt	Boligtype i alt	Boligtype i alt	Boligtype i alt	Boligtype i alt	Boligtype i alt
Geografi															
Oslo i alt	24 934	28 260	33 183	36 468	34 359	35 768	38 337	42 615	46 432	48 038	49 340	55 475	64 706	69 293	69 532
Bydel Gamle Oslo	26 256	29 285	35 143	38 161	34 845	35 452	37 584	42 045	46 181	48 421	50 255	57 329	68 730	73 181	74 605
Bydel Grünerløkka	27 416	30 813	36 628	39 312	36 440	37 314	40 084	44 456	48 990	51 078	53 173	60 548	71 505	75 909	76 413
Bydel Sagene	28 350	31 211	37 054	40 293	37 390	38 812	41 273	46 450	51 625	53 733	55 362	63 422	74 206	79 348	80 766
Bydel St Hanshaugen	29 801	34 420	40 496	44 302	40 942	41 759	44 969	50 247	54 898	57 328	58 065	65 631	76 307	82 436	81 649
Bydel Frogner	31 258	37 014	42 561	46 223	44 180	45 150	49 273	54 552	58 995	61 100	61 839	68 158	79 721	87 326	87 923
Bydel Ullern	28 214	31 688	36 315	41 137	40 508	40 156	44 085	48 314	52 006	52 393	53 452	58 411	66 381	72 944	74 109
Bydel Vestre Aker	26 334	29 924	33 678	37 784	36 709	36 779	40 989	44 686	47 770	48 269	48 787	53 137	60 682	66 065	67 186
Bydel Nordre Aker	28 021	31 814	36 184	39 999	39 323	39 529	43 725	49 154	53 039	54 668	55 117	59 917	70 426	74 778	78 073
Bydel Bjerke	21 832	23 346	28 216	31 308	29 765	29 902	32 512	35 720	39 278	40 481	41 745	46 945	53 847	59 001	58 834
Bydel Grorud	18 393	20 177	23 417	25 990	24 437	25 098	27 079	29 464	33 027	34 412	35 157	40 138	46 629	50 248	49 224
Bydel Stovner	17 657	19 008	22 075	25 057	23 666	23 599	24 821	27 583	30 364	31 420	32 063	35 510	40 255	44 165	42 394
Bydel Alna	18 751	20 490	24 067	26 859	25 429	25 798	27 731	30 837	33 679	35 636	35 895	41 021	47 562	51 370	49 042
Bydel Østensjø	21 650	23 237	27 183	30 228	28 619	29 088	32 045	35 174	38 339	40 111	40 856	45 599	54 106	56 946	56 556
Bydel Nordstrand	22 779	25 426	29 457	33 170	31 724	32 558	34 963	38 873	41 822	43 515	44 584	48 959	56 377	61 543	60 350
Bydel Søndre Nordstrand	16 137	17 471	20 503	23 463	22 866	22 729	24 273	26 468	28 626	30 302	30 204	34 226	39 156	43 009	41 235
Sentrum	33 385	38 922	49 643	51 255	47 175	48 903	49 714	57 192	61 794	59 825	58 354	65 391	81 145	87 539	85 210
Marka	23 078	20 457	34 733	36 228	32 457	32 752	40 226	49 426	38 650	41 385	-	46 265	43 232	52 762	62 388

The data were exported to Excel file from the website.

Location Data

Location data is data describing places and venues, such as their geographical location, category, working hours, full address, and so on. For this project **Foursquare** location data was be used.

For getting the coordinates of boroughs **Geopy** geocoding web services was used. Unfortunately not all data were correct. This happened because there are several locations in Norway having this name. Ti fix the data **Google Maps** and Nearby functions were used to get the correct coordinates which were fixed manually directly in the resulting dataframe.

Methodology

1. Preparing the Notebook to execute the code following **libraries were imported**:
    - *import numpy as np*
    - *import pandas as pd*
    - *import json*
    - *from geopy.geocoders import Nominatim*
    - *from pandas.io.json import json_normalize*
    - *import matplotlib.cm as cm*
    - *import matplotlib.colors as colors*
    - *from sklearn.cluster import KMeans*
    - *!pip install folium*
    - *import folium*
    - *import requests*
    - *import matplotlib as mpl*
    - *import matplotlib.pyplot as plt*
2. First, the dataset with names of the boroughs was created
3. Them coordinates were build using **Geopy** and the boroughs were put to the map using **Folium**
4. Next step was to explore the boroughs using the **Foursquare**. Using developers account access to data was secured.
5. Then top 5 most common venues were analyzed. And based on that new dataframe with top 10 venues for each borough was made.
6. As a next step **k-Means Clustering** was used for grouping the boroughs into groups.
7. Then housing prices per square meter were plotted using **matplotlib**

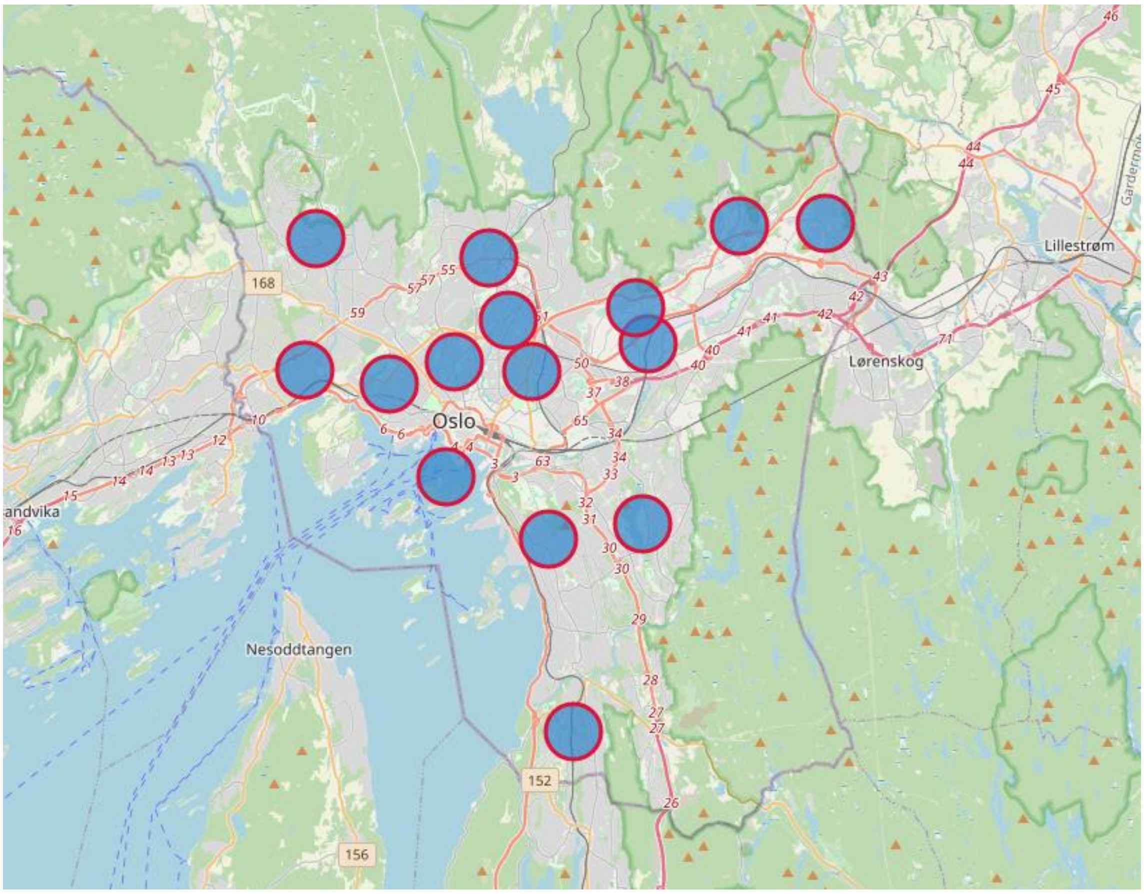
Results

The dataset with coordinates was build using **Geopy**. After visual examination it was noticed that the data for at least one borough were not correct. Nordstrand is a name which in Geopy belong to another Norwegian location. After further investigation it was found out that there are several “Nordstrand” in Norway. So it was decided to find the correct coordinates using **Google maps** and correct dataframe manually. As a result following dataframe was developed:

	Borough	Latitude	Longitude
0	Alna	59.932417	10.835276
1	Bjerke	59.941395	10.829208
2	Frogner	59.922224	10.706649
3	Gamle Oslo	59.899237	10.734767
4	Grorud	59.961424	10.880549
5	Grünerløkka	59.925471	10.777421
6	Nordre Aker	59.953638	10.756412
7	Nordstrand	54.487378	8.865286
8	Sagene	59.938273	10.765849
9	St. Hanshaugen	59.927950	10.738958
10	Stovner	59.962140	10.922823
11	Søndre Nordstrand	59.835944	10.798496
12	Ullern	59.925818	10.665132
13	Vestre Aker	59.958300	10.670319
14	Østensjø	59.887563	10.832748

	Borough	Latitude	Longitude
0	Alna	59.932417	10.835276
1	Bjerke	59.941395	10.829208
2	Frogner	59.922224	10.706649
3	Gamle Oslo	59.899237	10.734767
4	Grorud	59.961424	10.880549
5	Grünerløkka	59.925471	10.777421
6	Nordre Aker	59.953638	10.756412
7	Nordstrand	59.883600	10.785700
8	Sagene	59.938273	10.765849
9	St. Hanshaugen	59.927950	10.738958
10	Stovner	59.962140	10.922823
11	Søndre Nordstrand	59.835944	10.798496
12	Ullern	59.925818	10.665132
13	Vestre Aker	59.958300	10.670319
14	Østensjø	59.887563	10.832748

This allowed to put the boroughs to the map using **Folium**:



Using the **Foursquare** in total the iteration over all boroughs gives us 543 venues in 137 unique categories:

```
In [19]: oslo_venues = getNearbyVenues(names=df_oslo['Borough'],latitudes=df_oslo['Latitude'],longitudes=df_oslo['Longitude'])
         print(oslo_venues.shape)
         oslo_venues.head()
```

```
(543, 7)
```

Out[19]:

	Borough	Borough Latitude	Borough Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Alna	59.932417	10.835276	BROBEKK STORCASH AS	59.932201	10.831548	Grocery Store
1	Alna	59.932417	10.835276	Meny	59.936573	10.826252	Grocery Store
2	Alna	59.932417	10.835276	Serrano's	59.925757	10.846810	Spanish Restaurant
3	Alna	59.932417	10.835276	Stoff Og Stil	59.929213	10.846513	Arts & Crafts Store
4	Alna	59.932417	10.835276	Kiwi	59.937958	10.826332	Grocery Store

```
print('There are {} uniques categories.'.format(len(oslo_venues['Venue Category'].unique())))
```

```
There are 137 uniques categories.
```

After that rows were grouped by boroughs and by taking the mean of the frequency of occurrence of each category:

```
In [26]: oslo_grouped = oslo_arranged.groupby('Borough').mean().reset_index()
         oslo_grouped
```

Out[26]:

	Borough	Advertising Agency	Art Museum	Arts & Crafts Store	Asian Restaurant	Athletics & Sports	Auto Workshop	Automotive Shop	BBQ Joint	Bakery	Bar
0	Alna	0.000000	0.0000	0.043478	0.000000	0.000000	0.043478	0.043478	0.00	0.043478	0.000000
1	Bjerke	0.000000	0.0000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.000000
2	Frogner	0.000000	0.0100	0.000000	0.030000	0.000000	0.000000	0.000000	0.00	0.070000	0.020000
3	Gamle Oslo	0.000000	0.0000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.000000
4	Grorud	0.000000	0.0000	0.000000	0.090909	0.090909	0.000000	0.000000	0.00	0.000000	0.000000
5	Grünerløkka	0.000000	0.0125	0.012500	0.025000	0.000000	0.000000	0.000000	0.00	0.037500	0.037500
6	Nordre Aker	0.045455	0.0000	0.000000	0.090909	0.000000	0.000000	0.000000	0.00	0.136364	0.000000
7	Nordstrand	0.000000	0.0000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.090909	0.000000
8	Sagene	0.000000	0.0000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.048193	0.048193
9	St. Hanshaugen	0.000000	0.0000	0.000000	0.030000	0.000000	0.000000	0.000000	0.01	0.080000	0.030000
10	Stovner	0.000000	0.0000	0.000000	0.000000	0.200000	0.000000	0.000000	0.00	0.000000	0.000000
11	Søndre Nordstrand	0.000000	0.0000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.000000
12	Ullern	0.000000	0.0000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.000000
13	Vestre Aker	0.000000	0.0000	0.000000	0.000000	0.037037	0.000000	0.000000	0.00	0.000000	0.000000
14	Østensjø	0.000000	0.0000	0.000000	0.000000	0.111111	0.000000	0.000000	0.00	0.000000	0.000000

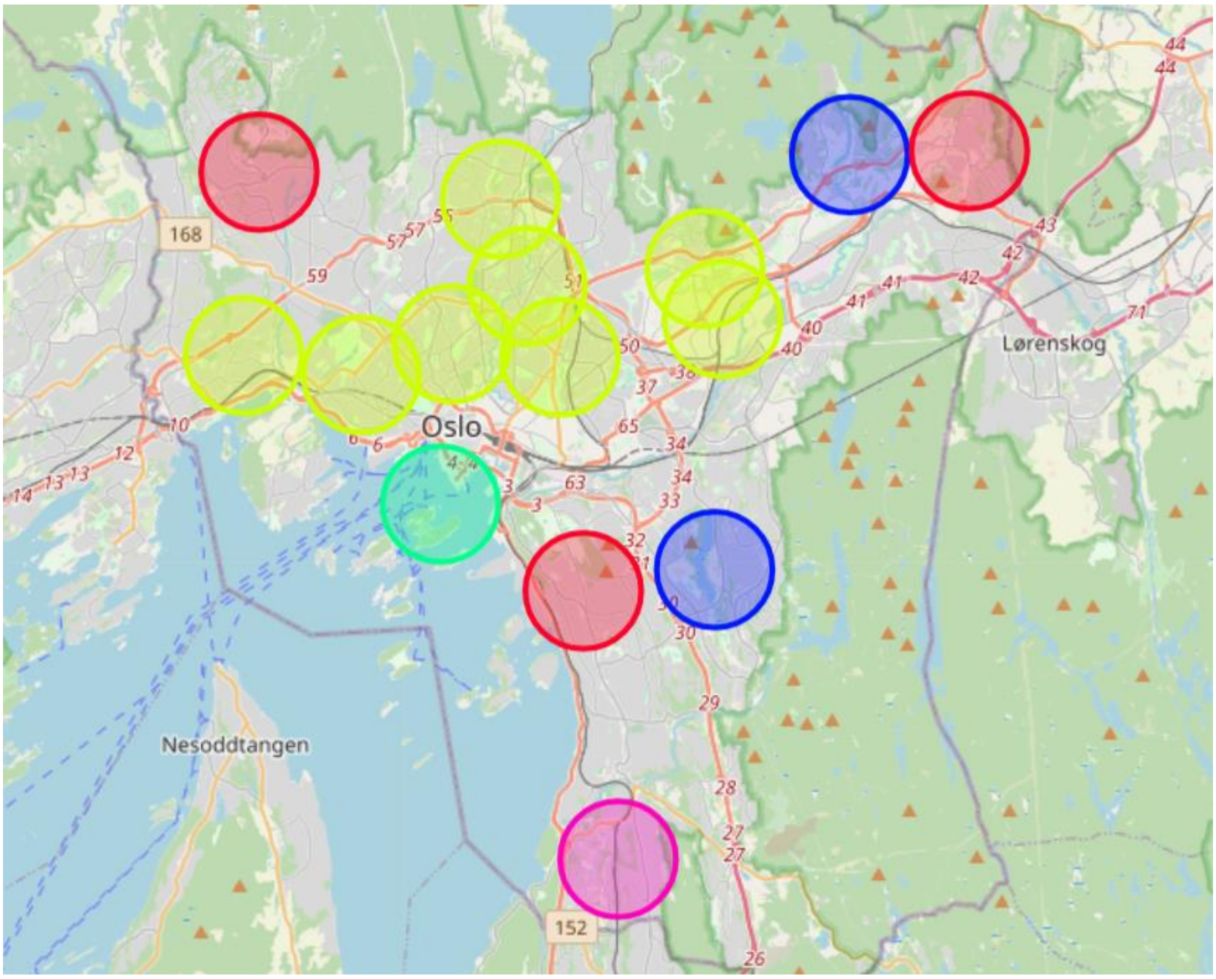
	Borough	Basketball Court	Bathing Area	Bay	Beach	Beer Bar	Beer Garden	Beer Store	Boarding House	Boat or Ferry
0	Alna	0.000000	0.000000	0.000000	0.000000	0.00	0.00	0.0000	0.00	0.00000
1	Bjerke	0.000000	0.000000	0.000000	0.000000	0.00	0.00	0.0000	0.00	0.00000
2	Frogner	0.000000	0.000000	0.000000	0.000000	0.00	0.00	0.0000	0.01	0.00000
3	Gamle Oslo	0.000000	0.043478	0.043478	0.043478	0.00	0.00	0.0000	0.00	0.26087
4	Grorud	0.000000	0.000000	0.000000	0.000000	0.00	0.00	0.0000	0.00	0.00000
5	Grünerløkka	0.000000	0.000000	0.000000	0.000000	0.00	0.00	0.0125	0.00	0.00000
6	Nordre Aker	0.000000	0.000000	0.000000	0.000000	0.00	0.00	0.0000	0.00	0.00000
7	Nordstrand	0.000000	0.000000	0.000000	0.000000	0.00	0.00	0.0000	0.00	0.00000
8	Sagene	0.000000	0.000000	0.000000	0.000000	0.00	0.00	0.0000	0.00	0.00000
9	St. Hanshaugen	0.000000	0.000000	0.000000	0.000000	0.01	0.01	0.0100	0.00	0.00000
10	Stovner	0.000000	0.000000	0.000000	0.000000	0.00	0.00	0.0000	0.00	0.00000
11	Søndre Nordstrand	0.000000	0.000000	0.000000	0.000000	0.00	0.00	0.0000	0.00	0.00000
12	Ullern	0.000000	0.000000	0.000000	0.000000	0.00	0.00	0.0000	0.00	0.00000
13	Vestre Aker	0.000000	0.000000	0.000000	0.000000	0.00	0.00	0.0000	0.00	0.00000
14	Østensjø	0.055556	0.000000	0.000000	0.000000	0.00	0.00	0.0000	0.00	0.00000

Then top 5 most common venues were analyzed. And based on that new dataframe with top 10 venues for each borough was made:

	Borough	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	Alna	Metro Station	Grocery Store	Furniture / Home Store	Pet Store	Bus Station	Spanish Restaurant	Flower Shop	Hotel	Bakery	Automotive Shop
1	Bjerke	Grocery Store	Gym / Fitness Center	Hotel	Supermarket	Park	Café	Farm	Fast Food Restaurant	Pizza Place	Wine Shop
2	Frogner	Bakery	Café	Coffee Shop	Hotel	Indian Restaurant	Scandinavian Restaurant	Burger Joint	Pub	Pizza Place	Grocery Store
3	Gamle Oslo	Boat or Ferry	Pier	Scandinavian Restaurant	Castle	Bay	Other Nightlife	Seafood Restaurant	Café	Mexican Restaurant	Burger Joint
4	Grorud	Metro Station	Grocery Store	Convenience Store	Wine Shop	Asian Restaurant	Athletics & Sports	Pizza Place	Bus Station	Supermarket	Donut Shop
5	Grünerløkka	Grocery Store	Café	Botanical Garden	Coffee Shop	Bar	Bakery	Gym / Fitness Center	Park	Bus Station	Sushi Restaurant
6	Nordre Aker	Bakery	Gym	Asian Restaurant	Metro Station	Shopping Mall	Sushi Restaurant	Grocery Store	Gastropub	Hotel	Pizza Place
7	Nordstrand	Grocery Store	Metro Station	Soccer Field	Pizza Place	Soccer Stadium	Pet Store	Gas Station	Bakery	Light Rail Station	Stadium
8	Sagene	Café	Sushi Restaurant	Coffee Shop	Park	Grocery Store	Pizza Place	Bar	Bakery	Indian Restaurant	Brewery
9	St. Hanshaugen	Bakery	Café	Coffee Shop	Pizza Place	Scandinavian Restaurant	Park	Indian Restaurant	Restaurant	Asian Restaurant	Clothing Store
10	Stovner	Gas Station	Athletics & Sports	Grocery Store	Golf Course	Shopping Mall	Yoga Studio	Falafel Restaurant	Electronics Store	Eastern European Restaurant	Donut Shop
11	Søndre Nordstrand	Grocery Store	Shopping Mall	Fast Food Restaurant	Gym	Stadium	Train Station	Dog Run	Falafel Restaurant	Electronics Store	Eastern European Restaurant
12	Ullern	Bus Station	Market	Metro Station	Gourmet Shop	Burger Joint	Electronics Store	Convenience Store	Harbor / Marina	Light Rail Station	Flower Shop
13	Vestre Aker	Grocery Store	Ski Area	Restaurant	Café	Metro Station	Soccer Field	Lake	Museum	Gas Station	Athletics & Sports
14	Østensjø	Metro Station	Athletics & Sports	Shopping Mall	Yoga Studio	Supermarket	Grocery Store	Lake	Convenience Store	Pizza Place	Burger Joint

As a next step **k-Means Clustering** was used for predicting categorical classed labels. I have tried to find the optimal number of clusters by simply trying different numbers of clusters and examining them visually on having something meaningful in common.

After trying 3, 4, 5 and 6 clusters I have decided that optimal number of clusters is 5. On a map they looks like that:



Cluster 1

```
oslo_merged.loc[oslo_merged['Cluster Labels'] == 0, oslo_merged.columns[[0] + list(range(4, oslo_merged.shape[1]))]]
```

	Borough	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
7	Nordstrand	Grocery Store	Stadium	Pet Store	Metro Station	Light Rail Station	Soccer Stadium	Bakery	Soccer Field	Department Store	Deli / Bodega
10	Stovner	Grocery Store	Department Store	Ski Area	Video Game Store	Athletics & Sports	Golf Course	Shopping Mall	Gas Station	Diner	Electronics Store
13	Vestre Aker	Grocery Store	Ski Area	Metro Station	Soccer Field	Restaurant	Café	Scandinavian Restaurant	Theme Park Ride / Attraction	Disc Golf	Museum

Cluster 2

```
oslo_merged.loc[oslo_merged['Cluster Labels'] == 1, oslo_merged.columns[[0] + list(range(4, oslo_merged.shape[1]))]]
```

	Borough	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	Alna	Metro Station	Grocery Store	Furniture / Home Store	Bus Station	Pet Store	Bakery	Market	Bookstore	Kids Store	Electronics Store
1	Bjerke	Grocery Store	Café	Pizza Place	Supermarket	Metro Station	Farm	Fast Food Restaurant	Hotel	Gym / Fitness Center	Deli / Bodega
2	Frogner	Café	Bakery	Coffee Shop	Indian Restaurant	Scandinavian Restaurant	Hotel	Pizza Place	Wine Shop	Burger Joint	Pub
5	Grünerløkka	Grocery Store	Café	Gym / Fitness Center	Sushi Restaurant	Bar	Park	Bakery	Coffee Shop	Botanical Garden	Asian Restaurant
6	Nordre Aker	Gym	Grocery Store	Bus Stop	Shopping Mall	Bakery	Metro Station	Advertising Agency	Theme Park	Coffee Shop	Café
8	Sagene	Café	Sushi Restaurant	Grocery Store	Coffee Shop	Park	Bakery	Pizza Place	Bar	Brewery	Indian Restaurant
9	St. Hanshaugen	Bakery	Café	Coffee Shop	Scandinavian Restaurant	Park	Pizza Place	Indian Restaurant	Asian Restaurant	Gym / Fitness Center	Bar
12	Ullern	Bus Station	Market	Metro Station	Italian Restaurant	Rafting	Electronics Store	Burger Joint	Light Rail Station	Flower Shop	Coffee Shop

Cluster 3

```
oslo_merged.loc[oslo_merged['Cluster Labels'] == 2, oslo_merged.columns[[0] + list(range(4, oslo_merged.shape[1]))]]
```

	Borough	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
3	Gamle Oslo	Boat or Ferry	Pier	Castle	Scandinavian Restaurant	History Museum	Italian Restaurant	Market	Café	Seafood Restaurant	Beach

Cluster 4

```
oslo_merged.loc[oslo_merged['Cluster Labels'] == 3, oslo_merged.columns[[0] + list(range(4, oslo_merged.shape[1]))]]
```

	Borough	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	Grorud	Grocery Store	Metro Station	Supermarket	Wine Shop	Asian Restaurant	Athletics & Sports	Gym	Pizza Place	Soccer Field	Yoga Studio
14	Østensjø	Metro Station	Athletics & Sports	Shopping Mall	Pizza Place	Basketball Court	Skating Rink	Burger Joint	Lake	Supermarket	Soccer Field

Cluster 5

```
oslo_merged.loc[oslo_merged['Cluster Labels'] == 4, oslo_merged.columns[[0] + list(range(4, oslo_merged.shape[1]))]]
```

	Borough	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
11	Søndre Nordstrand	Grocery Store	Shopping Mall	Stadium	Athletics & Sports	Train Station	Fast Food Restaurant	Business Service	Disc Golf	Farm	Falafel Restaurant

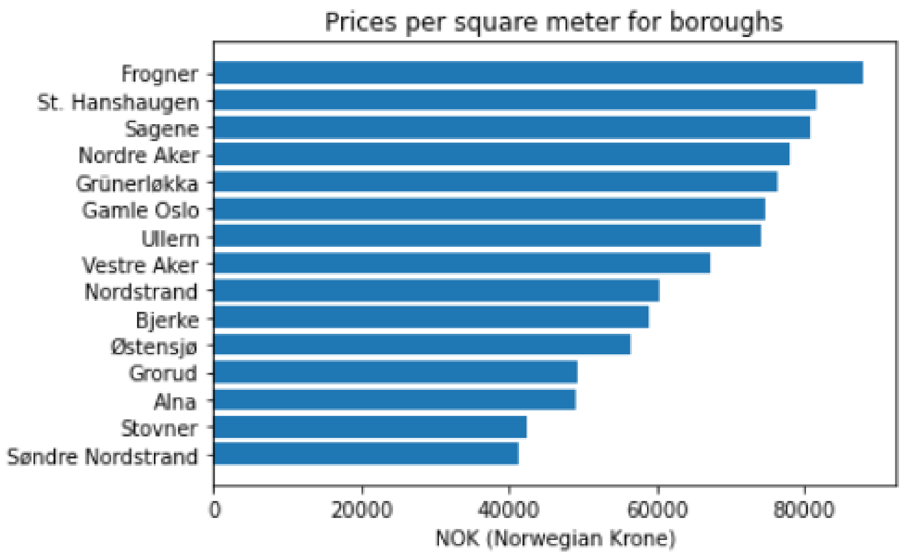
Pricing

After that it was decided to check if some considerations could be done using information about pricing in the boroughs. So new dataframe was created:

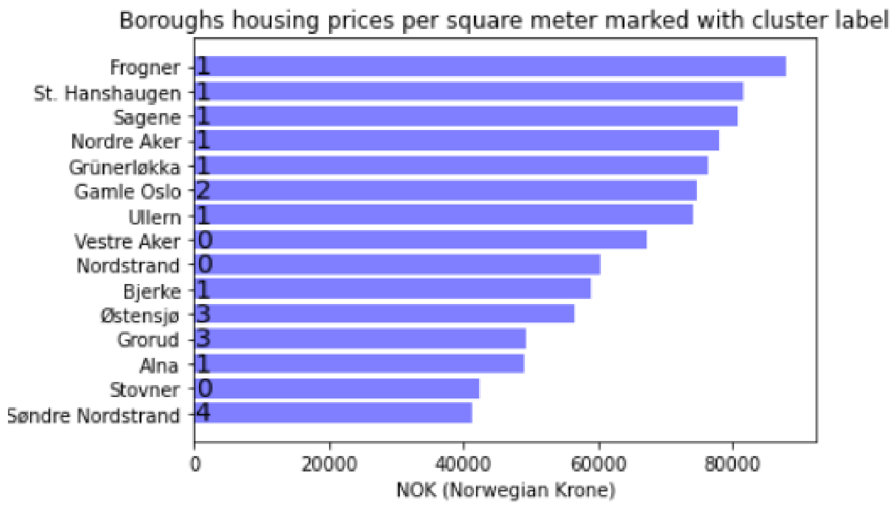
```
url = 'https://raw.githubusercontent.com/Julechka/Coursera_Capstone/master/Capstone-oslo-price.csv'
df_csv = pd.read_csv(url)
prices_oslo = pd.merge(prices_oslo, df_csv, on=['Borough'])
prices_oslo
```

	Borough	Cluster Labels	Price
0	Alna	1	49042
1	Bjerke	1	58834
2	Frogner	1	87923
3	Gamle Oslo	2	74605
4	Grorud	3	49224
5	Grünerløkka	1	76413
6	Nordre Aker	1	78073
7	Nordstrand	0	60350
8	Sagene	1	80766
9	St. Hanshaugen	1	81649
10	Stovner	0	42394
11	Søndre Nordstrand	4	41235
12	Ullern	1	74109
13	Vestre Aker	0	67186
14	Østensjø	3	56556

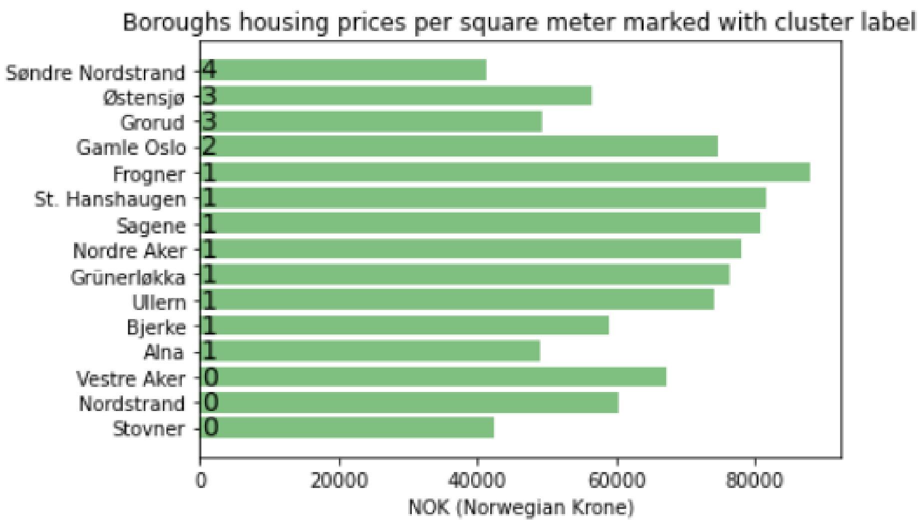
Visualizing these data using the horizontal plot:



Next, every bar was marked with the label mark:

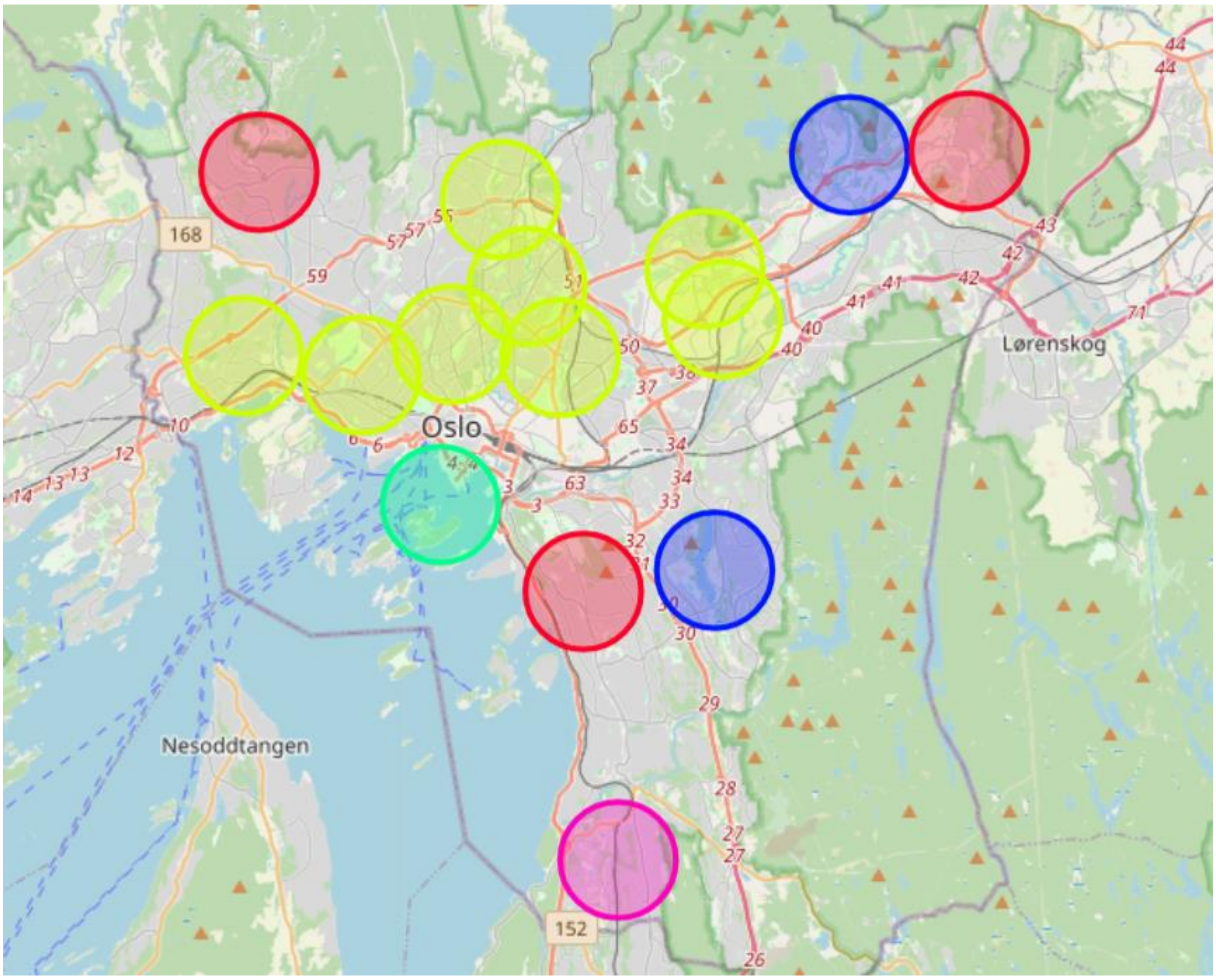


And, finally, grouped by these labels:

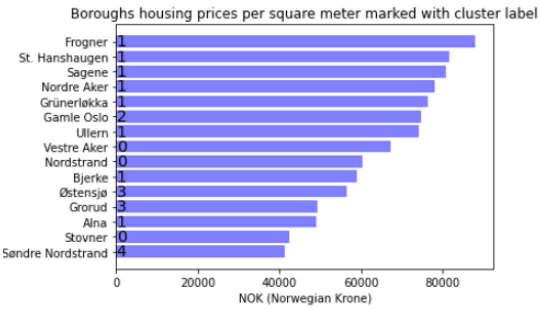


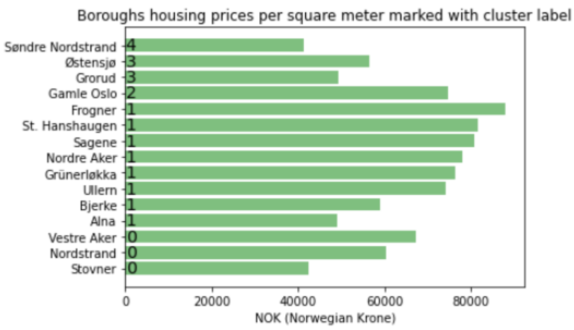
Discussion

After trying 3, 4, 5 and 6 clusters I have decided that optimal number of clusters is 5. The decision was made based on the observations of the developed clusters. On a map final cluster set visualized like that:



Pricing:





Let's discuss our findings cluster by cluster.

Cluster Label 0, Color on a map: red. Economy, healthy life style

Venues, forming Cluster 0 (red marker on a map) are mostly sport venues and grocery stores. When it comes to pricing, housing is not very expensive. One of the boroughs is even one of the cheapest areas (Stovner).

Cluster Label 1, Color on a map: yellow. Luxus, food lovers

This cluster has many cafes and restaurants nearby and it has most expensive housing.

Cluster Label 2, Color on a map: green. Travelers

This cluster consists of only one borough, which is the historical center with mostly tourist attractions, not many groceries and that's why very suitable for travelers mostly.

Cluster Label 3, Color on a map: blue. Budget, family

This cluster has a mix of sport venues, groceries and fast foods with metro station nearby. It's relatively cheap, not very central, but having convenient public transport connection to the center.

Cluster Label 4, Color on a map: magenta. Cheapest

This cluster represented by one borough, Søndre Nordstrand. It seems to be a very balanced in regards to shops and sport venues and it also has the cheapest housing. It also has least central proximity among all boroughs.

Conclusion

It will be very interesting to further investigate the boroughs by including into consideration schools proximity, types of housing and demographic factors.

But this research requires access to more advanced data sources and a lot of data wrangling. It could be a useful tool if to develop it further.

References

https://en.wikipedia.org/wiki/Oslo

https://www.finn.no/job

https://www.lifeinnorway.net/this-is-viken-norways-new-super-county/

http://statistikkbanken.oslo.kommune.no/webview/index.jsp?Geografisubset=30100%2C30101+-+30117&headers=r&stubs=Geografi&measure=common&virtualslice=Gjennomsnittligkvadratmeterpris_value&layers=Boligtype&layers=virtual&study=http%3A%2F%2F192.168.101.44%3A80%2Fobj%2FfStudy%2FNPS-omsatteboliger-kvadratmeterpris&Boligtypesubset=1&mode=cube&v=2&virtualsubset=Gjennomsnittligkvadratmeterpris_value&rsubset=2004+-+2018&Boligtypeslice=1&measuretype=4&cube=http%3A%2F%2Fstatistikkbanken.oslo.kommune.no%3A80%2Fobj%2FfCube%2FNPS-omsatteboliger-kvadratmeterpris_C1&top=yes

https://www.statista.com/

https://www.oslo.kommune.no/