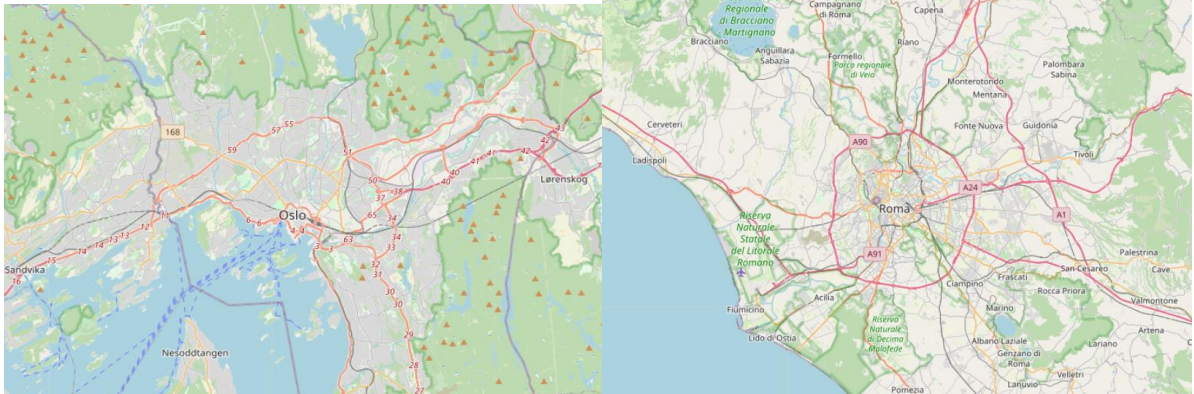


# Comparison of Rome and Oslo

Weronika Żak



*Oslo on the left, Rome on the right*

## 1. Introduction

### 1.1 Background

My idea is to compare the capitals of two different countries in Europe, that is Italy and Norway. Since the climate and temperature are different, it may be worth exploring what kind of venues or types of business can you find in Oslo and Rome and how much they differ.

### 1.2 Problem

The aim of this project is to answer questions:

- Do people living in western countries have different ways of spending time than in the colder North?
- What kind of business is the most profitable in these places?

### 1.3 Target Audience

This project may be interesting for those, who have money and an idea of some kind of business in Europe, but are not sure, where should they start it.

## 2. Data acquisition and cleaning

### 2.1 Data sources

A list of Oslo's and Rome's districts will be scraped with BeautifulSoup library from Wikipedia website. After that, I'm going to use Geocoder to look for the coordinates and Foursquare API

to get more details about the capitals' venues. The data gathered from Wikipedia had 154 Italian neighborhoods and 105 Norwegian.

## 2.2 Data cleaning

Some of the scraped results had an additional text in brackets, so I had to get rid of them. Moreover, after using Geocoder few neighborhoods' locations were too far away from the capitals, so I decided to drop these rows from the table, so they wouldn't bring in chaos. In the Oslo table, I chose the places to delete by name, and when cleaning the Rome table, I headed by latitude and longitude. After data cleaning there remained 104 Italian neighborhoods and 89 Norwegian.

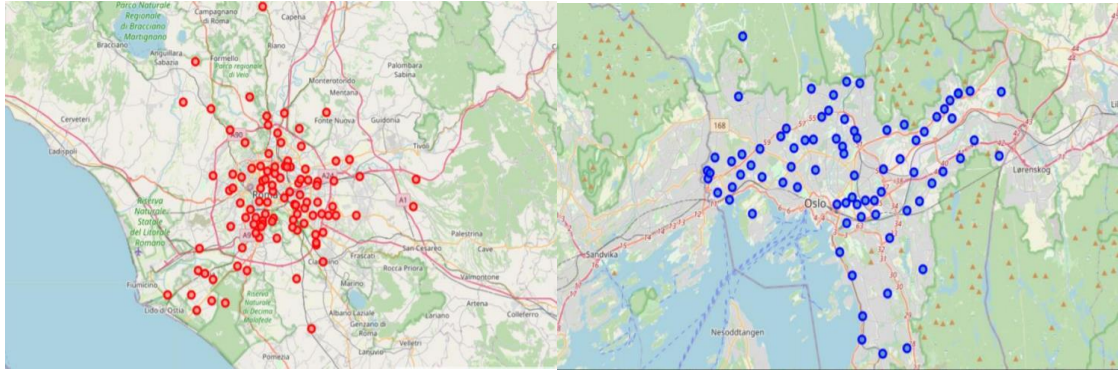
## 2.3 Feature selection

I will be leveraging on features in a reliable location information provider such as the Foursquare API to explore various types of venues and their categories available in each neighborhood. I will also retrieve and save details about each of the venues nearby (i.e. within 500m) in each of the neighborhoods. At the end of that task, the tables in the notebook consist of columns such as:

- District's name
- Neighbourhood
- Business name (for example store name)
- Business category
- Longitude
- Latitude
- Cluster label

	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
0	Alna	Alnabru	59.926682	10.836498	3	Bus Station	Paper / Office Supplies Store	Convenience Store	Food & Drink Shop	Food	Flower Shop
1	Alna	Ellingsrud	59.934191	10.920897	1	Metro Station	Warehouse Store	Gym / Fitness Center	Bed & Breakfast	Grocery Store	Yoga Studio
2	Alna	Furuset	59.941067	10.896399	0	Shopping Mall	Metro Station	Hockey Arena	Supermarket	Flower Shop	Fish Market
3	Alna	Haugerud	59.922116	10.854522	7	Gym	Metro Station	Grocery Store	Flower Shop	Fish Market	Fast Food Restaurant
4	Alna	Hellerud	59.910067	10.829839	5	Moving Target	Metro Station	Yoga Studio	Farm	Food & Drink Shop	Food
5	Alna	Lindeberg	59.933155	10.882706	7	Grocery Store	Furniture / Home Store	Motorcycle Shop	Metro Station	Food	Flower Shop
6	Alna	Trosterud	59.927182	10.865258	2	Furniture / Home Store	Hotel	Construction & Landscaping	Grocery Store	Fast Food Restaurant	Café
7	Alna	Tveita	59.914031	10.842241	2	Grocery Store	Metro Station	Convenience Store	Video Game Store	Supermarket	Bus Stop
8	Bjerke	Linderud	59.940963	10.838420	2	Wine Shop	Bakery	Fast Food Restaurant	Café	Grocery Store	Gym / Fitness Center
9	Bjerke	Tonsenhagen	59.947696	10.827078	1	Trail	Bus Station	Pool	Grocery Store	Yoga Studio	Flower Shop
10	Bjerke	Årvoll	59.945281	10.809868	1	Supermarket	Chinese Restaurant	Grocery Store	Convenience Store	Farm	Bus Station
11	Bjerke	Veitvet	59.944530	10.846939	1	Bowling Alley	Shopping Mall	Soccer Field	Supermarket	Flower Shop	Fish Market
12	Bjerke	Økern	59.928261	10.806278	2	Convenience Store	Auto Workshop	Park	Bus Stop	Sushi Restaurant	Metro Station

### 3. Methodology



Data scraping from the Wikipedia page that contains the up-to-date details of Rome and Oslo boroughs with neighborhoods has been used.

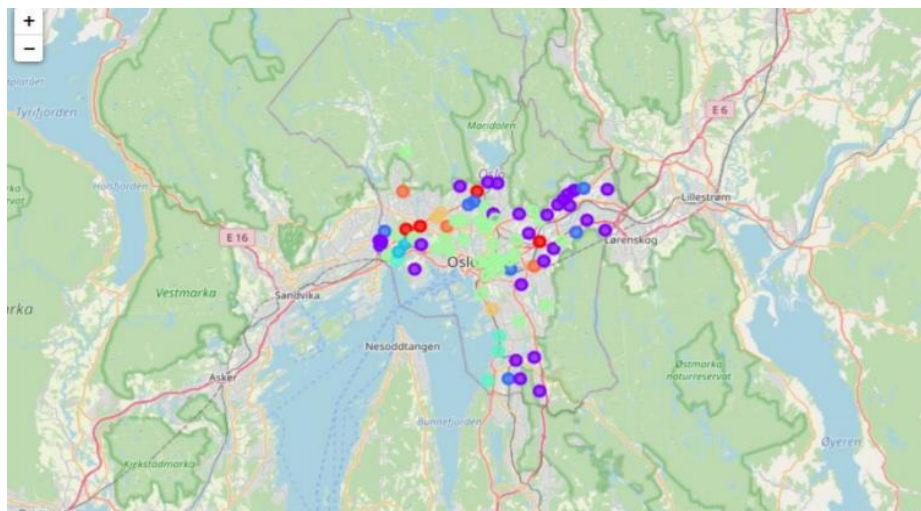
Furthermore, the OpenCage Geocoder API has been used for retrieving the coordinates and locations of these neighborhoods. Getting longitude and Latitude is important as it will allow me to input this information into the location information provider such as Foursquare API to get more venues' details in these neighborhoods.

Machine learnings techniques such as the K-Means to segment and cluster these neighborhoods will also be used, so that they can be grouped together. It will be helpful to understand their similarities. This is critical as we need to recommend the regions of the neighborhoods of the choice in our recommendation so that particular business can easily replicate their business model across neighborhoods easily.

Finally, with all these methodologies, we will then be able to come up with the best recommendation for the management of any kind of five businesses (food providing services, sport, entertainment, etc.

### 4. Results

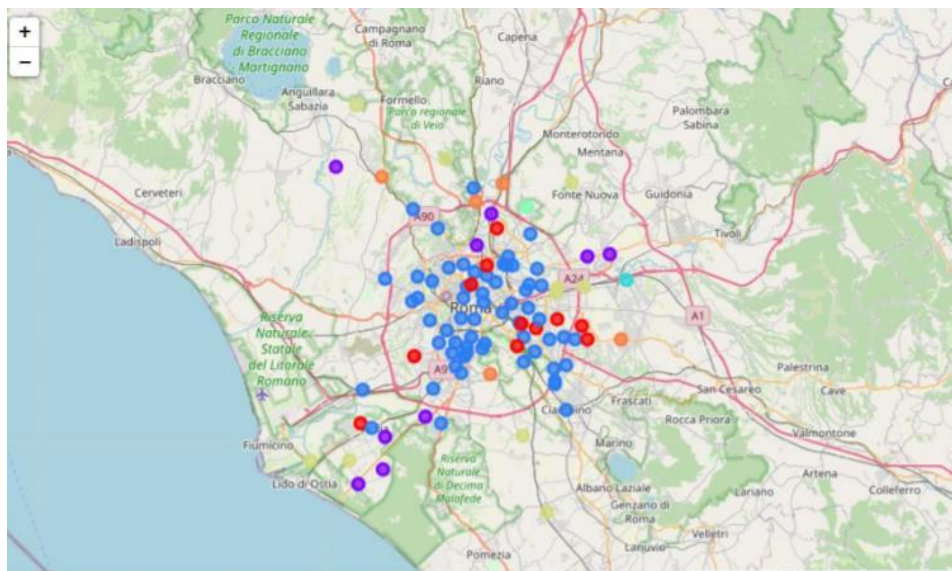
#### 4.1. Oslo Clustering



Oslo consists of myriads of stores and shops. There are also many restaurant, cultural, and sport-related places. On the other hand, there aren't many hotels, tourist lodgings or entertainment places.

- **Cluster 1:** e. g. Ammerud, Tåsen – shopping district, many markets and stores, few hotels
- **Cluster 2:** e. g. Bygdøy, Kjelsås– many gyms, parks and gardens. Consists of smaller shops and restaurants.
- **Cluster 3:** e.g. Økern, Uranienborg – tourist district, hotels, bakeries, cafes, restaurants.
- **Cluster 4:** e.g. Korsvoll, Frysja – smaller stores, bus stations.
- **Cluster 5:** e.g. Bestum, Ullern – smaller businesses and stores
- **Cluster 6:** e.g. Hellerud, Vækerø – entertainment district, shops
- **Cluster 7:** e.g. Nordstrand, Ljan – sport facilities, cultural activities, family friendly neighborhoods
- **Cluster 8:** e.g. Kringsjå, Lysejordet – tourists district, many grocery stores.

## 4.2 Rome Clustering



Located in the south the city Rome seems to have more restaurants, cafes, and food-related places. On the other hand, there aren't many sport-related places like gyms or yoga studios. What's surprising, looks like there aren't many cultural places like galleries, museums, libraries, etc. either.

- **Cluster 1:** e. g. Alessandrina, Tufello – cultural activities, parks, plazas, zoos
- **Cluster 2:** e. g. Infernetto, Malafede– Italian Restaurants, food-related business mainly, few supermarkets and stores
- **Cluster 3:** e.g. Trionfale, Aventino– hotels and restaurant, district for workers.
- **Cluster 4:** e.g. Lunghezza – women stores, gardens, park.
- **Cluster 5:** e.g. Bufalotta – spa, sport activities.
- **Cluster 6:** e.g. La Rustica, Santa Palomba– tourist district, hotels and restaurants, few sport places
- **Cluster 7:** e.g. Grottaperfetta, La Storta– good place for coffee lovers, family friendly neighborhood, bakeries and zoos.

## 5. Discussion

Before starting a business, some further data analysis of the optimal location of a shop may be required. The location of stores still changes so it's important to rerun the retrieving neighborhoods' nearby venues part of the code once in a while to update data. The results below apply to mid-2020.

Business recommendation based on clustering:

- In Oslo, there are more sport-related activities than in Rome, where it's more focused on Restaurants and Cafes.
- For the food business, it is recommended to look for central neighborhoods in Italy or the suburbs of Oslo.
- The entertainment business is most profitable in the south suburbs of Rome and the northern part of Oslo.
- In terms of lodging and stay tourists are most welcomed in the western and eastern part of Rome and on the east from the centre of Oslo.
- There are many shops all over Oslo so if someone wants to open another supermarket they might consider the western areas. On the other hand, there are few places with a supermarket in Rome, so it's less risky business.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
2	Furuset	Shopping Mall	Metro Station	Hockey Arena	Supermarket	Flower Shop	Fish Market
30	Ammerud	Metro Station	Grocery Store	Athletics & Sports	Supermarket	Soccer Field	Department Store
42	Tåsen	Metro Station	Shopping Mall	Grocery Store	Flower Shop	Fish Market	Fast Food Restaurant
44	Berg	Metro Station	Shopping Mall	Grocery Store	Flower Shop	Fish Market	Fast Food Restaurant
61	Mortensrud	Shopping Mall	Grocery Store	Metro Station	Convenience Store	Bus Station	Flower Shop
73	Ullernåsen	Trail	Hotel	Bus Station	Metro Station	Electronics Store	Food
74	Montebello	Metro Station	Hotel	Yoga Studio	Food	Flower Shop	Fish Market

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
13	Fidene	Italian Restaurant	Grocery Store	Zoo	Electronics Store	Food	Flower Shop
24	Tiburtino Nord	Italian Restaurant	Ice Cream Shop	Dessert Shop	Ethiopian Restaurant	Food Truck	Food
29	Settecamini	Italian Restaurant	Café	Plaza	Cocktail Bar	Zoo	Food
65	Mezzocammino	Italian Restaurant	Flower Shop	Park	Café	Gym / Fitness Center	Zoo
70	Malafede	Italian Restaurant	Supermarket	Park	Zoo	Electronics Store	Flower Shop
75	Castel Fusano	Italian Restaurant	Food Truck	Furniture / Home Store	Electronics Store	Food	Flower Shop
76	Infernetto	Italian Restaurant	Golf Course	Zoo	Electronics Store	Food	Flower Shop
93	Santa Maria di Galeria	Italian Restaurant	Zoo	Electronics Store	Food	Flower Shop	Flea Market
94	Tor di Quinto	Italian Restaurant	Light Rail Station	Zoo	Ethiopian Restaurant	Food Truck	Food

## 6. Conclusion

Located in south Rome seems to have more restaurants, cafes, and food-related places. On the other hand, there aren't many sport-related places like gyms or yoga studios. What's surprising, looks like there aren't many cultural places like galleries, libraries, etc. either.

Oslo consists of myriads of stores and shops. There are also many restaurant, cultural, and sport-related places. On the other hand, there aren't many hotels, tourist lodging or entertainment places.