Rent Prices & Nearby Venues Data Analysis of Munich

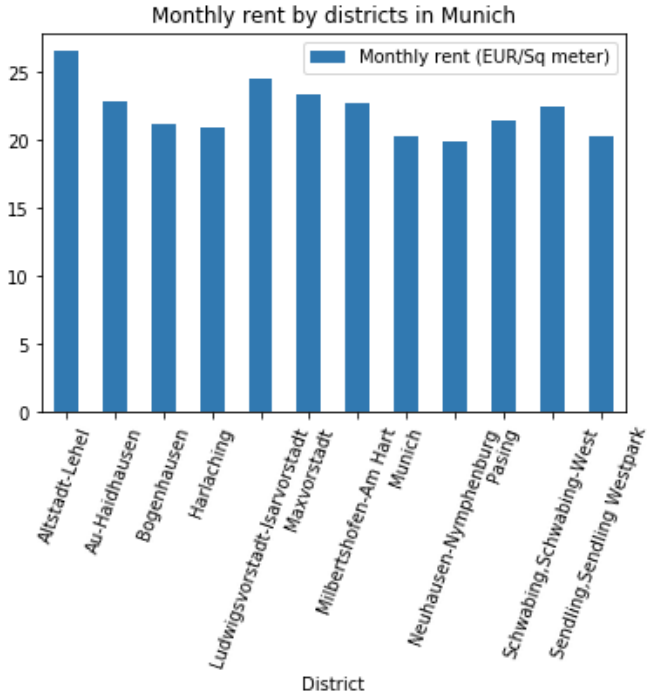
Introduction

Munich is the third-largest city in Germany with about 1.56 million inhabitants. It is divided into 25 districts and it has the most populated density per km2 in Germany (4,500 people per km2). Together with London, Paris and Berlin, the Munich property market ranks as one of the most important in Europe. As in 2020 demand exceeded available properties, Munich remains Europe's favourite locations among investors[1].

This report intends to provide information to investors that want to buy an accommodation in Munich considering rent prices and location attractiveness. Based on these two factors, Munich's districts will be segmented and clustered using Python and machine learning k-means. The report includes tables and a map with the plotted clusters.

Data section

The report contains the following data:
"Bayern Postal Codes" from geonames.org[2] using Pandas library web scrapping function. "Average rent of apartments in Munich, Germany for the first half of 2019, by district"[3] downloaded from Statista and imported into Github project's repository. "Postal Codes in Munich" [4] from official Munich website imported into Github project's repository. Foursquare locations API to collect venues in Munich.



Matplotlib barchart

Methodology section

GitHub repository was used as the main database for this project. The master data called "dataset 3" was created using two different sources from geonames.org, Average_apart_rent_Munich.csv and Postal_Code_Munich(3).csv. Dataset 3 contains districts, Average price Monthly Rents, latitude and longitude of each postal code. It contains 50 postal codes.

	Postal Code	Combined Districts	Monthly rent (EUR/Sq meter)	Latitude	Longitude
0	80331	Altstadt-Lehel	26.5700000000	48.135	11.571
1	80333	Altstadt-Lehel//Maxvorstadt	25.0100000000	48.145	11.567
2	80335	Altstadt-Lehel//Ludwigsvorstadt-Isarvorstadt//...	24.8366666667	48.143	11.555
3	80336	Altstadt-Lehel//Ludwigsvorstadt-Isarvorstadt//...	23.7733333333	48.135	11.559
4	80337	Ludwigsvorstadt-Isarvorstadt//Sendling,Sendlin...	22.3750000000	48.122	11.545
5	80469	Altstadt-Lehel//Ludwigsvorstadt-Isarvorstadt//...	23.7733333333	48.13	11.573
6	80538	Altstadt-Lehel//Schwabing,Schwabing-West	24.5150000000	48.14	11.588
7	80539	Altstadt-Lehel//Maxvorstadt	25.0100000000	48.15	11.583
8	80634	Neuhausen-Nymphenburg	19.9900000000	48.149	11.533
9	80636	Maxvorstadt//Neuhausen-Nymphenburg	21.7200000000	48.156	11.541
10	80637	Neuhausen-Nymphenburg	19.9900000000	48.172	11.528
11	80638	Neuhausen-Nymphenburg	19.9900000000	48.16	11.493
12	80639	Neuhausen-Nymphenburg	19.9900000000	48.153	11.515
13	80686	Sendling,Sendling Westpark	20.2600000000	48.138	11.507
14	80687	Pasing	21.4600000000	48.141	11.502
15	80689	Pasing	21.4600000000	48.129	11.489
16	80796	Schwabing,Schwabing-West	22.4600000000	48.15	11.583
17	80797	Maxvorstadt//Schwabing,Schwabing-West	22.9550000000	48.161	11.559
18	80798	Maxvorstadt//Schwabing,Schwabing-West	22.9550000000	48.15	11.583
19	80799	Maxvorstadt//Schwabing,Schwabing-West	22.9550000000	48.15	11.583
20	80801	Maxvorstadt//Schwabing,Schwabing-West//Schwabi...	22.7900000000	48.15	11.583

See notebook for all postal codes.

Foursquare API was used to explore venues in a radius of 500meters of each poste code (given their latitude and longitude) and 1674 venues were found.

```
(1674, 7)
```

	Postal Code	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	80331	48.135	11.571	Asamkirche (St. Johann Nepomuk)	48.1350534503	11.5697462772	Church
1	80331	48.135	11.571	Japanische Feinkost Mikado (美門日本食料品店)	48.1361261400	11.5687318300	Gourmet Shop
2	80331	48.135	11.571	The High	48.1331010263	11.5729389662	Cocktail Bar
3	80331	48.135	11.571	Ringlers	48.1340970000	11.5683020000	Sandwich Place
4	80331	48.135	11.571	TeeGschwendner	48.1353980000	11.5694550000	Tea Room

The above data were then grouped by postal codes and the mean of the frequency of occurrence of each category was created. The final data frame displays the top 10 venues for the 50 postal codes.

	Postal Code	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	80331	Café	German Restaurant	Hotel	Coffee Shop	Plaza	Bavarian Restaurant	Italian Restaurant	Cosmetics Shop	Cocktail Bar	Clothing Store
1	80333	Café	History Museum	Nightclub	Burger Joint	Restaurant	Plaza	Italian Restaurant	Art Museum	Asian Restaurant	Movie Theater
2	80335	Hotel	Bakery	Bar	Middle Eastern Restaurant	Coffee Shop	Drugstore	Post Office	Office	Italian Restaurant	Ice Cream Shop
3	80336	Hotel	Middle Eastern Restaurant	German Restaurant	Café	Camera Store	Asian Restaurant	Bakery	Mexican Restaurant	Nightclub	Indie Movie Theater
4	80337	Vietnamese Restaurant	Café	Italian Restaurant	Gastropub	Doner Restaurant	Supermarket	Food & Drink Shop	Bosnian Restaurant	Caucasian Restaurant	Restaurant
5	80469	Café	Bar	Vietnamese Restaurant	Asian Restaurant	Cocktail Bar	Pizza Place	Ice Cream Shop	Afghan Restaurant	Italian Restaurant	Gay Bar
6	80538	Italian Restaurant	German Restaurant	Hotel	Bar	Plaza	Pastry Shop	Art Museum	Burger Joint	Outdoor Sculpture	Snack Place
7	80539	Café	Bar	Italian Restaurant	Ice Cream Shop	Restaurant	Sushi Restaurant	Thai Restaurant	Plaza	Doner Restaurant	Bagel Shop
8	80634	Hotel	Café	German Restaurant	Drugstore	Bakery	Indian Restaurant	Supermarket	Ice Cream Shop	Sushi Restaurant	Vietnamese Restaurant
9	80636	Bakery	German Restaurant	Indian Restaurant	Italian Restaurant	Hotel	Dessert Shop	Café	Restaurant	Ramen Restaurant	Organic Grocery
10	80637	German Restaurant	Hotel	Tram Station	Supermarket	Taverna	Metro Station	Café	Italian Restaurant	Stadium	Light Rail Station

See notebook for all postal codes

The previous table shows that several postal codes have common venues. Therefore, the unsupervised learning K-means algorithm will be used to cluster the Munich's postal codes. The K-means is vastly used for clustering in many data science applications, especially useful to quickly discover insights from data. Below is the merged table with cluster labels for each postal codes.

	Postal Code	Combined Districts	Monthly rent (EUR/Sq meter)	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
0	80331	Altstadt-Lehel	26.5700000000	48.135	11.571	4	Café	German Restaurant	Hotel	Coffee Shop
1	80333	Altstadt-Lehel//Maxvorstadt	25.0100000000	48.145	11.567	2	Café	History Museum	Nightclub	Burger Joint
2	80335	Altstadt-Lehel//Ludwigsvorstadt-Isarvorstadt//...	24.8366666667	48.143	11.555	4	Hotel	Bakery	Bar	Middle Eastern Restaurant
3	80336	Altstadt-Lehel//Ludwigsvorstadt-Isarvorstadt//...	23.7733333333	48.135	11.559	4	Hotel	Middle Eastern Restaurant	German Restaurant	Café
4	80337	Ludwigsvorstadt-Isarvorstadt//Sendling,Sendlin...	22.3750000000	48.122	11.545	4	Vietnamese Restaurant	Café	Italian Restaurant	Gastropub
5	80469	Altstadt-Lehel//Ludwigsvorstadt-Isarvorstadt//...	23.7733333333	48.13	11.573	2	Café	Bar	Vietnamese Restaurant	Asian Restaurant
6	80538	Altstadt-Lehel//Schwabing,Schwabing-West	24.5150000000	48.14	11.588	4	Italian Restaurant	German Restaurant	Hotel	Bar
7	80539	Altstadt-Lehel//Maxvorstadt	25.0100000000	48.15	11.583	2	Café	Bar	Italian Restaurant	Ice Cream Shop
8	80634	Neuhausen-Nymphenburg	19.9900000000	48.149	11.533	4	Hotel	Café	German Restaurant	Drugstore
9	80636	Maxvorstadt//Neuhausen-Nymphenburg	21.7200000000	48.156	11.541	4	Bakery	German Restaurant	Indian Restaurant	Italian Restaurant
10	80637	Neuhausen-Nymphenburg	19.9900000000	48.172	11.528	4	German Restaurant	Hotel	Tram Station	Supermarket

See notebook for all postal codes

The number of venues is calculated for each cluster:

Cluster Labels	Venue
0	18
1	129
2	629
3	1
4	897

The mean for monthly rents in EUR for each cluster was computed.

```
cluster 0 :  21.593333333333334
cluster 1 :  21.33230769230769
cluster 2 :  23.436111111111103
cluster 3 :  20.89
cluster 4 :  22.32636363636365
```

Each cluster has been analysed and classified as follows:
CLUSTER 0 : // Avg rent price: 21.59EUR (EUR/Sq meter) // 18 venues // Transportations, restaurants, fast foods and farmer markets
CLUSTER 1: // Avg rent price: 21.33EUR(EUR/Sq meter) // 129 venues // Transportations, restaurants, supermarket, bakery, banks
CLUSTER 2: // Avg rent price: 23.43EUR (EUR/Sq meter) // 624 venues // Cafes, bars and restaurants
CLUSTER 3: Outlier
CLUSTER 4: // Avg rent price: 22.32EUR (EUR/Sq meter) // 898 venues // Multiple Social Venues, Accommodations

Example of Cluster :

CLUSTER 2: // Avg rent price: 23.43EUR (EUR/Sq meter) // 624 venues // Cafes, bars and restaurants ¶

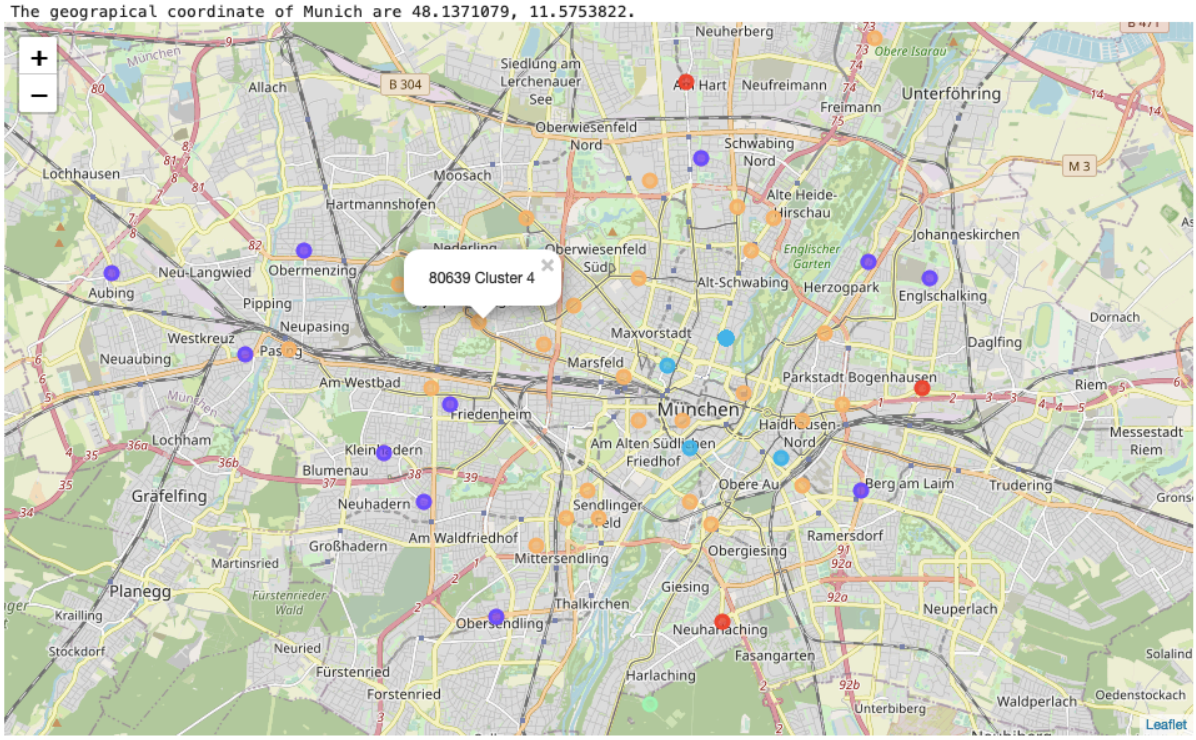
```
df4.loc[df4['Cluster Labels'] == 2, df4.columns[[0] + list(range(5, df4.shape[1]))]]
```

	Postal Code	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	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	80333	2	Café	History Museum	Nightclub	Burger Joint	Restaurant	Plaza	Italian Restaurant	Art Museum	Asian Restaurant	Movie Theater
5	80469	2	Café	Bar	Vietnamese Restaurant	Asian Restaurant	Cocktail Bar	Pizza Place	Ice Cream Shop	Afghan Restaurant	Italian Restaurant	Gay Bar
7	80539	2	Café	Bar	Italian Restaurant	Ice Cream Shop	Restaurant	Sushi Restaurant	Thai Restaurant	Plaza	Doner Restaurant	Bagel Shop
16	80796	2	Café	Bar	Italian Restaurant	Ice Cream Shop	Restaurant	Sushi Restaurant	Thai Restaurant	Plaza	Doner Restaurant	Bagel Shop
18	80798	2	Café	Bar	Italian Restaurant	Ice Cream Shop	Restaurant	Sushi Restaurant	Thai Restaurant	Plaza	Doner Restaurant	Bagel Shop
19	80799	2	Café	Bar	Italian Restaurant	Ice Cream Shop	Restaurant	Sushi Restaurant	Thai Restaurant	Plaza	Doner Restaurant	Bagel Shop
20	80801	2	Café	Bar	Italian Restaurant	Ice Cream Shop	Restaurant	Sushi Restaurant	Thai Restaurant	Plaza	Doner Restaurant	Bagel Shop
22	80803	2	Café	Bar	Italian Restaurant	Ice Cream Shop	Restaurant	Sushi Restaurant	Thai Restaurant	Plaza	Doner Restaurant	Bagel Shop
42	81667	2	Italian Restaurant	Bar	Café	German Restaurant	Indian Restaurant	Plaza	Hotel	Ice Cream Shop	French Restaurant	Thai Restaurant

Results

Munich's Latitude and Longitude was imported with API Geolocator and
Python folium library was used to visualize the city on a map. Clusters were plot on the map.
When selecting a specific cluster on the map the following information appears:
Red → « Postal Code, Cluster 0 »
Violet → « Postal Code, Cluster 1 »
Blue → « Postal Code, Cluster 2 »
Orange → « Postal Code, Cluster 4 »

The geograpical coordinate of Munich are 48.1371079, 11.5753822.



Discussion and recommendations

Districts with postal codes belonging to clusters 2 and 4 tend to be slightly more expensive than clusters 0 and 1. As shown on the map, clusters 2 and 4 tend to be more central. Multiplying previous clusters' average prices by 50 could give investors an idea of how much they will be able to rent their accommodations, for example a 50sqm flat. Clusters 2 and 4 respectively 1171,5EUR/month and 1116EUR/month. Clusters 0 and 1 respectively 1079,5EUR/month and 1066,5EUR/month. Clusters 2 and 4 have respectively 629 and 897 venues, much more than clusters 1 and 0 i.e. 18 and 129 venues. However, clusters have different categories of venues. Cluster 2 has a more intense density of cafe, bars and restaurants than cluster 4 and might be more appropriate for students. Therefore, in the area corresponding to cluster 2, investors should rather buy small single room flats. Accommodations belonging to Postal codes from cluster 4 will benefit from a more diverse type of venues (café and restaurants but also supermarkets, bakeries, banks, clothing stores, drugstore, offices, yoga studio…) and might be better for young workers or families. In those areas, it might be a better investment to buy accommodations with more rooms. Accommodations in Cluster 1 will have all the necessary venues and will be quieter than accommodations in 2 and 4. Cluster 0 has few venues but might be the quietest area and the less congested. Cluster 0 and 1 might be more suitable for families, for example small houses with garden.

Conclusion

Munich's districts were segmented and clustered using Python and the k-means machine learning. Data for districts, average price monthly rents, latitude and longitude of each postal codes were web scrapped. Foursquare API was used to explore venues in a radius of 500meters of each poste code. Unsupervised learning K-means algorithm was used to cluster Munich's postal codes. Previous analyses help investors to decide where to buy an accommodation. Investors can identify clusters and have a better idea about the location price and venues density and category.

Further analyses could include to web scrape specialised websites such as "ImmobilienScout24" or "Immowelt.de" to gather more data about prices and include them as a component for clustering. Furthermore, optimizing clustering could be done by finding a better k for the k-means with the Elbow Method.

References

[1] https://www.muenchen.de/rathaus/wirtschaft_en/munich-business-location/economic-data.html

[2] https://www.geonames.org/postal-codes/DE/BY/bayern.html

[3] https://www.statista.com/statistics/800552/rental-costs-in-munich-germany-by-district/

[4] https://www.muenchen.de/int/en/living/postal-codes.html