Analysis of rental apartments in different boroughs in Berlin and Brandenburg

David Janzen

March 21, 2021

1. Introduction

1.1 Background

All over the world people are searching for apartments. A reason to look for a new apartment is, for example, at the beginning of a new stage of life. With the study such a period of life begins and the task is now to find a cheap apartment in a good location. At least in Berlin these are very expensive compared to other places in Germany. So it is difficult to find a suitable one. Desirable is to find a favorable apartment with the right number of rooms and the appropriate space that also is located in the right area. To guaranty good mobility it is preferable to live near the center of Berlin, so you can reach every place as fast as possible.

1.2 Problem and Interest

The goal of this analysis is to give an overview of the cold rent for apartments in the boroughs of Berlin and also in Brandenburg. An important factor is for example the number of rooms. It is desirable to know in which borough you are likely to move. This analysis is especially useful for those who are currently looking for an apartment. It is interesting to know what is nearby the apartment too. If you like to go to restaurants you maybe want to know which kind you will find most around your apartment.

2. Data acquisition, pre-processing and cleaning

2.1 Data sources

The data with the information about the apartments, which were used for the analysis, are taken from immobilienscout24. Here apartments in Berlin within a radius of 50 km were looked at. In order not to limit the search, the filtering of the price, the rooms, and the area was omitted. This means that all apartments in Berlin within a radius of 50 km were searched for, regardless of their price, number of rooms, or area.

The data for the edges of the boroughs was taken from GitHub Gist and the information about the boroughs with their respective neighborhoods was taken from Wikipedia.

2.2 Data acquisition

To get the data from immobilienscout24, each page of the search query was temporarily saved and its content was parsed using regular expressions. The regular expressions were modified so that almost all apartments with their information could be listed and saved.

However, in the "Information" column, only this information was taken, which is located on the respective overview page. If there was further information, which is only visible when you click on the respective apartment, this was not taken over. In other words, the information in the column "Information" contains only the information that can be seen when one is on a page with the results (the respective apartments).

Furthermore, after the coordinates of the apartments were determined (see section 2.3), the venues located within 500 m of the respective apartment were determined by the foursquare API. However, the limit of 100 venues per search query for the free version applies here. This information was saved to a file and the procedure repeated for a few days, as there is also a limit to how many search queries can be made per day for the free version.

2.3 Pre-Processing and data cleaning
First, the entries that are duplicated are discarded. Then the information about the neighborhoods is written into an array, which is later responsible for the division of the apartments into boroughs. This classification is complicated by the inconsistency of the data. Sometimes the borough was mentioned, but sometimes only the neighborhood. Most of the time it said at the end that the apartment was in Berlin and sometimes it didn't, even though it was. There were also entries where there was additional information in parentheses. All these cases had to be taken into account. The apartments that still did not provide information about the boroughs were discarded since this only occurred a few times. Finally, the coordinates where the apartments are located were determined. This was done with the library GeoPy. In the case of one apartment, however, clearly wrong coordinates were determined, since these were not in Germany. However, in some cases, there was the problem that no coordinates could be determined. Here it was tried to determine the centers of the boroughs and then to give these apartments these coordinates. Finally, the prices per square meter were determined, added to the table, and finally saved. The table finally has 2555 entries with different apartments.

3. Exploratory Data Analysis

3.1 Information that is important for the interpretation of the data
It is important to know that these are not the entire apartments in Berlin/Brandenburg, but only those that are currently advertised on immobilienscout24. In the following, it is therefore not possible to make any statements about the distribution of apartments in the boroughs, but only to provide information about the apartments that are still available in the boroughs. Furthermore, not all apartments that are available are used, since others that are not listed on immobilienscout24 are not included in the analysis.

3.2 Distribution of n-room apartments in a borough
Figure 1 shows the distribution of n-room apartments in each borough. Each borough offers apartments with a different number of rooms. The graph shows the percentage of apartments with a certain number of rooms in a borough. A striking feature is the high proportion of 2-room apartments on offer. Except for Neukölln, Charlottenburg-Wilmersdorf, and Marzahn-Hellersdorf, the proportion is always higher than for other numbers of rooms. The average of 2-room apartments is 37%. Both Friedrichshain-Kreuzberg (48.6%) and Reinickendorf (46.8%) have the largest share of detached 2-room apartments, while Charlottenburg-Wilmersdorf (26.4%) and Neukölln (27.1%) have the smallest.
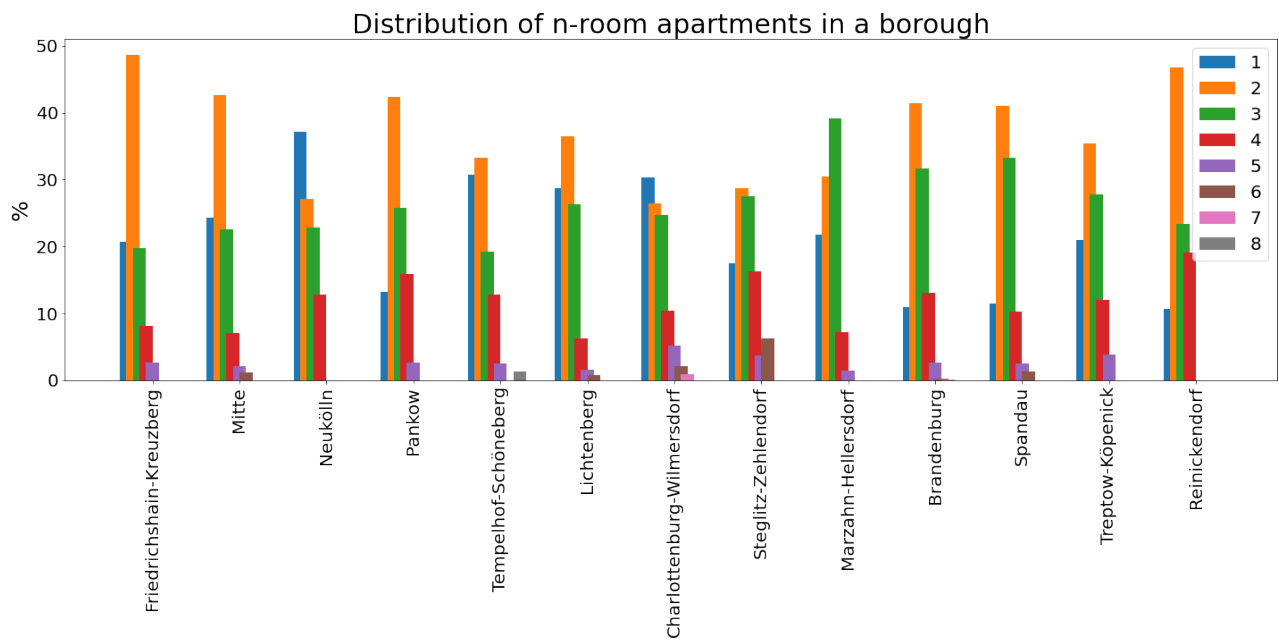
Figure 1: Distribution of n-room apartments in a borough. The colors correspond to the number of rooms.

3.3 Average price of n-room apartments

Figure 2 shows the average cold rent of n-room apartments in the boroughs. It is only intended to show the price segment in which apartments are available. The big outlier in Charlottenburg-Wilmersdorf for 7-room apartments can be seen. However, it should be noted that there are only two apartments with 9900 € each. An overview for 1 and 2-room apartments is given in Figure 3. Here one recognizes that one pays on the average 400€ for a 1-room apartment and 600€ for a 2-room apartment in the lower price segment. Neukölln, Marzahn-Hellersdorf and Brandenburg are the "boroughs" to be named here. Although Brandenburg is not a borough, it was included in the evaluation as an overview. It is interesting to see that although Neukölln is in the lower price category for 1-room apartments, it is in the upper-middle for 2-room apartments.
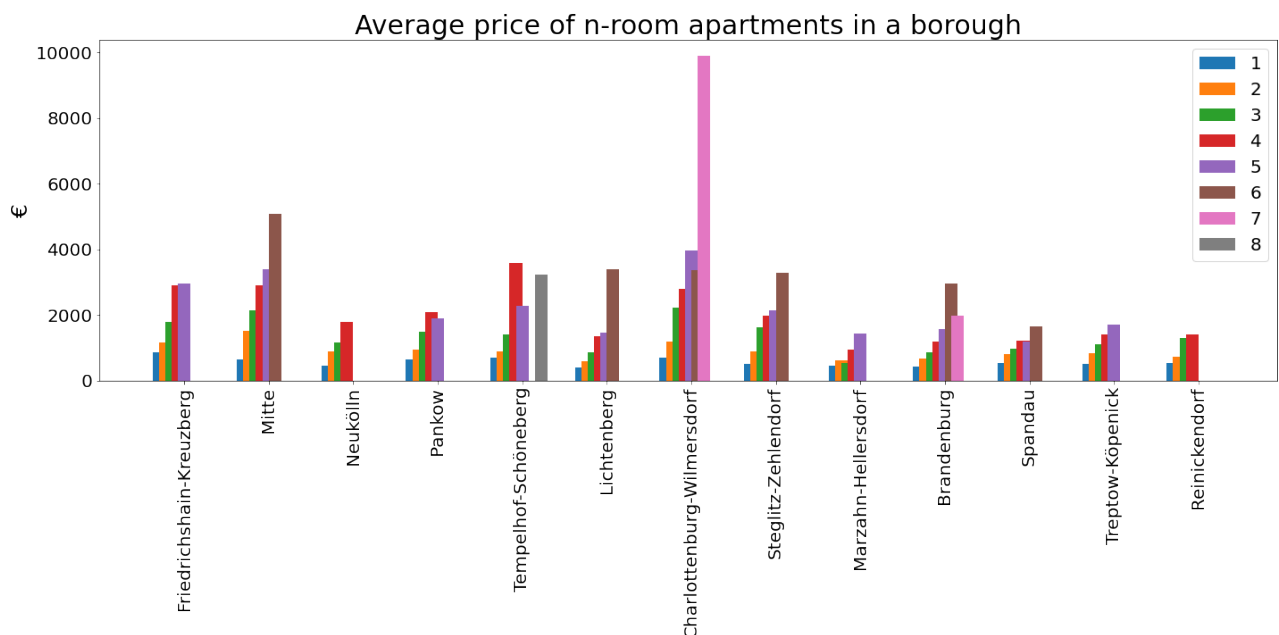

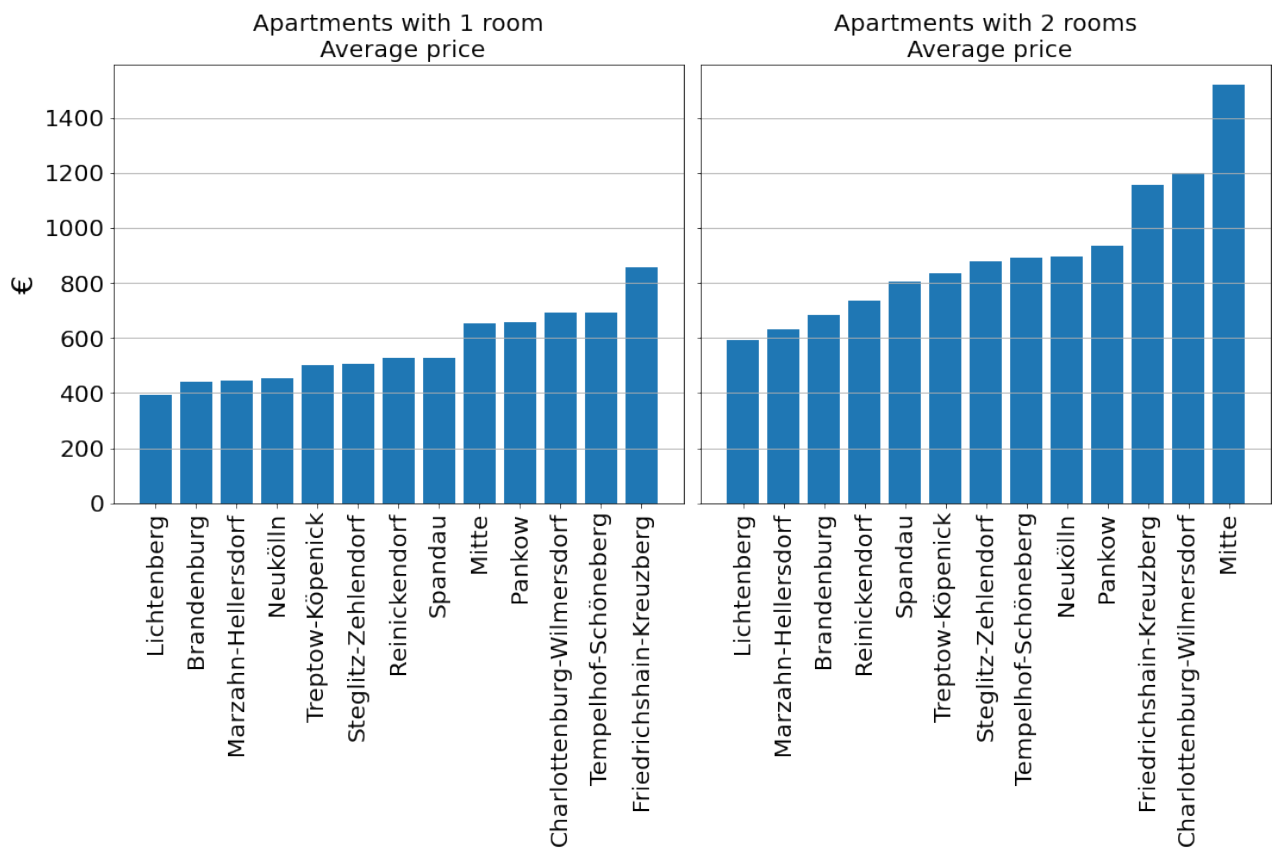*Figure 2: Average price of n-room apartments in a borough. The colors correspond to the number of rooms.*

Figure 3: Average price of 1-room (left) and 2-room (right) apartments in a borough.
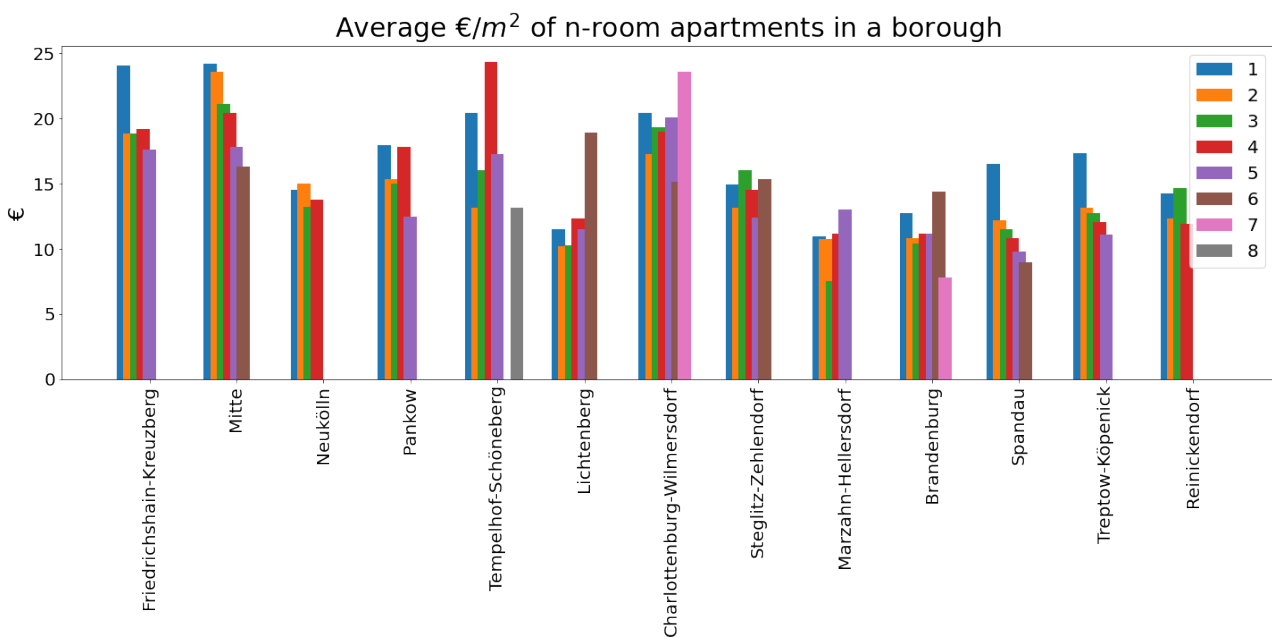

Figure 4: Average price per square meter of n-room apartments in a borough. The colors correspond to the number of rooms.

Since the price of apartments depends not only on the number of rooms but also on the area offered, figure 4 shows the average price per square meter of n-room apartments in the boroughs. Extraordinary attention should be paid to the fact that mostly the 1-room apartments have a higher price per square meter than multi-room apartments. In Friedrichshain-Kreuzberg and Mitte, the 1-room apartments are the most expensive. In general, Friedrichshain-Kreuzberg, Mitte, Tempelhof-Schöneberg and Charlottenburg-Wilmersdorf seem to be the more expensive areas. In many boroughs, the transition from a 1-room apartment to a 2-room apartment, when looking at the price per square meter, is steeper than for 2- to 3-room apartments. This means that whether you have a 2-room apartment or a 3-room apartment plays a marginal role in terms of price per square meter. They do not differ too much.

3.4 Distribution and number of apartments in a borough

Figure 5 shows the distribution of apartments in the respective boroughs. It can be seen that the distribution of the apartments on offer is not homogeneous and that there are areas in the boroughs that have a higher density of offers. Figure 6 shows the number of apartments on offer in the boroughs. The two major peaks of Brandenburg and Mitte are striking. There are very many free-standing apartments there, but Charlottenburg-Wilmersdorf and Treptow-Köpenick also offer a large number of apartments.
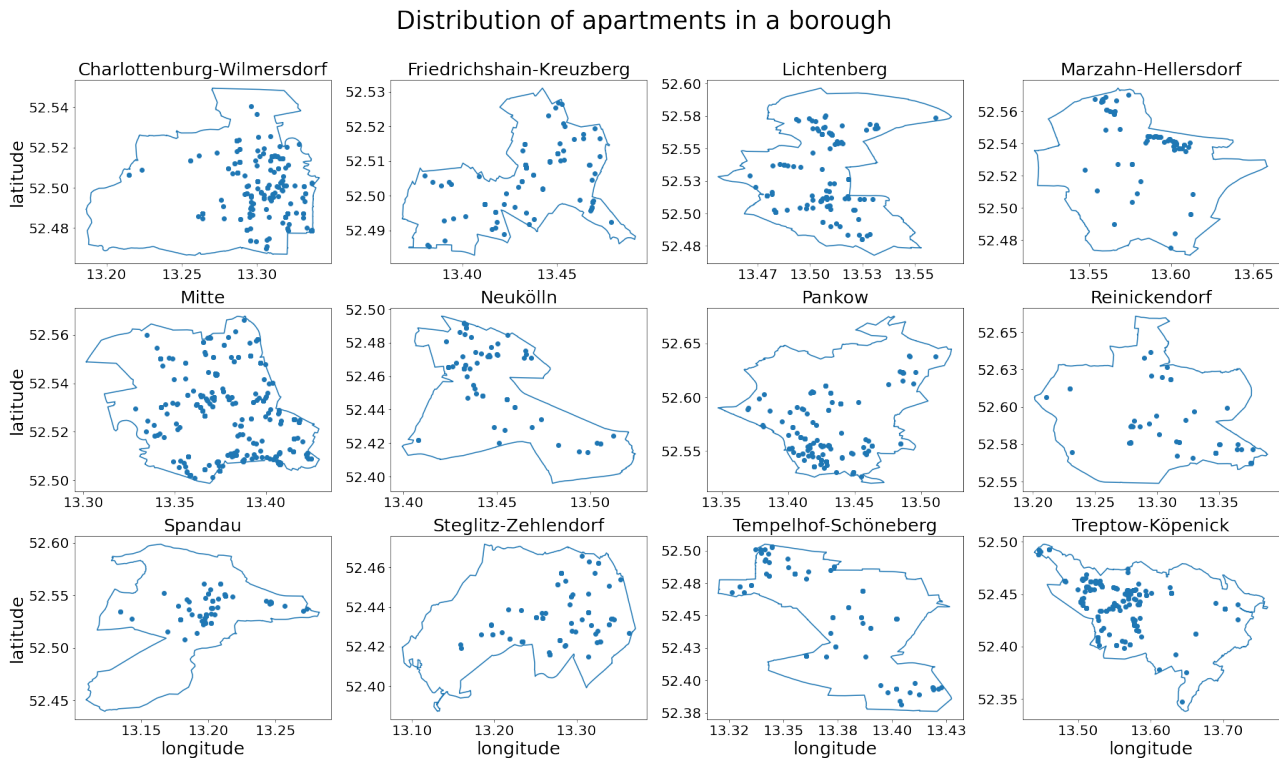

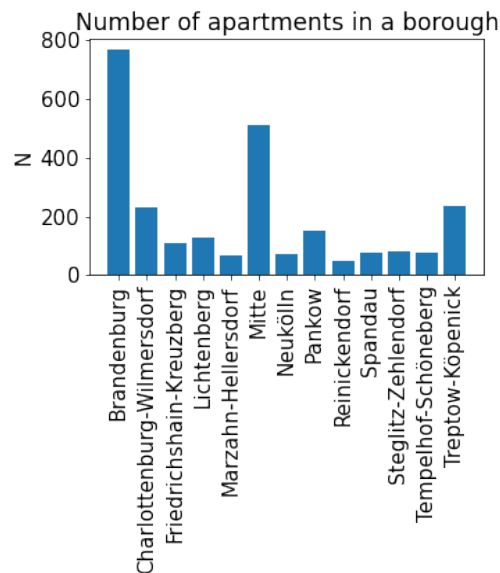*Figure 5: Distribution of apartments in a borough.*

Figure 6: Total number of apartments in a borough.

3.5 Summary and interpretation of the results

In most cases, far more 2-room apartments are on offer. Except in Neukölln, Charlottenburg-Wilmersdorf and Marzahn-Hellersdorf. The reason could be that there are people who either live alone and prefer 1-room apartments, or families who need more than two rooms. 4-8 rooms there are fewer, because the demand is probably not so large anyway, for it. Thus, apartments with fewer rooms are offered rather.

There are a lot of apartments on offer in Brandenburg and they are also cheap. The reason for this is that many people have a job in Berlin and thus do not want to live far outside in Brandenburg, because the way to work thus takes very long. Therefore, the apartments in Brandenburg are probably not as popular as those in Berlin and thus they are offered cheaper. Another reason why there are so many free apartments in Brandenburg is its size. Probably there is a not insignificant part of people who commute between Berlin and Brandenburg but live in houses in Brandenburg. However, this is not further illuminated here.

In Mitte there are also many apartments on offer, but they are very expensive compared to other places in Germany. The reason for the high prices is their location. Mitte is very popular, because from here you can quickly get to all places and there are many sights. But there are probably still so many detached apartments, because of the high apartment prices. Charlottenburg-Wilmersdorf has similar prices per square meter as the borough Mitte, but far fewer apartments to offer, even though it has a larger area. Since only the apartments currently on offer are considered here, and no information is available about the inhabited apartments, an interpretation here is very difficult. The difference, however, is that in Charlottenburg-Wilmersdorf the 1-3 room apartments are more or less equally distributed. A slightly higher number of 1-room apartments can be seen, but there are still a similar number of 1,2 and 3-room apartments, in contrast to Mitte. In Mitte, however, 2-room apartments dominate. This could be one reason why there are far more apartments available in Mitte than in Charlottenburg-Wilmersdorf. However, the distribution of n-room apartments looks similar to that of Mitte when one considers Friedrichshain-Kreuzberg, which is only slightly lower in price than Mitte and has a similar location. Another attempt to explain the many detached apartments in Mitte is that there are many tourists in this borough and people who are not from Berlin but still have an apartment there for their entertainment. People who live permanently in Berlin could therefore move to the surrounding boroughs. However, since just many leisure activities because of COVID-19 can not be realized, it may be that the people who want to be entertained in Mitte have canceled their apartments and now many are vacant.

If you are looking for a 1-room apartment that is affordable and not too far from the edge, you might find it in Lichtenberg. Here is the minimum of the average price for 1-room apartments. Marzahn-Hellersdorf is a little lower with the price per square meter, but it is also a little further out.

4. Venues and Clustering

4.1 Most frequent venues
As already described in chapter 2.2, the venues that are located within a radius of 500 m were determined by the Foursquare API. Afterward, one hot encoding was used to bring the data into a table and to finally determine the 10 most frequent venues for the boroughs, which can be seen in Table 1. It should be noted that these are not the venues that occur most frequently throughout the borough. This information is not given. It is only about the most frequent venues in the borough, which are located in the vicinity of the offered apartments.

It can be seen that the majority of the restaurants that can be found in the vicinity of the apartments are Italian restaurants. Directly behind them come the German restaurants. Understandably, there are many stores, but also cafes are still very common.

4.2 Clustering
If we let the k-means clustering algorithm search for three clusters, we can see in Figure 7 that the boroughs located in the center and south of Berlin (Charlottenburg-Wilmersdorf, Mitte, Friedrichshain-Kreuzberg, Tempelhof-Schöneberg and Neukölln) form a cluster (yellow) and the boroughs around them (Steglitz-Zehlendorf, Spandau, Reinickendorf, Pankow, Lichtenberg and Treptow-Köpenick) form another cluster (red). Marzahn-Hellersdorf forms the third cluster (purple). In the appendix there are more figures with different number of clusters. However, the same clusters usually remain in the center. It should also be noted that the cluster for Brandenburg is not shown because it is located around Berlin. Other cluster sizes are shown in the appendix.
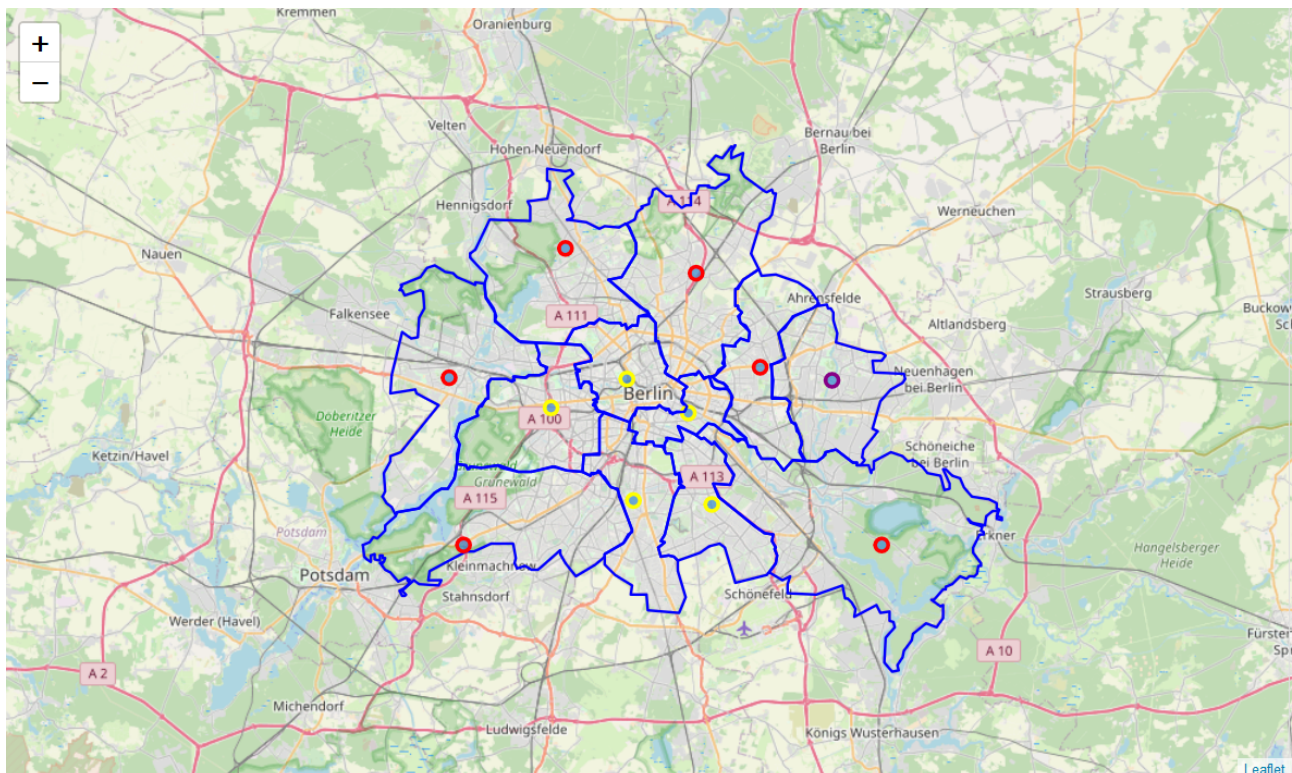

Figure 7: Three Clusters of venues in the boroughs of Berlin. The colors indicate the different clusters.

	1. most frequent venue	2. most frequent venue	3. most frequent venue	4. most frequent venue	5. most frequent venue	6. most frequent venue	7. most frequent venue	8. most frequent venue	9. most frequent venue	10. most frequent venue
Brandenburg	Hotel	German Restaurant	Supermarket	Café	Vacation Rental	Bakery	Italian Restaurant	Clothing Store	Drugstore	French Restaurant
Charlottenburg-Wilmersdorf	Hotel	Italian Restaurant	Supermarket	Café	Bakery	German Restaurant	Vietnamese Restaurant	Plaza	Chinese Restaurant	Trattoria/Osteria
Friedrichshain-Kreuzberg	Café	Bar	Coffee Shop	Italian Restaurant	Bakery	Nightclub	Hotel	Supermarket	Vegetarian / Vegan Restaurant	Pizza Place
Lichtenberg	Supermarket	Tram Station	Drugstore	Bakery	Café	Bus Stop	Park	Fast Food Restaurant	Shopping Mall	Hotel
Marzahn-Hellersdorf	Supermarket	Tram Station	Bus Stop	Bank	Bakery	Gym / Fitness Center	Multiplex	Shopping Mall	Climbing Gym	Drugstore
Mitte	Hotel	Coffee Shop	Café	Bakery	Italian Restaurant	Hotel Bar	Supermarket	Restaurant	Bar	Gym / Fitness Center
Neukölln	Bar	Café	Coffee Shop	Supermarket	Bakery	Bus Stop	Italian Restaurant	Breakfast Spot	Plaza	Turkish Restaurant
Pankow	Café	Supermarket	Bakery	Tram Station	Italian Restaurant	Coffee Shop	Vietnamese Restaurant	Organic Grocery	Park	Bar
Reinickendorf	Supermarket	Bakery	Restaurant	Café	Bus Stop	Italian Restaurant	German Restaurant	Indian Restaurant	Ice Cream Shop	Clothing Store
Spandau	Supermarket	Bus Stop	Pizza Place	Drugstore	Clothing Store	Italian Restaurant	Ice Cream Shop	Fast Food Restaurant	Hotel	Bakery
Steglitz-Zehlendorf	Café	Supermarket	Bus Stop	Italian Restaurant	Park	Drugstore	German Restaurant	Bakery	Bank	Organic Grocery
Tempelhof-Schöneberg	Supermarket	Café	Italian Restaurant	German Restaurant	Bus Stop	Hotel	Bakery	Plaza	Drugstore	Vietnamese Restaurant
Treptow-Köpenick	Supermarket	Café	Tram Station	Bakery	German Restaurant	Drugstore	Italian Restaurant	Ice Cream Shop	Pizza Place	Park

Table 1. Most frequent venues

5. Conclusion

In this study, I examined the prices of rental apartments in terms of their geographic location. Both the number of apartments offered and their price per square meter were analyzed. It was found out that there are a lot of detached 2-room apartments. It was also shown that apartments in the center of Berlin are more expensive than those on the outskirts. Far more favorably it is however outside of Berlin, thus in Brandenburg. This is also where most of the apartments are still available. But also in the borough Mitte, many are not yet occupied. Furthermore, the price per square meter for 1-room apartments is usually more expensive than for multi-room apartments. In Berlin and Brandenburg, Italian restaurants are very common. But also German restaurants are still relatively often represented. A supermarket is usually not far away. If you look at the venues, there are two main clusters and one secondary cluster, when you look at three clusters. The first cluster is the boroughs located in the center and the southern area and the second cluster is the boroughs located around the center. The third cluster consists of only one borough.

6. Future directions

Since in this study only apartments were considered, which were offered on immobilienscout24, only limited statements about the popularity of boroughs could be made. Future projects can use even more listings from different providers to be able to make a more accurate statement. In addition, it would be helpful to look not only at the apartments currently on offer but also at those that are occupied. For example, it has been shown that it is mostly the 2-room apartments that are vacant. Now, however, the question arises whether this is because few people are living in 2-room apartments or whether far more 2-room apartments have been built in Berlin. It would be interesting to know the distribution of the total number of apartments built. This could be especially important for the further planning of new housing. Based on these results alone, it would make sense to build fewer new 2-room apartments and more 1 and 3-room apartments. The knowledge about the number of families and single persons and their number of rooms, the price and their location could also bring a great gain in knowledge.
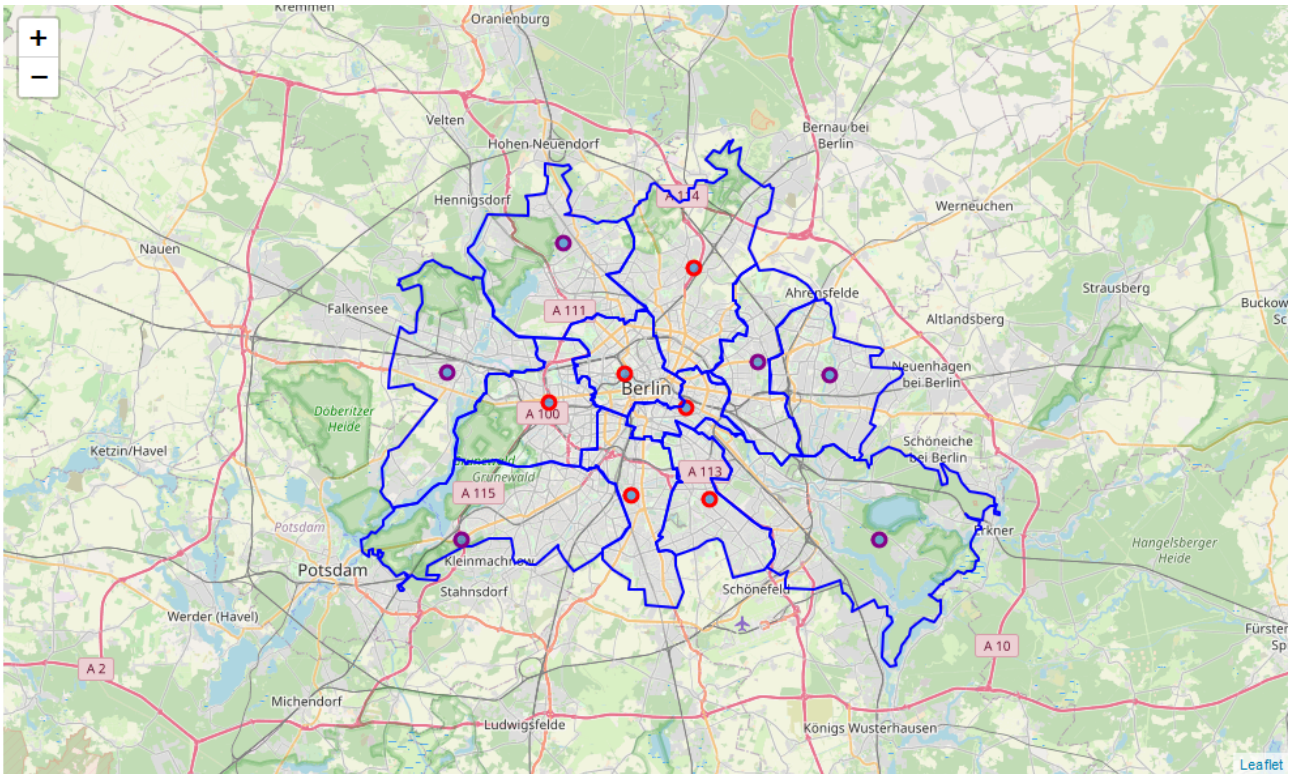
Appendix

Figure 8: Two Clusters of venues in the boroughs of Berlin. The colors indicate the different clusters.
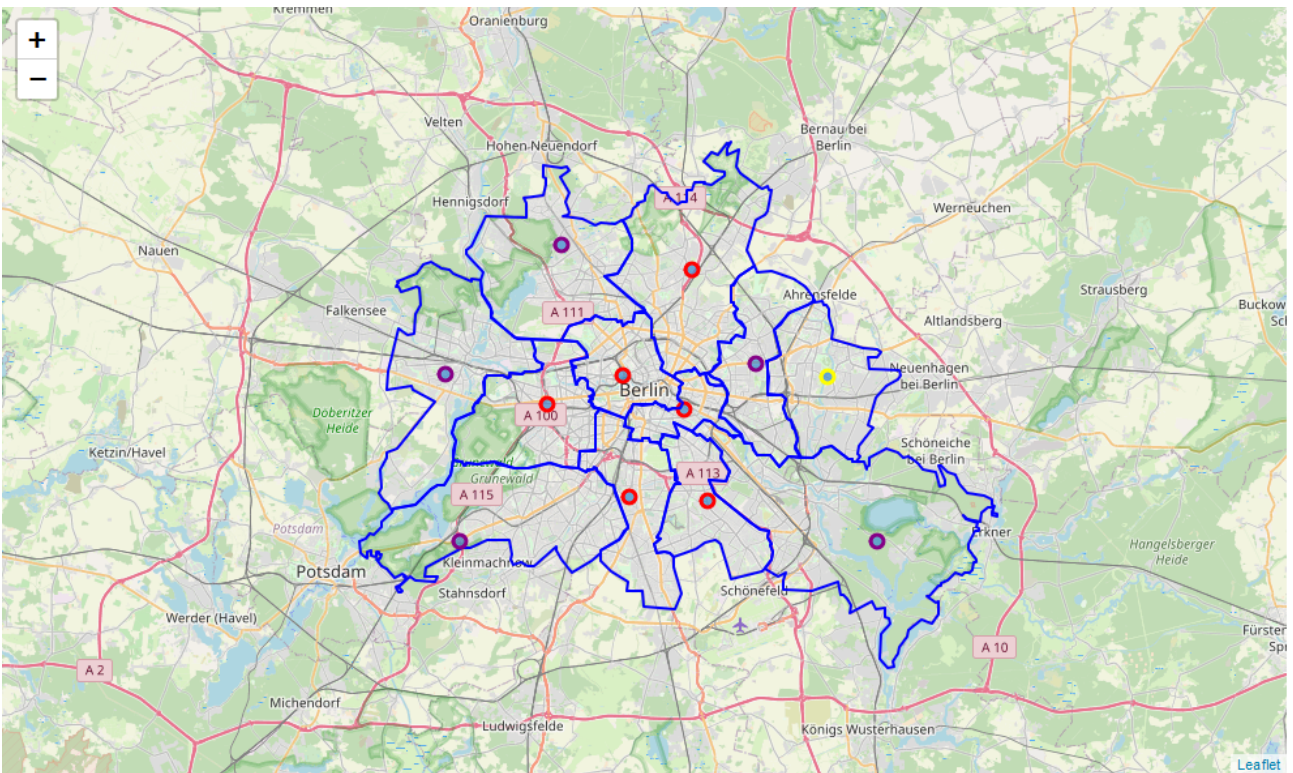

Figure 9: Four Clusters of venues in the boroughs of Berlin. The colors indicate the different clusters.
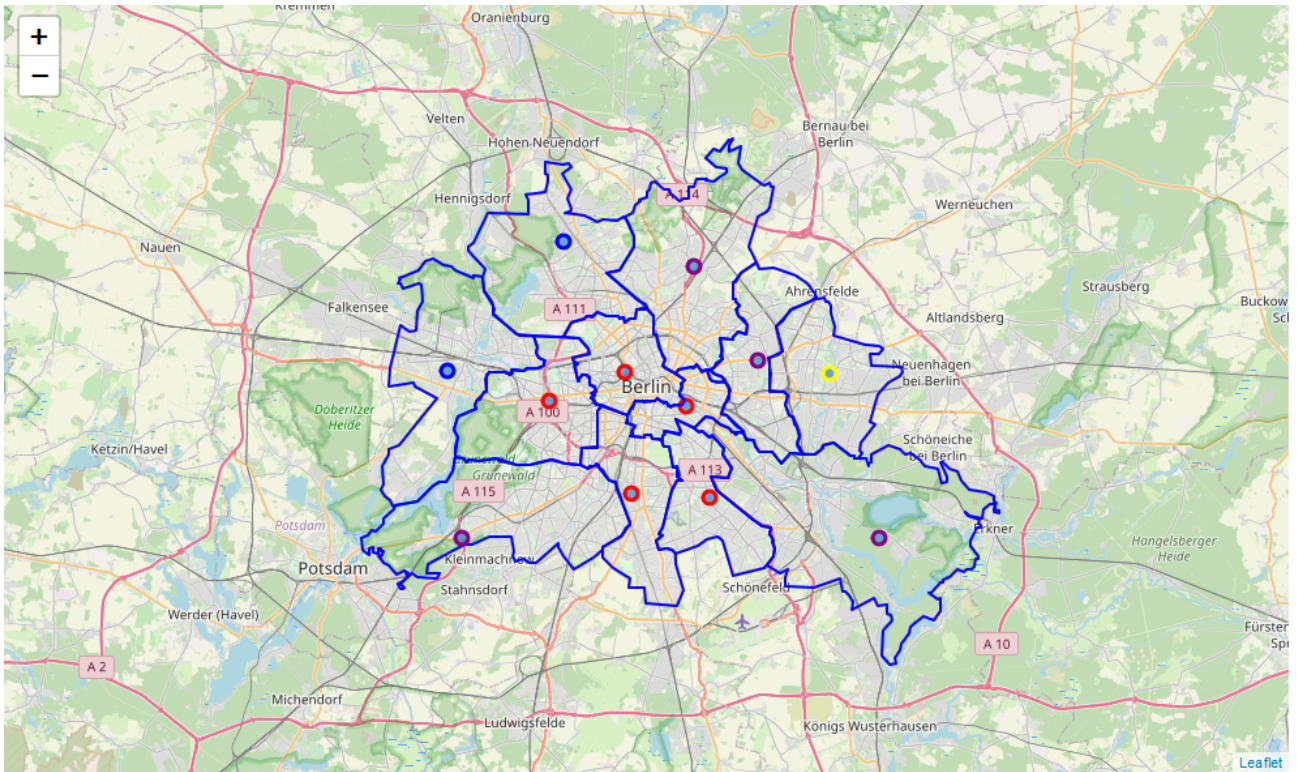
Figure 10: Five Clusters of venues in the boroughs of Berlin. The colors indicate the different clusters.