

Most Expensive Neighborhoods and Areas in Berlin

Dr. Lei Zhang

May 22, 2021



1 Table of contents

- Introduction: Business Problem
- Data
- Methodology
- Analysis
- Results and Discussion
- Conclusion

2 Introduction: Business Problem

There is a saying about Berlin: "arm aber sexy". "arm" is a German adjective meaning "poor". Why is Berlin "arm"? Berlin was once a very rich city, but it became poor because it, as the capital of Germany, much more involved in the German history than any other city in Germany. It was tragically occupied and divided. However, after the fall of the Berlin Wall, Berlin is recovering very fast. Now Berlin, as the biggest city in Germany with more than 3.7 million residents, as a

city full of histories, as an international multicultural metropolis, becomes one of the sexist city in Germany.

In this report, we will try to find out where in Berlin have the highest rent *in the next a few years*. Specifically, this report will be targeted to stakeholders interested in investing an apartment house in Berlin.

We will try to detect

- locations that currently have the highest rent;
- locations that have or are expected to have the fastest rent growth.

We will use our data to generate a few most promising neighborhoods and areas based on these criteria. A ranking list of the regions will be provided and the perspectives will be analysed, so that best possible final location can be chosen by stakeholders.

3 Data

Based on definition of our problem, factors that will influence our decision are:

- the current regional average monthly rents per square meter for an apartment exclusive of heating and other additional costs in neighborhoods/areas of Berlin;
- the growth of the regional average monthly rents per square meter for an apartment exclusive of heating and other additional costs in neighborhoods/areas of Berlin.

While it is easy to find out the average rent for each borough, it is not easy to find out the average rent of a specific area, e.g. the average rent within a radius of 500 m from Berlin-Alexanderplatz. Also it is relatively easy to collect data showing the growth of the average rent of each borough over the years, but it is not easy to get those data for each neighborhood/area. We are going to analyse the collected data for Berlin boroughs and some indirectly related data for the neighborhoods and areas to estimate the more specific rents and their growth.

Following data sources will be needed to extract/generate the desired information:

- general information about Berlin boroughs and neighborhoods will be extracted from the Wikipedia site [Verwaltungsgliederung Berlins](#) through web scraping;
- average rent of each borough of Berlin from 2009 to 2020 will be extracted from pdf documents on the government website [Berliner Wohnungsmarktbericht](#);
- demographic data about the population of each borough of Berlin from 2010 to 2019 will be extracted from pdf documents on the government website [Berliner Wohnungsmarktbericht](#);
- coordinate of Berlin center and the center of each borough of Berlin will be obtained using the [Nominatim API](#);
- **FourSquare API** will be used to get the details and types of venues in the vicinity of a neighborhood of Berlin.

3.1 General Information about Boroughs and Neighborhoods of Berlin

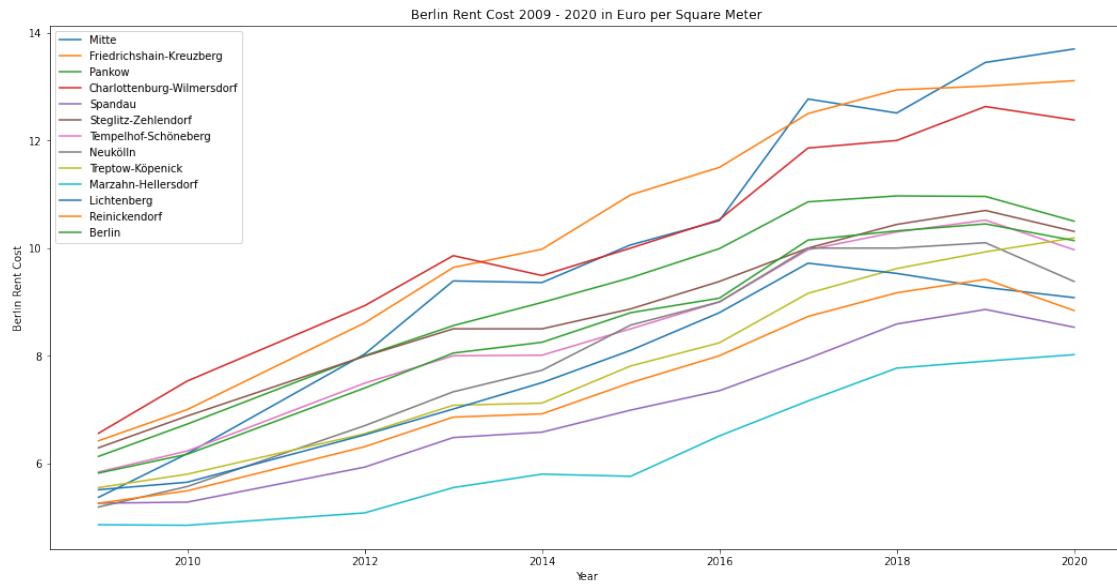
We first use webscraping to get some general information about boroughs and neighborhoods of Berlin from the Wikipedia site [Verwaltungsgliederung Berlins](#). Here is the table we get:

	Borough		Neighborhood	Area	Population	Density
0	Mitte	Mitte, Moabit, Hansaviertel, Tiergarten, Wedd...	3947.0	385.748	9.773	
1	Friedrichshain-Kreuzberg	Friedrichshain, Kreuzberg	2016.0	290.386	14.404	
2	Pankow	Prenzlauer Berg, Weißensee, Blankenburg, Heine...	10301.0	409.335	3.974	
3	Charlottenburg-Wilmersdorf	Charlottenburg, Wilmersdorf, Schmargendorf, Gr...	6472.0	343.592	5.309	
4	Spandau	Spandau, Haselhorst, Siemensstadt, Staaken, Ga...	9191.0	245.197	2.668	
5	Steglitz-Zehlendorf	Steglitz, Lichterfelde, Lankwitz, Zehlendorf, ...	10250.0	310.071	3.025	
6	Tempelhof-Schöneberg	Schöneberg, Friedenau, Tempelhof, Mariendorf, ...	5309.0	350.984	6.611	
7	Neukölln	Neukölln, Britz, Buckow, Rudow, Gropiusstadt	4493.0	329.917	7.343	
8	Treptow-Köpenick	Alt-Treptow, Plänterwald, Baumschulenweg, Joha...	16842.0	273.689	1.625	
9	Marzahn-Hellersdorf	Marzahn, Biesdorf, Kaulsdorf, Mahlsdorf, Helle...	6174.0	269.967	4.373	
10	Lichtenberg	Friedrichsfelde, Karlshorst, Lichtenberg, Falk...	5229.0	294.201	5.626	
11	Reinickendorf	Reinickendorf, Tegel, Konradshöhe, Heiligensee...	8946.0	266.408	2.978	

Then we manually input the rent of each borough of Berlin from 2009 to 2020 from the pdf files on the website [Berliner Wohnungsmarktbericht](#).

	Year	Mitte	Friedrichshain-Kreuzberg	Pankow	Charlottenburg-Wilmersdorf	Spandau	Steglitz-Zehlendorf	Tempelhof-Schöneberg	Neukölln	Treptow-Köpenick	Marzahn-Hellersdorf	Lichtenberg	Reinickendorf	Berlin	
0	2009	5.37		6.42	6.13		6.56	5.26	6.29	5.84	5.19	5.55	4.86	5.51	5.26 5.82
1	2010	6.18		7.00	6.73		7.53	5.28	6.88	6.23	5.57	5.80	4.85	5.65	5.49 6.17
2	2012	8.03		8.61	8.00		8.93	5.93	7.99	7.49	6.70	6.55	5.08	6.53	6.31 7.40
3	2013	9.39		9.64	8.56		9.86	6.48	8.50	8.00	7.33	7.08	5.55	7.01	6.86 8.05
4	2014	9.36		9.98	8.99		9.49	6.58	8.50	8.01	7.73	7.12	5.80	7.50	6.92 8.25
5	2015	10.06		10.99	9.45		10.00	6.99	8.87	8.50	8.57	7.81	5.76	8.10	7.50 8.80
6	2016	10.51		11.50	9.99		10.53	7.35	9.38	9.00	9.00	8.24	6.51	8.80	8.00 9.07
7	2017	12.77		12.50	10.86		11.86	7.95	10.00	9.97	10.00	9.16	7.16	9.72	8.73 10.15
8	2018	12.51		12.94	10.97		12.00	8.59	10.44	10.30	10.00	9.62	7.77	9.53	9.17 10.32
9	2019	13.45		13.01	10.96		12.63	8.86	10.70	10.52	10.10	9.93	7.90	9.27	9.42 10.45
10	2020	13.70		13.11	10.50		12.38	8.53	10.31	9.97	9.38	10.19	8.02	9.08	8.84 10.14

We can take a preview of the visualized data.

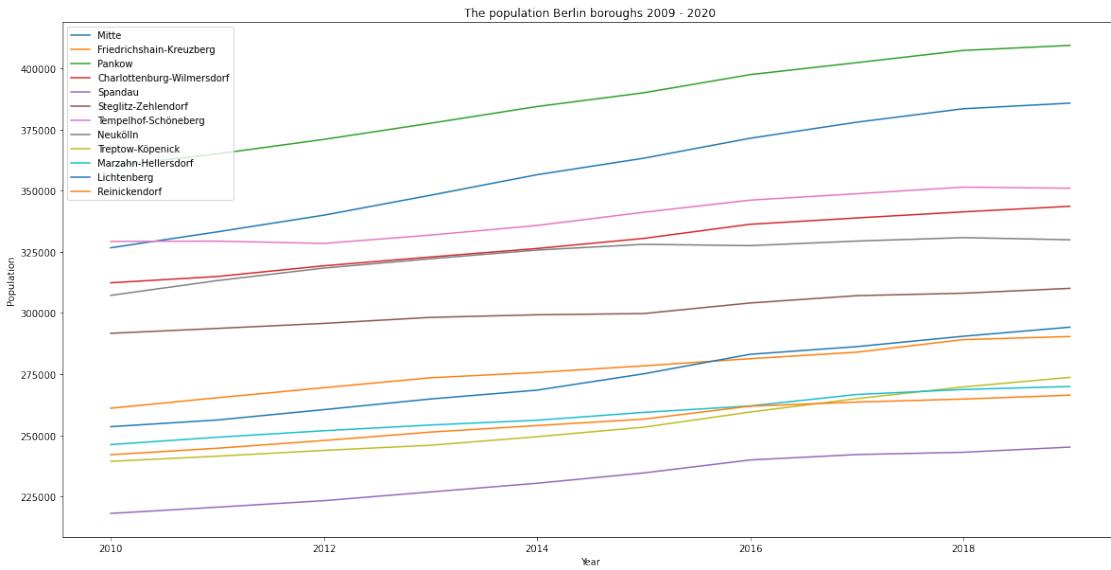


The green line is the average rent of the whole of Berlin. Apparently, the three Boroughs Berlin-Mitte, Berlin-Friedrichshain-Kreuzberg, Berlin-Charlottenburg-Wilmersdorf are above average, and the others are below average.

Then we collect the demographic data about the population of each borough of Berlin from 2010 to 2019 from the same source.

	Year	Mitte	Friedrichshain-Kreuzberg	Pankow	Charlottenburg-Wilmersdorf	Spandau	Steglitz-Zehlendorf	Tempelhof-Schöneberg	Neukölln	Treptow-Köpenick	Marzahn-Hellersdorf	Lichtenberg	Reinickendorf			
0	2010	326650		261090	360017		312340	218094	291700		329245	307204	239374	246225	253539	242084
1	2011	333152		265361	365021		314911	220645	293692		329361	313245	241487	249232	256280	244727
2	2012	339974		269471	370937		319289	223305	295746		328428	318356	243844	251879	260505	247887
3	2013	348102		273517	377532		322870	226868	298200		331822	322153	245951	254226	264858	251325
4	2014	356506		275691	384367		326354	230419	299268		335767	325716	249440	256173	268465	254000
5	2015	363236		278393	389976		330468	234630	299765		341161	328062	253333	259373	275142	256617
6	2016	371407		281323	397406		336249	239942	304086		346108	327522	259524	262015	283121	261919
7	2017	377965		283974	402289		338831	242143	307076		348739	329387	264999	266684	286246	263597
8	2018	383457		289120	407289		341327	243080	308077		351429	330786	269775	268739	290493	264826
9	2019	385748		290386	409335		343592	245197	310071		350984	329917	273689	269967	294201	266408

Here is the visualization:

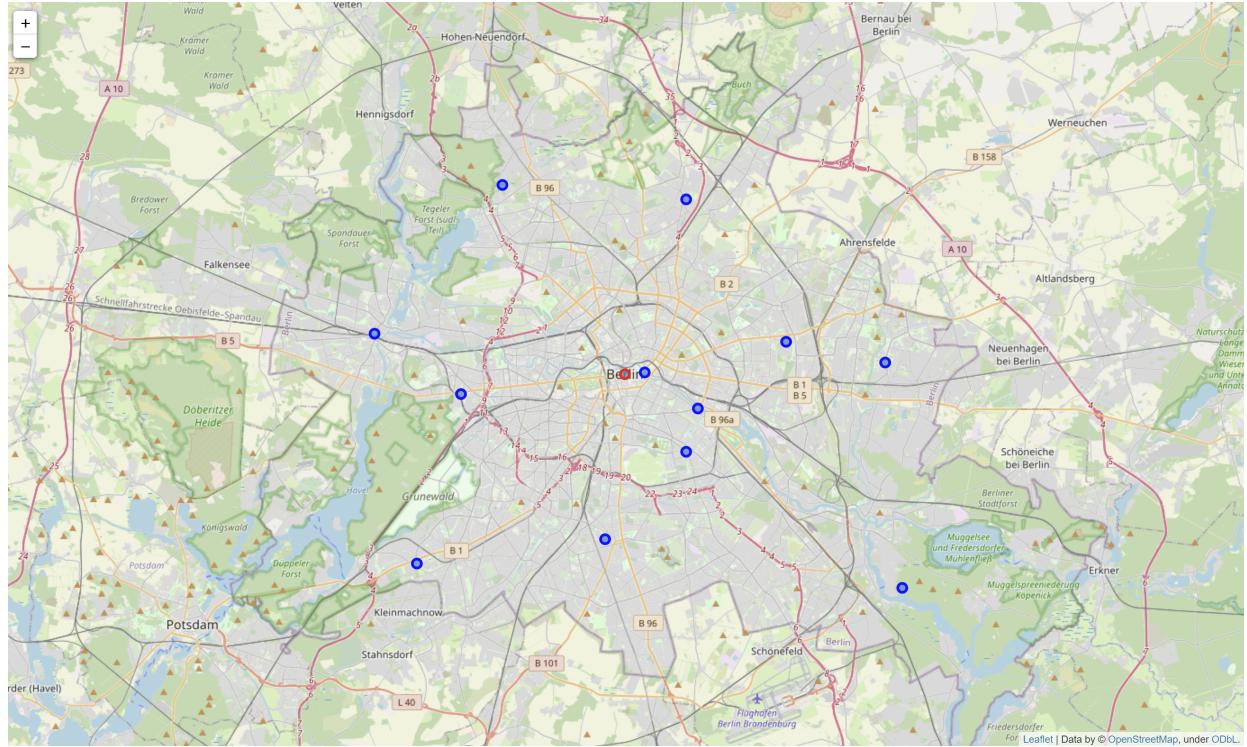


3.2 City, Borough, Neighborhood Coordinates

We get latitude & longitude coordinates of the center of Berlin and the center of each borough of Berlin using the **Nominatim API** and store them into a Pandas DataFrame.

	Borough	Latitude	Longitude
0	Mitte	52.517885	13.404060
1	Friedrichshain-Kreuzberg	52.501115	13.444285
2	Pankow	52.597917	13.435316
3	Charlottenburg-Wilmersdorf	52.507856	13.263952
4	Spandau	52.535788	13.197792
5	Steglitz-Zehlendorf	52.429205	13.229974
6	Tempelhof-Schöneberg	52.440603	13.373703
7	Neukölln	52.481150	13.435350
8	Treptow-Köpenick	52.417893	13.600185
9	Marzahn-Hellersdorf	52.522523	13.587663
10	Lichtenberg	52.532161	13.511893
11	Reinickendorf	52.604763	13.295287
12	Berlin	52.517037	13.388860

Then we plot these locations into a map to get an initial impression.



The middle red circle is the center of Berlin. It is the crosspoint of **Friedrichstraße** and **Unter den Linden**. The blue ones are the center of the boroughs.

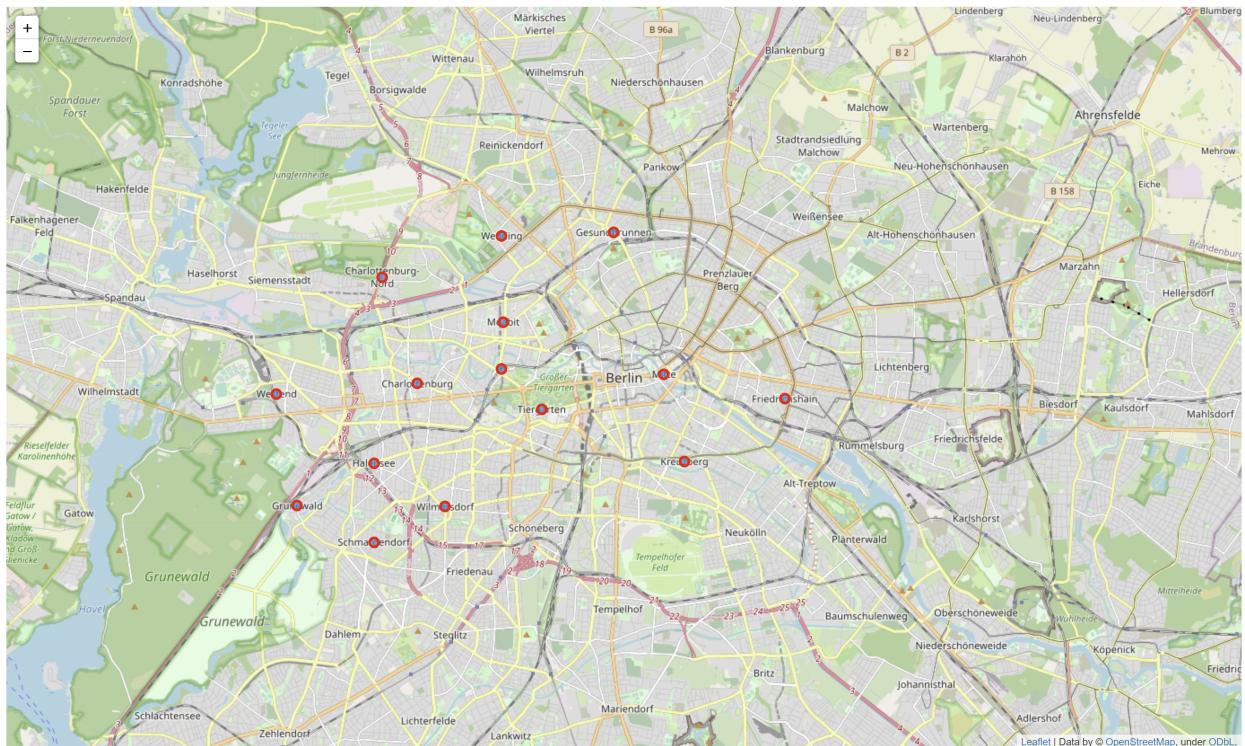
This amplifies our data:

	Borough	Neighborhood	Area	Population	Density	Latitude	Longitude
0	Mitte	Mitte, Moabit, Hansaviertel, Tiergarten, Wedd...	3947.0	385.748	9.773	52.517885	13.404060
1	Friedrichshain-Kreuzberg	Friedrichshain, Kreuzberg	2016.0	290.386	14.404	52.501115	13.444285
2	Pankow	Prenzlauer Berg, Weißensee, Blankenburg, Heine...	10301.0	409.335	3.974	52.597917	13.435316
3	Charlottenburg-Wilmersdorf	Charlottenburg, Wilmersdorf, Schmargendorf, Gr...	6472.0	343.592	5.309	52.507856	13.263952
4	Spandau	Spandau, Haselhorst, Siemensstadt, Staaken, Ga...	9191.0	245.197	2.668	52.535788	13.197792
5	Steglitz-Zehlendorf	Steglitz, Lichterfelde, Lankwitz, Zehlendorf, ...	10250.0	310.071	3.025	52.429205	13.229974
6	Tempelhof-Schöneberg	Schöneberg, Friedenau, Tempelhof, Mariendorf, ...	5309.0	350.984	6.611	52.440603	13.373703
7	Neukölln	Neukölln, Britz, Buckow, Rudow, Gropiusstadt	4493.0	329.917	7.343	52.481150	13.435350
8	Treptow-Köpenick	Alt-Treptow, Plänterwald, Baumschulenweg, Joha...	16842.0	273.689	1.625	52.417893	13.600185
9	Marzahn-Hellersdorf	Marzahn, Biesdorf, Kaulsdorf, Mahlsdorf, Helle...	6174.0	269.967	4.373	52.522523	13.587663
10	Lichtenberg	Friedrichsfelde, Karlshorst, Lichtenberg, Falk...	5229.0	294.201	5.626	52.532161	13.511893
11	Reinickendorf	Reinickendorf, Tegel, Konradshöhe, Heiligensee...	8946.0	266.408	2.978	52.604763	13.295287

Now, we'll get the coordinates for the neighborhoods. We will only focus on the neighborhoods that we are interested in, namely those whose boroughs have an above average rent.

	Borough	Neighborhood	Latitude	Longitude
0	Mitte	Mitte	52.517885	13.404060
1	Mitte	Moabit	52.530102	13.342542
2	Mitte	Hansaviertel	52.519123	13.341872
3	Mitte	Tiergarten	52.509778	13.357260
4	Mitte	Wedding	52.550123	13.341970
5	Mitte	Gesundbrunnen	52.550920	13.384846
6	Friedrichshain_Kreuzberg	Friedrichshain	52.512215	13.450290
7	Friedrichshain_Kreuzberg	Kreuzberg	52.497644	13.411914
8	Charlottenburg_Wilmersdorf	Charlottenburg	52.515747	13.309683
9	Charlottenburg_Wilmersdorf	Wilmersdorf	52.487115	13.320330
10	Charlottenburg_Wilmersdorf	Schmargendorf	52.478902	13.292996
11	Charlottenburg_Wilmersdorf	Grunewald	52.487347	13.263754
12	Charlottenburg_Wilmersdorf	Westend	52.513399	13.255842
13	Charlottenburg_Wilmersdorf	Charlottenburg-Nord	52.540525	13.296266
14	Charlottenburg_Wilmersdorf	Halensee	52.497226	13.292999

Again, we will plot the neighborhoods into a map to get some impression.



3.3 Foursquare

Now we have our location candidates. We'll then use Foursquare API to get info on the venues in each neighborhood. In particular, we are interested in the nearby hotels.

This concludes the data gathering phase - we're now ready to use this data for analysis to produce the report!

4 Methodology

The goal of this project is to find out the neighborhoods and areas of Berlin that would have the highest rent in the next a few years. Therefore, two factors are extremely import. One is the current average rent, the other is the expected growth of the rent. From the preliminary look on the data that we have collected, it is clear that currently Berlin-Mitte has the highest average rent - 13.70 EUR every square meter, Berlin-Friedrichshain-Kreuzberg comes with 13.11 EUR the second, and Berlin-Charlottenburg-Wilmersdorf with 12.38 EUR is also above average. Then we will come to the growth. It is quite clear from the line chart that none of the boroughs of Berlin whose average rents are below the Berlin-average could have a chance to climbe up to the top. Thus it is reasonable for us to only focus on these three above mentioned boroughs.

In first step we will use **simple linear regression models** to estimate the growth of the average rent of each of the three boroughs and Berlin as a whole from 2009 to 2020. We visualize the results and put them together to compare.

The second step of our analysis will be to use **simple linear regression models** to estimate the growth of the population of each of the three boroughs and Berlin as a whole from 2009 to 2020. We also visualize the results and compare them.

In third step we will create **clusters of locations based on the patterns their nearby venues (k-means clustering and density-based clustering)** for all the neighborhoods in the three boroughs having in mind that areas with similar patterns of nearby venues should have similar average rent. We will analyse the similarities among each cluster.

In the fourth step we will use the patterns that we discovered to estmitie which areas of Berlin should have the highest rent in the next a few years.

In the last step we will use the data collected from [Berliner Mietspiegel](#) to test our estimation.

5 Analysis

In this simple linear regression model we use the years as independent variable to estimate the dependent variables - the average rent of each borough.

Berlin-Coefficients: [[0.44519635]]

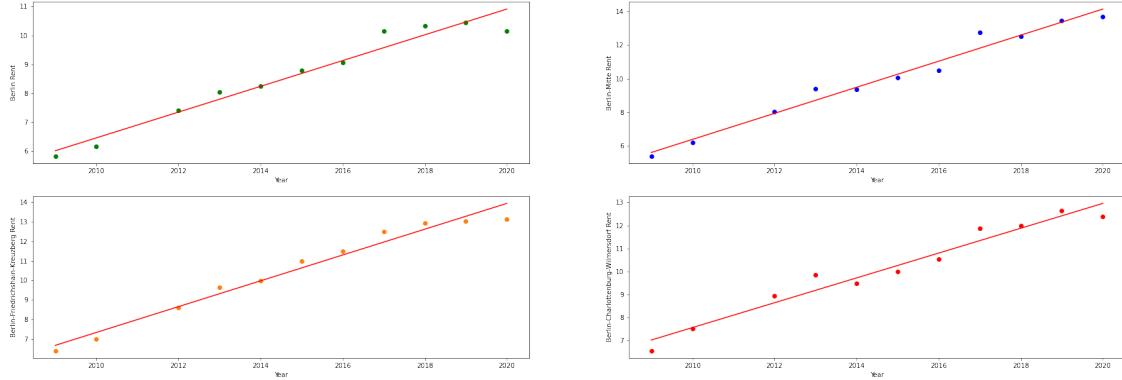
Berlin-Mitte-Coefficients: [[0.77734222]]

Berlin-Friedrichshain-Kreuzberg-Coefficients: [[0.65827489]]

Berlin-Charlottenburg-Wilmersdorf-Coefficients: [[0.5398317]]

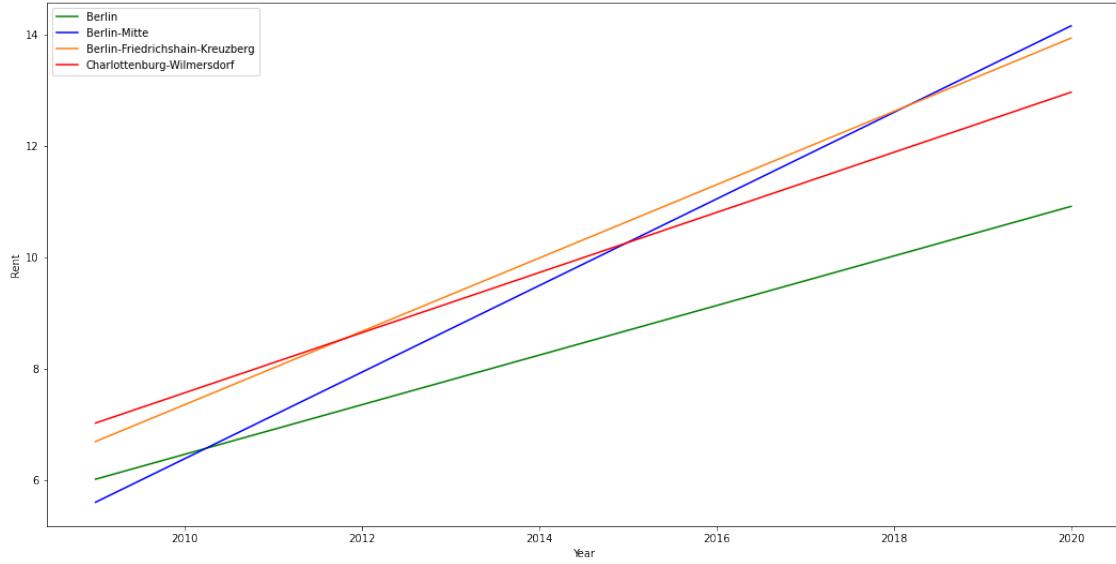
Here the coefficients indicate the rate of growth - the higher the number is, the faster it growth.

Let's visualize the result.



The models fit the samples quite well and no sign of overfitting. However, since the values only differ marginally, it is hard to compare them visually, even side by side!

So, we put them together!



Clearly, Berlin-Mitte has not only the highest rent, but also the fastest rent growth! Berlin-Friedrichshain-Kreuzberg in the second place with respect to both current rent and rent growth.

Next, we move to population growth.

Berlin-Mitte-Coefficients: 6973.436363636361

Berlin-Friedrichshain-Kreuzberg-Coefficients: 3203.7090909090903

Berlin-Charlottenburg-Wilmersdorf-Coefficients: 3685.70303030303

Clearly, Berlin-Mitte has the fastest population growth, and Berlin-Charlottenburg-Wilmersdorf is the second. But here one has to pay attention that population growth is not equal to population density growth, and the latter reflex more about the rent growth. Let's also consider the

population density growth by dividing the current data by the areas.

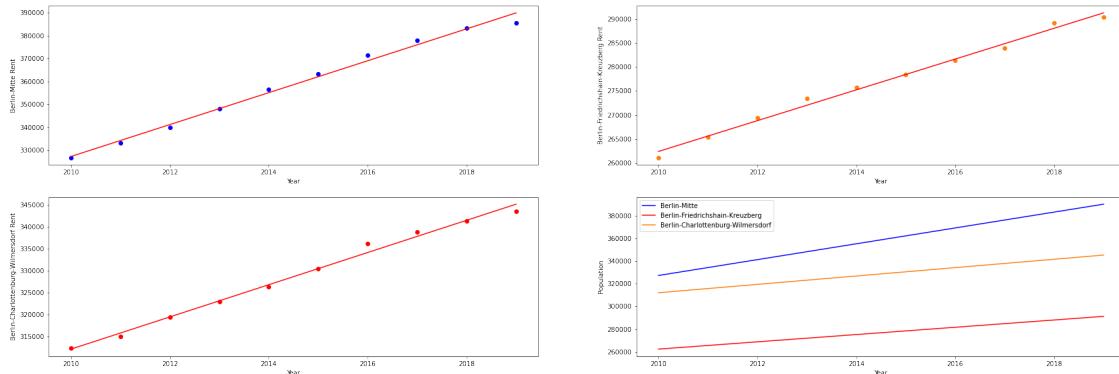
Berlin-Mitte-Coefficients: 1.7667687772070833

Berlin-Friedrichshain-Kreuzberg-Coefficients: 1.5891414141414126

Berlin-Charlottenburg-Wilmersdorf-Coefficients: 0.5694843989961417

Now our ranking comes back. It is consistent with the ranking of rent growth.

Next, let's visualize the population growth.



Now let's use **density-based clustering DBSCAN** to cluster the nearby venues. This clustering is only based on the locations of the venues, not base on the types of the venues. The advantage of DBSCAN is that it is robust against the outliers and you don't have to specify the number k which is an important parameter in the k-means clustering.

Here are the labels. The venues with label “-1” are the outliers.

```
set(labels)
```

```
{-1, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10}
```

The outliers are in:

Wedding
Schmargendorf
Grunewald
Westend
Charlottenburg-Nord

Let's now visualize the clusters on the maps.

The clusters and the coordinates of their centroids are:

Cluster 0, Mitte, coordinates: 52.52022415596093, 13.40376360485069.

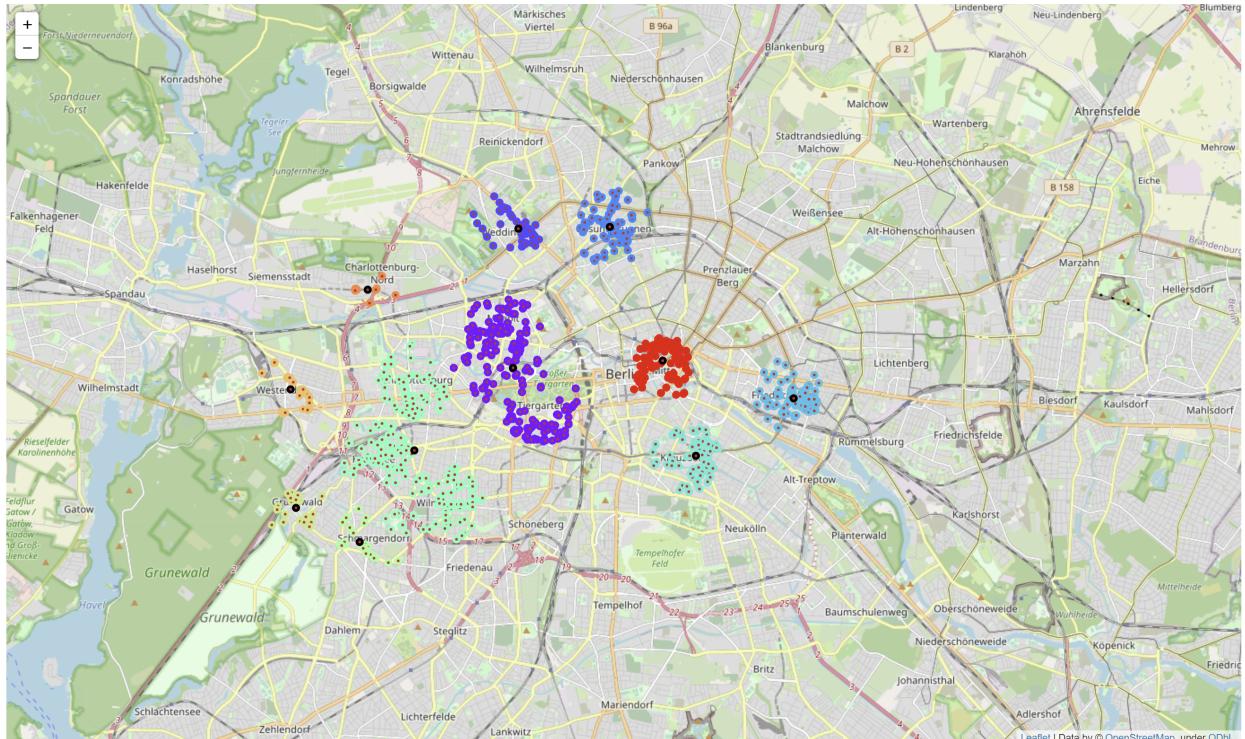
Cluster 1, Moabit, Hansaviertel, Tiergarten, Charlottenburg, coordinates:
52.5186084276777, 13.346268722550933.

Cluster 2, Wedding, coordinates: 52.55082119062809, 13.348438845426461.

Cluster 3, Gesundbrunnen, coordinates: 52.55122848288395, 13.383479520616007.

Cluster 4, Friedrichshain, coordinates: 52.51134647226065, 13.453956962818445.

Cluster 5, Kreuzberg, coordinates: 52.498114097915995, 13.416404598823647.
 Cluster 6, Charlottenburg, Wilmersdorf, Schmargendorf, Halensee, coordinates:
 52.49938592173205, 13.30867343748716.
 Cluster 7, Schmargendorf, coordinates: 52.47803669430278, 13.287706600213536.
 Cluster 8, Grunewald, coordinates: 52.485954325142174, 13.263261895735477.
 Cluster 9, Westend, coordinates: 52.51353499019267, 13.261082077240117.
 Cluster 10, Charlottenburg-Nord, coordinates: 52.53666861510058,
 13.29063831106081.

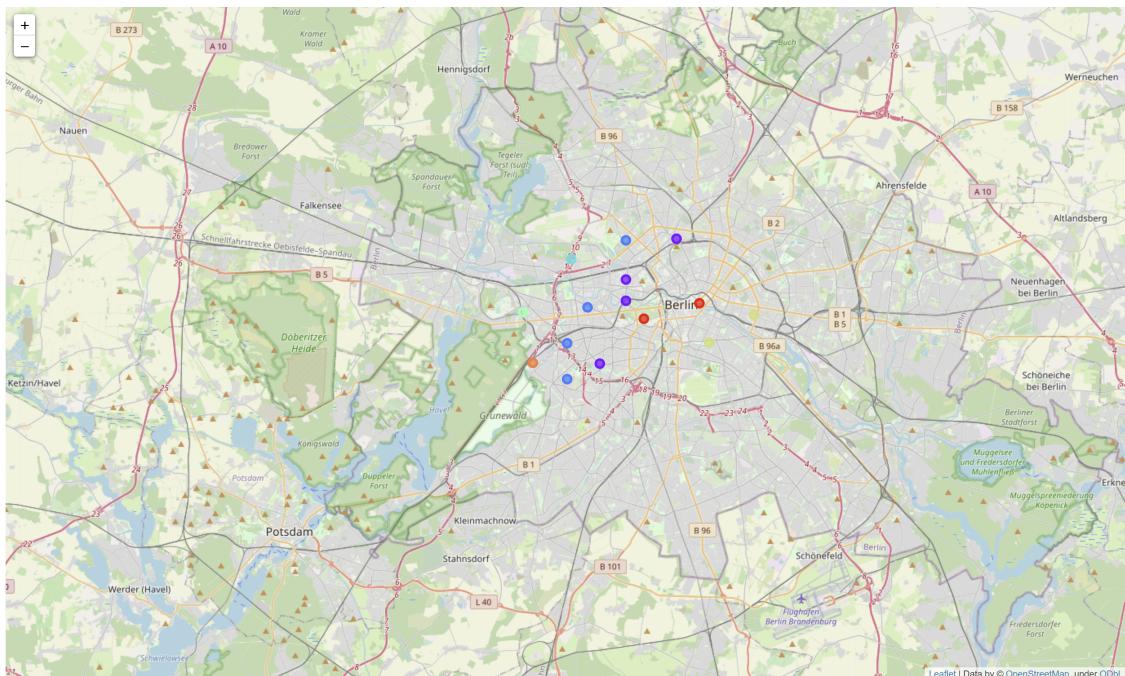


The result is no surprising. Since the clustering is based on the locations of the venues, the centroids of the clusters are basically the centroids of the neighborhoods, with exceptions cluster 1 and cluster 6 where large neighborhoods are clustered together. But we can still read some information out of this clustering: cluster 7, 8, 9, 10 have very few venues in them. In Berlin, this means simply that these are forests. Living in forests can also be expensive, but you would not expect that the rent rises sharply there, because forest stays as forest unless there is a forest fire. Also, neighborhoods with outliers, *i.e.* those who get “-1” in their labels, do not look promising, because in crowded areas, the venues are usually clustered together.

Now let's cluster the venues based on their types. We take the mean of the values according to each neighborhood to see which kind of venue is more popular in the neighborhood.

	Borough	Neighborhood	Latitude	Longitude	Cluster Labels
0	Mitte	Mitte	52.517885	13.404060	0
1	Mitte	Moabit	52.530102	13.342542	1
2	Mitte	Hansaviertel	52.519123	13.341872	1
3	Mitte	Tiergarten	52.509778	13.357260	0
4	Mitte	Wedding	52.550123	13.341970	2
5	Mitte	Gesundbrunnen	52.550920	13.384846	1
6	Friedrichshain_Kreuzberg	Friedrichshain	52.512215	13.450290	5
7	Friedrichshain_Kreuzberg	Kreuzberg	52.497644	13.411914	5
8	Charlottenburg_Wilmersdorf	Charlottenburg	52.515747	13.309683	2
9	Charlottenburg_Wilmersdorf	Wilmersdorf	52.487115	13.320330	1
0	Charlottenburg_Wilmersdorf	Schmargendorf	52.478902	13.292996	2
1	Charlottenburg_Wilmersdorf	Grunewald	52.487347	13.263754	6
2	Charlottenburg_Wilmersdorf	Westend	52.513399	13.255842	4
3	Charlottenburg_Wilmersdorf	Charlottenburg-Nord	52.540525	13.296266	3
4	Charlottenburg_Wilmersdorf	Halensee	52.497226	13.292999	2

Now we see that Berlin-Mitte-Mitte and Berlin-Mitte-Tiergarten are in the same cluster - cluster 0, Berlin-Friedrichshain and Berlin-Kreuzberg are in the same cluster 5. Based on what we have observed so far, there are good reasons to believe that cluster 0 is the top-tier, followed by cluster 5, and cluster 1 seems to be also competitive. Let's visualize the result.



Let's look into the pattern of the venues in each cluster.

	Borough	Neighborhood	Latitude	Longitude	Cluster Labels
0	Mitte	Mitte	52.517885	13.404060	0
1	Mitte	Moabit	52.530102	13.342542	1
2	Mitte	Hansaviertel	52.519123	13.341872	1
3	Mitte	Tiergarten	52.509778	13.357260	0
4	Mitte	Wedding	52.550123	13.341970	2
5	Mitte	Gesundbrunnen	52.550920	13.384846	1
6	Friedrichshain_Kreuzberg	Friedrichshain	52.512215	13.450290	5
7	Friedrichshain_Kreuzberg	Kreuzberg	52.497644	13.411914	5
8	Charlottenburg_Wilmersdorf	Charlottenburg	52.515747	13.309683	2
9	Charlottenburg_Wilmersdorf	Wilmersdorf	52.487115	13.320330	1
0	Charlottenburg_Wilmersdorf	Schmargendorf	52.478902	13.292996	2
1	Charlottenburg_Wilmersdorf	Grunewald	52.487347	13.263754	6
2	Charlottenburg_Wilmersdorf	Westend	52.513399	13.255842	4
3	Charlottenburg_Wilmersdorf	Charlottenburg-Nord	52.540525	13.296266	3
4	Charlottenburg_Wilmersdorf	Halensee	52.497226	13.292999	2

We see that in cluster 0, hotel is the most popular venue type and there is the museum venue type, especially in Berlin-Mitte-Mitte. This tells us that these two places lie at the heart of the city's tourist attractions. Thus very high average rent is expected in each of these neighborhoods. We also see that Berlin-Mitte-Hansaviertel from cluster 1 has hotel as its most popular venues, although there are no museums.

It seems that the density of hotels is a good indicator of the average rent of an area. We should look into that.

There are 208 hotels nearby.

There are 55 Hotels in the vicinity of Berlin-Mitte.

There are 12 Hotels in the vicinity of Berlin-Moabit.

There are 15 Hotels in the vicinity of Berlin-Hansaviertel.

There are 29 Hotels in the vicinity of Berlin-Tiergarten.

There are 2 Hotels in the vicinity of Berlin-Wedding.

There are 10 Hotels in the vicinity of Berlin-Gesundbrunnen.

There are 29 Hotels in the vicinity of Berlin-Friedrichshain.

There are 15 Hotels in the vicinity of Berlin-Kreuzberg.

There are 9 Hotels in the vicinity of Berlin-Charlottenburg.

There are 11 Hotels in the vicinity of Berlin-Wilmersdorf.

There are 2 Hotels in the vicinity of Berlin-Schmargendorf.

There are 2 Hotels in the vicinity of Berlin-Grunewald.

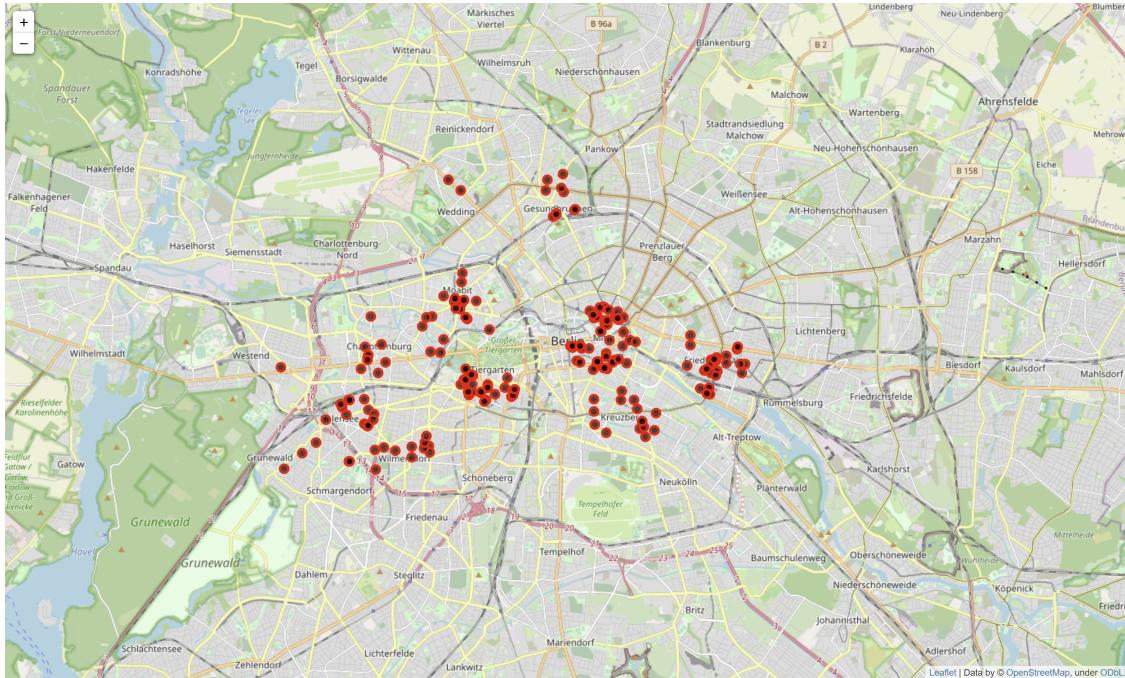
There are 1 Hotels in the vicinity of Berlin-Westend.

There are 0 Hotels in the vicinity of Berlin-Charlottenburg-Nord.

There are 16 Hotels in the vicinity of Berlin-Halensee.

As expected, Berlin-Mitte has the most hotels in its vicinity. Although Berlin-Tiergarten and Berlin-Friedrichshain have the same number of hotels in their vicinities, considering the small size of Berlin-Tiergarten, hotels are obviously denser in Tiergarten than in Friedrichshain.

Let's visualize these observations.



Let's use DBSCAN again to cluster the hotels based on their locations.

```
set(labels)
```

```
{-1, 0, 1, 2, 3, 4}
```

The clusters and the coordinates of their centroids are:

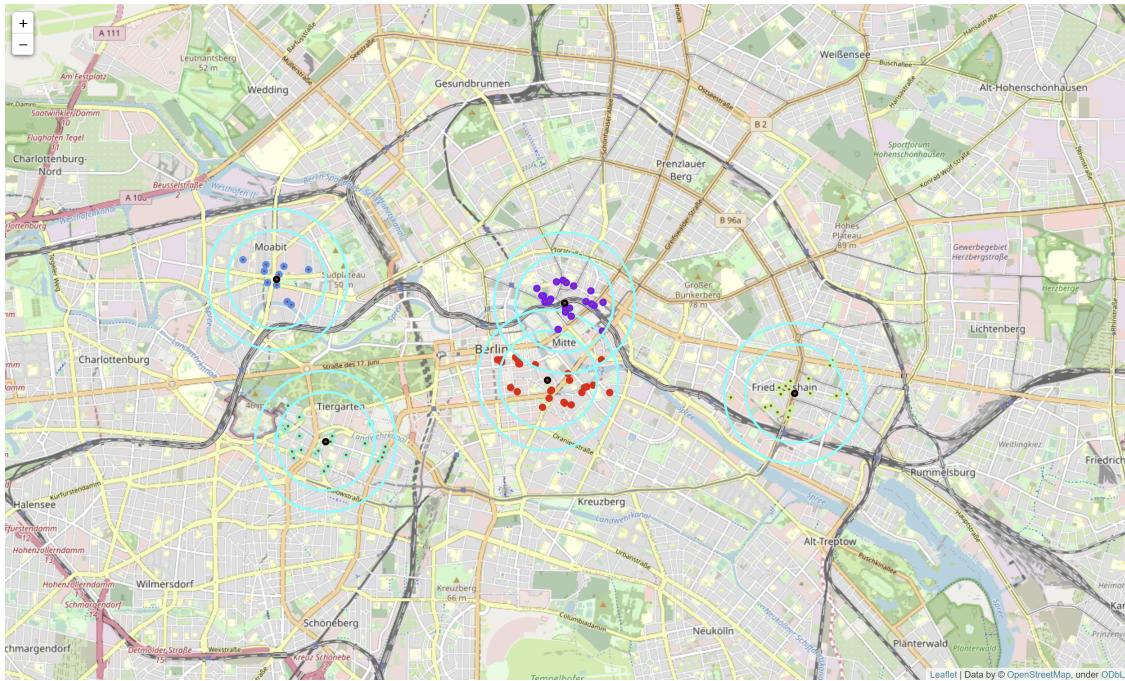
Cluster 0, Mitte, coordinates: 52.51311007061527, 13.400613842047896.

Cluster 1, Mitte, coordinates: 52.52292197640304, 13.404153473698498.

Cluster 2, Moabit, Hansaviertel, coordinates: 52.525921857800796,
13.34378139897381.

Cluster 3, Tiergarten, coordinates: 52.505297478714574, 13.35403627597626.

Cluster 4, Friedrichshain, coordinates: 52.51138271709561, 13.452391736530407.



Now we can present a ranking list of promising areas with highest rent for the next a few years to the stakeholders.

Recommended Berlin areas for further analysis

No.1 is in Berlin-Mitte with center coordinates (52.51311007061527, 13.400613842047896).

Within 700 m you can find:
 Hackescher Markt,
 Berlin-Alexanderplatz,
 Auguststraße,
 Museumsinsel.

No.2 is in Berlin-Mitte with center coordinates (52.52292197640304, 13.404153473698498).

Within 700 m you can find:
 Gendarmenmarkt,
 Leipziger Straße,
 Museumsinsel.

No.3 is in Berlin-Tiergarten with center coordinates (52.505297478714574, 13.35403627597626).

Within 700 m you can find:
 Kurfürstenstraße,
 Tiergarten,
 Nollendorfplatz.

Within 1000 m you can find:
Wittenbergplatz,
Zoologischer Garten.

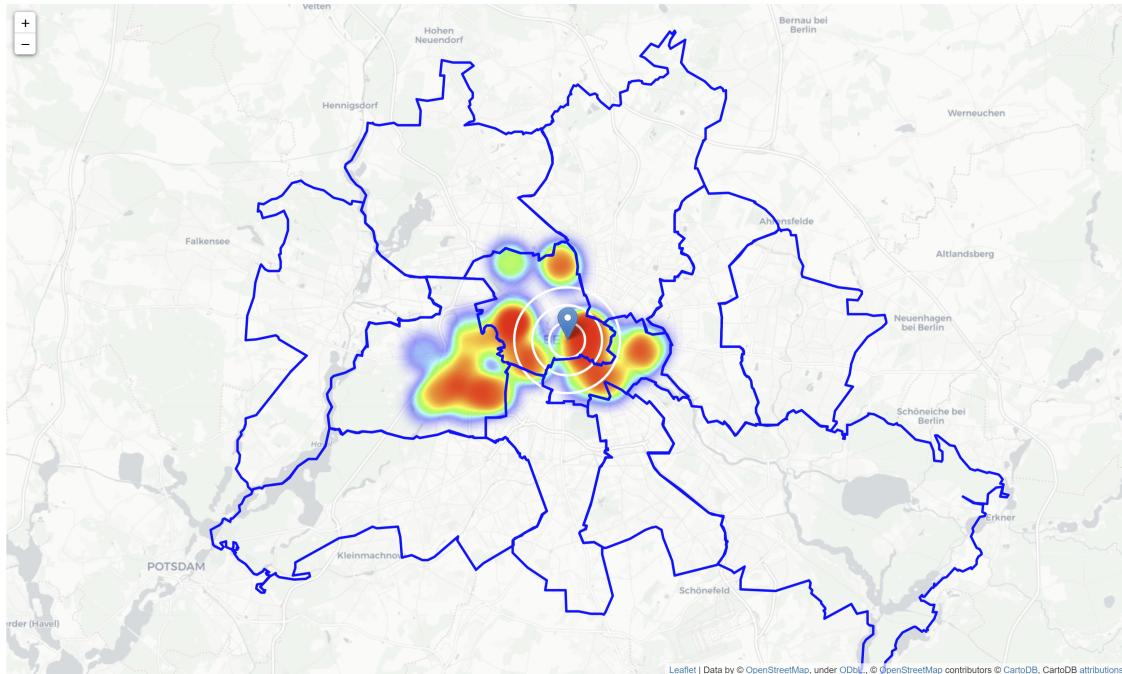
No.4 is in Berlin-Friedrichshain with center coordinates (52.51138271709561, 13.452391736530407).

Within 700 m you can find:
Karl-Marx-Allee,
Frankfurter Tor,
Warschauer Straße.

No.5 is in Berlin-Hansaviertel and Berlin-Moabit with center coordinates (52.51138271709561, 13.452391736530407).

Within 700 m you can find:
Turmstraße,
Kleiner Tiergarten.
Within 1000 m you can find:
Schloss Bellevue.

Let's create a map showing **heatmap / density of hotels** and try to extract some meaningful info from that. Also, let's show borders of Berlin boroughs on our map and a few circles indicating distance of 1km, 2km and 3km from the center of Berlin.



This concludes our analysis.

6 Results and Discussion

We have found 5 center locations which are expected to have the highest average rent in the vicinity in the next a few years. Four of the zones are located in the Berlin borough Mitte, one of them belongs to the Berlin borough Friedrichshain-Kreuzberg.

We can then evaluate our findings using the [Wohnlagenkarte2021](#). This is a heat map of the living conditions of areas of Berlin, the darker the color is, the better the living conditions of that area is supposed to be. To compare the heat map with our findings, one has to keep in mind that an area with "good living condition" does not have to be expensive. For example, a huge area southwest from the center of Berlin are marked as "good living conditions" by Berliner Mietspiegel. Most of this huge area belongs to the Berlin borough Steglitz-Zehlendorf, which is a former American sector before the fall of the Berlin Wall. However, as one can see, the average rent of this borough is below that of Berlin. Another very important point about this borough is that since most part of this borough is covered by forests, although the living conditions there are considered to be good by some people, the living conditions there will not become "better" very fast, so one does not expect a sharp increase of average rent in this borough. This is witnessed by our analysis on the rent growth and population growth.

The former British sector Charlottenburg-Wilmersdorf has some similarities with Steglitz-Zehlendorf. Compare with Steglitz-Zehlendorf it is much closer to the city center. Indeed, the avenue Kurfürstendamm which is located at the heart of Charlottenburg-Wilmersdorf is one of the busiest shopping and business avenue of the metropolis, and it was once the business center of West-Berlin. One similarity between Charlottenburg-Wilmersdorf and Steglitz-Zehlendorf is that they all belong to the former West-Berlin before the fall the Berlin Wall. Since the city west is already highly developed, the city east seems to be developing faster than its west counterpart. This means, in general, one expects that the average rent of the city east grows faster than that of the city west. This is partially the reason why the average rent of the former business center of West-Berlin - the area round Kurfürstendamm - grows slower than that of today's city center like Berlin-Alexanderplatz.

Since we were trying to find neighborhoods and areas in Berlin which have the *highest* average rent in the next a few years, we omitted those boroughs of Berlin whose average rents are below the average of the whole of Berlin. These include a huge borough: Berlin-Pankow, a former Soviet sector in the city east. However, this is not completely fair, because Pankow has a very central neighborhood - Prenzlauer Berg which is developing very fast in recent years. As one can see from the heat map of the Berliner Mietspiegel, Prenzlauer Berg is also a very promising area of Berlin. Although one can expect that no areas of this neighborhood would have a chance against the top 2 in our list, some areas of Prenzlauer Berg *may* take down the last two in our ranking list.

Lastly, we want to remark that the simple linear regression model we used to estimate the growth of the average rent in each borough of Berlin can only be used for ranking purpose, it is *not* suitable for a serious numeric estimation. The point is that there are so many important factors which could influence the average rent, but we did not take into account any of them. For example, on 23.02.2020 a new law **Berliner Mietendeckel** came into force in our socialist state Berlin. The law provides that all the rent in the rental agreement must be freezed for at least 5 years, and the new rent in a new rental agreement should not surpass the old one which was contracted before 18.06.2019. This law could cause a headache for the landlords in Berlin, because many tenants in Berlin are paying very cheap rents which were contracted decades ago and there is almost no way of ending the contract unless the tenants do not pay. The Berliner Mietendeckel together with

the corona-pandemic dragged down the growth of the average rent in 2020 in Berlin, as we can see from our analysis. However, on 25.03.2021 the Federal Constitutional Court (Bundesverfassungsgericht) concluded that the Berliner Mietendeckel is unconstitutional, and is therefore void by law. This new conclusion of the Federal Constitutional Court plus Germany's slow recovery from the pandemic could cause the average rent in Berlin grow sharply as before. But there are many other factors to be considered as well, so the simple linear regression model is only under the assumption that the current political and economic situations are *stable*.

7 Conclusion

The purpose of this project is to identify the areas in Berlin which would have the highest rent in the next a few years. We started by studying the boroughs using the data from the Investitionsbank Berlin (IBB). We first concluded that only three boroughs of Berlin, namely Mitte, Friedrichshain-Kreuzberg, Charlottenburg-Wilmersdorf, have above average rents, so we narrow ourselves down to these three boroughs of Berlin. Then we analyse the rent growth and population growth of the three boroughs. We have concluded that Mitte has the highest average rent and rent/population/population-density growth, Friedrichshain-Kreuzberg has the second highest average rent and rent/population-density growth, while Charlottenburg-Wilmersdorf has the third highest average rent, second population growth, third rent/population-density growth.

Next, we look into the boroughs to find out which neighborhoods and regions should have the highest average rent. To do this, we employ the Foursquare API to check the patterns of the venues of those neighborhoods having in mind that neighborhoods with similar pattern of venues should be in the same rent class. By analysing the patterns we found out that the density of nearby hotels is a good indicator of the average rent level. Then we use clustering (DBSCAN) to cluster the nearby hotels of the neighborhoods in the three boroughs. By taking the centroids of each cluster, we made a ranking list of recommended Berlin areas for further analysis. These are:

1. Location (52.52292197640304, 13.404153473698498) in Berlin-Mitte near Berlin-Alexanderplatz;
2. Location (52.52292197640304, 13.404153473698498) in Berlin-Mitte near Gendarmenmarkt;
3. Location (52.505297478714574, 13.35403627597626) in Berlin-Tiergarten near Kurfürstenstraße;
4. Location (52.51138271709561, 13.452391736530407) in Berlin-Friedrichshain near Frankfurter Tor;
5. Location (52.51138271709561, 13.452391736530407) in Berlin-Hansaviertel and Berlin-Moabit near Turmstraße.

Needless to say, No.1 and No.2 which are located at the center of the Berlin metropolis are expected to have the highest average rents in the next a few years, and they grow very fast. But the property price is also expected to be very high.

No.3, which located at the west side of the center of Berlin, bridges the city west and the city east. From its central position in the city, one can also expect a very solid rent growth rate. The property price there would also be very high.

No.4 located at the city east, is one of the fastest developing regions of Berlin after the fall of the Berlin Wall. For stakeholders who are short of capitals, this is surely a wonderful area of investment with very high cost/performance ratio.

No.5 is located at the city west. The southern part of this area has very attractive living conditions, and it is well-known as a multicultural neighborhood of Berlin. It is definitely worth for an investment.