

Rents and social places density in Berlin, Germany

Ioannis Karatasios

March, 2019

1 Introduction

1.1 Background

Berlin is the capital and largest city of Germany by both area and population. More than 3.5 million people are spread across 12 boroughs. However, compared to the rest of Germany, the rents for flats are extremely high and the trend goes up. In 2011, the square meter (for a 30m² flat) was € 8.36 which is very low compared to nowadays €15.31¹. Depending on the location, the differences in rents are enormous. Berlin Mitte (center of Berlin) is considered the most expensive borough.

Life in Berlin is very colorful. When considering moving move to Berlin, not only rent may matter, but also the diverse venue types. Different boroughs provide different kinds of activities and lifestyle. For example, the borough Treptow-Kpenick is known for its nature but prices for rents are higher than in the very east/west boroughs Spandau and Marzahn.

However, in this work, the focus is on providing data on rents in Berlin and cluster boroughs according to their similarity in terms of venues. Moreover, we will predict how prices will change within the next years with (simple) simple regression.

1.2 Problem

When people decide to move to Berlin, they might want have an overview on rents in Berlin as difference in prices is enormous. Moreover, it would be interesting to see the kind of venues the boroughs are known for. By combining both data, rents and venue categories, we will create a map that has each borough colored according to its prices for flats and clustered according to top venue categories.

1.3 Interest

The data and the map is helpful for people who plan to move to Berlin but especially for real estate agencies that help their customers pick good borough candidates. Besides money, the real estate agency may focus on customers' interest.

2 Data acquisition and cleaning

2.1 Data acquisition

To tackle both problems, we will need data on rents in Berlin, borough coordinates and boundaries. The site <https://www.wohnungsboerse.net/mietspiegel-Berlin/2825> has current and historical data. I will scrape the site to obtain current prices for rents (March 2019) and historical data. The latter is used to show how prices in rents almost doubled since 2011 and will keep rising. Customers of real estate agencies have to be aware of this fact as it might have a positive effect on their choice: either now or never! However, the data can be used in combination with linear regression to predict how rents will change within the next 2-3 years.

A kml shape file² containing a number of administrative areas in Berlin is provided by the statistical office of Berlin and Brandenburg. However, instead of transforming the shape file into GeoJSON format and extracting boroughs and boundaries, I will use the github repository of

¹<https://www.wohnungsboerse.net/mietspiegel-Berlin/2825>

²<https://www.statistik-berlin-brandenburg.de/produkte/opendata/geometrien0D.asp?Kat=6301>

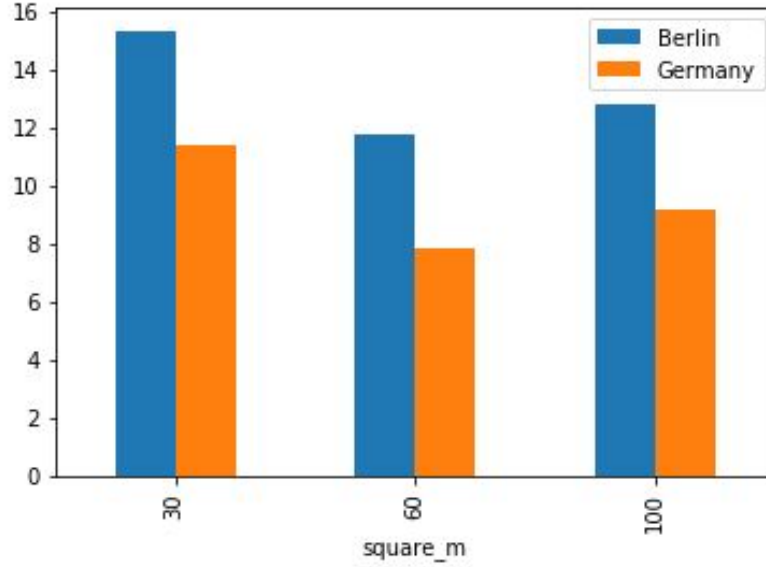


Figure 1: Rent index per square meter in 2019. Berlin and Germany in comparison.

m-hoerz³ which has a GeoJSON file containing names and boundaries of Berlin's 12 boroughs. Thank you m-hoerz!

Once borough details are extracted and details on rents are scraped, I will use Google Maps to obtain center coordinates (latitude / longitude) of each borough and Berlin. Then, to obtain the most common venues in each borough, I will use the Foursquare API.

The data will allow for predictions on rent prices for the next years and for creating a map that:

- displays boroughs of Berlin in form of markers and clustered according to venue density
- visualizes differences in rents (choropleth map) according to prices of March 2019.

2.2 Data Cleaning

The source <https://www.wohnungsboerse.net/mietspiegel-Berlin/2825> contains data on rents for flats and houses. I will focus on flats and exclude houses. Furthermore, I will compare small ($30m^2$), medium ($60m^2$), and big ($100m^2$) flats and historical data from 2011 to present. Prices are given in Euro and are per square meter. All other information will be discarded.

m-hoerz's github repository has a GeoJSON file including boroughs and their shape. Further cleaning is thus not required.

3 Descriptive/Exploratory Analysis

3.1 Overview on average rents in Berlin and Germany

To begin with, Berlin is compared to the rest of Germany very expensive (see Figure 1). Smaller flats ($30m^2$) are most expensive. In Berlin, the price per square meter is 15.32, and 11.39 in the rest of Germany. Furthermore, medium size flats ($60m^2$) are cheaper than big size flats ($100m^2$).

The rents for flats increased enormously in the past years. While the square meter for small size flats was 8.36 € in 2011, it increased to 13.93 €. Living in the city has become very

³<https://github.com/m-hoerz/berlin-shapes>

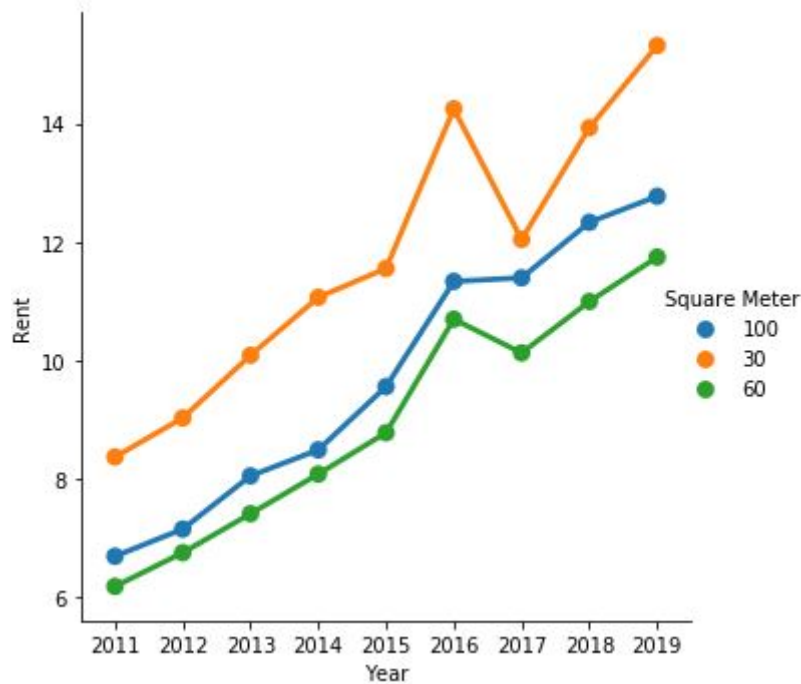


Figure 2: Rent index for flats in Berlin from 2011 - 2019

expensive! Historical and current data (March 2019) can be seen on figure 2. There is a positive trend regarding all three types of flat sizes.

3.2 Overview on average rents in Berlin's boroughs

Berlin is divided in 12 boroughs: Mitte, Friedrichshain-Kreuzberg, Pankow, Charlottenburg-Wilmersdorf, Spandau, Steglitz-Zehlendorf, Tempelhof-Schneberg, Neukölln, Treptow-Kpenick, Marzahn-Hellersdorf, Lichtenberg, Reinickendorf. We might assume, locations closer to the center are more expensive. The differences in price can be seen in figure 3. Marzahn-Hellersdorf is the cheapest borough, where one square meter costs on average 8.22 €. It is located in the far east of Berlin. In contrast, Mitte (center of Berlin), has the highest rates. The square meter costs 18.25 €.

Diagrams are good, but figures even better. Especially for customers. I used the provided GeoJSON file to create a map and display boundaries of the 12 boroughs. The data on rents is used to colorize each borough, which can be seen in figure 4. Far east and west Berlin are cheap. The more you go to the center of Berlin, the more the rents increase. It is obvious, that Berlin Mitte is the most expensive borough.

3.3 Borough Clustering according to venues

For this step, I used the Foursquare API to retrieve 100 nearby venues in a radius of 1000m - for each borough (see Figure 5). Results show that boroughs Treptow-Kpernick and Pankow have the lowest amount of venues, less than 10 in a radius of 1000m of its center. They are located in the very north and south west of Berlin and are thus pretty far away from the center. Mitte and Friedrichsheim-Kreuzberg in contrast, have the highest amount of venues. Seems legit as both boroughs are close to the center of the city.

In order to cluster boroughs, we need the optimal k - number of clusters. The elbow method was less helpful in this case as there is no clear 'elbow' visible (see Figure 6). However, I have chosen $k = 4$ as it seems a good choice.

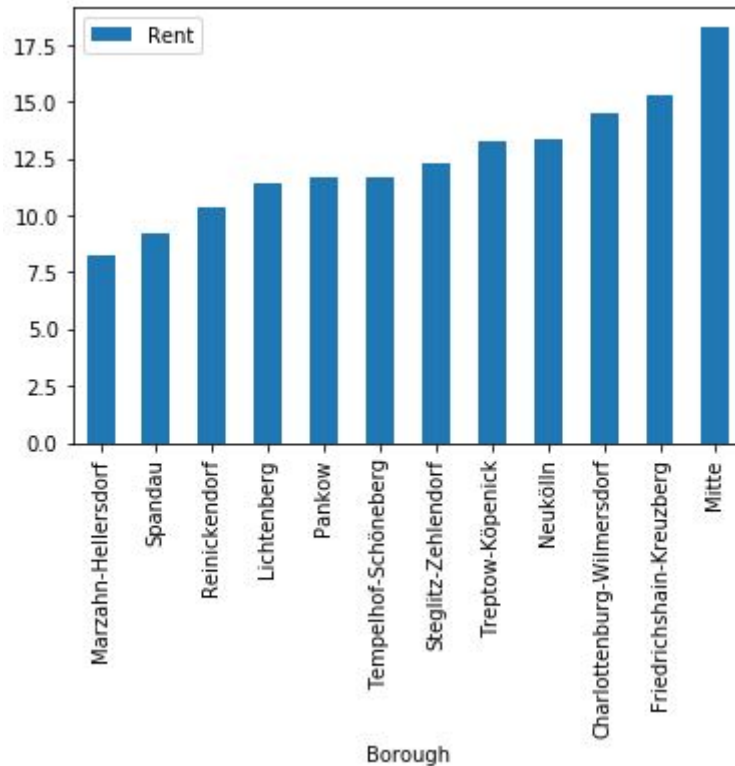


Figure 3: Rent prices in Berlin's boroughs (March, 2019)

Clustering is fine but what does it tell us? And why are some boroughs more expensive than others apart from their location to the center? Well, after the clusters had been formed, I printed the top 5 venues for each cluster with the goal of finding suitable names and/or descriptions for clusters:

Cluster 0 Bus Stop, Department Store, History Museum, Historic Site, Sculpture Garden

Cluster 1 Hotel, Forest, Scenic Lookout, Mountain, Yoga Studio

Cluster 2 Supermarket, Hotel, Gas Station, Bus Stop, Yoga Studio

Cluster 3 Coffee Shop, Hotel, Restaurant, Cafe, Falafel Restaurant

Now we can give better descriptions to clusters. :

Cluster 0 History & Museum (Orange)

Cluster 1 Nature & Relax (Red)

Cluster 2 Supermarket & Food (Blue)

Cluster 3 Cafe & Restaurant (Green)

Now, we combine both information, rents and clusters to generate a choropleth map:

4 Predictive Modeling

I was curious how rents will change within the next 2-3 years. Current trends have been previously shown in figure 1. Linear Regression helped drawing a straight line that can be used

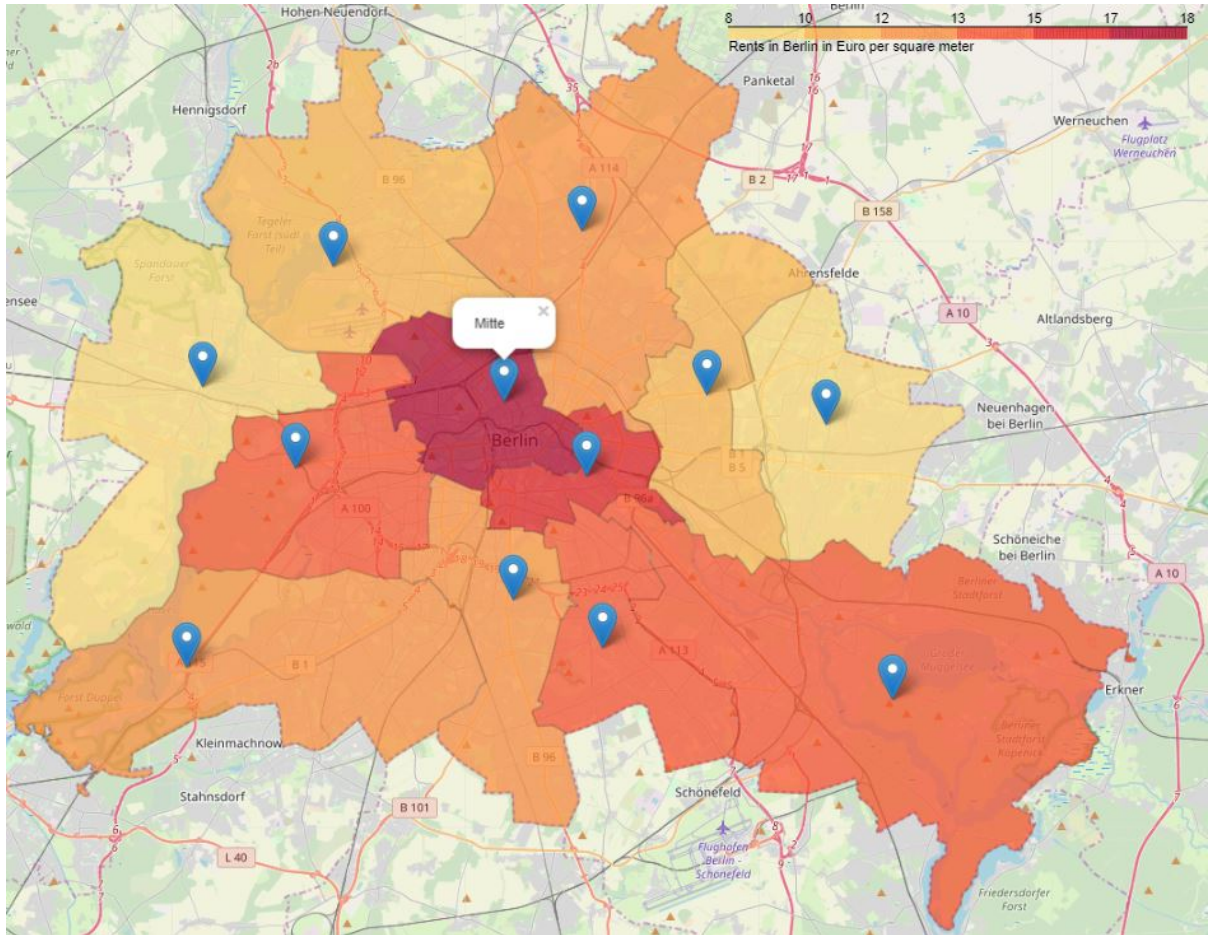


Figure 4: Boroughs are colored according to their prices on rents.

for simple predictions. Results are shown in figure 8. The prices for rents are almost double compared to 2011. If the trend remains as is it, positive, prices per square meter will increase to 18,35 € for small, 15,42 € for medium, and 17,16 € for big size flats by 2024.

5 Conclusion

Berlin is a beautiful and big city, but also expensive. If someone plans to move to Berlin, he first should consider that now it is the best time since prices for rents will increase in the next years. We looked at historical and current data to make simple predictions. Our maps show that Berlin Mitte, the center of Berlin, is the most expensive borough. Boroughs located far east and west/north west are much cheaper. Customers of real estate agencies have to be aware of these facts.

Foursquare was used to obtain 100 venues in a radius of 1000m for each borough. Latter have been grouped into 4 clusters according to venue density. The results show that 8 of 12 boroughs fall under cluster 3: Cafe & Restaurant. Pankow (north) and Marzahn-Hellersdorf (far east) are known for Supermarkets and food in general. Neukölln and Treptow-Köpenick are in mid-range when it comes to rents. However, both boroughs differ in terms of venues. The venue topic of Neukölln is History & Museum. Treptow-Köpenick instead, is known for Nature & Relax.

Real estate agencies can show these information to their customers when it comes to choosing boroughs.

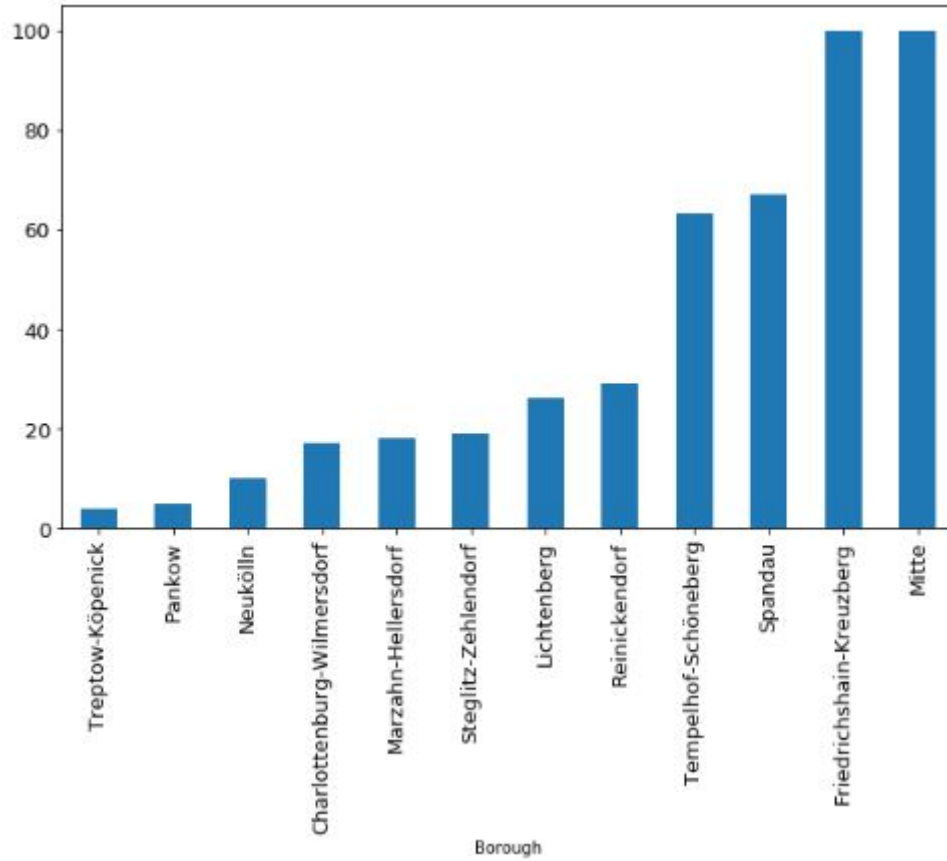


Figure 5: Boroughs sorted according to amount of venues

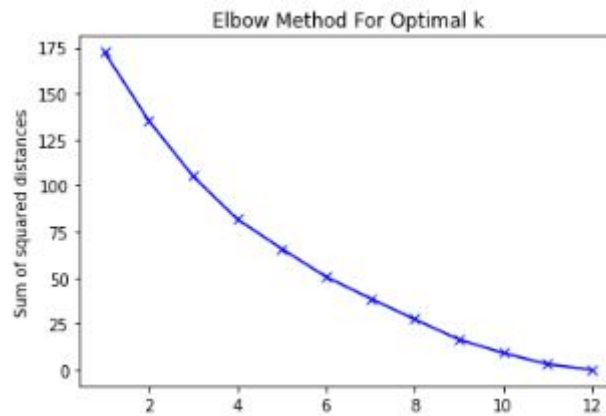


Figure 6: Elbow method to find optimal cluster size for clustering Berlin's boroughs

6 Future Directions

The data is restricted to renting flats in Berlin. But, can be extended to cover purchasing flats and houses in general. Besides money customers are willing to pay and interest in venues, real estate agencies may want to use more features to suit their customers' needs, e.g., crime rate of different boroughs, or school rating.

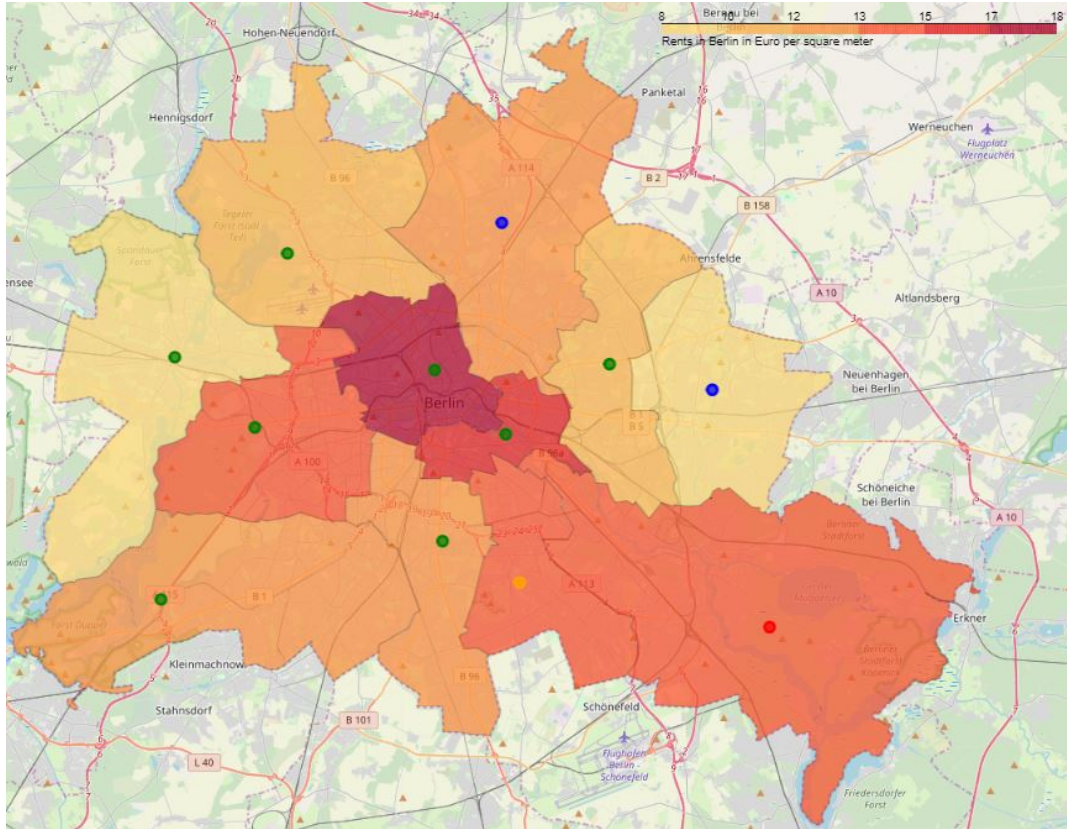


Figure 7: Boroughs are colored according to their average rents and clustered according to venue density.

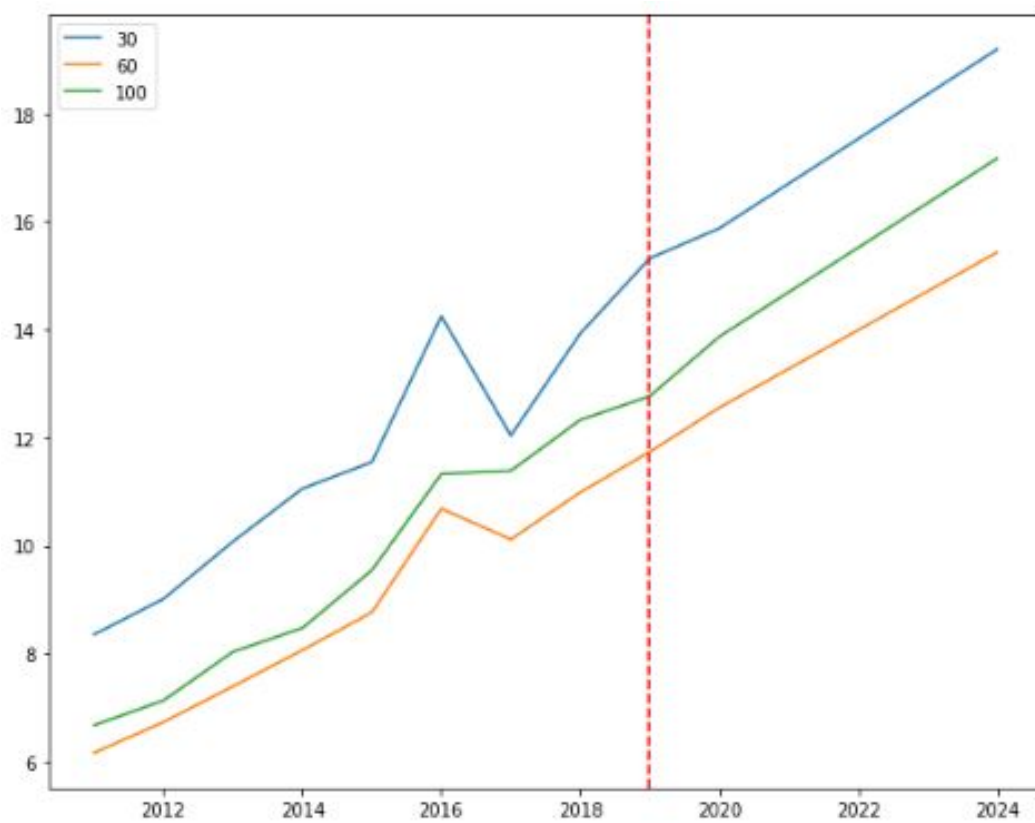


Figure 8: Predictions on rents for the coming years. Price is given in Euro and per square meter