Real Estate Pricing Analysis in Germany

Introduction

Germany is one of the biggest and most financially powerful countries in Europe . The population of the country is radically increasing annually due to its sustainable economic and industrial development. Since 2012 more than 1.000.000 people every year immigrate to Germany, mostly due to their need for financial growth.
The biggest percentage of the immigrants decide to move to big german cities such as Berlin, Hamburg, Munich etc. , cities that are more multicultural and offer more job opportunities to foreigners. This situation leads to an enormous increase concerning the real estate demand, and thus a research in that field becomes very interesting.

Real estate companies, Industrial companies, even individuals become more and more interested in house pricing, since it turns out to be a quite sufficient life investment. A closer look to the current house pricing situation in Germany appears to be very appealing to individuals and companies which have a tendency to economic growth.

Data

For the purpose of this survey a dataset from kaggle.com (Germany Real Estate Housing Price) was utilized. It contains data from April 2020 concerning the average housing price in different German States (Bundesländer) , the respective postal code and the current conditions of each house – appartment (number of rooms, year of construction etc.)
The postal code in this case is very important so that we know in which area of the relevant german state the house is located.
Except for the postal code the benefits of Foursquare were utilized, since the quantity and the kind of venues that are closely located have a strong impact in the configuration of the prices. Venues such as Hotels, Museums,

Restaurants, Bars as we already know , are features that are capable to transform an area to a very attractive one.

Methodology

This survey is a Machine Learning problem and more precisely a Clustering problem. The machine learning method which was used is the K- Means machine learning method. For acquiring information concerning the venues located close to each area the data platform Foursquare was used. The purpose of the project is to create an appropriate number of clusters based on the price, the quantity and the kind of venues located in each Bundesland. In the end of this project we will be able to see the difference between the german states according to the housing price and what kind of attractive destinations each area has to offer. Additionally, we acquire a general view of the most visitable venues in each Bundesland.

Results

We created an interactive map of Germany (see Figure 1) including the clusters found by the clustering method. For the best choice of clusters the elbow method was utilized (see Figure 2).

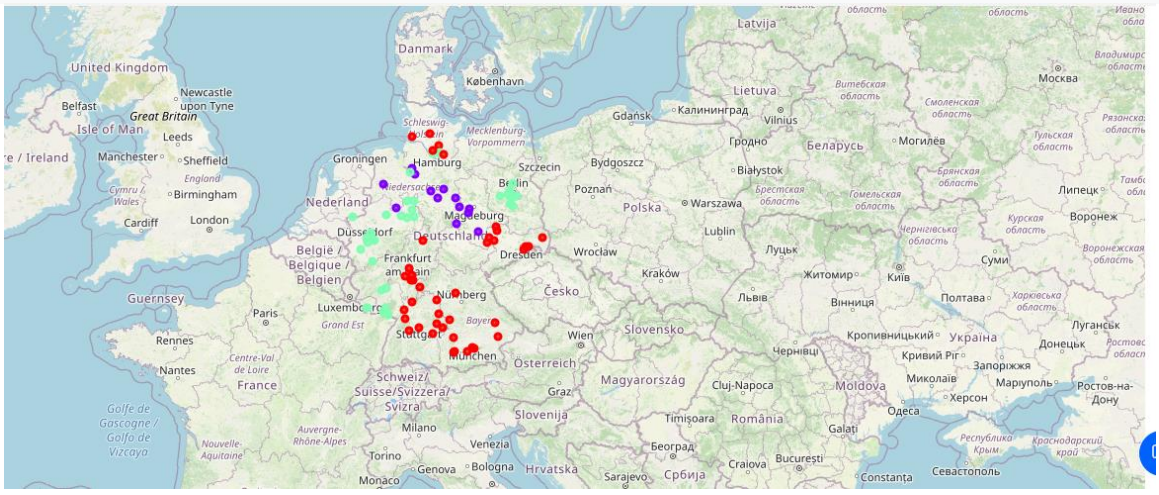

Figure 1 – Interactive map of German States and its relevant clusters based on the housing price

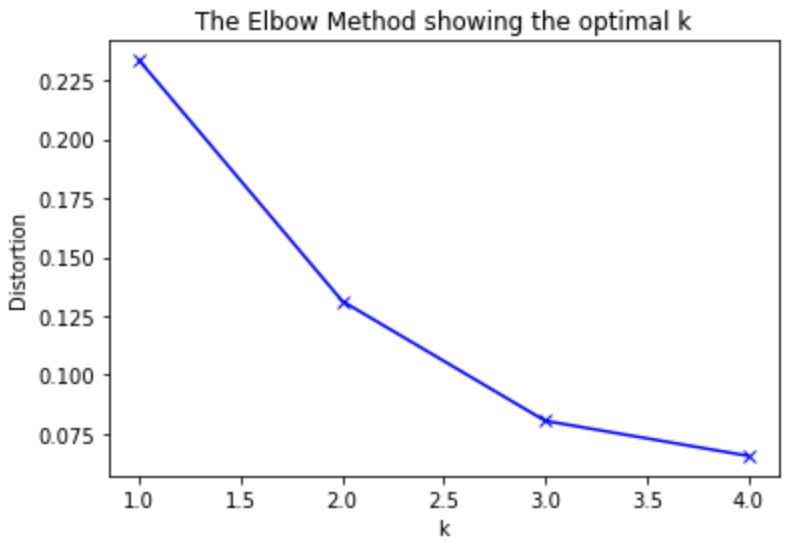

Figure 2 – Elbow Method (Illustrates the most efficient number of clusters)

We created bins in order to create price intervals and categorize them as follows:

- “Low Price” (price < 1.000.000 €)
- “Medium Price” (1.000.000 € < price < 3.000.000 €)
- “High Price” (price > 3.000.000 €)

Surprisingly, according to our results the biggest percentage of the housing prices around Germany belongs to the “Low Price” category , a small percentage to “Medium Price” and an extremely small one to “High Price”.

Additionally, the most common venues are quite the same in every german state with non remarkable differences. Although as it was expected quantity of each venue differs from state to state .

Conclusion

Housing price in Germany is still in an affordable level even in a demanding state such as Berlin in which the german capital is located, taking into account the average financial situation in Germany. Undoubtedly, the level of prices depends on the location in each state (See Bayern – appears the highest price) and the kind of venues around the area (Bayern-Hotels and Restaurants- Highest Price).