Identifying Yoga-friendly Neighbourhoods in Frankfurt a.M, Germany

IBM Applied Data Science Capstone Project

Dr. Meryem Tanarhte

Oktober, 8th 2021

1. Introduction

1.1. Problem definition

Frankfurt am Main is a world city, and internationally renowned as a financial hub and an important centre for industry, services and exhibition space. Frankfurt am Main is the largest city in the FrankfurtRheinMain region with a population of 760,000.
Frankfurt is a city of research and also a business location well known in Germany for its very high quality of life.
Kumud Schramm, spokeswoman for Germany's Yoga Teachers' Association, says that more than three million Germans do yoga on a regular basis, and there are more than 8,000 teachers in the country who serve them. "We're really experiencing a boom in yoga" Schramm recently told Focus magazine. "Our lives are hectic, demanding, and take us to the limits of our physical and psychological limits. People are seeking quiet, relaxation, and a look inside themselves. That's what they get with yoga."
The Yoga community is very large in Frankfurt, with hundreds of courses offered by fitness studios and personal yoga trainers. Opening a yoga studio requires to find a suitable neighborhood in this vibrant city taking into account several variables such as competition, finding new customers and being close to open-air spaces.

In this project we will try to solve this problem by using location data and machine learning techniques to identify „yoga-friendly" neighborhoods in Frankfurt a.M, Germany where we can open a Yoga studio business. To determine which neighborhoods are "yoga-friendly" we will focus on neighborhoods with high number of offices to be an easy-access to the employees in need of relaxation and sport options after working hours, universities to reach students and public parks and gardens to be able to show-case the yoga techniques and organize yoga classes in open-air areas.

1.2. Stakeholders Interest

This project would be of interest to yoga teachers looking for new neighborhoods to start their businesses. Additionally, employees and students can benefit from this information to be able to choose from a large offer of yoga studios.

2. Data and Methods

2.1. Data overview

For this project we need data about:

- Neighbourhood boundaries
- Types of venues needed to open a yoga studio (offices, universities and parks)
- The number of yoga studios venues in each neighbourhood

The data that we will use for this analysis is a combination of an Excel file that has been prepared for the purpose of the analysis from multiple sources (Frankfurt_districts.xls) and the location/venue information in Foursquare using Foursquare API.

2.2. Methodology

To identify the most suitable neighborhood for our yoga studio, we used the machine learning method "k-means clustering".

K-means clustering was chosen because it meets two conditions that make k-means clustering ideal:

- We do not have labeled data
- We know how many clusters we would like: "Yoga-friendly", "Not yoga-friendly" and "Less Yoga-Friendly".

3. Results

First, the current residence neighborhoods of Frankfurt city and their boundaries extracted from the Excel file (Frankfurt_districts.xls) are shown in Figure 1

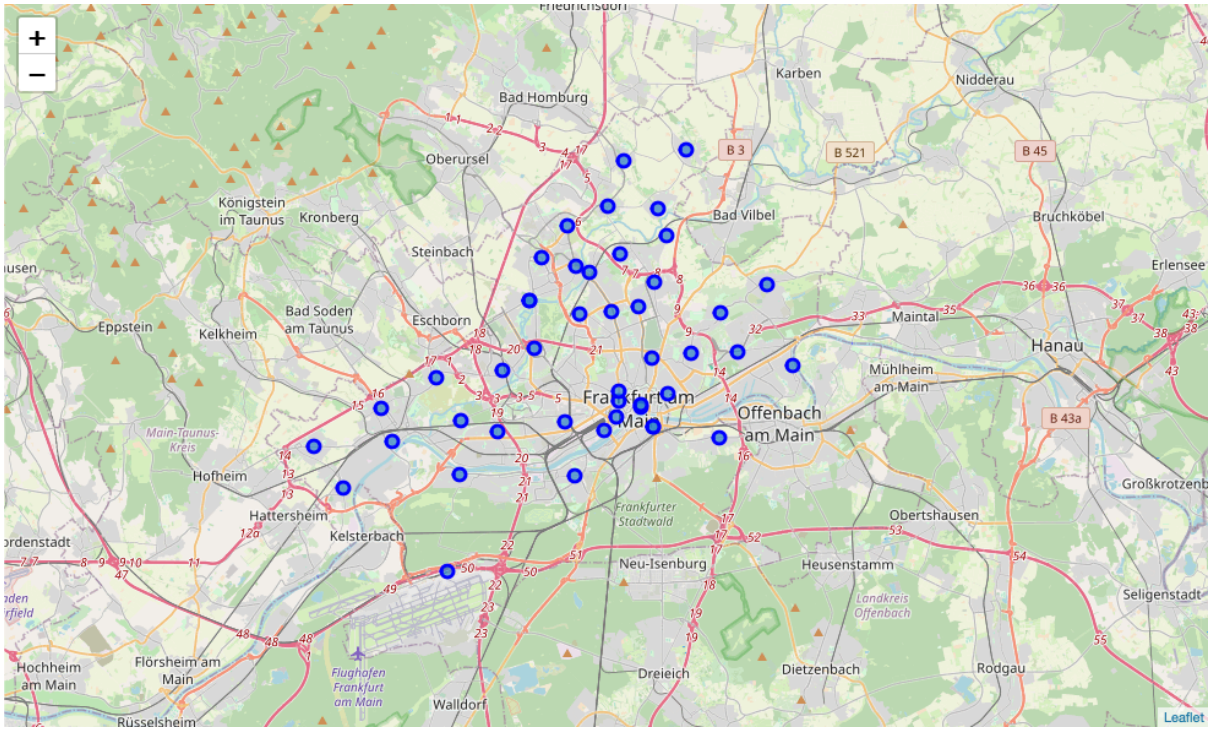

Figure 1: Map showing Frankfurt Neighborhoods.

Figure 2 shows the existing Yoga studios and their distribution across the neighborhoods in Frankfurt City

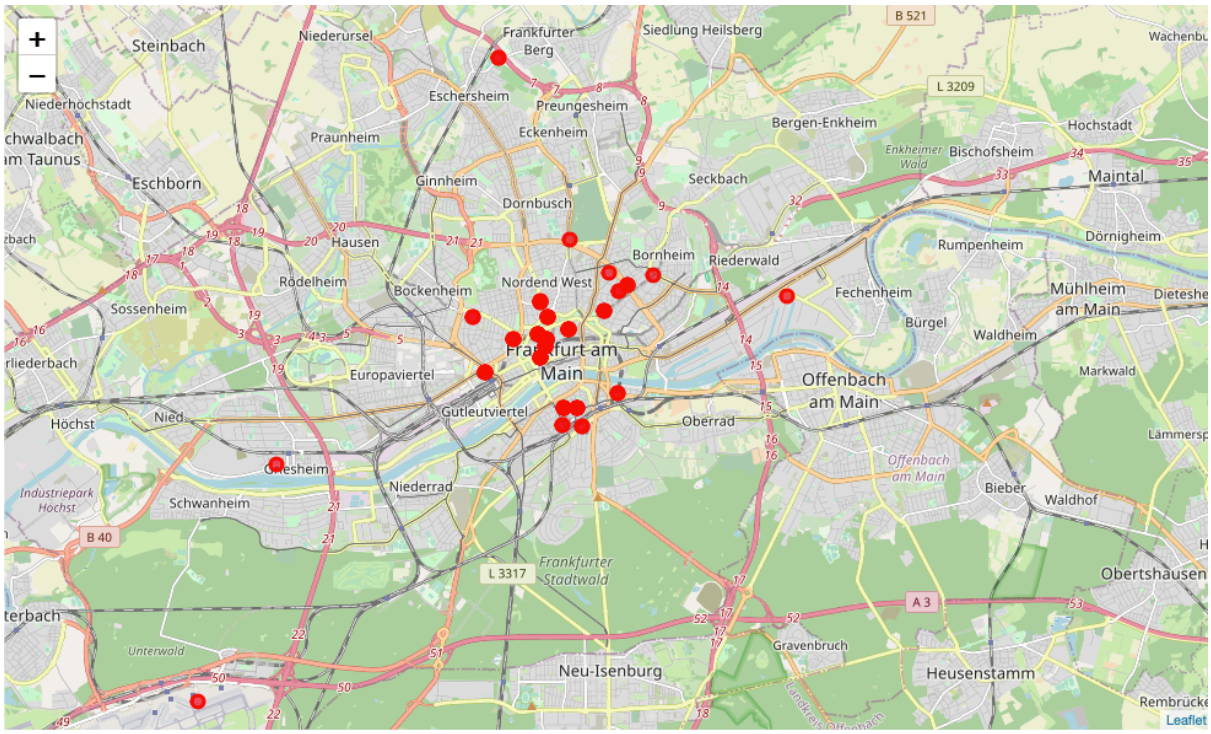

Figure 2: Map of the yoga studios in Frankfurt

In total Frankfurt has **73 yoga studios** located mainly in the city center.

Using Foursquare API, we obtain data about:
- the number of offices to be an easy-access to the employees in need of relaxation and sport options after working hours,
- universities to reach students
- public parks and gardens to be able to show-case the yoga techniques and organize yoga classes in open-air areas.

We found **1559 offices, 27 university building and 269 parks** across the city neighborhoods.

After obtaining the data from Foursquare API, we apply K-means to the dataset. We used 3 as our cluster numbers to represent the following categories: Yoga-friendly, Not Yoga-Friendly and Cautious. In the latter case, a better assessment of the market is needed.
To ensure we have the optimal number of categories, we plot the inertia (sum of squared distances) for each data point.

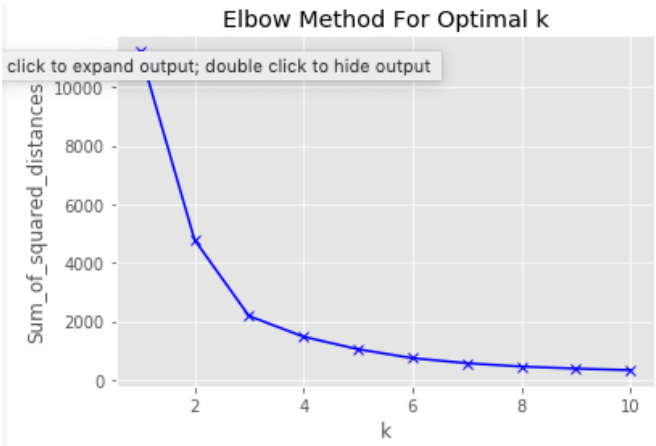

Figure 3: Elbow Method for Optimal k

We found the **elbow point for the data points as 3**.

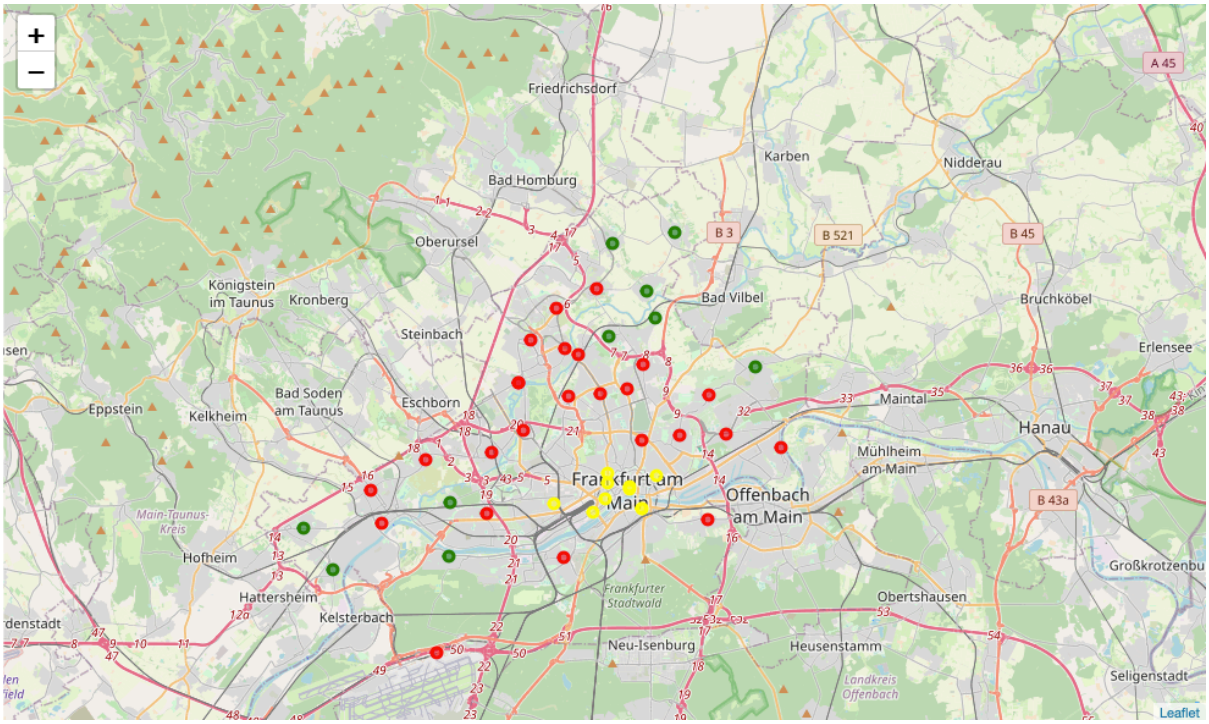

Figure 4: Neighborhoods clusters (Green: Yoga-friendly, Red: Not yoga-friendly, Yellow: Less Yoga-friendly)

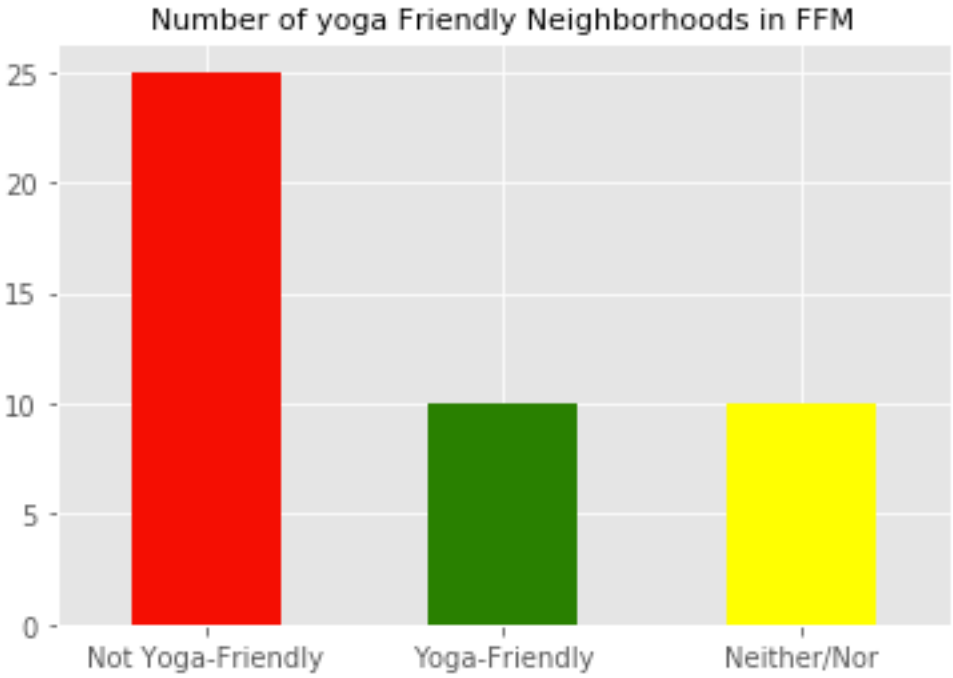

Figure 5: Boxplot of the number of yoga-friendly based on the cluster color.

Based on the results there appears to be plenty of 10 Yoga-Friendly neighborhoods to choose from in Frankfurt. There also appears to be more than 25 neighborhoods that need to be avoided. With most of these being outside of the city center. The rest of

neighborhoods (10) are neither good not bad for the yoga business (less yoga-friendly). Additional market assessment is needed.

4. Discussion and further ideas

The yoga market in Frankfurt has been booming in the last decade, and the city has been expanding with new offices, universities and housing.
The "less-friendly" neighborhoods have been changing and could be actually "yoga-friendly" because of the city expansion.

This study could be improved with:

1. Adding more venues for analysis such as pools, playgrounds for family yoga-activities or libraries.
2. Add boundaries for each neighborhood. Currently only using a 1-mile radius to define the boundaries of the neighborhoods assigned latitudes and longitudes. This could cause some overlapping of neighborhoods while searching venues.
3. Filter the results by budget as the cost of rental of yoga studios can be very expensive in Frankfurt.

5. Conclusion

After evaluating the neighborhoods of different postal codes in Frankfurt we obtained Yoga-friendly activities within a mile of the each district. Further clustering the results in order to get the most yoga-friendly neigborhoods based on all of our criteria, taking into account nearby offices, universities and parks/gardens.

The results suggest there are only few yoga-friendly neighborhoods to choose from. There are 10 only-friendly Postal Codes located all across Frankfurt. So stakeholders would not have a problem finding the right neighborhood based on their own preferences, needs & budget.

Additional market analysis based on budget constraints should be made to choose other from neighborhoods that are neither "yoga-friendly" nor "not yoga-friendly".