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

IBM Applied Data Science Capstone Project

Dr. Meryem Tanarhte

04/15/2021

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 our 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 Amount 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_neighborhoods.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: non-suitable, very suitable and less suitable.

3. Results

4. Discussion

5. Conclusion