Introduction

Berlin is the capital and largest city of Germany by both area and population inhabitants make it the most populous city proper of the European Union. The city is one of Germany's 16 federal states. It is surrounded by the state of Brandenburg, and contiguous with Potsdam, Brandenburg's capital. It is a world city of culture, politics, media and science. Its economy is based on high-tech firms and the service sector, encompassing a diverse range of creative industries. Moving to berlin is one of my childhood dreams.

Business Problem

The objective of this capstone project is to analyse and select the best neighbourhood in the city of Berlin, Germany to relocate. Using data science methodology and machine learning techniques like clustering, this project aims to provide solutions to answer the business question:

If you are going to move to Berlin, where is the best place to rent an apartment?

Target Audience of this project

This project is particularly useful to students and immigrants to help them in choosing the right place neighbourhood that meet their different needs. It's also beneficial for travellers and tourists as it gives them general idea about the places in berlin and the cost of living.

Data

To solve the problem, we will need the following data:

- List of neighbourhoods in Berlin and the average rent price of each neighbourhood.
- Latitude and longitude coordinates of those neighbourhoods.
- Venue categories data.

Sources of data and methods to extract them:

A large data 96 columns for Berlin's AirBnB listings-summery are available for free in Kaggle, from this data set we can extract apartments rent price, coordinates and neighbourhood. This will be the base for rent cost analysis. After that, using Foursquare API we can obtain the venue categories data for clustering.

Links:

Kaggle Data:

https://www.kaggle.com/brittabettendorf/berlin-airbnb-data#listings_summary.csv

Foursquare:

https://foursquare.com/developers/apps