COURSERA / IBM
Data Science
Professional Certificate

Capstone Project

Starting an Indonesian Restaurant in Berlin

Thom Versteeg

Coursera Capstone Project

27.5.2019

1. Introduction / Business Problem

Berlin, the capital of Germany, is a vibrant city with over three million inhabitants, a big tourism sector. As such it has a large and diverse selection of restaurants and cafés. Dining out is a big part of the culture and as such, the city is an interesting location to establish a new restaurant. However, due to the heavy competition, the exact location of the restaurant can make or break the business.

My client is planning to open a restaurant which serves Indonesian Cuisine. This particular cuisine is fairly rare in Berlin and has potential to become very popular. However, similar cuisines, such as the Vietnamese, Thai and Korean cuisine are well established. In order to choose a location for his restaurant, my client wants to know which neighbourhood has the biggest potential for a restaurant of this kind of food to be successful. His search is limited to the most popular boroughs of the city: *Mitte, Friedrichshain-Kreuzberg, Neukölln* and *Prezslauerberg.* In this report, I present an advice for the ideal location to establish a new Indonesian restaurant in Berlin.

2. Data

Three types of data are required for this project:

1. Names and coordinates of the neighbourhoods of Berlin
2. Venues and foot traffic in Berlin areas

2.1 Names and coordinates of the neighbourhoods of Berlin

A list of the names of the different neighbourhoods of berlin was scraped from Wikipedia (https://de.wikipedia.org/wiki/Liste_der_Bezirke_und_Ortsteile_Berlins). The coordinates were added using the *Nominatim* module of the Python *Geopy* library.

2.2 Venues and foot traffic in Berlin areas

The Foursquare API was used to retrieve information about the most popular spots in berlin. The popular spots returned depends on the highest foot traffic at the time when the call is made. We may get different popular venues at different times of the day. Therefore, the API call was made at 8 PM, during dinner rush hour.