

Presentation

Applied Data Science Capstone

Capstone Project - The Battle of Neighborhoods

André Schlinke

2019/03/06

Outline

- 1 Introduction
- 2 Data Acquisition and Cleaning
- 3 Exploratory Data
- 4 Clustering
- 5 Conclusion and Outlook

Introduction

Berlin:

- is the capital of Germany.
- has a lot of history and culture.
- is a multicultural metropol.
- has over 10 million tourists visiting the city every year.
- has a rich culinary scene with cooking styles from all over the world.

⇒ However, tourists as well as locals may find it difficult to find a restaurant serving local German food in Berlin.

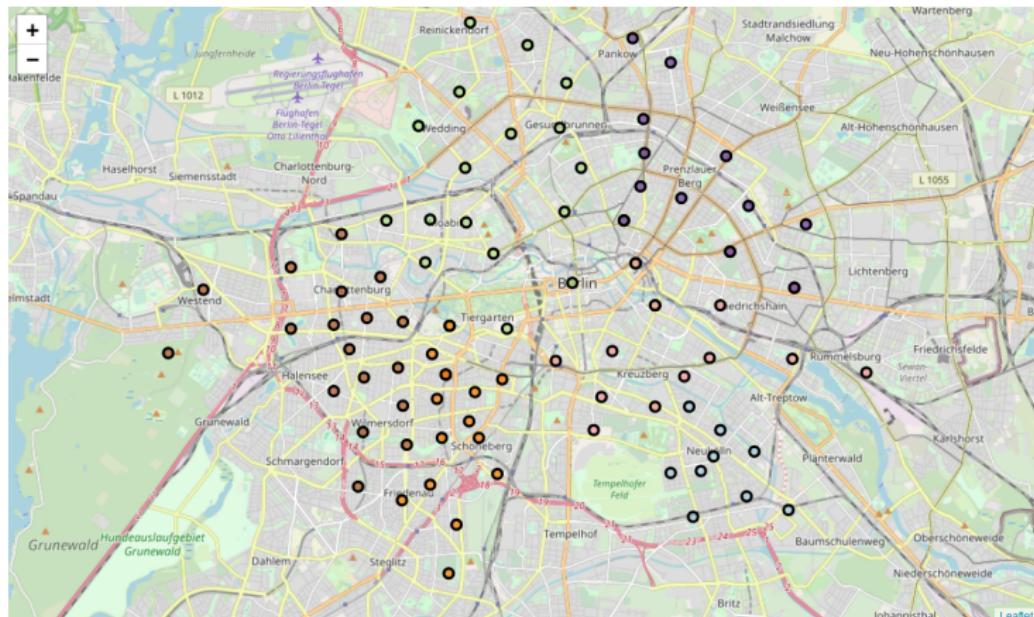
Data acquisition

- CSV-file containing a hand-picked selection of postal codes, corresponding neighborhoods, and boroughs. Content was extracted by several wikipedia articles.
- Pgeocode library to extract the latitude and longitude per each postal code.
- Explore venues with Foursquare API

Data cleaning

- Postal codes which are part of multiple neighborhoods were summarized to one postal code sample with the several neighborhoods being separated by comma.
- Only interested in venues with 'Restaurant' in category name.
- Drop duplicates of restaurants.

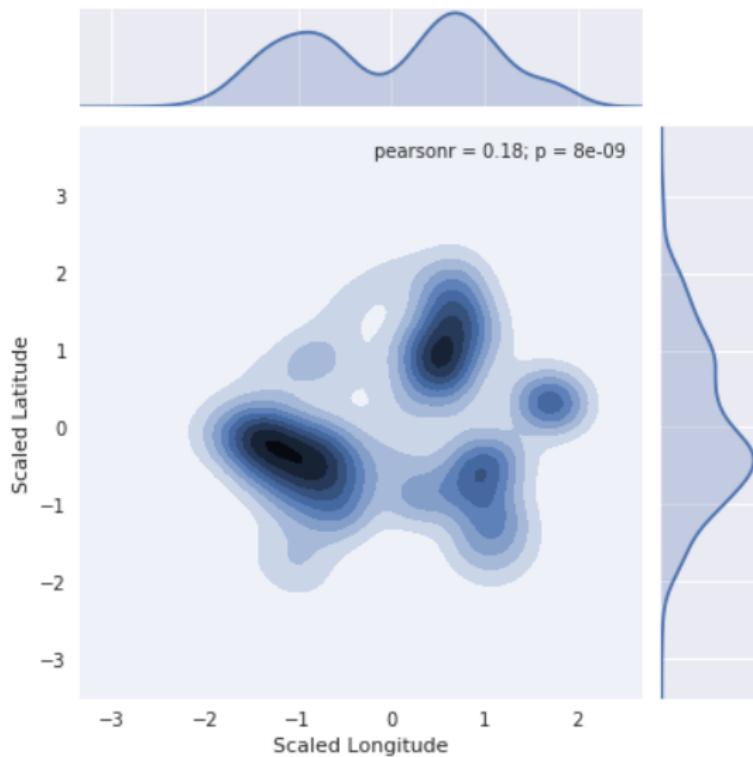
Location of all postal codes examined



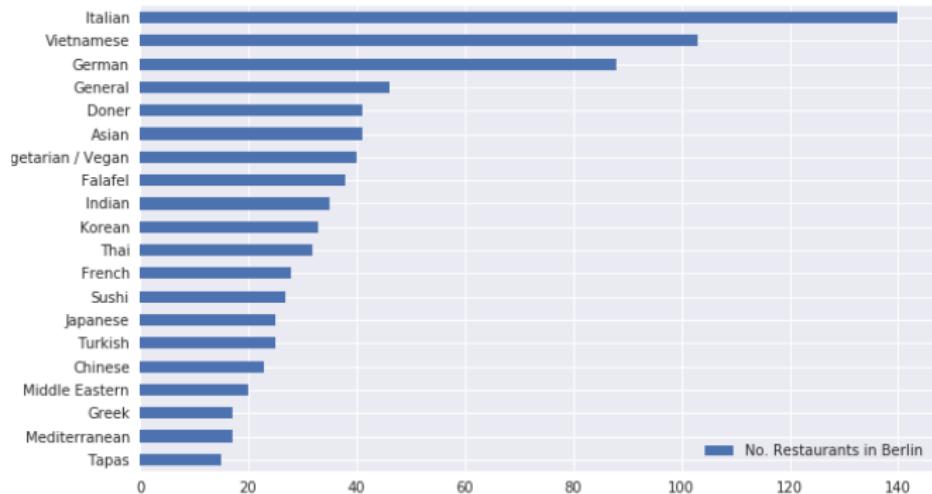
Locations of all restaurants found in Berlin



Scaled distribution of all restaurants found in Berlin

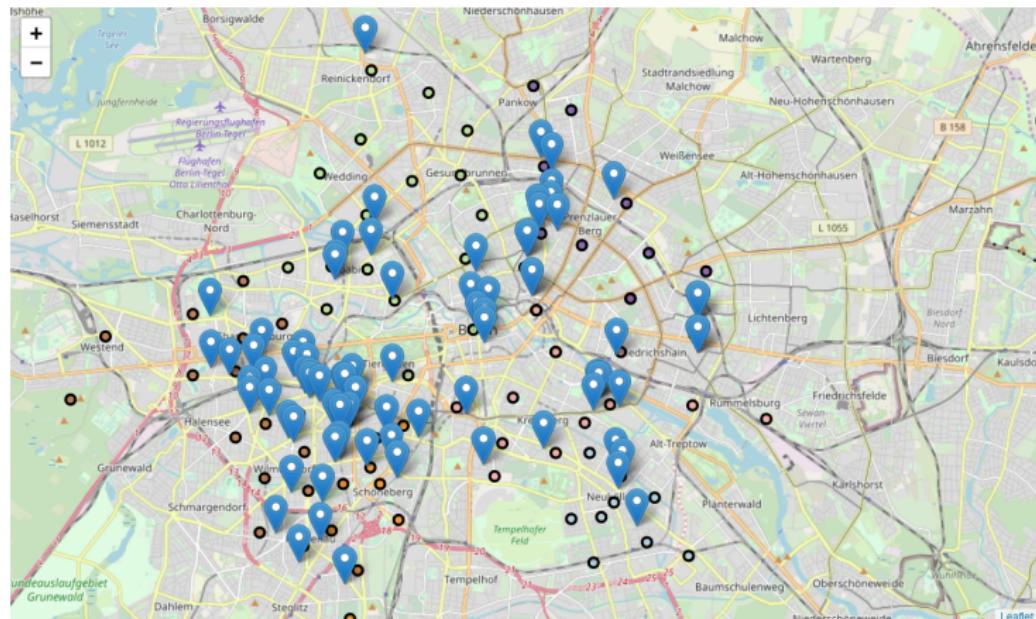


Excerpt of total no. of restaurants per category (top 20)

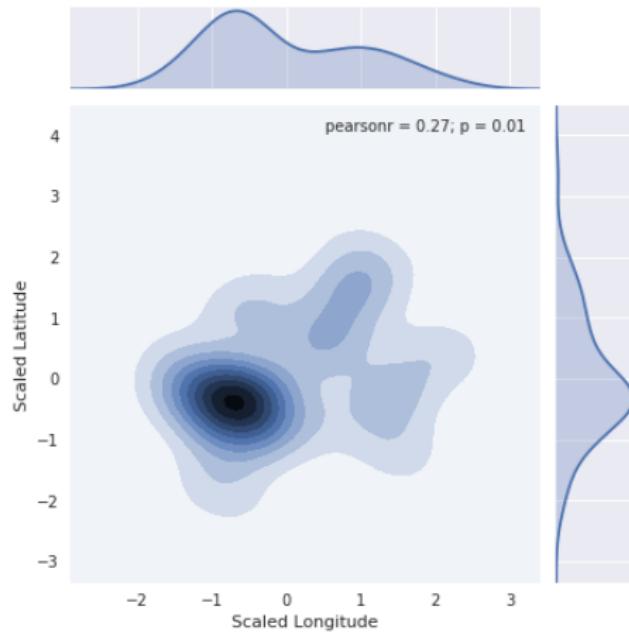


- German restaurants are on the third place in Berlin with a total of 88 restaurants. There are only more Vietnamese and Italian restaurants with Italian being the most with a total number of 140 restaurants.

Locations of all German restaurants found in Berlin



Scaled distribution of all German restaurants found in Berlin



⇒ German restaurant centered in southwest area of the city.

Percentage of German restaurants per borough

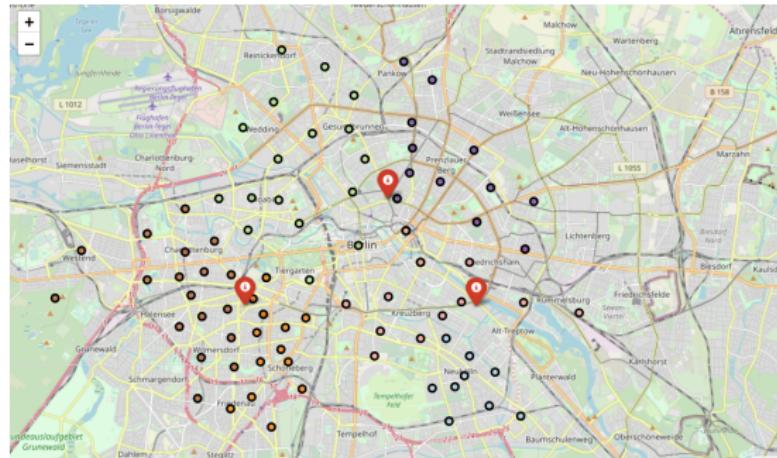
Borough	No. of Restaurants	No. of German Restaurants	Percentage of German Restaurants per Borough
Tempelhof-Schöneberg	149	18	12.1%
Charlottenburg-Wilmersdorf	225	26	11.6%
Mitte	170	17	10.0%
Friedrichshain-Kreuzberg	189	12	6.3%
Pankow	188	11	5.9%
Neukölln	88	4	4.5%

⇒ Neukölln has the lowest concentration of German restaurants with 4.5% while Tempelhof-Schöneberg has the highest with 12.1%.

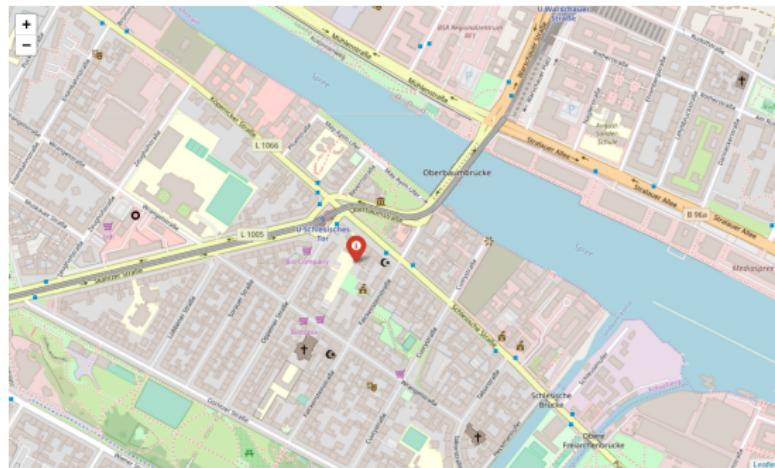
Data cleaning

- Clustering algorithm k-means to determine 3 hot-spots.
- Train clustering algorithm on latitude and longitude values of non-German restaurants.
- Out coming three cluster centers are then considered to be our recommendations.

Location of the three recommendations

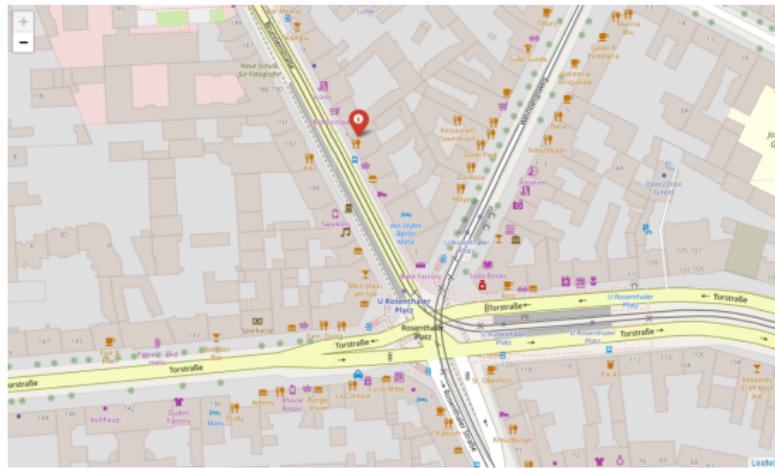


Location of the three recommendations



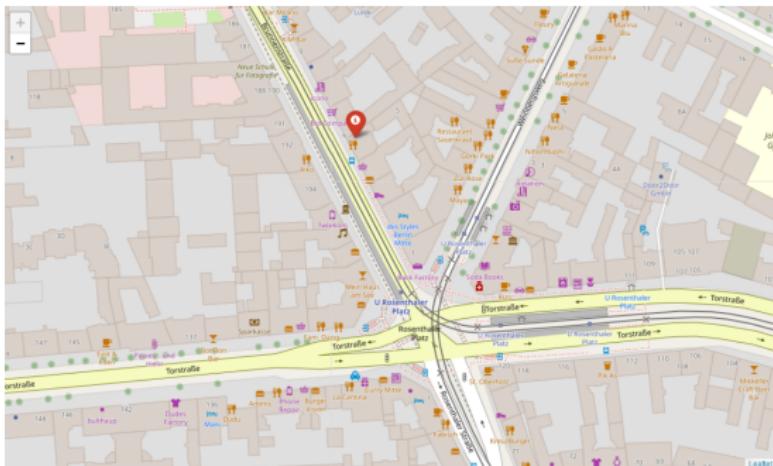
- Schlesisches Tor is in Kreuzberg and is one of the go to areas if you want to party. There are a lot of nightclubs and restaurants.

Location of the three recommendations



- Rosenthaler Platz in Prenzlauer Berg is one of my favorite spots to go out to eat. Here are a number of very good Asian food restaurants.

Location of the three recommendations



- Lietzenburger Strae is at the border of Schöneberg and Wilmersdorf. As we discussed earlier, the competition in this area for German restaurants is higher, nevertheless German restaurants seem to be pretty successful here.

Conclusion and outlook

- German restaurants a third most common restaurants in Berlin after Italian and Vietnamese restaurants.
- They are mostly located in the southwest of Berlin.
- Neukölln is the Borough the the least amount and concentration of German restaurants.
- Three recommendation for opening a new restaurant:
 - Schlesisches Tor is in Kreuzberg.
 - Rosenthaler Platz in Prenzlauer Berg.
 - Lietzenburger Strae at the border of Schöneberg and Wilmersdorf.
- Future studies could examine different types of restaurants more closely as well as more neighborhoods could be analyzed as well.