IBM Data Science Capstone Project
The Battle of Neighborhoods

Marc Schneider

28.01.2020

1 Introduction

The city of Frankfurt am Main is at the center of the Rhine-Main Metropolitan Region in Germany. It covers an area of 248.31 km² and has a population of 753,056 inhabitants. Frankfurt is known to be a global hub for commerce, culture and tourism. For example, the European Central Bank, the Frankfurt Stock Exchange, the Messe Frankfurt and the Frankfurt Airport are some of the busiest and most influential institutions in Europe. As these institutions attract people from all kinds of backgrounds, Frankfurt itself is culturally, ethically and religiously diverse, i.e. a big part of the population has a migration background.

As a city's demography gets more and more diverse, the amount of new venues offering foods from all over the world increases as well. According to an analysis by *Sommer in Hamburg*[1] the five most popular cuisines in Germany are Italian, German, Greek, Chinese and Turkish. It is not too far fetched that a multicultural city like Frankfurt offers an array of venues to try dishes from these cuisines. A useful tool for finding such venues in a city is *Foursquare*[2]. *Foursquare* is an app which lets users search, rate and recommend places to other users.

The goal of this project is to use the *Foursquare* API to explore the culinary landscape of Frankfurt and answer questions like:

- What is the most popular cuisine in Frankfurt?

- Are some cuisines over- or underrepresented?

- Are there correlations between the demography and the venues?

Who can benefit from this analysis? On the one hand, this report might reveal some good places for anyone who wants to try a new venue which serves one of the five mentioned cuisines. On the other hand, this report can give recommendations to future founders who want to open a new restaurant but who are uncertain what kind of dishes are popular and/or underrepresented.

---

[1] https://sommer-in-hamburg.de/kultur/kochen-essen-kulinarik/beliebteste-laender-kuechen/

[2] https://foursquare.com/

2 Data

The data used for this analysis is separated into four parts. In the following sections, they will be introduced briefly.
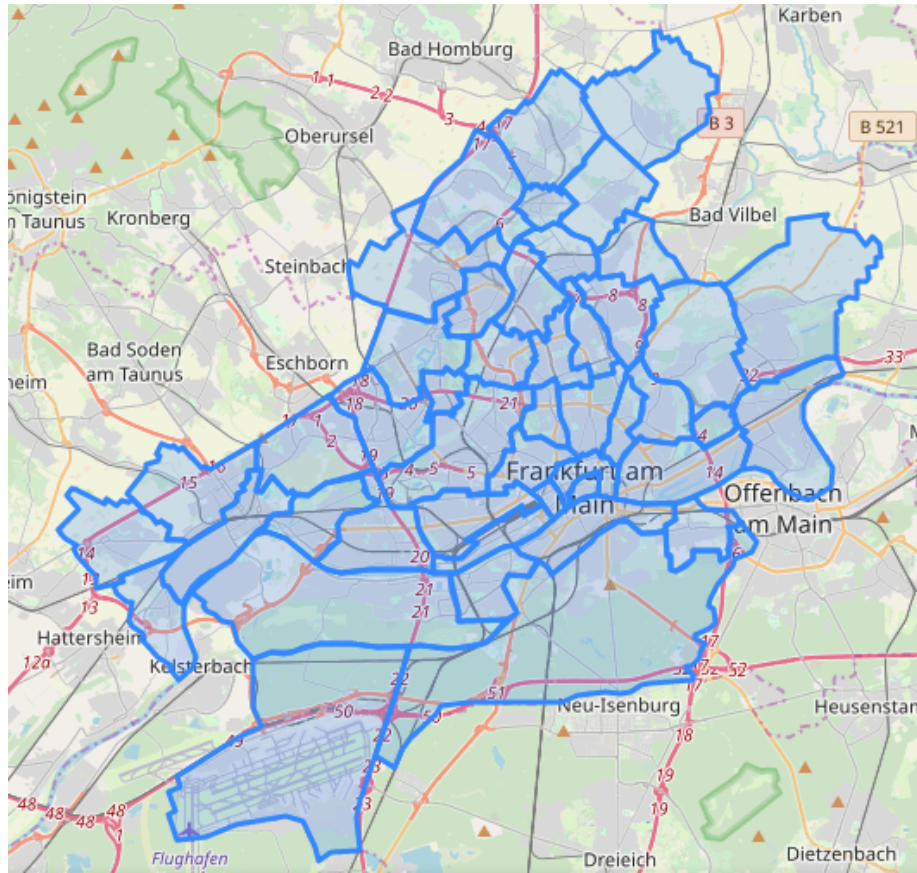
2.1 GeoData

Figure 1: Map of districts

The first part of the data is a GeoJSON file which contains the borders of the 46 districts in Frankfurt. These borders can be seen in figure 1. They will be used to assign the venues found by the *Foursquare* API to a district. The original file can be found on GitHub[3].

---

[3]`https://github.com/codeforamerica/click_that_hood/blob/master/public/data/frankfurt-main.geojson`

2.2 Census data

	District	Area in square km	Inhabitants	Female	Male	Germans	Foreigners	Foreigners in percent	Inhabitants per square km	Longitude	Latitude
0	Altstadt	0.506	4151	2015	2136	2604	1547	37.3	8204	8.681539	50.110590
1	Innenstadt	1.491	6605	3088	3517	3514	3091	46.8	4430	8.682205	50.112817
2	Bahnhofsviertel	0.542	3561	1378	2183	1638	1923	54.0	6570	8.667658	50.107810
3	Westend-Süd	2.497	18822	9593	9229	13655	5167	27.5	7538	8.661165	50.117950
4	Westend-Nord	1.632	10198	5321	4877	7214	2984	29.3	6249	8.666019	50.129776

Figure 2: Census data

The second part of the data consists of census data about each district in Frankfurt. The data dates back to 31.12.2018 and contains information like the number of inhabitants, the area or the amount of Germans and foreigners in each district. It does not contain the latitude and longitude of each district but both of these can be derived from the GeoJSON data referenced in section 2.1. After translating and cleaning, the data is stored in a *pandas* DataFrame (DF) as shown in figure 2. It has a shape of $(46, 11)$, one row for each district. The original data in German can be found here[4].

2.3 Overview of venues

The third part of the data is generated by using the *Foursquare* API. The API is used to find all venues for each of the five most popular cuisines in Germany (Italian, German, Greek, Chinese and Turkish). Figure 3 shows where these venues are located on the map. The resulting DF has a shape of $(625, 7)$ and stores the $Name, ID, Address, Latitutde, Longitude, Cuisine$ and $District$ of a venue. It is shown in figure 4.

---

[4]`https://www.frankfurt.de/sixcms/media.php/738/01_Bev%C3%B6lkerung_Ende2018.pdf`
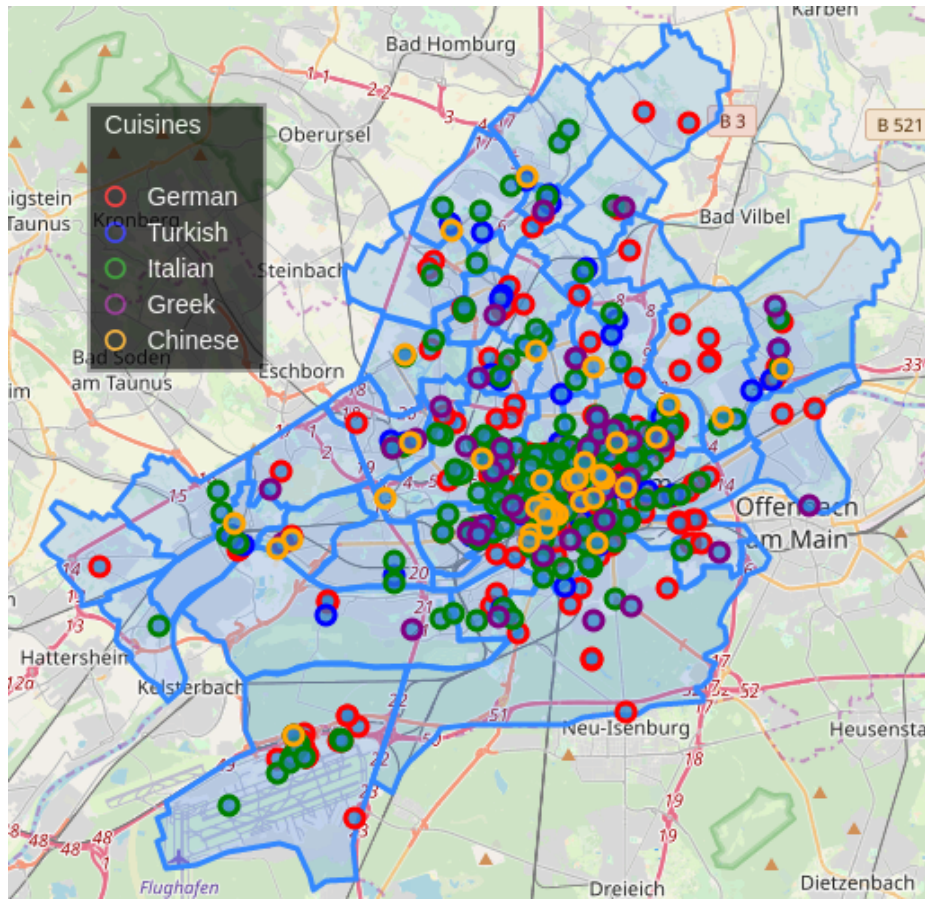
Figure 3: Venues on map

	Name	ID	Address	Lat	Lng	Cuisine	District
0	Paulaner am Dom	4be18ea74ed5c9b61ba1ccce	Domplatz 6	50.110876	8.685925	DE	Altstadt
1	Heimat – Essen und Weine	4cfd2a882c1aa090410e057a	Berliner Str. 70	50.111125	8.678286	DE	Altstadt
2	Römer Pils Brunnen	4b6ac454f964a52072df2be3	Töngesgasse 19	50.113450	8.683873	DE	Altstadt
3	Max on One	4e9dc4dff5b95ad93faa8788	Thurn-und-Taxis-Platz 2	50.115110	8.680233	DE	Innenstadt
4	Klosterhof	4b058851f964a52053be22e3	Weißfrauenstr. 3	50.109259	8.677224	DE	Altstadt

Figure 4: Venues DataFrame

2.4 Venue details

The fourth and final part of the data is also generated with the *Foursquare* API. Each venue contained in the first DF of section 2.3 is queried again to get the venue's details. The API returns information about a venue including ratings, tips, and photos. The results are stored in a separate DF shown in figure 5. It has a shape of $(615, 8)$. Note that this second DF has 10 less rows than the previous DF. The reason for that is that some venues serve dishes from multiple cuisines. When joining both DFs this becomes apparent. Figure 6 shows this for three exemplary venues.

	ID	Likes	Listed	Photos	Price	Rating	Rating Signals	Tips
0	4be18ea74ed5c9b61ba1ccce	293	260	425	2.0	8.5	433.0	97
1	4cfd2a882c1aa090410e057a	43	136	49	2.0	8.7	59.0	16
2	4b6ac454f964a52072df2be3	60	44	74	2.0	7.8	99.0	25
3	4e9dc4dff5b95ad93faa8788	16	14	44	2.0	8.6	17.0	3
4	4b058851f964a52053be22e3	109	127	147	2.0	7.8	169.0	42

Figure 5: Venue details DataFrame

	Name	ID	Address	Lat	Lng	Cuisine	District
102	Café Diesseits	4b05884df964a52031bd22e3	Konrad-Broßwitz-Str. 1	50.124916	8.639161	DE	Bockenheim
418	Café Diesseits	4b05884df964a52031bd22e3	Konrad-Broßwitz-Str. 1	50.124916	8.639161	IT	Bockenheim
48	Café Karin	4b05884ef964a5205cbd22e3	Großer Hirschgraben 28	50.111621	8.678097	DE	Altstadt
544	Café Karin	4b05884ef964a5205cbd22e3	Großer Hirschgraben 28	50.111621	8.678097	GR	Altstadt
545	Dorade am Main	4b1e51c2f964a520b41824e3	Dorade am Main Carl-von-Noorden Platz 5	50.097209	8.664123	GR	Sachsenhausen-Nord
251	Dorade am Main	4b1e51c2f964a520b41824e3	Dorade am Main Carl-von-Noorden Platz 5	50.097209	8.664123	TR	Sachsenhausen-Nord

Figure 6: Duplicate venues