

Capstone Project – Italian Restaurant in Berlin

INTRODUCTION

- **Background:**

- In this project we will try to find an optimal location for a restaurant. Specifically, this report will be targeted to stakeholders interested in opening an **Italian restaurant** in **Berlin**, Germany.

- **Problem:**

- Since there are lots of restaurants in Berlin we will try to detect **locations that are not already crowded with restaurants**

- **Interest:**

- We are also particularly interested in **areas with no Italian restaurants in vicinity**. We would also prefer locations **as close to city center as possible**, assuming that first two conditions are met

DATA ACQUISITION And CLEANING

- Data:

Based on definition of our problem, factors that will influence our decision are:

- number of existing restaurants in the neighborhood (any type of restaurant)
- number of and distance to Italian restaurants in the neighborhood, if any
- distance of neighborhood from city center

- Following data sources will be needed to extract/generate the required information:

- centers of candidate areas will be generated algorithmically and approximate addresses of centers of those areas will be obtained using **Google Maps API reverse geocoding**
- number of restaurants and their type and location in every neighborhood will be obtained using **Foursquare API**
- coordinate of Berlin center will be obtained using **Google Maps API geocoding** of well known Berlin location (Alexanderplatz)

METHODOLOGY

- Detect areas of Berlin that have low restaurant density, particularly those with low number of Italian restaurants. We will limit our analysis to area ~6km around city center
- In first step we have collected the required **data: location and type (category) of every restaurant within 6km from Berlin center** (Alexanderplatz). We have also **identified Italian restaurants** (according to Foursquare categorization).
- Second step in our analysis will be calculation and exploration of '**restaurant density**' across different areas of Berlin - we will use **heatmaps** to identify a few promising areas close to center with low number of restaurants in general (*and* no Italian restaurants in vicinity) and focus our attention on those areas
- In third and final step we will focus on most promising areas and within those create **clusters of locations that meet some basic requirements** established in discussion with stakeholders: we will take into consideration locations with **no more than two restaurants in radius of 250 meters**, and we want locations **without Italian restaurants in radius of 400 meters**
- We will present map of all such locations but also create clusters (using **k-means clustering**) of those locations to identify general zones / neighborhoods / addresses which should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders

Exploratory Data Analysis

- Let's perform some basic explanatory data analysis and derive some additional info from our raw data. First let's count the **number of restaurants in every area candidate**.

```
Average number of restaurants in every area with radius=300m: 5.1401098901098905
```

```
] :
```

	Address	Latitude	Longitude	X	Y	Distance from center	Restaurants in area
0	Ringbahnstraße 50, 12099 Berlin	52.470194	13.388575	390541.280176	5.814557e+06	5992.495307	5
1	09R/27L, 12101 Berlin	52.470314	13.397404	391141.280176	5.814557e+06	5840.376700	0
2	09R/27L, 12049 Berlin	52.470434	13.406234	391741.280176	5.814557e+06	5747.173218	0
3	Oderstraße 174, 12049 Berlin	52.470552	13.415063	392341.280176	5.814557e+06	5715.767665	0
4	Warthestraße 23A, 12051 Berlin	52.470670	13.423893	392941.280176	5.814557e+06	5747.173218	1
5	Altenbraker Str. 15, 12053 Berlin	52.470788	13.432722	393541.280176	5.814557e+06	5840.376700	7
6	Karl-Marx-Straße 212, 12055 Berlin	52.470904	13.441552	394141.280176	5.814557e+06	5992.495307	6
7	Hessenring 34, 12101 Berlin	52.474683	13.375159	389641.280176	5.815077e+06	5855.766389	0
8	Kleineweg 125, 12101 Berlin	52.474804	13.383989	390241.280176	5.815077e+06	5604.462508	0
9	09L/27R, 12101 Berlin	52.474924	13.392820	390841.280176	5.815077e+06	5408.326913	0

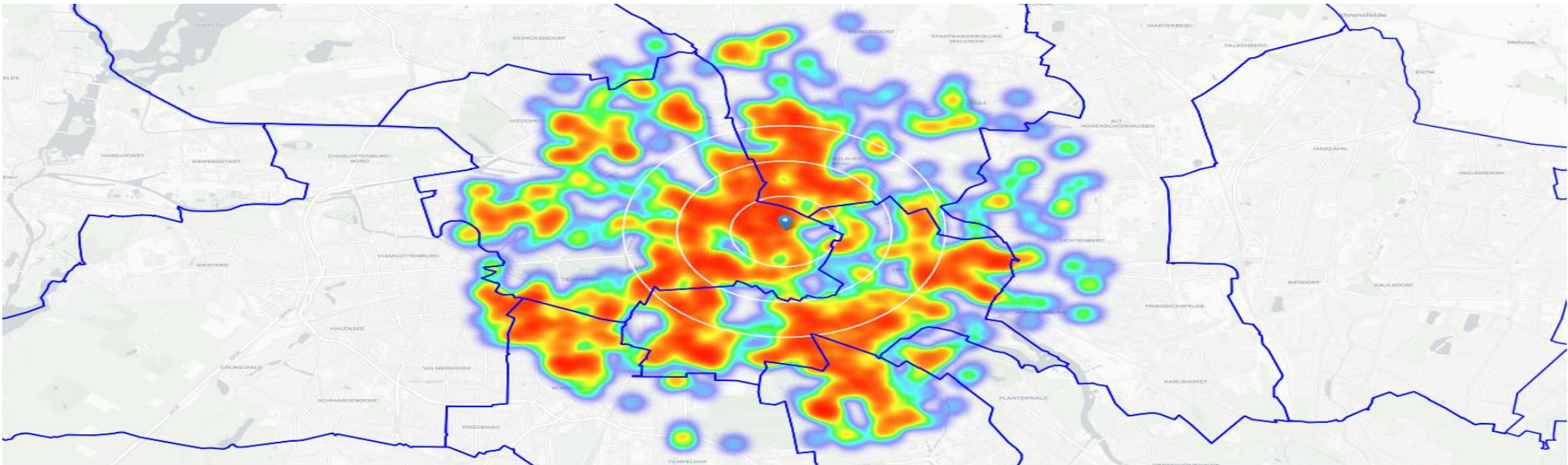
Analysis

- OK, now let's calculate the **distance to nearest Italian restaurant from every area candidate center** (not only those within 300m - we want distance to closest one, regardless of how distant it is).

	Address	Latitude	Longitude	X	Y	Distance from center	Restaurants in area	Distance to Italian restaurant
0	Ringbahnstraße 50, 12099 Berlin	52.470194	13.388575	390541.280176	5.814557e+06	5992.495307	5	264.408532
1	09R/27L, 12101 Berlin	52.470314	13.397404	391141.280176	5.814557e+06	5840.376700	0	830.999331
2	09R/27L, 12049 Berlin	52.470434	13.406234	391741.280176	5.814557e+06	5747.173218	0	1269.038823
3	Oderstraße 174, 12049 Berlin	52.470552	13.415063	392341.280176	5.814557e+06	5715.767665	0	829.067436
4	Warthestraße 23A, 12051 Berlin	52.470670	13.423893	392941.280176	5.814557e+06	5747.173218	1	575.681166
5	Altenbraker Str. 15, 12053 Berlin	52.470788	13.432722	393541.280176	5.814557e+06	5840.376700	7	293.966217
6	Karl-Marx-Straße 212, 12055 Berlin	52.470904	13.441552	394141.280176	5.814557e+06	5992.495307	6	317.390305
7	Hessenring 34, 12101 Berlin	52.474683	13.375159	389641.280176	5.815077e+06	5855.766389	0	776.047531
8	Kleineweg 125, 12101 Berlin	52.474804	13.383989	390241.280176	5.815077e+06	5604.462508	0	378.018237
9	09L/27R, 12101 Berlin	52.474924	13.392820	390841.280176	5.815077e+06	5408.326913	0	635.252552

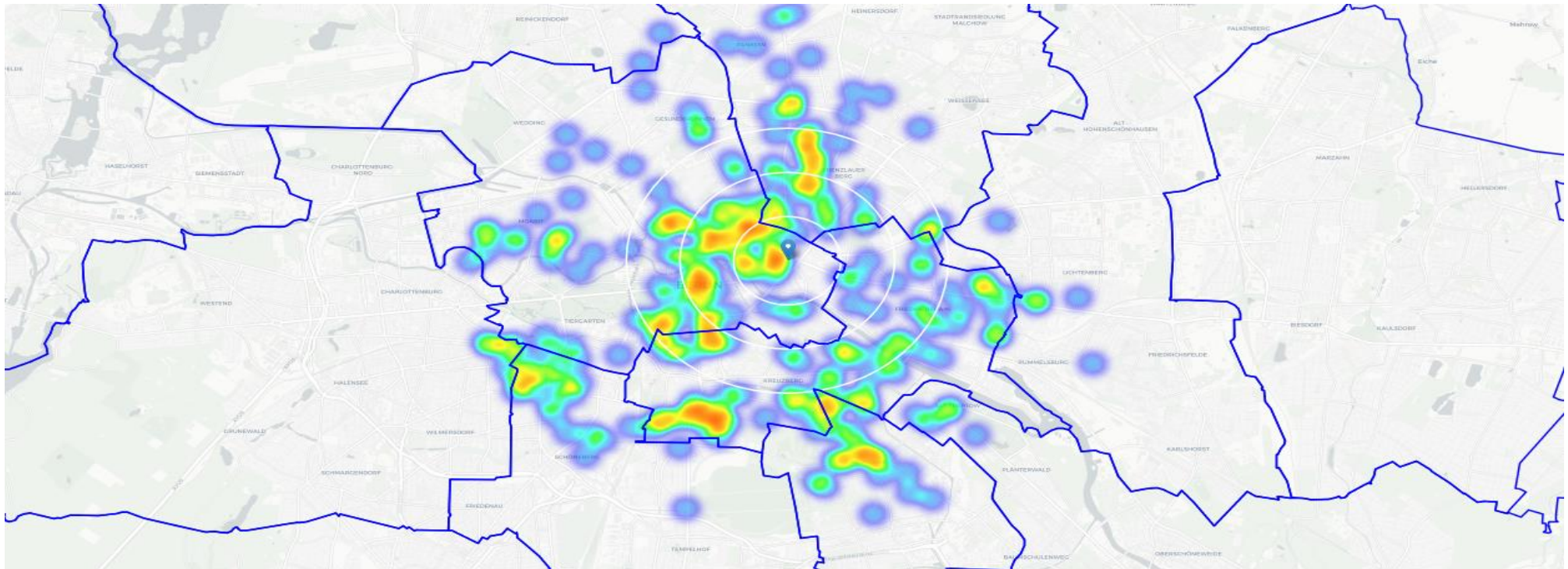
Analysis

- **on average Italian restaurant can be found within ~500m** from every area center candidate. That's fairly close, so we need to filter our areas carefully!
- Let's create a map showing **heatmap / density of restaurants** and try to extract some meaningful info from that. Also, let's show **borders of Berlin boroughs** on our map and a few circles indicating distance of 1km, 2km and 3km from Alexanderplatz



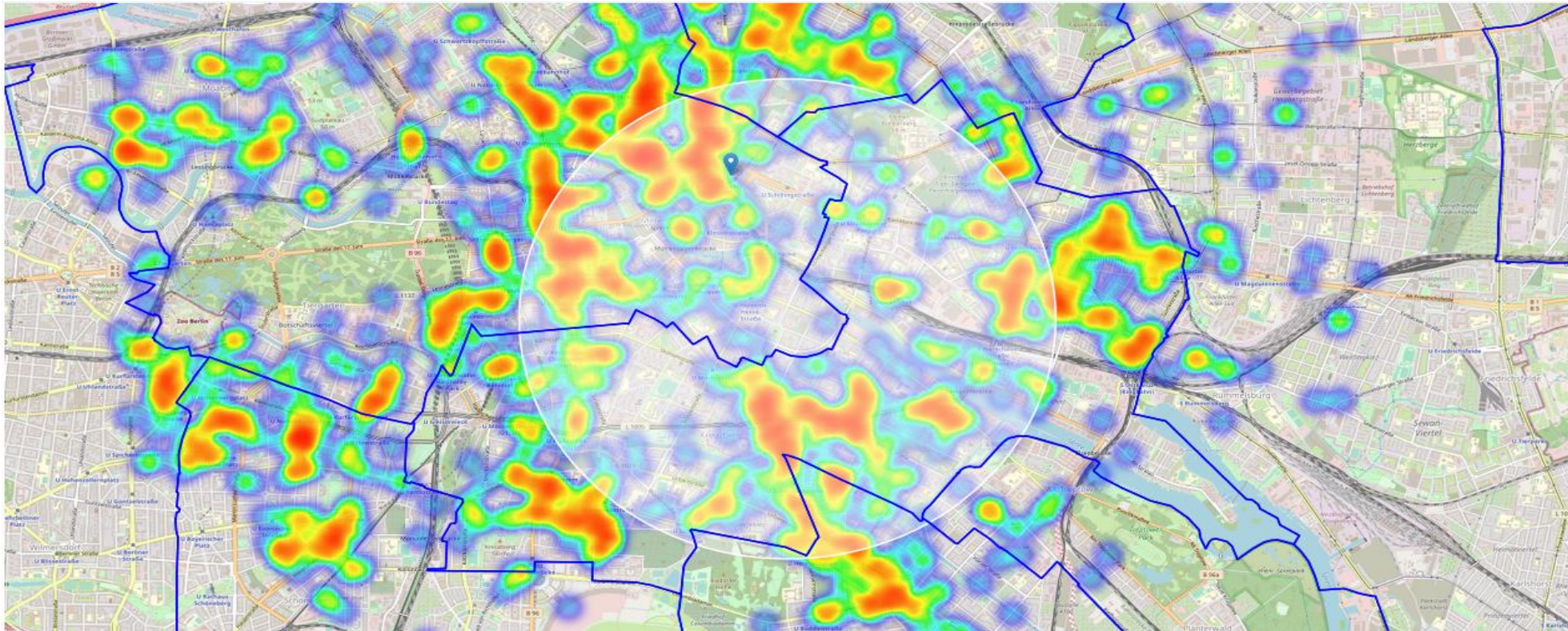
Analysis

- Looks like a few pockets of low restaurant density closest to city center can be found **south, south-east and east from Alexanderplatz.**
- Let's create another heatmap map showing **heatmap/density of Italian restaurants** only.



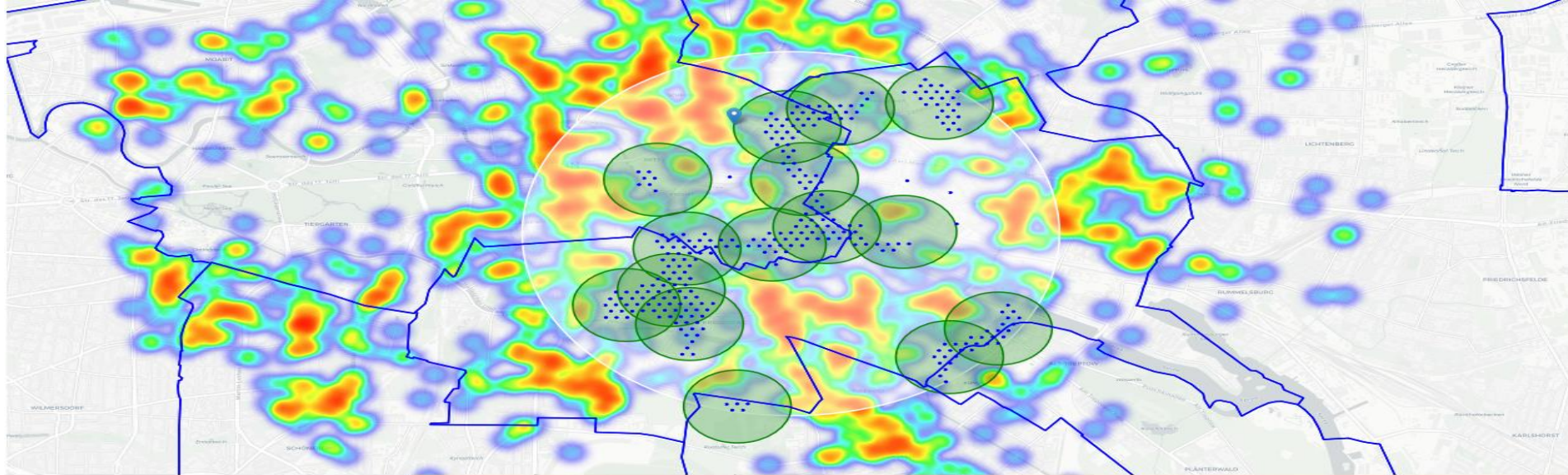
Analysis

- Let's define new, more narrow region of interest, which will include low-restaurant-count parts of Kreuzberg and Friedrichshain closest to Alexanderplatz



Analysis

- we're interested only in **locations with no more than two restaurants in radius of 250 meters, and no Italian restaurants in radius of 400 meters.**
- Looking good. We now have a bunch of locations fairly close to Alexanderplatz (mostly in Kreuzberg, Friedrichshain and south-east corner of Mitte boroughs), and we know that each of those locations has no more than two restaurants in radius of 250m, and no Italian restaurant closer than 400m. Any of those locations is a potential candidate for a new Italian restaurant, at least based on nearby competition



Analysis

- This concludes our analysis. We have created 15 addresses representing centers of zones containing locations with low number of restaurants and no Italian restaurants nearby, all zones being fairly close to city center (all less than 4km from Alexanderplatz, and about half of those less than 2km from Alexanderplatz).
- Although zones are shown on map with a radius of ~500 meters (green circles), their shape is actually very irregular and their centers/addresses should be considered only as a starting point for exploring area neighborhoods in search for potential restaurant locations.
- Most of the zones are located in Kreuzberg and Friedrichshain boroughs, which we have identified as interesting due to being popular with tourists, fairly close to city center and well
- connected by public transport

```
=====
Addresses of centers of areas recommended for further analysis
=====

Vor dem Schlesischen Tor 2, 10997 Berlin      => 3.7km from Alexanderplatz
Stallschreiberstraße 46, 10969 Berlin        => 1.7km from Alexanderplatz
Platz der Vereinten Nationen 25, 10249 Berlin => 1.0km from Alexanderplatz
Köpenicker Str. 40, 10179 Berlin             => 1.6km from Alexanderplatz
Dog Park, Hasenheide 82, 10967 Berlin        => 3.9km from Alexanderplatz
Prinzenstraße 1, 10969 Berlin                => 2.8km from Alexanderplatz
Breite Str. 1, 10178 Berlin                  => 1.0km from Alexanderplatz
Landsberger Allee 37, 10249 Berlin           => 1.9km from Alexanderplatz
Holzmarktstraße 55, 10179 Berlin             => 1.0km from Alexanderplatz
Berolinastraße 14, 10178 Berlin              => 0.5km from Alexanderplatz
Neuenburger Str. 16, 10969 Berlin            => 2.7km from Alexanderplatz
Michaelkirchpl. 23, 10179 Berlin             => 1.7km from Alexanderplatz
Lobeckstraße 62, 10969 Berlin                => 2.3km from Alexanderplatz
Wiener Straße 58, 10999 Berlin               => 3.8km from Alexanderplatz
Str. der Pariser Kommune 8, 10243 Berlin     => 2.2km from Alexanderplatz
```

Results and Discussion

- Our analysis shows that although there is a great number of restaurants in Berlin (~2000 in our initial area of interest which was 12x12km around Alexanderplatz), there are pockets of low restaurant density fairly close to city center
- Highest concentration of restaurants was detected north and west from Alexanderplatz, so we focused our attention to areas south, south-east and east, corresponding to boroughs Kreuzberg, Friedrichshain and south-east corner of central Mitte borough.
- Another borough was identified as potentially interesting (Prenzlauer Berg, north-east from Alexanderplatz), but our attention was focused on Kreuzberg and Friedrichshain which offer a combination of popularity among tourists, closeness to city center, strong socio-economic dynamics *and* a number of pockets of low restaurant density.
- Result of all this is 15 zones containing largest number of potential new restaurant locations based on number of and distance to existing venues - both restaurants in general and Italian restaurants particularly.
- This, of course, does not imply that those zones are actually optimal locations for a new restaurant! Purpose of this analysis was to only provide info on areas close to Berlin center but not crowded with existing restaurants (particularly Italian) - it is entirely possible that there is a very good reason for small number of restaurants in any of those areas, reasons which would make them unsuitable for a new restaurant regardless of lack of competition in the area.

Conclusion

- Purpose of this project was to identify Berlin areas close to center with low number of restaurants (particularly Italian restaurants) in order to aid stakeholders in narrowing down the search for optimal location for a new Italian restaurant.
- . By calculating restaurant density distribution from Foursquare data we have first identified general boroughs that justify further analysis (Kreuzberg and Friedrichshain), and then generated extensive collection of locations which satisfy some basic requirements regarding existing nearby restaurants.
- Clustering of those locations was then performed in order to create major zones of interest (containing greatest number of potential locations) and addresses of those zone centers were created to be used as starting points for final exploration by stakeholders.
- Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.