Geo spatial data analysis for opening a food Restaurant or an office in Berlin

Prabhu Kumar Reddy Appalapuri M.Sc
prabhu.appalapuri@gmail.com

Table of Content

Abstract

Geospatial analysis is a process of data gathering and manipulation of data such as GPS, historical data. These are described in terms of geographic coordinates or a location in terms of a street address, postal code. In the following report, the data is composed of GPS coordinates, location data of several entities(i.e company) from Berlin. Data exploration and clustering were applied to the data for recommending a better place to open a restaurant/office in Berlin. These place recommendations were made based on the highest number of companies and to have a reasonable rental price.

1 Introduction & Problem statement

During the daytime, especially during lunch hours, office areas provide huge opportunities for restaurants. An average meal price i.e one lunch meal is about 5€. The shops are usually always full during lunch hours (11 am to 2.30 pm). Given this scenario, I will be showing the benefits and pitfalls of opening a restaurant in highly dense office places including office delivery. However, I am unaware of the profit, I do believe there will be huge benefits by opening a restaurant in the dense area of companies. I will be covering the top 7 places in Berlin.

Additionally, by understanding the type of companies that are located in each area will result in valuable information for opening a new office. Such as finding an area relatively less cost for opening an office.

1.1 Target Audience

Probably, the following types of clients or a group of people would be interested in this project.

1. For data scientists, who want to do exploratory data analysis techniques to obtain necessary data, to analyze, and to tell a story out of it.
2. Business personnel who wants to invest or open a restaurant. This analysis will be a comprehensive guide to start or expand restaurants targeting the large pool of office places during lunch hours.
3. Furthermore, the analysis of company locations in Berlin will hugely benefit the business personnel for opening their new office.

2 Data preparation

The required data is gathered through various resources such as wikipedia, firmendb, suche-postleitzahl and miet-spiegel web sources.

2.1 Description of the data

To solve the above problems, a location-based data set is important. However, neither it is not available directly on the internet nor from the Foursquare website. Hence, I decided to scrape the required data.There are 2 data sets:

1. Company-related data with following columns:
   - Company name: In Germany, a company is characterized as mbH, GmbH, AG, AG &Co.
   - Address: It is composed of a street name, GPS coordinates, zip code, neighborhood
   - Category: It is a type of company i.e software company, construction company e.t.c
2. Berlin geographical data set is composed of the following columns:
   - Zip Code
   - Neighborhood
   - District

2.1.1 Load dataset having district and its neighborhoods

Berlin neighborhoods and boroughs were extracted from wiki with a BeautifulSoup library. The following has looks like this.

```
berlin_neighborhoods = pd.read_csv("data/berlin_places.csv")
berlin_neighborhoods = berlin_neighborhoods[["Ortsteil","Bezirk"]]
berlin_neighborhoods["Ortsteil"] = berlin_neighborhoods["Ortsteil"].str.strip()
berlin_neighborhoods["Bezirk"] = berlin_neighborhoods["Bezirk"].str.strip()
```

	Ortsteil	Bezirk	Bezirkgeo
0	Mitte	Mitte	[52.5176896, 13.4023757]
1	Moabit	Mitte	[52.5176896, 13.4023757]
2	Hansaviertel	Mitte	[52.5176896, 13.4023757]
3	Tiergarten	Mitte	[52.5176896, 13.4023757]
4	Wedding	Mitte	[52.5176896, 13.4023757]

2.1.2 Load the dataset Neighborhood and its postal codes

```
berlin_postalcodes = pd.read_excel("data/Bundesland Berlin.xlsx")
berlin_postalcodes[berlin_postalcodes.Ortsteil=="Wedding"]
```

	PLZ	Ortsteil
241	13347	Wedding
243	13349	Wedding
244	13351	Wedding
246	13353	Wedding
249	13357	Wedding
251	13359	Wedding
256	13405	Wedding
259	13407	Wedding
262	13409	Wedding

Berlin neighbourhood zipcodes are freely available from the source "suche-postleitzahl.org". The postal code dataset is shown below.

2.1.3 Load the Company Dataset

```
profiles = pd.read_csv("data/company_profile.csv")
profiles = profiles[["url","location","info","branch"]]
profiles['location'] = profiles['location'].apply(lambda x: "{:.3f}".format(x) if not pd.isnull(x) else x)
```

```
# Grouping values based on company name, i.e a single row per company
profiles['location'] = profiles['location'].astype(str)
profiles['info'] = profiles['info'].astype(str)
cleaned_profiles = profiles.groupby(["url"])["info"].agg([('info', ','.join)])
cleaned_location = profiles.groupby(["url"])["location"].agg([('location', ','.join)])
profiles["branch"] = profiles["branch"].astype(str)
cleaned_branch = profiles.groupby(["url"])["branch"].agg([('branch', ','.join)])
cleaned_branch["branch"] = cleaned_branch["branch"].apply(lambda x : [x.split(",")[0]])
print(cleaned_branch.head())
print(cleaned_location.head())
print(cleaned_profiles.head())
```

```
                                                                    branch
url
(KA) Kraft Automobile GmbH                     [Autohandel und Kfz-Handel (Nutzfahrzeuge]
(know:bodies) gesellschaft für integrierte komm...        [Public-Relations-Beratung]
07schanksysteme gmbh                                 [Herstellung von Messinstrumenten]
0815-Industries KG                                           [Werbung und Marketing]
1 Berlin x Hausverwaltung GmbH + Co. KG                                 [Verwaltung]
                                                              location
url
(KA) Kraft Automobile GmbH                     52.479,13.336,nan,nan,nan,nan
(know:bodies) gesellschaft für integrierte komm...  52.518,13.287,nan,nan,nan,nan
07schanksysteme gmbh                           52.542,13.355,nan,nan,nan,nan
0815-Industries KG                             52.564,13.474,nan,nan,nan,nan
1 Berlin x Hausverwaltung GmbH + Co. KG                    nan,nan,nan,nan
                                                                          info
url
(KA) Kraft Automobile GmbH                     nan,nan,(KA) Kraft Automobile GmbH,Wexstrasse ...
(know:bodies) gesellschaft für integrierte komm...  nan,nan,(know:bodies) gesellschaft für integri...
07schanksysteme gmbh                           nan,nan,07schanksysteme gmbh,Sprengelstrasse 1...
0815-Industries KG                             nan,nan,0815-Industries KG,Feldtmannstrasse 15...
1 Berlin x Hausverwaltung GmbH + Co. KG        1 Berlin x Hausverwaltung GmbH + Co. KG,Königi...
```

The data frames namely "cleaned_branch", "cleaned_location" and "cleaned_profiles" were preprocessed by removing unnecessary column values for example "None", "Nan" values.

```
# Droping data which does't have Latitude an longitude
company_data = company_data.dropna()
print(company_data.shape)
company_data.head(8)
```

(5660, 7)

	Name	Street	Zipcode	City	Lat	Lon	Branch
0	(KA) Kraft Automobile GmbH	Wexstrasse 15	10715	Berlin	52.479	13.336	Autohandel und Kfz-Handel (Nutzfahrzeuge
1	(know:bodies) gesellschaft für integrierte kom...	Sophie-Charlotten-Strasse 103	14059	Berlin	52.518	13.287	Public-Relations-Beratung
2	07schanksysteme gmbh	Sprengelstrasse 15	13353	Berlin	52.542	13.355	Herstellung von Messinstrumenten
3	0815-Industries KG	Feldtmannstrasse 152	13088	Berlin	52.564	13.474	Werbung und Marketing
7	1-2-3 Beschläge GmbH	Colditzstrasse 33	12099	Berlin	52.455	13.396	Herstellung von Schlössern und Beschlägen
8	1-2-3 Gebäudemanagement Berlin GmbH	Fredericiastrasse 28	14059	Berlin	52.511	13.282	Reinigung von Gebäuden
9	1-2-3 Marriage UG (haftungsbeschränkt)	Elisabethstrasse 35	12307	Berlin	52.390	13.387	Dienstleistungen a.n.g.
10	1. maXXwill UG (haftungsbeschränkt)	Hubertusstrasse 8	12163	Berlin	52.460	13.326	Grosshandel mit Computern

2.1.4 Load the dataset Average Rental Price per Ortsteil

Average rental price per are is available from the "mietspiegel-Berlin". These rental prices were shown below.

```
avgprice = pd.read_csv("data/prices.csv")
# preprocess data
avg1 = avgprice["text"][0]
avg1 = list(filter(None, [item.strip() for item in avg1.split("\n")]))

avg2 = avgprice["text"][1]
avg2 = list(filter(None, [item.strip() for item in avg2.split("\n")]))
```

```
def dataframe(x):
    y = np.array(x)
    y = y.reshape(-1,2)
    df = pd.DataFrame(y[1:], columns=y[0])
    df["Ortsteil"] = df["STADTTEIL"].apply(lambda x : x.split(" (")[0])
    return df
```

```
df_avg1 = dataframe(avg1)
df_avg1.head()
```

	STADTTEIL	€/m²	Ortsteil
0	Lichterfelde (Steglitz)	11,20 €	Lichterfelde
1	Mahlsdorf (Hellersdorf)	10,36 €	Mahlsdorf
2	Mariendorf (Tempelhof)	10,90 €	Mariendorf
3	Marienfelde (Tempelhof)	19,82 €	Marienfelde
4	Märkisches Viertel	8,22 €	Märkisches Viertel

3 Methodology

In this work, data analysis, data visualization and clustering were used. As an initial step all necessary libraries such as Wikipedia, bs4, sklearn. However, the data needed to preprocess before the data exploration begins. After data preprocessing of company dataset will be grouped by neighborhood area to get the total number of companies per area. During the first step, company data will be grouped per zip code. Nevertheless, the zip codes which are having above 50 companies will be considered to get a better understanding of nearby shops(i.e venues). For the second step, the grouped data per zip code will be grouped again. At this step from the dataframe, an index number of that row will be considered as the label of the neighborhood. For visualization purposes, these labels will be used for analyzing the clusters of companies that are located in each area.

For opening a restaurant or an office, rental price and surrounding places to a company will play a major role to thrive in the business . The average rental prices were gathered only for neighborhoods with a minimum of 50 companies. Thus the rental price of the dataframe are ordered in an ascending manner. All these information, then used for ranking of each neighborhood. The ranking will be done based on highest companies per neighborhood and to have the lowest price. From the Foursquare API, list of all nearby shops i.e venues will be extracted and categorised according to the rank of the neighborhood's. Also, it helps to determine the companies which are similar in terms of nearby venues to the city center and the other neighborhood

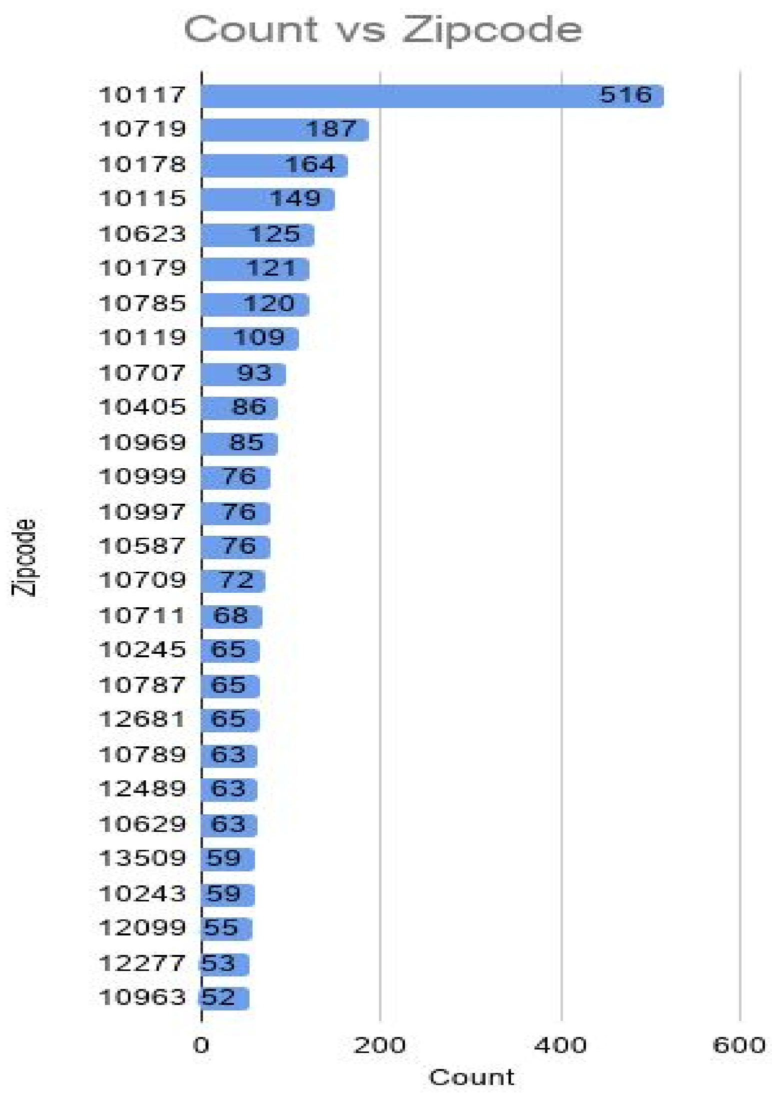
As a part of the report, I will walk you through each step of this project and address them separately. These answers will justify a better place to open a restaurant or to open an office for my stakeholders.

4 Evaluations

4.1 Companies per zip code

The following table shows the number of companies per zip code in Berlin. Also, Geopy library was used for better visualization of the cluster of companies.

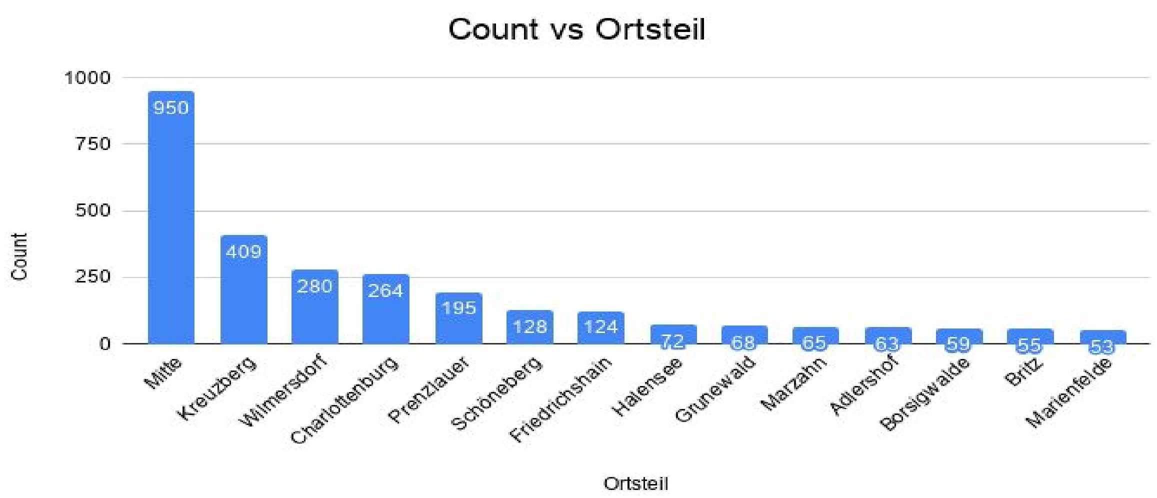
```
df_agg = company_data[['Ortsteil','Zipcode', 'Name']].groupby(['Ortsteil','Zipcode']).count()
df_agg = df_agg.reset_index()
df_agg = df_agg.sort_values(by = "Name", ascending=False)
df_agg_50 = df_agg[df_agg["Name"]>50]
print("Shape :", df_agg_50.shape)
df_agg_50
```



4.2 Companies per neighborhood

```
grouped = df_50_loc.groupby('Ortsteil')["Count"].sum().reset_index()
grouped = grouped.sort_values('Count', ascending=False)
grouped = grouped.reset_index()
grouped
```

	Index	Ortsteil	Count
0	10	Mitte	950
1	7	Kreuzberg	409
2	13	Wilmersdorf	280
3	3	Charlottenburg	264
4	11	Prenzlauer Berg	195
5	12	Schöneberg	128
6	4	Friedrichshain	124
7	6	Halensee	72
8	5	Grunewald	68
9	9	Marzahn	65
10	0	Adlershof	63
11	1	Borsigwalde	59
12	2	Britz	55
13	8	Marienfelde	53



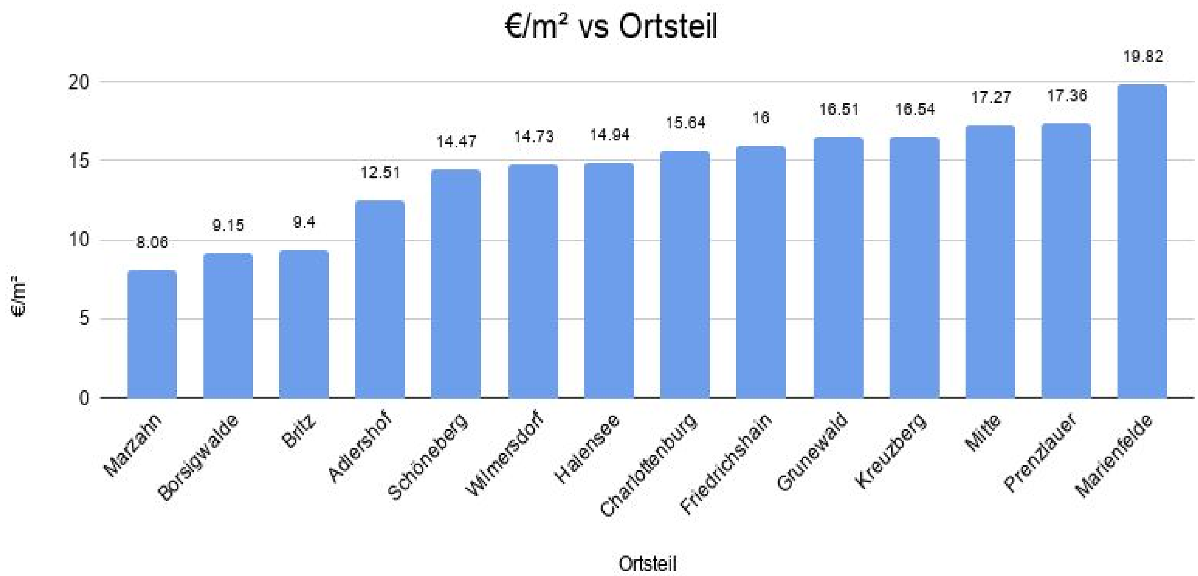
After having data analysis of company data, it describes, neighborhood "Mitte" has the highest number of companies. "Kreuzberg" has the second-highest number of companies. However, "Schöneberg" and "Friedrichshain" are in the top 7 positions. "Mitte" area is situated in central Berlin and mostly in its old town, it is traversed by the river Spree. Also, most of the city tourist attractions are situated in Mitte. Hence, it does sound like an ideal location for a restaurant or an office there. However, it might be expensive. Let's continue with the further analysis of average rental prices and which types of venues have existed in each neighborhood.

4.3 Average rental price per area

```
best_location_prices = merged.loc[grouped.Ortsteil.values.tolist()]["€/m²"]
best_location_prices = best_location_prices.reset_index()
best_location_prices = best_location_prices.groupby("Ortsteil")["€/m²"].mean().reset_index()
best_location_prices = best_location_prices.sort_values("€/m²")

best_location_prices
```

	Ortsteil	€/m²
9	Marzahn	8.06
1	Borsigwalde	9.15
2	Britz	9.4
0	Adlershof	12.51
12	Schöneberg	14.47
13	Wilmersdorf	14.73
6	Halensee	14.94
3	Charlottenburg	15.64
4	Friedrichshain	16
5	Grunewald	16.51
7	Kreuzberg	16.54
10	Mitte	17.27
11	Prenzlauer Berg	17.36
8	Marienfelde	19.82



4.4 Rank the neighborhoods

Probably, it is the best approach so far for opening a restaurant or an office based on statistics of data. Rule of thumb, having the lowest price, maximizing the number of companies around the restaurant helps to thrive a business. Therefore this logic helps to rank each location to select the top 7 places for any business in an ideal place.

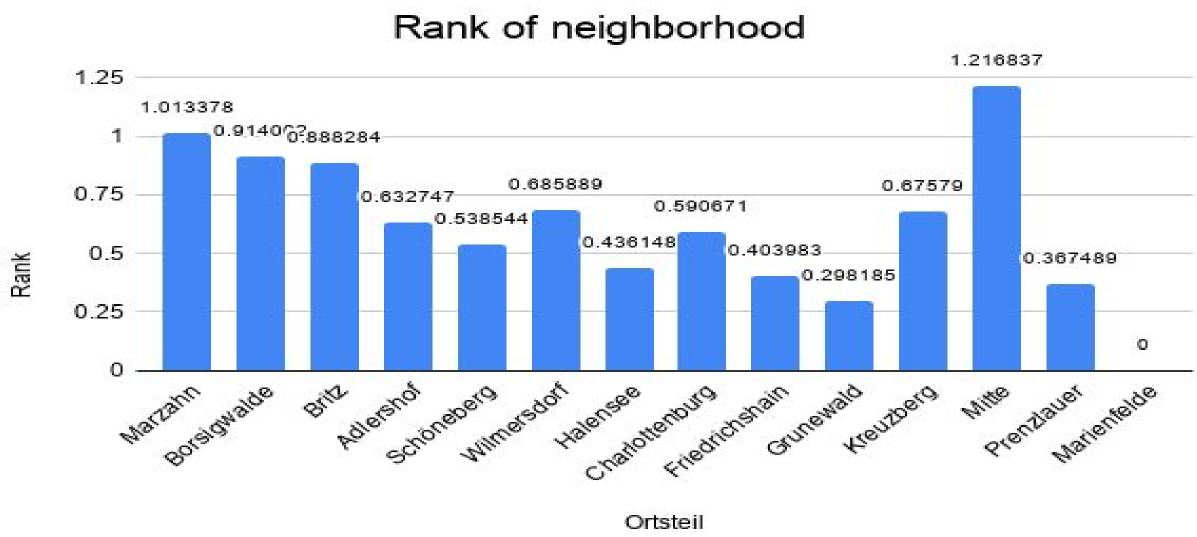
```
from sklearn.preprocessing import minmax_scale
```

```
sc = minmax_scale(avg_rp_nei[["Count", "€/m²"]])
sc_df = pd.DataFrame(sc, columns=["Wei_Count", "Wei_price"])
sc_df["Wei_price"] = -(sc_df["Wei_price"]-1)
sc_df.loc[13,"Wei_price"] = 0
```

```
avg_rp_nei = pd.concat([avg_rp_nei,sc_df], axis =1)
avg_rp_nei["Rank"] = sc_df.sum(axis=1)
avg_rp_nei
```

	index	Ortsteil	Count	€/m²	Wei_Count	Wei_price	Rank
0	9	Marzahn	65	8.06	0.013378	1	1.013378
1	1	Borsigwalde	59	9.15	0.006689	0.907313	0.914002
2	2	Britz	55	9.4	0.00223	0.886054	0.888284
3	0	Adlershof	63	12.51	0.011148	0.621599	0.632747
4	12	Schöneberg	128	14.47	0.083612	0.454932	0.538544
5	13	Wilmersdorf	280	14.73	0.253066	0.432823	0.685889
6	6	Halensee	72	14.94	0.021182	0.414966	0.436148

7	3	Charlottenburg	264	15.64	0.235229	0.355442	0.590671
8	4	Friedrichshain	124	16	0.079153	0.32483	0.403983
9	5	Grunewald	68	16.51	0.016722	0.281463	0.298185
10	7	Kreuzberg	409	16.54	0.396878	0.278912	0.67579
11	10	Mitte	950	17.27	1	0.216837	1.216837
12	11	Prenzlauer Berg	195	17.36	0.158305	0.209184	0.367489
13	8	Marienfelde	53	19.82	0	0	0



4.5 Grouping similar services of companies

In top 7 neighborhoods, a minimum of 10 companies with a particular branch were considered to maintain consistency for recommendation of a place.

```
# Filtering company dataset for top 7 locations.
top_7_cp_data = company_data.merge(top_7, left_on="Ortsteil", right_on="Ortsteil")
top_7_cp_data.shape
```

```
(2054, 14)
```

```
x = top_7_cp_data.groupby(["Branch"])["Rank"].count().reset_index()
x[x["Rank"]>10].sort_values("Rank", ascending=False)["Rank"].sum()
```

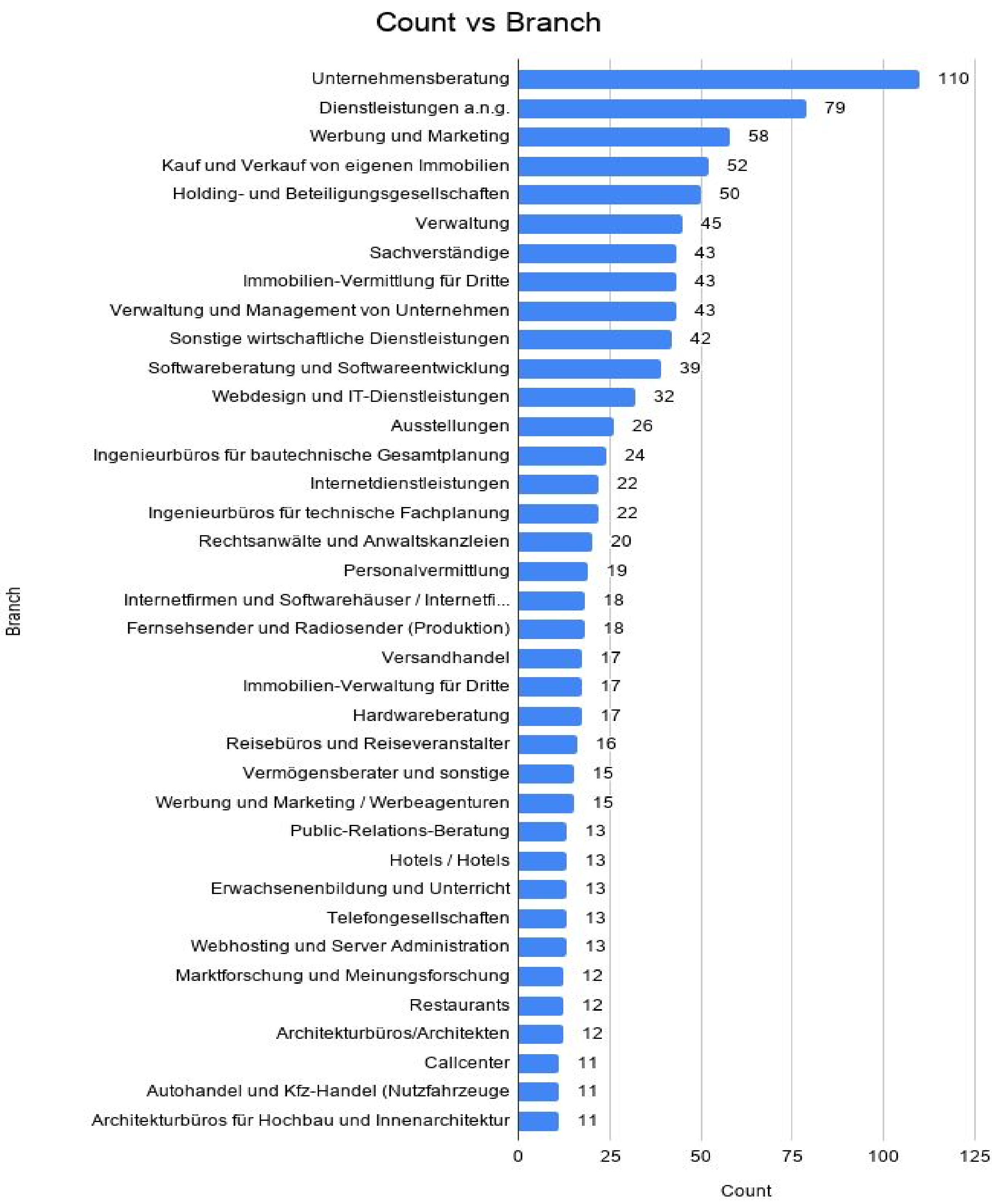
```
1036
```

```
services_df = x[x["Rank"]>10].sort_values("Rank", ascending=False)
services_df = services_df.reset_index(drop=True)
```

```
services_df = services_df.reset_index()
services_df
```

	index	Branch	Count
0	0	Unternehmensberatung	110
1	1	Dienstleistungen a.n.g.	79
2	2	Werbung und Marketing	58
3	3	Kauf und Verkauf von eigenen Immobilien	52
4	4	Holding- und Beteiligungsgesellschaften	50
5	5	Verwaltung	45
6	6	Sachverständige	43
7	7	Immobilien-Vermittlung für Dritte	43
8	8	Verwaltung und Management von Unternehmen und ...	43
9	9	Sonstige wirtschaftliche Dienstleistungen	42
10	10	Softwareberatung und Softwareentwicklung	39
11	11	Webdesign und IT-Dienstleistungen	32
12	12	Ausstellungen	26
13	13	Ingenieurbüros für bautechnische Gesamtplanung	24
14	14	Internetdienstleistungen	22
15	15	Ingenieurbüros für technische Fachplanung	22
16	16	Rechtsanwälte und Anwaltskanzleien	20
17	17	Personalvermittlung	19
18	18	Internetfirmen und Softwarehäuser / Internetfi...	18
19	19	Fernsehsender und Radiosender (Produktion)	18
20	20	Versandhandel	17
21	21	Immobilien-Verwaltung für Dritte	17
22	22	Hardwareberatung	17
23	23	Reisebüros und Reiseveranstalter	16
24	24	Vermögensberater und sonstige Vermögensberatung	15
25	25	Werbung und Marketing / Werbeagenturen	15
26	26	Public-Relations-Beratung	13
27	27	Hotels / Hotels	13
28	28	Erwachsenenbildung und Unterricht	13
29	29	Telefongesellschaften	13
30	30	Webhosting und Server Administration	13
31	31	Marktforschung und Meinungsforschung	12
32	32	Restaurants	12
33	33	Architekturbüros/Architekten	12
34	34	Callcenter	11

35	35	Autohandel und Kfz-Handel (Nutzfahrzeuge	11
36	36	Architekturbüros für Hochbau und Innenarchitektur	11



4.6 Exploring nearby venues to each company

Here, I am going to use the knowledge of venue_data and the top 7 areas having a minimum of 10 companies which belongs to a particular branch in each area. I will walk you through the top 7 areas having different types of venue categories.

Now that I have valuable information on each company i.e rental price, it's rank wise neighborhood preference, nearby venue categories, zip Code, and neighborhood name. Based on this information, we will do further analysis. Also, we can cluster the companies that are similar in spatial data analysis. This may give a broad idea for opening a restaurant, whether a similar restaurant already opened or not. In terms of company services (i.e branch wise) which types of venue categories have existed. Eventually, these kinds of information reveal an ideal location for having a new office/restaurant or relocation of an existing branch. Nearby venues were extracted with in a range of 800m to each company.

To get nearby venues, I have used Foursquare API. This API is limited to 950 calls per day. If you want to have more API calls, then upgrade your account. After analyzing data there are 354 unique categories available.

Here, I am going to use the knowledge of venue data for each company from Foursquare API and the top 7 areas having a minimum of 10 companies which belongs to a particular branch in each area.

Now that I have valuable information on each company i.e rental price, it’s rank wise neighborhood preference, nearby venue categories, zip code, and neighborhood name. Based on this information, we will do further analysis. Also, we can cluster the company’s that are similar in spatial data analysis. This may give a broad idea for opening a restaurant, whether a similar restaurant already opened or not. In terms of company services(i.e branch wise) which types of venue categories have existed. Eventually, these kinds of information reveal an ideal location for having a new office/restaurant or relocation of an existing branch.

```python
# Let's find out how many unique categories can be curated from all the returned venues
print('There are {} uniques categories.'.format(len(venue_data['Venue Category'].unique())))
print(venue_data['Venue Category'].unique())
```

```
There are 354 uniques categories.
```

```
# Assigning company zip code and Ortsteil
venue_data_ortsteil = venue_data.merge(services_cp_data[["Name", "Zipcode", "Ortsteil","€/m²","Branch","Rank"]]
                                      left_on="Name",right_on="Name")
venue_data_ortsteil.head()
```

A sample data frame for a single company is shown below.

Name	(KA) Kraft Automobile GmbH				
Company Latitude	52.479	52.479	52.479	52.479	52.479
Company Longitude	13.424	13.424	13.424	13.424	13.424
Venue	Bieberbau	EDEKA Schmidt	Süßkramdealer	Rudolph-Wilde-Park	Zig Zag Jazz Club
Venue Latitude	52.47964	52.476985	52.477275	52.482571	52.475245
Venue Longitude	13.333733	13.332797	13.330189	13.339849	13.34024
Venue Category	German Restaurant	Supermarket	Candy Store	Park	Jazz Club
Zipcode	10715	10715	10715	10715	10715
Ortsteil	Wilmersdorf	Wilmersdorf	Wilmersdorf	Wilmersdorf	Wilmersdorf
€/m²	14.73	14.73	14.73	14.73	14.73
Branch	Autohandel und Kfz-Handel (Nutzfahrzeuge				
Rank	0.685889	0.685889	0.685889	0.685889	0.685889

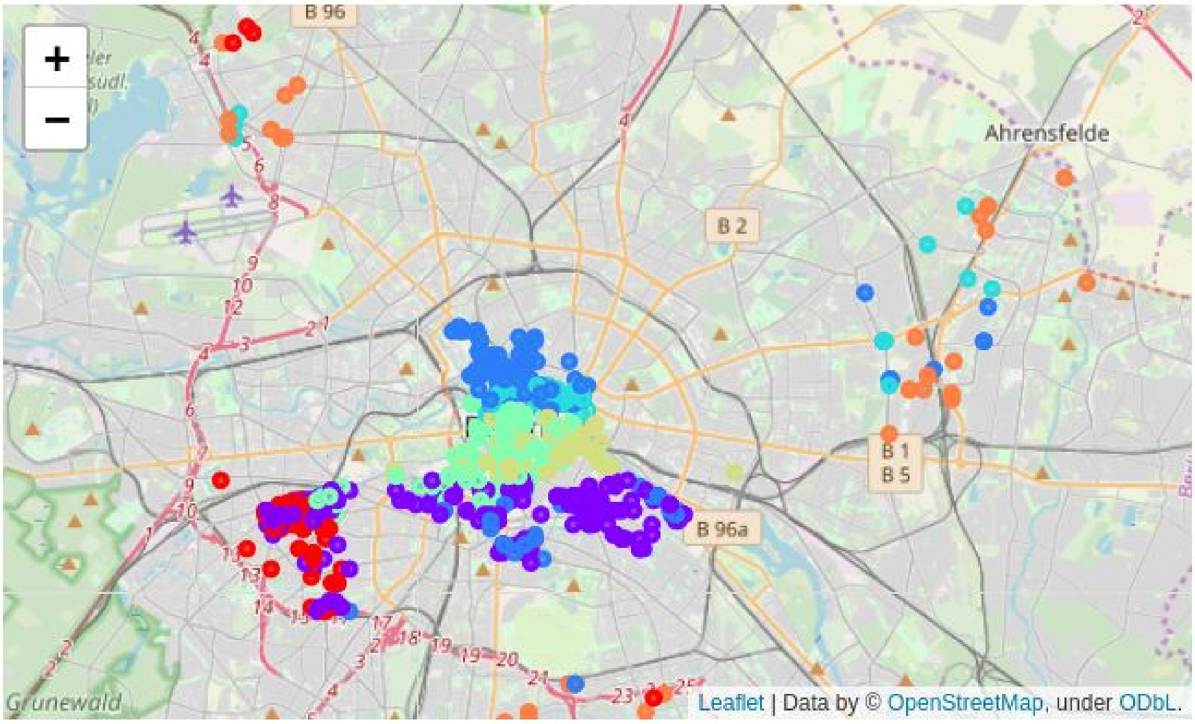
After extraction of venue data has needs to preprocess and shown as in below table. Nearest venue considered as a top most venue to each company.

Name	(KA) Kraft Automobile GmbH	1000eyes GmbH	12designer GmbH	2001 Medizin + Service GmbH	213 Gesellschaft für Besseres Wohnen mbH
1st Most Common Venue	Café	Hotel	Café	Hotel	Hotel
2nd Most Common Venue	Supermarket	Clothing Store	Bakery	German Restaurant	Coffee Shop
3rd Most Common Venue	Greek Restaurant	Café	Hotel	Italian Restaurant	Clothing Store
4th Most	Plaza	German	Bar	Movie Theater	Ice Cream Shop

Common Venue		Restaurant			
5th Most Common Venue	Italian Restaurant	Movie Theater	Nightclub	Dessert Shop	Café
6th Most Common Venue	Organic Grocery	Cocktail Bar	Ice Cream Shop	Café	Italian Restaurant
7th Most Common Venue	Food & Drink Shop	Italian Restaurant	Vietnamese Restaurant	Clothing Store	Park
8th Most Common Venue	German Restaurant	Zoo Exhibit	Coffee Shop	Bookstore	Sandwich Place
9th Most Common Venue	Bistro	Restaurant	Shoe Store	Japanese Restaurant	Art Gallery
10th Most Common Venue	Gas Station	Art Museum	Middle Eastern Restaurant	Furniture / Home Store	Vietnamese Restaurant
11th Most Common Venue	Metro Station	Burger Joint	Beer Garden	French Restaurant	Tea Room
12th Most Common Venue	Mexican Restaurant	French Restaurant	German Restaurant	Middle Eastern Restaurant	Optical Shop
13th Most Common Venue	Middle Eastern Restaurant	Furniture / Home Store	Italian Restaurant	Modern European Restaurant	Breakfast Spot
14th Most Common Venue	Garden	Gym / Fitness Center	Falafel Restaurant	Cocktail Bar	Bookstore
15th Most Common Venue	Fountain	Modern European Restaurant	Rock Club	Restaurant	German Restaurant

4.7 Clustering companies based on nearby venues

For every company, 15 venues were shortlisted. Based on the following information such as Hotel, Bar and Nightclub e.t.c were the 1st most venue to most of the companies. The following image shows the companies that are having similar venues nearby. KMeans algorithms has used to cluster the similar type of companies into a 7 clusters. Below figure show you the similarities with a colour variation.



Clustering companies based on nearby venues

5 Results and Discussion

From the data exploration, the business problem has been answered such as selecting an ideal place according to the rental prices per neighborhood. In addition, we were able to see similar venue properties per company. For this work, I have approached to get data from web resources like Wikipedia, Firmendb, python libraries like Geopy, and Foursquare API in order to set up a very realistic data-analysis scenario. We have found out that:

1. “Mitte” area has the highest number 950 companies and then followed by “Kreuzberg” with the 409 companies. In the top 3rd, 4th, 5th places were occupied by “Wilmersdorf” with 280, “Charlottenburg with 264 and “Prenzlauer Berg” with 195.
2. The average rental prices for the top 5 areas are as follows: Marzahn has the lowest at 8.06 euros/sqm and in the second place “Borsigwalde” with 9.15 euros/sqm. The rest of the places that are having lesser than 15 euros/sqm are “Britz” having 9.40 euros/sqm, Adlershof 12.51 euros/sqm, Schöneberg 14.47 euros/sqm, Wilmersdorf 14.73 euros/sqm and Halensee 14.94 euros/sqm.
3. So far, we have seen places with the highest number of companies and the lowest rental price areas, which may conclude for the best place for opening a restaurant/ an office. However, which you thought in your mind may not an ideal place. The following results will explain why it is!
4. Ranking of neighborhoods has done based on having the highest companies and to have the lowest price of the area which will benefit a lot. The top 7 rank wise places are as follows fro m highest to lowest rank: Mitte, Marzahn, Borsigwalde, Britz, Wilmersdorf, Kreuzberg, Adlershof.
5. In terms of company services, “Unternehmensberatung” occupied the highest place whereas “Architekturbüros für Hochbau und Innenarchitektur” has the lowest number of companies.
6. Bar, Nightclub, Café, Supermarket, and Italian Restaurant are the most common venues which are located near to most of the companies.

6 Conclusion

Finally, I would recommend to open a restaurant/office in the following places *Marzahn, Borsigwalde, Britz, Wilmersdorf* or in *Kreuzberg* area. These places do have a nice infrastructure for start-ups.