APPLIED DATA SCIENTE CAPSTONE PROJECT
BERLIN - Where to open a new restaurant: neighborhood analysis

Introduction and Business issue

After COVID pandemic, more problematic waves have passed and once the vaccination process is fully advanced, a well-known Spanish restaurant chain company is planning to open a new restaurant in Berlin. What will be the best place to do it?

During 2020 tourism was one the main industries affected by lockdown decisions around the world. According to UN data, during 2020 international arrivals are estimated to have dropped to 381 million, down from 1.461 billion $ in 2019 — a 74% decline. In countries whose economies are heavily reliant on tourism as the south of Europe (Italy, Portugal, Greece or Spain), the precipitous drop in visitors was, and remains, devastating.

Berlin was not left out of this huge crisis. It is the capital and the biggest city of Germany, the second most populous city in the European Union, Berlin has nearly 3,6 million residents from more than 190 countries with a population density of 4,200 people per km², the city is divided into 12 boroughs, 95 neighborhoods. Also, it is considered a top European destination – ranked third after London and Paris.

During 2020 even though the world is facing the Coronavirus crisis, Berlin welcomed almost 5 million tourists in the whole year 2020, which represents a decrease of 65% of the same period in 2019. At the beginning of 2021, between January and April 400,000 tourists have visited Berlin, and it is expected these figures could rise as vaccination process improves and frontiers are widely opened. There are more than 1.300 restaurants in Berlin.

To face this issue, we can solve this problem by creating a map and information chart that shows the real distribution of these restaurants in Berlin and clustering each area according to the density of the place. We will need to find a method to use Foursquare location data where machine learning to help us make decisions for the Spanish restaurant company.

In this project, I will try to use Foursquare location data and clustering methods to divide regions into different groups based on their restaurant location information.

Description of the data

For this project, data needed is as follows:

1. **Berlin neighborhood data:** list of Boroughs and neighborhoods and their latitudes and longitudes.

   - Data source: https://en.wikipedia.org/wiki/Boroughs_and_neighborhoods_of_Berlin
   - Description: We will discard the Berlin area (district) table through Wikipedia. Then using geocoder class of the Geopy to get coordinates (latitude and longitude) of these 12 main areas.

Localities   [edit]

As of 2012, the twelve boroughs are made up of a total of 96 officially recognized localities (*Ortsteile*). Almost all of them are further subdivided i[...]
(34.9 km² or 13.5 sq mi), the smallest one is Hansaviertel (53 ha or 130 acres). The most populated is Neukölln (154,127 inhabitants in 2009), th[...]

(01) Mitte

Locality	Area in km²	Population as of 2008	Density inhabitants per km²	Map
(0101) Mitte	10.70	79,582	7,445	
(0102) Moabit	7.72	69,425	8,993	
(0103) Hansaviertel	0.53	5,889	11,111	
(0104) Tiergarten	5.17	12,486	2,415	
(0105) Wedding	9.23	76,363	8,273	
(0106) Gesundbrunnen	6.13	82,729	13,496	

(02) Friedrichshain-Kreuzberg

Locality	Area in km²	Population as of 2008	Density inhabitants per km²	Map
(0201) Friedrichshain	9.78	114,050	11,662	
(0202) Kreuzberg	10.40	147,227	14,184	

(03) Pankow

2. Restaurants in each neighborhood in Berlin:

- Data source: Foursquare API
- Description: By using this API, we will obtain all venues in each community. We can filter these places to get only restaurants.

Methodology

In this section, I will describe the data analysis and how I used the data to yield the results.

First of all, we get information about boroughs and neighborhood of Berlin scrapped from Wikipedia: https://en.wikipedia.org/wiki/Boroughs_and_neighborhoods_of_Berlin.

For this, I used the pandas read function. I had to clean the resulting data frame in terms of unnecessary information or data that could not be handled in a data frame, such as picture data of the coat of arms of each district. The result is a nice data frame:

	neighborhood_id	neighborhood	borough	city
0	0101	Mitte	Mitte	Berlin
1	0102	Moabit	Mitte	Berlin
2	0103	Hansaviertel	Mitte	Berlin
3	0104	Tiergarten	Mitte	Berlin
4	0105	Wedding	Mitte	Berlin
...	...	...	...	...
91	1207	Waidmannslust	Reinickendorf	Berlin
92	1208	Lübars	Reinickendorf	Berlin
93	1209	Wittenau	Reinickendorf	Berlin
94	1210	Märkisches Viertel	Reinickendorf	Berlin
95	1211	Borsigwalde	Reinickendorf	Berlin

96 rows × 4 columns

Now the aim is to add coordinates details for each of 96 neighborhoods. I used the *nominatim* function to add geospatial data to the data frame, that is the latitude and the longitude seen on the right side of the following table.

	neighborhood_id	neighborhood	borough	city	neighborhood_coord	Latitude	Longitude
0	0101	Mitte	Mitte	Berlin	(39.98020495, -7.905590887431517)	39.980205	-7.905591
1	0102	Moabit	Mitte	Berlin	(52.5301017, 13.3425422)	52.530102	13.342542
2	0103	Hansaviertel	Mitte	Berlin	(52.5191234, 13.3418725)	52.519123	13.341872
3	0104	Tiergarten	Mitte	Berlin	(50.3409222, 6.956329)	50.340922	6.956329
4	0105	Wedding	Mitte	Berlin	(52.550123, 13.34197)	52.550123	13.341970
...	...	...	...	...	...	...	...
91	1207	Waidmannslust	Reinickendorf	Berlin	(52.6080354, 13.3225327)	52.608035	13.322533
92	1208	Lübars	Reinickendorf	Berlin	(52.6146467, 13.3530197)	52.614647	13.353020
93	1209	Wittenau	Reinickendorf	Berlin	(52.5912366, 13.3233195)	52.591237	13.323320
94	1210	Märkisches Viertel	Reinickendorf	Berlin	(52.5993123, 13.3565324)	52.599312	13.356532
95	1211	Borsigwalde	Reinickendorf	Berlin	(52.5840624, 13.3022741)	52.584062	13.302274

Then by using the folium package with this data frame, a map is created with all Berlin's neighborhood:
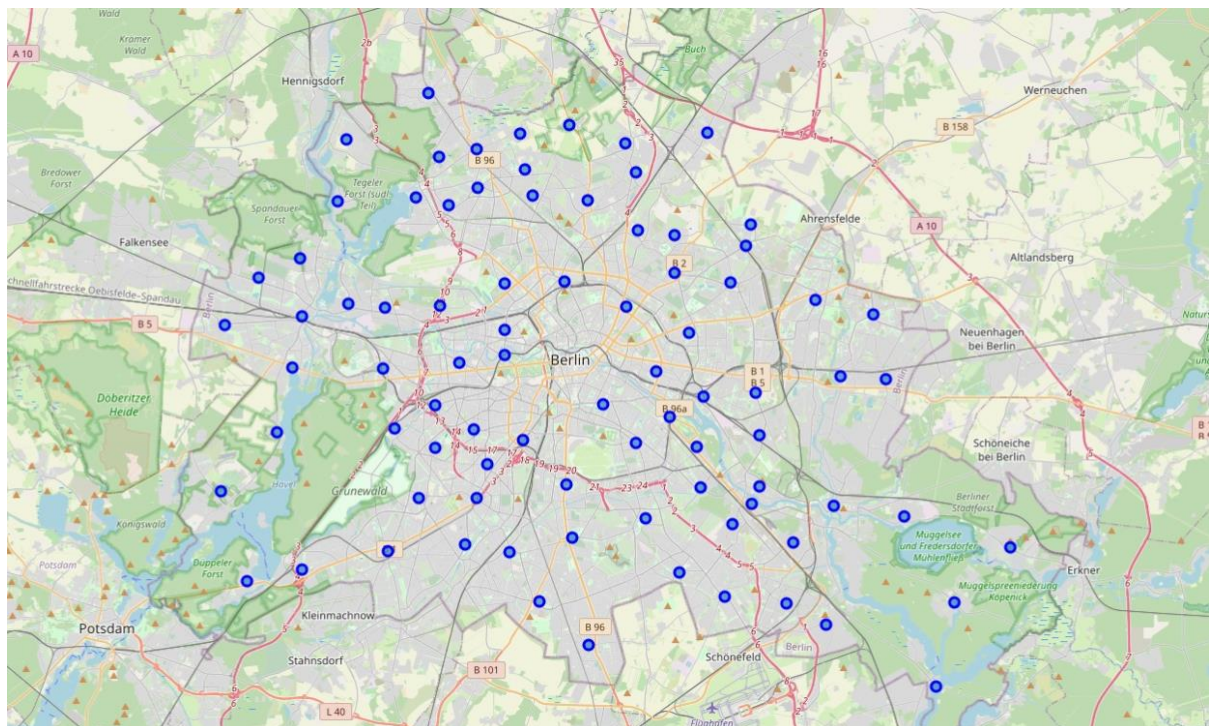
Now by using Foursqare API, the aim of this part is getting details using data exploratory analysis to extract valuable information and insights about all these 96 different neighborhoods. The aim is getting rich information which could help us to make the rights decisions.

Then, retrieved the foursquare data for all venues on foursquare with less than 1000 meters from each center of each city district, as indicated as blue dots in the map above.

The result was a list of 6247 venues all over Berlin city. Out of these 6247 venues, 1364 where restaurants. These 364 restaurants come from 56 unique restaurant categories, such as Italian, German, Greek or Vietnamese.

To find clusters of restaurant types in the different city districts, I first transformed the data frame with the restaurant venues, associated to city districts, by one-hot encoding (0/1), as seen in the picture below.

	Neighborhood	African Restaurant	American Restaurant	Argentinian Restaurant	Asian Restaurant	Austrian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Caucasian Restaurant	...	Swiss Restaurant	Syrian Restaurant	Szechuan Restaurant	Tapas Restaurant	Thai Restaurant	Turkish Home Cooking Restaurant	Turkish Restaurant	Vegetarian / Vegan Restaurant	Vietnamese Restaurant	Yemeni Restaurant
1	Moabit	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	1	0	0
2	Moabit	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
3	Moabit	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
4	Moabit	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	1	0
5	Moabit	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0

5 rows × 57 columns

Next step was grouping to show the frequency of each category of restaurants in each city district.

	Neighborhood	African Restaurant	American Restaurant	Argentinian Restaurant	Asian Restaurant	Austrian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Caucasian Restaurant	...	Swiss Restaurant	Syrian Restaurant	Szechuan Restaurant	Tapas Restaurant	Thai Restaurant	Turkish Home Cooking Restaurant	Turkish Restaurant	Vegetarian / Vegan Restaurant	Vietnamese Restaurant	Yemeni Restaurant
0	Adlershof	0.0	0.000000	0.000000	0.000000	0.000000	0.0	0.0	0.000000	0.000000	...	0.0	0.0	0.0	0.000000	0.000000	0.0	0.000000	0.000000	0.000000	0.0
1	Alt-Hohenschönhausen	0.0	0.000000	0.000000	0.000000	0.000000	0.0	0.0	0.000000	0.000000	...	0.0	0.0	0.0	0.000000	0.000000	0.0	0.000000	0.000000	0.363636	0.0
2	Alt-Treptow	0.0	0.000000	0.000000	0.000000	0.000000	0.0	0.0	0.000000	0.000000	...	0.0	0.0	0.0	0.055556	0.111111	0.0	0.000000	0.111111	0.055556	0.0
3	Altglienicke	0.0	0.000000	0.000000	0.000000	0.000000	0.0	0.0	0.000000	0.000000	...	0.0	0.0	0.0	0.000000	0.000000	0.0	0.000000	0.000000	0.000000	0.0
4	Baumschulenweg	0.0	0.000000	0.000000	0.055556	0.000000	0.0	0.0	0.000000	0.000000	...	0.0	0.0	0.0	0.000000	0.000000	0.0	0.055556	0.055556	0.111111	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
83	Wilhelmsruh	0.0	0.037037	0.037037	0.000000	0.000000	0.0	0.0	0.000000	0.000000	...	0.0	0.0	0.0	0.000000	0.037037	0.0	0.037037	0.000000	0.037037	0.0
84	Wilhelmstadt	0.0	0.000000	0.058824	0.000000	0.058824	0.0	0.0	0.000000	0.000000	...	0.0	0.0	0.0	0.000000	0.000000	0.0	0.000000	0.000000	0.117647	0.0
85	Wilmersdorf	0.0	0.000000	0.000000	0.083333	0.000000	0.0	0.0	0.041667	0.000000	...	0.0	0.0	0.0	0.000000	0.000000	0.0	0.000000	0.000000	0.125000	0.0
86	Wittenau	0.0	0.032258	0.064516	0.000000	0.000000	0.0	0.0	0.000000	0.000000	...	0.0	0.0	0.0	0.000000	0.032258	0.0	0.000000	0.000000	0.000000	0.0
87	Zehlendorf	0.0	0.000000	0.000000	0.000000	0.000000	0.0	0.0	0.000000	0.052632	...	0.0	0.0	0.0	0.000000	0.052632	0.0	0.000000	0.000000	0.000000	0.0

Finally, I used this information to create a data frame in which you can see the most common restaurant venue types for each city district:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Adlershof	German Restaurant	Greek Restaurant	Italian Restaurant	Sushi Restaurant	Middle Eastern Restaurant	Korean Restaurant	Indian Restaurant	Restaurant	Russian Restaurant	Modern European Restaurant
1	Alt-Hohenschönhausen	Vietnamese Restaurant	German Restaurant	Indian Restaurant	Italian Restaurant	Greek Restaurant	African Restaurant	Russian Restaurant	Middle Eastern Restaurant	Modern European Restaurant	New American Restaurant
2	Alt-Treptow	Falafel Restaurant	Vegetarian / Vegan Restaurant	Thai Restaurant	Italian Restaurant	Lebanese Restaurant	German Restaurant	Dumpling Restaurant	Spanish Restaurant	Russian Restaurant	Tapas Restaurant
3	Altglienicke	Greek Restaurant	Italian Restaurant	Korean Restaurant	Chinese Restaurant	Sushi Restaurant	German Restaurant	Restaurant	African Restaurant	Schnitzel Restaurant	Modern European Restaurant
4	Baumschulenweg	Falafel Restaurant	Vietnamese Restaurant	Italian Restaurant	German Restaurant	Chinese Restaurant	Indian Restaurant	Fast Food Restaurant	Sushi Restaurant	Doner Restaurant	Mexican Restaurant
5	Biesdorf	German Restaurant	African Restaurant	Lebanese Restaurant	Mexican Restaurant	Middle Eastern Restaurant	Modern European Restaurant	New American Restaurant	Persian Restaurant	Peruvian Restaurant	Portuguese Restaurant
6	Blankenburg	German Restaurant	African Restaurant	Lebanese Restaurant	Mexican Restaurant	Middle Eastern Restaurant	Modern European Restaurant	New American Restaurant	Persian Restaurant	Peruvian Restaurant	Portuguese Restaurant
7	Blankenfelde	Mexican Restaurant	Greek Restaurant	German Restaurant	Restaurant	African Restaurant	Schnitzel Restaurant	Middle Eastern Restaurant	Modern European Restaurant	New American Restaurant	Persian Restaurant
8	Bohnsdorf	Chinese Restaurant	Greek Restaurant	African Restaurant	Seafood Restaurant	Mexican Restaurant	Middle Eastern Restaurant	Modern European Restaurant	New American Restaurant	Persian Restaurant	Peruvian Restaurant
9	Bonsgwalde	German Restaurant	Restaurant	Italian Restaurant	Eastern European Restaurant	Indian Restaurant	Falafel Restaurant	Argentinian Restaurant	Thai Restaurant	Seafood Restaurant	Japanese Restaurant
10	Britz	Italian Restaurant	German Restaurant	Vietnamese Restaurant	Fast Food Restaurant	Middle Eastern Restaurant	Restaurant	Vegetarian / Vegan Restaurant	Turkish Restaurant	Turkish Home Cooking Restaurant	Falafel Restaurant
11	Charlottenburg	Italian Restaurant	German Restaurant	Korean Restaurant	Vietnamese Restaurant	Asian Restaurant	Thai Restaurant	Japanese Restaurant	Sushi Restaurant	Greek Restaurant	French Restaurant
12	Charlottenburg-Nord	Persian Restaurant	Argentinian Restaurant	Turkish Restaurant	Italian Restaurant	Restaurant	Schnitzel Restaurant	Seafood Restaurant	Indian Restaurant	Russian Restaurant	Thai Restaurant
13	Dahlem	Italian Restaurant	Doner Restaurant	German Restaurant	Sushi Restaurant	Greek Restaurant	Chinese Restaurant	Indian Restaurant	French Restaurant	Swiss Restaurant	Korean Restaurant
14	Falkenhagener Feld	Turkish Restaurant	Italian Restaurant	Greek Restaurant	Restaurant	Fast Food Restaurant	African Restaurant	Schnitzel Restaurant	Middle Eastern Restaurant	Modern European Restaurant	New American Restaurant
15	Fennpfuhl	Italian Restaurant	Falafel Restaurant	Thai Restaurant	Vegetarian / Vegan Restaurant	Middle Eastern Restaurant	Tapas Restaurant	Modern European Restaurant	Restaurant	South American Restaurant	Syrian Restaurant
16	Französisch Buchholz	Mexican Restaurant	Asian Restaurant	Greek Restaurant	Doner Restaurant	Seafood Restaurant	Middle Eastern Restaurant	Modern European Restaurant	New American Restaurant	Persian Restaurant	Peruvian Restaurant
17	Friedenau	Italian Restaurant	Korean Restaurant	Sushi Restaurant	Vietnamese Restaurant	Thai Restaurant	Chinese Restaurant	Indian Restaurant	Greek Restaurant	German Restaurant	African Restaurant
18	Friedrichsfelde	Italian Restaurant	Doner Restaurant	Vietnamese Restaurant	Vietnamese Restaurant	German Restaurant	Greek Restaurant	Restaurant	Italian Restaurant	African Restaurant	Portuguese Restaurant
19	Friedrichshagen	Seafood Restaurant	Asian Restaurant	Cajun / Creole Restaurant	Greek Restaurant	German Restaurant	African Restaurant	Middle Eastern Restaurant	Modern European Restaurant	New American Restaurant	Persian Restaurant

Now with all this information, it is possible to run a machine learning algorithm, in this case a k-mean clustering algorithm.

Results

In this case, we decide to use 5 cluster to group the results. That's the formula used:

```python
# import k-means from clustering stage
from sklearn.cluster import KMeans
```

```python
# set number of clusters (5)
kclusters = 5

Berlin_grouped_clustering = Berlin_grouped.drop('Neighborhood', 1)

# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(Berlin_grouped_clustering)

# check cluster labels generated for each row in the dataframe
kmeans.labels_[0:10]
```
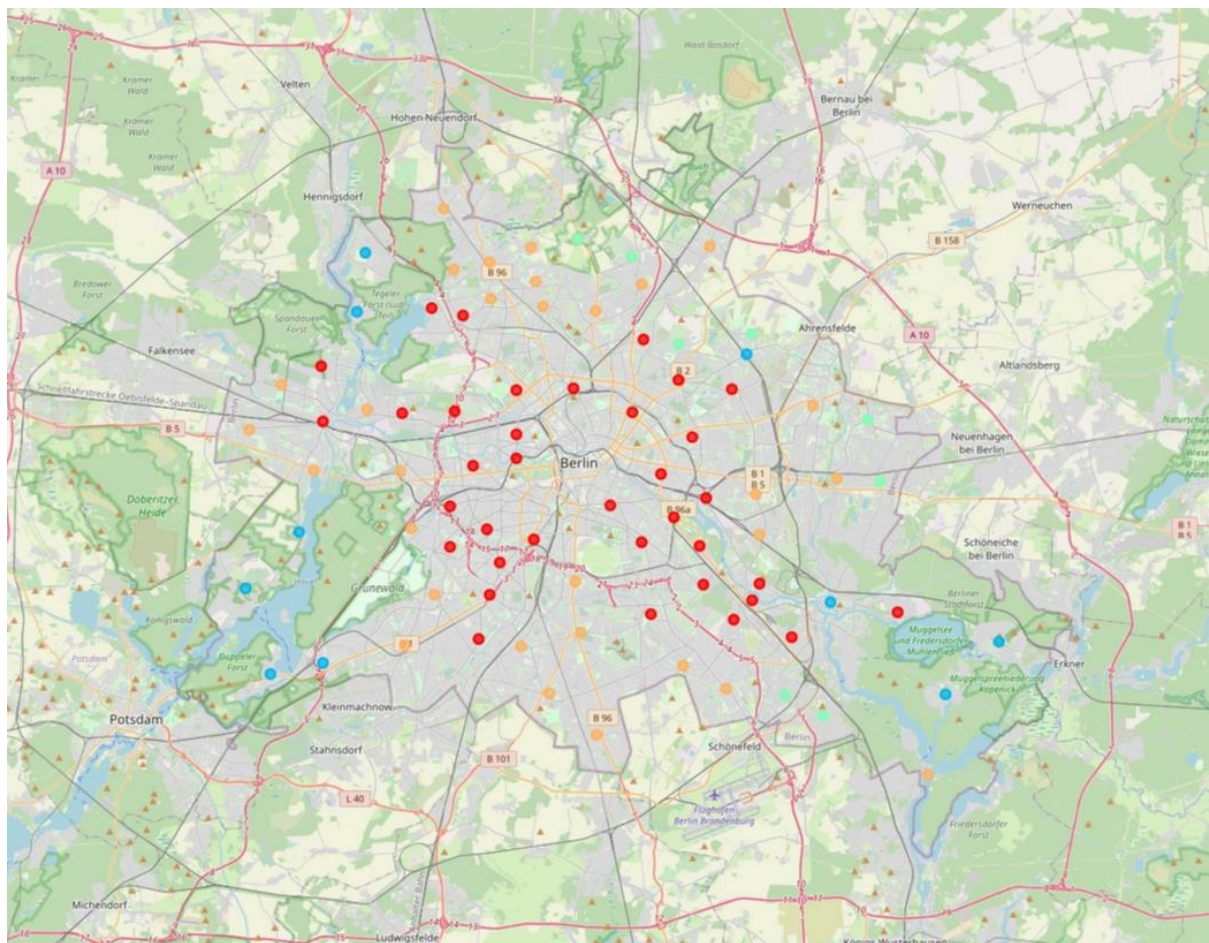
And here below the result in a data frame:

	Cluster Labels	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6
0	0	Adlershof	German Restaurant	Greek Restaurant	Italian Restaurant	Sushi Restaurant	Middle Eastern Restaurant	
1	0	Alt-Hohenschönhausen	Vietnamese Restaurant	German Restaurant	Indian Restaurant	Italian Restaurant	Greek Restaurant	
2	0	Alt-Treptow	Falafel Restaurant	Vegetarian / Vegan Restaurant	Thai Restaurant	Italian Restaurant	Lebanese Restaurant	
3	3	Altglienicke	Greek Restaurant	Italian Restaurant	Korean Restaurant	Chinese Restaurant	Sushi Restaurant	
4	0	Baumschulenweg	Falafel Restaurant	Vietnamese Restaurant	Italian Restaurant	German Restaurant	Chinese Restaurant	
5	1	Biesdorf	German Restaurant	African Restaurant	Lebanese Restaurant	Mexican Restaurant	Middle Eastern Restaurant	Mod
6	1	Blankenburg	German Restaurant	African Restaurant	Lebanese Restaurant	Mexican Restaurant	Middle Eastern Restaurant	Mod
7	3	Blankenfelde	Mexican Restaurant	Greek Restaurant	German Restaurant	Restaurant	African Restaurant	
8	3	Bohnsdorf	Chinese Restaurant	Greek Restaurant	African Restaurant	Seafood Restaurant	Mexican Restaurant	M
9	0	Borsigwalde	German Restaurant	Restaurant	Italian Restaurant	Eastern European Restaurant	Indian Restaurant	
10	0	Britz	Italian Restaurant	German Restaurant	Vietnamese Restaurant	Fast Food Restaurant	Middle Eastern Restaurant	
11	0	Charlottenburg	Italian Restaurant	German Restaurant	Korean Restaurant	Vietnamese Restaurant	Asian Restaurant	
12	0	Charlottenburg-Nord	Persian Restaurant	Argentinian Restaurant	Turkish Restaurant	Italian Restaurant	Restaurant	
13	4	Dahlem	Italian Restaurant	Doner Restaurant	German Restaurant	Sushi Restaurant	Greek Restaurant	
14	4	Falkenhagener Feld	Turkish Restaurant	Italian Restaurant	Greek Restaurant	Restaurant	Fast Food Restaurant	
15	0	Fennpfuhl	Italian Restaurant	Falafel Restaurant	Thai Restaurant	Vegetarian / Vegan Restaurant	Middle Eastern Restaurant	
16	3	Französisch Buchholz	Mexican Restaurant	Asian Restaurant	Greek Restaurant	Doner Restaurant	Seafood Restaurant	M
17	0	Friedenau	Italian Restaurant	Korean Restaurant	Sushi Restaurant	Vietnamese Restaurant	Thai Restaurant	
18	4	Friedrichsfelde	Italian Restaurant	Doner Restaurant	Argentinian Restaurant	Vietnamese Restaurant	German Restaurant	
19	0	Friedrichshagen	Seafood Restaurant	Asian Restaurant	Cajun / Creole Restaurant	Greek Restaurant	German Restaurant	

What we see in the table are the city districts and their most common venues, and they now have been assigned five different cluster labels from 0 to 4.

We can now use the cluster labels to show the city districts marked with a cluster-specific color on a map, again using folium:

As you can see there are 96 bubbles for the 96 city districts, with five different colors for the five different clusters.

Now these five clusters of restaurant type can be shown in the map of Berlin city. With regards to this clusters, they have been named according to the restaurant concentration the data shows.

- **Cluster 1 - Multicultural cluster**

	Neighborhood	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	
1	Moabit	13.342542	0	Vietnamese Restaurant	Turkish Restaurant	Chinese Restaurant	African Restaurant	
2	Hansaviertel	13.341872	0	Seafood Restaurant	Doner Restaurant	Kebab Restaurant	Italian Restaurant	
4	Wedding	13.341970	0	Italian Restaurant	Turkish Restaurant	Restaurant	Chinese Restaurant	
5	Gesundbrunnen	13.384846	0	Seafood Restaurant	Vietnamese Restaurant	Sushi Restaurant	Modern European Restaurant	
6	Friedrichshain	13.450290	0	Italian Restaurant	Falafel Restaurant	Vegetarian / Vegan Restaurant	Thai Restaurant	
7	Kreuzberg	13.411914	0	Turkish Restaurant	Italian Restaurant	Vietnamese Restaurant	Spanish Restaurant	
8	Prenzlauer Berg	13.428565	0	Vietnamese Restaurant	Falafel Restaurant	Japanese Restaurant	Indian Restaurant	
9	Weißensee	13.463002	0	German Restaurant	Indian Restaurant	Greek Restaurant	Restaurant	
11	Heinersdorf	13.437015	0	Vietnamese Restaurant	Italian Restaurant	Falafel Restaurant	German Restaurant	
21	Charlottenburg	13.309683	0	Italian Restaurant	German Restaurant	Korean Restaurant	Vietnamese Restaurant	

- **Cluster 2 - East Berlin German restaurant**

	Neighborhood	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	
10	Blankenburg	10.955199	1	German Restaurant	African Restaurant	Lebanese Restaurant	Mexican Restaurant	M
71	Biesdorf	6.305603	1	German Restaurant	African Restaurant	Lebanese Restaurant	Mexican Restaurant	M

- **Cluster 3 - German cluster**

	Neighborhood	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	
3	Tiergarten	6.956329	2	German Restaurant	Eastern European Restaurant	Restaurant	Seafood Restaurant	
32	Gatow	13.180134	2	German Restaurant	Seafood Restaurant	Italian Restaurant	Schnitzel Restaurant	
33	Kladow	13.140052	2	German Restaurant	Italian Restaurant	Greek Restaurant	African Restaurant	
42	Nikolassee	13.198145	2	German Restaurant	Italian Restaurant	Restaurant	Asian Restaurant	
43	Wannsee	13.158937	2	German Restaurant	Italian Restaurant	Asian Restaurant	Austrian Restaurant	
64	Köpenick	13.576413	2	German Restaurant	Italian Restaurant	Greek Restaurant	Middle Eastern Restaurant	
66	Rahnsdorf	13.701802	2	German Restaurant	Seafood Restaurant	Fast Food Restaurant	Schnitzel Restaurant	
68	Müggelheim	13.661954	2	American Restaurant	German Restaurant	Seafood Restaurant	Vegetarian / Vegan Restaurant	
81	Neu-Hohenschönhausen	13.514065	2	German Restaurant	Indian Restaurant	Greek Restaurant	African Restaurant	
87	Konradshöhe	13.223198	2	Turkish Restaurant	German Restaurant	Restaurant	African Restaurant	
88	Heiligensee	13.229579	2	German Restaurant	Italian Restaurant	Restaurant	African Restaurant	

- **Cluster 4 – Greek cluster**

	Neighborhood	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
13	Stadtrandsiedlung Malchow	13.463285	3	Greek Restaurant	Restaurant	Fast Food Restaurant
15	Blankenfelde	13.388447	3	Mexican Restaurant	Greek Restaurant	German Restaurant
17	Französisch Buchholz	13.428110	3	Mexican Restaurant	Asian Restaurant	Greek Restaurant
60	Altglienicke	13.542646	3	Greek Restaurant	Italian Restaurant	Korean Restaurant
62	Bohnsdorf	13.570665	3	Chinese Restaurant	Greek Restaurant	African Restaurant
73	Mahlsdorf	13.613162	3	Greek Restaurant	Fast Food Restaurant	Italian Restaurant
74	Hellersdorf	13.604774	3	Greek Restaurant	Fast Food Restaurant	Falafel Restaurant
80	Wartenberg	13.517582	3	Greek Restaurant	Fast Food Restaurant	African Restaurant

- **Cluster 5 – Italian cluster**

	Neighborhood	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
12	Karow	13.486276	4	Italian Restaurant	Fast Food Restaurant	African Restaurant	Schnitzel Restaurant
14	Pankow	13.435316	4	Italian Restaurant	Mexican Restaurant	Asian Restaurant	Thai Restaurant
18	Niederschönhausen	13.401397	4	Mexican Restaurant	Italian Restaurant	Asian Restaurant	Modern European Restaurant
20	Wilhelmsruh	13.362206	4	Italian Restaurant	German Restaurant	Greek Restaurant	Restaurant
24	Grunewald	13.263754	4	Italian Restaurant	German Restaurant	Chinese Restaurant	Vietnamese Restaurant
25	Westend	13.255842	4	Italian Restaurant	German Restaurant	Vietnamese Restaurant	Doner Restaurant
29	Haselhorst	13.231156	4	German Restaurant	Italian Restaurant	Fast Food Restaurant	Argentinian Restaurant
31	Staaken	13.143367	4	Fast Food Restaurant	Turkish Restaurant	Italian Restaurant	Chinese Restaurant
35	Falkenhagener Feld	13.166894	4	Turkish Restaurant	Italian Restaurant	Greek Restaurant	Restaurant
36	Wilhelmstadt	13.191452	4	Italian Restaurant	German Restaurant	Restaurant	Vietnamese Restaurant
39	Lankwitz	13.345486	4	Chinese Restaurant	Italian Restaurant	Fast Food Restaurant	German Restaurant
40	Zehlendorf	13.258930	4	Italian Restaurant	German Restaurant	Greek Restaurant	Restaurant
41	Dahlem	13.281098	4	Italian Restaurant	Doner Restaurant	German Restaurant	Sushi Restaurant
46	Tempelhof	13.386448	4	Italian Restaurant	Restaurant	Dumpling Restaurant	Korean Restaurant
47	Mariendorf	13.390028	4	Italian Restaurant	Restaurant	Doner Restaurant	Eastern European Restaurant
48	Marienfelde	13.366592	4	Italian Restaurant	Fast Food Restaurant	Chinese Restaurant	Mexican Restaurant
49	Lichtenrade	13.402040	4	Italian Restaurant	Fast Food Restaurant	Greek Restaurant	Asian Restaurant

As you can see it is possible to define clusters of certain cuisines in Berlin.

Discussion

As it is shown in this report and notebook, these results have been achieved using a piece of all knowledges getting during all previous modules: typical ways of scraping, cleaning, handling, transforming and visualizing data, all the tools are simply there.

Even more of the tasks are very similar to ones previously done, I just should get to know the available open source packages and learn how to use them. Also, internet have a lot of sources of information to solve problems arise and move forward in the project.

Conclusion

Based on all results given by all the analysis done, I have achieved the goal presented at the outset of this blogpost: Spanish restaurant chain company can check where city district is the best one to locate its new restaurant based on the kind of cuisine they want to offer to clients.