



The Restaurant Battle of Neighborhoods in Cologne



Business Statement

- Cologne, the city the author lives in, attracts a large number of tourists, not least due to its famous cathedral, the trade fairs and conventions, such as the gamescom, and its vibrant party scene. For tourists, finding the right place to eat can be a challenge, though. German dishes include a lot of meat, often pork, which many people do not want to eat for health-related, religious, cultural or moral reasons. This is just one motive for giving tourists a good overview about what to eat where.

Data Description

- I will, as requested by the assignment task, use foursquare data about restaurants in Cologne. Foursquare is a US tech company from New York focusing on location data. Their technology and data powers apps such as Apple's Maps, Uber, Twitter and many other household names.
- Here is an example of a vegetarian restaurant in Cologne on foursquare: <https://de.foursquare.com/v/sattgr%C3%BCn/5c33306cc824ae002c2b414c>. I will use foursquare data such as the restaurant name, ID, location and category of food (vegetarian, Italian etc.).
- Also, I will use the overview of districts/city parts of Cologne from Wikipedia: https://en.wikipedia.org/wiki/Districts_of_Cologne

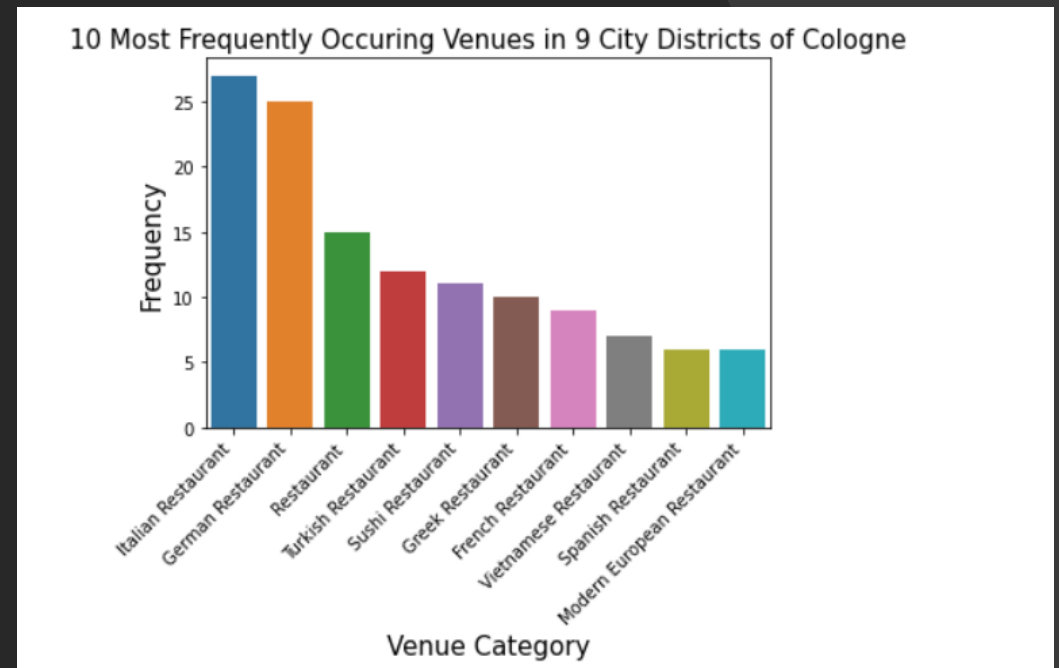


Methodology



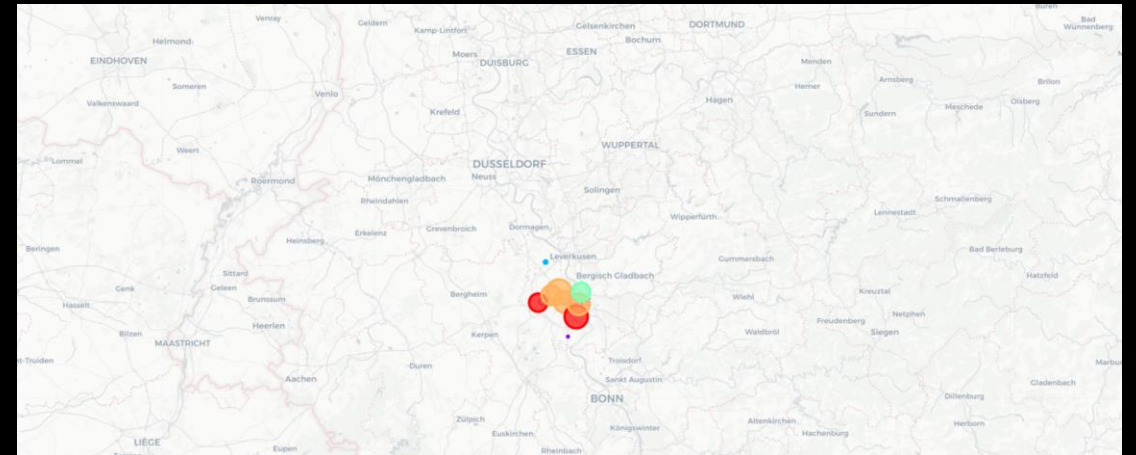
- I scraped data from Wikipedia to create a dataframe with the city districts of Cologne: https://en.wikipedia.org/wiki/Districts_of_Cologne.
- For this, I used the pandas read function. I had to clean the resulting data frame in terms of unnecessary information or data that could not be handled in a data frame, such as picture data of the coat of arms of each district. The result is a nice data frame:

- I plotted a bar chart with the frequency of the 10 most frequently occurring restaurants in the whole city, using seaborn/matplotlib packages. We can see that Italian, German and Turkish restaurants are the most frequently occurring restaurants in Cologne, which seems pretty reasonable, as Germany has a relatively high proportion of people with Italian and Turkish roots, and these cuisines being excellent and highly appreciated by large parts of the population - think about pizza, pasta or kebab!



Results

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



Cluster 1

]: Cologne_merged.loc[Cologne_merged['Cluster Labels'] == 0, Cologne_merged.columns[[1] + list(range(5, Cologne_merged.shape[1]))]]

]:

	City parts	District Councils	Latitude	Longitude	Cluster Labels	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
2	Braunsfeld, Junkersdorf, Klettenberg, Lindenth...	Bezirksamt Lindenthal Aachener Straße 220, 509...	50.935935	6.871246	0	German Restaurant	Sushi Restaurant	Restaurant	Italian Restaurant	Greek Restaurant	American Restaurant	Indian Restaurant	Mexican Restaurant
6	Eil, Elsdorf, Ensen, Finkenber, Gremberghoven...	Bezirksamt PorzFriedrich-Ebert-Ufer 64–70, D-5...	50.906705	6.999129	0	German Restaurant	Italian Restaurant	Restaurant	Greek Restaurant	French Restaurant	Seafood Restaurant	Mexlcan Restaurant	Fast Food Restaurant

This is German Restaurant cluster

Cluster 2

: Cologne_merged.loc[Cologne_merged['Cluster Labels'] == 1, Cologne_merged.columns[[1] + list(range(5, Cologne_merged.shape[1]))]]

:

	City parts	District Councils	Latitude	Longitude	Cluster Labels	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 Common Venue
1	Bayenthal, Godorf, Hahnwald, Immendorf, Marien...	Bezirksamt Rodenkirchen Hauptstraße 85, D-5099...	50.865622	6.969718	1	Scandinavian Restaurant	German Restaurant	Kebab Restaurant	Italian Restaurant	Israeli Restaurant	Indian Restaurant	Hungarian Restaurant	Greek Restaurant	Fre Res

This is Scandinavian Restaurant cluster



Cluster 4

```
[: Cologne_merged.loc[Cologne_merged['Cluster Labels'] == 3, Cologne_merged.columns[[1] + list(range(5, Cologne_merged.shape[1]))]]
```

	City parts	District Councils	Latitude	Longitude	Cluster Labels	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
8	Buchforst, Buchheim, Delbrück, Dünwald, Flitt...	Bezirksamt Mülheim Wiener Platz 2a, D-51065 Köln	50.958147	7.013526	3	Turkish Restaurant	Asian Restaurant	Mediterranean Restaurant	Italian Restaurant	Vietnamese Restaurant	Greek Restaurant	Comfort Food Restaurant	Vegetarian / Vegan Restaurant	Seafood Restaurant

This is Turkish Restaurant cluster

Cluster 3

```
: Cologne_merged.loc[Cologne_merged['Cluster Labels'] == 2, Cologne_merged.columns[[1] + list(range(5, Cologne_merged.shape[1]))]]
```

	City parts	District Councils	Latitude	Longitude	Cluster Labels	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
5	Blumenberg, Chorweiler, Esch/Auweiler, Fühling...	Bezirksamt Chorweiler Pariser Platz 1, D-50765...	51.021167	6.898034	2	Fast Food Restaurant	Italian Restaurant	Sushi Restaurant	French Restaurant	Israeli Restaurant	Indian Restaurant	Hungarian Restaurant	Greek Restaurant	German Restaurant

This is fast food restaurant cluster


```
_merged.loc[Cologne_merged['Cluster Labels'] == 4, Cologne_merged.columns[[1] + list(range(5, Cologne_merged.shape[1]))]]
```

Parts	District Councils	Latitude	Longitude	Cluster Labels	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
dt-Nord, Altstadt-Deutz, Neustadt-	Bezirksksamt Innenstadt Brückenstraße 19, D-50...	50.937328	6.959234	4	Italian Restaurant	Vietnamese Restaurant	Modern European Restaurant	German Restaurant	Sushi Restaurant	French Restaurant	Israeli Restaurant	Ja...
ndorf, emünd/Mengenich, feld, N...	Bezirksamt Ehrenfeld Venloer Straße 419 – 421,...	50.951502	6.916529	4	Restaurant	Italian Restaurant	German Restaurant	Japanese Restaurant	Sushi Restaurant	Spanish Restaurant	Falafel Restaurant	Ke...
stöckchen, rrich, Mauenheim, ...	Bezirksamt NippesNeusser Straße 450,D-50733 Köln	50.958994	6.941777	4	Italian Restaurant	German Restaurant	Greek Restaurant	French Restaurant	Sushi Restaurant	Spanish Restaurant	Restaurant	Ke...
, Höhenberg, oldt/Gremberg, Mer...	Bezirksamt KalkKalker Hauptstraße 247–273,D-51...	50.931923	7.005806	4	Restaurant	Italian Restaurant	Greek Restaurant	Vietnamese Restaurant	Middle Eastern Restaurant	Asian Restaurant	Comfort Food Restaurant	De...

ilian Restaurant cluster

Discussion

- If I reflect the work necessary to create these results, what comes to my mind is that for typical ways of scraping, cleaning, handling, transforming and visualizing data, all the tools are simply there. We just have to get to know the available open source packages and learn how to use them.

Conclusion

- We have achieved the results. This is just one example of fantastic data science uses cases one can realize applying technology which is available for free today! What a time to be alive.

References

- https://en.wikipedia.org/wiki/Districts_of_Cologne
- <https://de.foursquare.com/v/sattgr%C3%BCn/5c33306cc824ae002c2b414c>