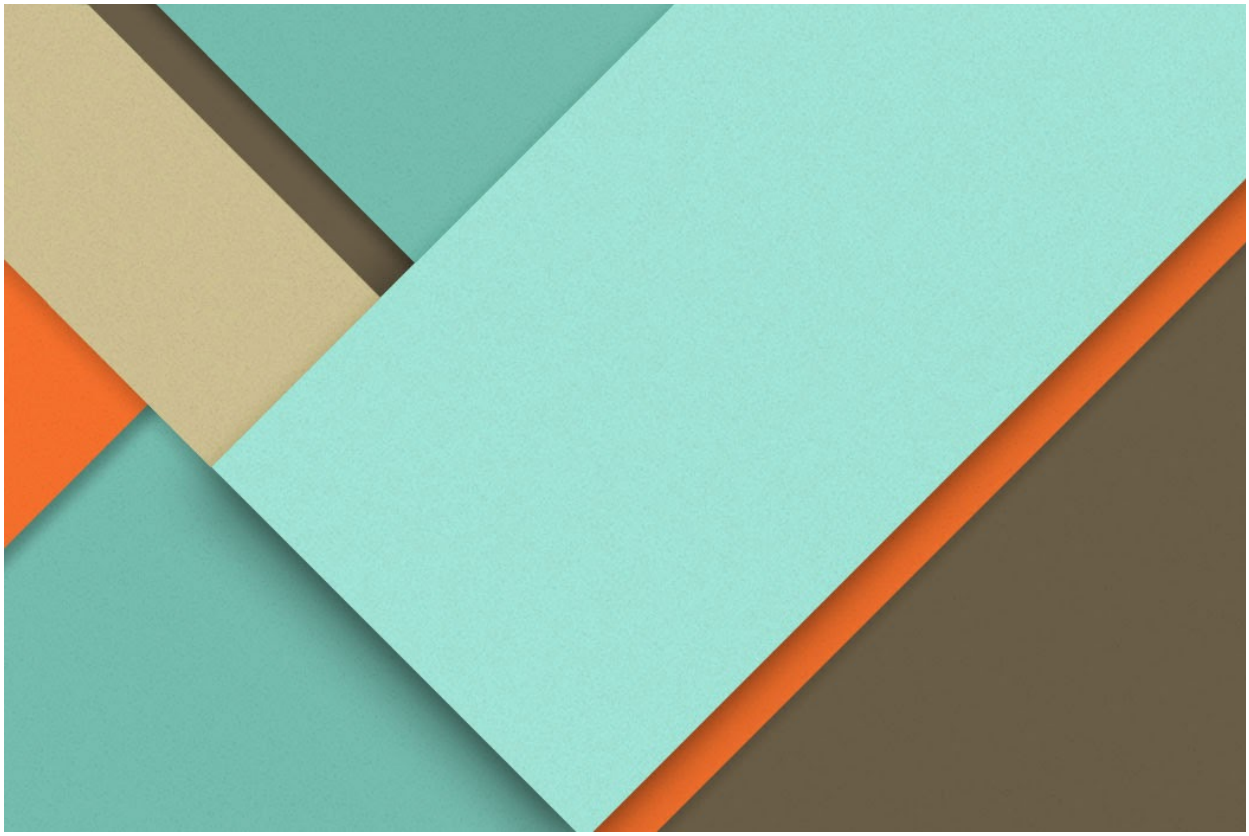


Germany country analysis

12.May.2020

—

Manuel Alemán

1. Introduction

The main focus of this job is to analyse the country of Germany in search of the best places to open a new restaurant, I decide to analyse Germany, because according to the OECD (Organization for Economic Co-operation and Development) Germany) Germany has one of the highest Work-Life Balance qualifications (8.4) compared to Mexico my home country that has the (1.1).

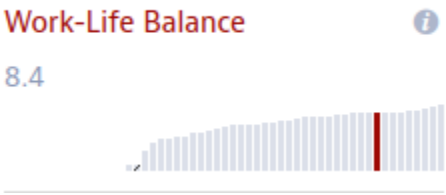


Image 1.1 Work-life balance Germany

Therefore the analysis in search of the best suitable place in Germany to open a new restaurant of typical food.

2. Data

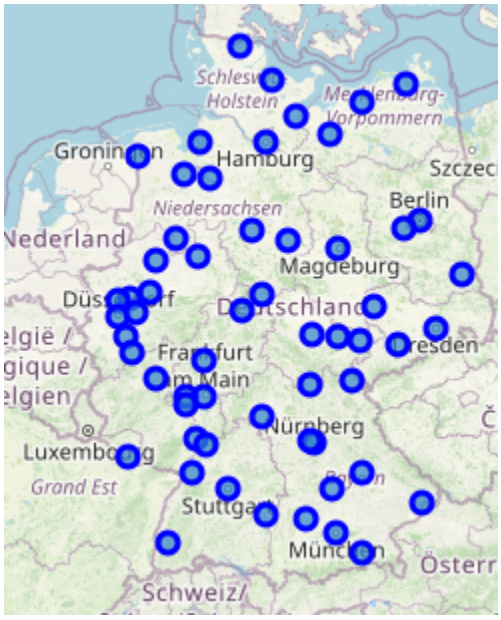
The data used was acquired from **http://www.simplemaps.com** consisting in the different cities in Germany which include the latitude and longitude of each city, making possible the interaction with the Foursquare API to obtain the different venues that each city has

2.1 Data cleaning

First I start with showing the top cities according to their population, after that it was necessary to delete columns that were not useful for this analysis, from the dataset downloaded from simplemaps we will only need the city name and the coordinates of the city.

	City	Latitude	Longitude
0	Berlin	52,5218	13,4015
1	Stuttgart	48,78	9,2
2	Frankfurt	50,1	8,675
3	Mannheim	49,5004	8,47
4	Hamburg	53,55	10

Once we have the new df with the coordinates it's time to print them in the map, the folium library was used for this.



3. Data Analysis

Now that all the cities are represented in the map, it's time to start the analysis using Foursquare and the coordinates in the dataset. I'll proceed to obtain the venues surrounding the city.

Once all the venues are obtained the one hot encoding is used to extract all the dummy variables that are necessary for the k-means clustering.

American Restaurant	Art Gallery	Art Museum	Asian Restaurant	Athletics & Sports	Re
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

Now that we have all the venues we can obtain the mean of all the venues to extract the mean of each venue in a certain city.

```
----Wuppertal----
                venue  freq
0   German Restaurant  0.25
1                Pool  0.25
2               Café  0.25
3    Sculpture Garden  0.25
4  Persian Restaurant  0.00


----Würzburg----
                  venue  freq
0           Supermarket  0.21
1  Fast Food Restaurant  0.14
2      Department Store  0.07
3               Bakery  0.07
4                 Park  0.07
```

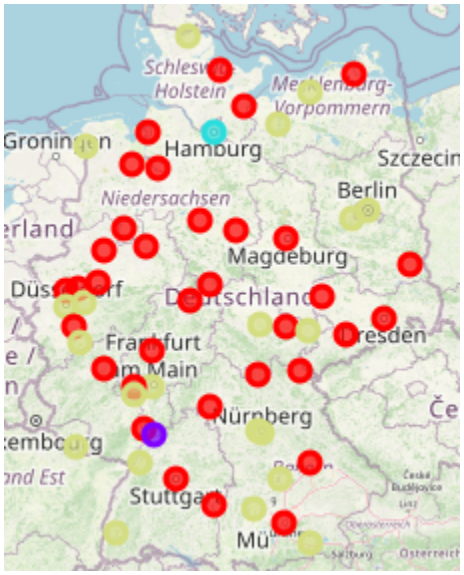
As we can see in the example the most common venue in Wuppertal are German Restaurants and in Würzburg are supermarkets, that means that probably those cities will be in different clusters.

I proceed to obtain the frequency of the data so I can make a table to make it more user friendly, showing the top 10 most commun venues in that city.

	Cluster Labels	City	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Mo Comm Ven
0	3	Augsburg	Steakhouse	Museum	Supermarket	Hotel	Bus Station
1	0	Berlin	Art Gallery	Hotel	Clothing Store	Ice Cream Shop	Coffee Shop
2	3	Bielefeld	Supermarket	Nightclub	Bar	Hotel	Asian Restaura

With this table I will proceed to cluster the information.

Now it's time to plot the data in the map to see the results and the different clusters created.



4. Results

This were the results obtained from the k-means clustering

I. Cluster 0

	City	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
11	Saarbrücken	Supermarket	Light Rail Station	Chinese Restaurant	Bakery	Shopping Mall	Zoo	Flea Market	Garden Center	Garden	Furniture / Home Store
18	Dortmund	Supermarket	Indie Theater	Mediterranean Restaurant	Flea Market	Bar	Gas Station	Garden Center	Garden	Furniture / Home Store	Frozen Yogurt Shop
40	Würzburg	Supermarket	Fast Food Restaurant	Bakery	Diner	Hotel	Gym / Fitness Center	Big Box Store	Park	Nightclub	Asian Restaurant
41	Oldenburg	Supermarket	Bavarian Restaurant	Doner Restaurant	Greek Restaurant	Bakery	Flea Market	Gas Station	Garden Center	Garden	Furniture / Home Store
46	Cottbus	Shoe Store	Supermarket	Plaza	Fast Food Restaurant	Garden Center	Garden	Furniture / Home Store	Frozen Yogurt Shop	French Restaurant	Franconian Restaurant
53	Stralsund	Gas Station	Supermarket	Harbor / Marina	Fast Food Restaurant	Zoo	Flea Market	Garden Center	Garden	Furniture / Home Store	Frozen Yogurt Shop
54	Coburg	Supermarket	Hotel	Multiplex	Thai Restaurant	Financial or Legal Service	Gas Station	Garden Center	Garden	Furniture / Home Store	Frozen Yogurt Shop

Here we have the supermarkets

II. Cluster 1

	City	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	Berlin	Art Gallery	Hotel	Clothing Store	Ice Cream Shop	Café	Coffee Shop	History Museum	Art Museum	Park	Indie Movie Theater
1	Stuttgart	Indian Restaurant	Bus Stop	Café	Mini Golf	Scenic Lookout	Zoo	Financial or Legal Service	Garden	Furniture / Home Store	Frozen Yogurt Shop
2	Frankfurt	Café	Bar	Bakery	Italian Restaurant	Drugstore	Park	Art Museum	Restaurant	Supermarket	Greek Restaurant
3	Mannheim	Café	Food & Drink Shop	Greek Restaurant	Plaza	Grocery Store	Music Venue	Light Rail Station	Turkish Restaurant	Trattoria/Osteria	Mediterranean Restaurant
4	Hamburg	Coffee Shop	Hotel	Vietnamese Restaurant	Cosmetics Shop	Burger Joint	Café	Clothing Store	German Restaurant	Restaurant	Cocktail Bar
5	Essen	Hotel	Bakery	Cocktail Bar	Drugstore	Italian Restaurant	Clothing Store	Coffee Shop	Seafood Restaurant	Cosmetics Shop	Dim Sum Restaurant
6	Duisburg	Italian Restaurant	Turkish Restaurant	Theater	Bar	Doner Restaurant	Financial or Legal Service	Gas Station	Garden Center	Garden	Furniture / Home Store

Tourist zone (hotels, restaurants).

III. Cluster 2

	City	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
10	Wuppertal	Pool	German Restaurant	Café	Sculpture Garden	Zoo	Financial or Legal Service	Garden	Furniture / Home Store	Frozen Yogurt Shop	French Restaurant
44	Göttingen	German Restaurant	Lake	Financial or Legal Service	Gastropub	Gas Station	Garden Center	Garden	Furniture / Home Store	Frozen Yogurt Shop	French Restaurant

As predicted Wüzburg and Wuppertal are in different clusters

IV. Cluster 3

	City	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
29	Freiburg im Breisgau	Hotel	Zoo	Fast Food Restaurant	Gas Station	Garden Center	Garden	Furniture / Home Store	Frozen Yogurt Shop	French Restaurant	Franconian Restaurant
56	Emden	Hotel	Steakhouse	Hostel	Fast Food Restaurant	Gas Station	Garden Center	Garden	Furniture / Home Store	Frozen Yogurt Shop	French Restaurant

5. Discussion

I used the K-means algorithm as part of this clustering study. For more detailed and accurate guidance, the data set can be expanded and the details of the neighborhood or street can also be drilled.

I also using the coordinates and with the help of Foursquare could obtain the venues from the different cities making possible the cluster analysis, as said before this analysis could be improve with a larger Dataset

6. Conclusion

In conclusion, people are looking forward to improving their quality of life, and looking to start a new business in cities with high levels of development, so data like this one can be a very effective tool to achieve a great outcome regarding where and what kind of business is the perfect one for a certain zone.

In this analysis we discover that the best zones to start a new business are in cluster 1, therefore, the recommendation It's to start the business there.

Using more tools like this could improve us as a society because this type of analysis can not only be done with commercial things, but can also be applied in many aspects of our society and, therefore, helps us to make better decisions.