

Battle of Cities in Germany - Which city in Germany we should choose to open our next bubble tea store

Table of contents

1. [Introduction and Business Problem](#)
2. [Data Section](#)
3. [Methodology](#)
4. [Results](#)
5. [Discussion Section](#)
6. [Conclusion](#)
7. [Appendix](#)

Introduction / Business Problem

The so-called "bubble tea" is gaining increasing popularity outside each asia in recent years. Originally the "bubble" in the name refers to the bubble resulted from blending black tea with ice and fruit syrups, but somehow it is mostly mistakenly understood as the tapioca or bobas as the topping in the drink.

In Düsseldorf, Germany we have been running two commercially successful bubble tea stores for more than ten years. Düsseldorf is one of the biggest city of Germany and the japanese center amongst german cities and is hosting a large varieties of japanese companies, intitutions and a big japanese population. In the recent decade Chinese and Koreans have been inhabiting this german city with great atmosphere of asian culture. That explains partially the reason why our bubble tea stores are popular. Meanwhile it should be pointed out that the majority of customers are still german, mostly young generation of german and children.

Now we have decided to expand our brand and to open more stores in Germany. All cities in Germany are under consideration. The core problem is to find a list of potential cities to open our store, where the business is more likely to succeed as we have experienced in Düsseldorf. We want to use data science to get a list to help our site search.

Data Section

To utilize the foursquare api on venues we need the information of longitude and latitude of major german cities. We adopt the following list of major german cities which encompasses all candidates in our understanding.
<https://www.latlong.net/category/cities-83-15.html> (<https://www.latlong.net/category/cities-83-15.html>)

	Place Name	Latitude	Longitude
0	Salzgitter, Lower Saxony, Germany	52.150002	10.333333
1	Hildesheim, Lower Saxony, Germany	52.150002	9.950000
2	Siegen, North Rhine-Westphalia, Germany	50.883331	8.016667
3	Erlangen, Bavaria, Germany	49.583332	11.016667
4	Remscheid, North Rhine-Westphalia, Germany	51.183334	7.200000
...
89	Kalbach, Germany	50.437061	9.639988
90	Baden-Baden, Baden-Württemberg, Germany	48.758339	8.243008
91	Heidelberg Old City, Baden-Württemberg, Germany	49.398750	8.672434
92	Lindau, Bavaria, Germany	47.557400	9.707209
93	Koblenz, Rhineland-Palatinate, Germany	50.360023	7.589907

By the information of longitude and latitude we could leverage Foursquare API (<https://developer.foursquare.com/developer/> (<https://developer.foursquare.com/developer/>)) to get the venue information including types, visits, ratings etc. We regard the core area for our site selecting for a city to be an area of radius of 10 kms of the city center.

A full description of the data we could get can be read in <https://developer.foursquare.com/docs/places-database/details/> (<https://developer.foursquare.com/docs/places-database/details/>).

Methodology

We scrape the data from the following list. <https://www.latlong.net/category/cities-83-15.html> (<https://www.latlong.net/category/cities-83-15.html>)

Place Name	Latitude	Longitude
Salzgitter, Lower Saxony, Germany	52.150002	10.333333
Hildesheim, Lower Saxony, Germany	52.150002	9.950000
Siegen, North Rhine-Westphalia, Germany	50.883331	8.016667
Erlangen, Bavaria, Germany	49.583332	11.016667
Remscheid, North Rhine-Westphalia, Germany	51.183334	7.200000
Jena, Thuringia, Germany	50.927223	11.586111
Bergisch Gladbach, North Rhine-Westphalia, Germany	51.099998	7.116667
Reutlingen, Baden-Württemberg, Germany	48.483334	9.216667
Bremerhaven, Bremen, Germany	53.549999	8.583333
Recklinghausen, North Rhine-Westphalia, Germany	51.584999	7.161944

we get the following data frame

	Place Name	Latitude	Longitude
0	Salzgitter, Lower Saxony, Germany	52.150002	10.333333
1	Hildesheim, Lower Saxony, Germany	52.150002	9.950000
2	Siegen, North Rhine-Westphalia, Germany	50.883331	8.016667
3	Erlangen, Bavaria, Germany	49.583332	11.016667
4	Remscheid, North Rhine-Westphalia, Germany	51.183334	7.200000
...
89	Kalbach, Germany	50.437061	9.639988
90	Baden-Baden, Baden-Württemberg, Germany	48.758339	8.243008
91	Heidelberg Old City, Baden-Württemberg, Germany	49.398750	8.672434
92	Lindau, Bavaria, Germany	47.557400	9.707209
93	Koblenz, Rhineland-Palatinate, Germany	50.360023	7.589907

94 rows × 3 columns

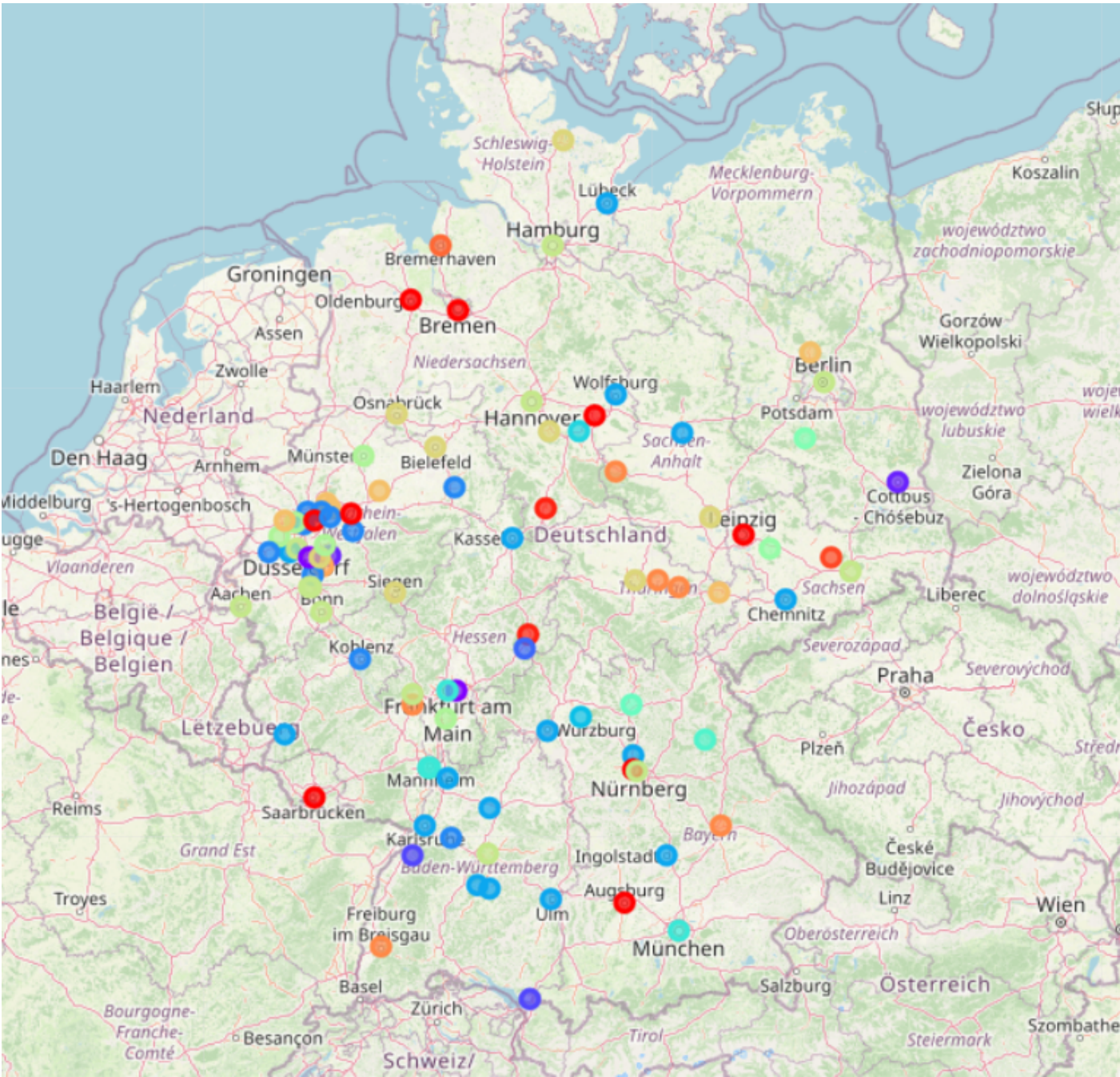
We use **FourSquare** to get the venue information. In particular we regard an area of radius 5 km as the core city part of our investigation, which would also be the range where we select our site.

The venue information will encompass the top venue. For instance the following is the top 30 venues in Düsseldorf. As one can see, many asian elements are involved.

```
array(['Café', 'Park', 'Coffee Shop', 'Japanese Restaurant',  
      'Supermarket', 'Beach', 'Scenic Lookout', 'Plaza', 'Restaurant',  
      'Pizza Place', 'Sushi Restaurant', 'Pedestrian Plaza', 'Trail',  
      'Ramen Restaurant', 'Italian Restaurant', 'Hotel', 'Grocery Store',  
      'Greek Restaurant', 'German Restaurant', 'Farmers Market',  
      'Ice Cream Shop', 'Asian Restaurant', 'Bar',  
      'Vietnamese Restaurant', 'Event Space', 'Juice Bar', 'Art Museum',  
      'Arts & Crafts Store', 'Pool', 'Bakery'], dtype=object)
```

So most popular venues reflect the interests, flavors, tendency. We use k-mean clustering to automatically categorize the major towns in Germany. To get a short list, k is chosen to be 23 here. We get a clustering of German major cities.

We use folium to visualize the cities in a map.



In particular the group which includes Düsseldorf is

	Place Name	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	...	91th Most Common Venue	92th Most Common Venue	93th Most Common Venue	94th Most Common Venue	95th Most Common Venue	96th Most Common Venue	97th Most Common Venue	98th Most Common Venue	99th Most Common Venue	100th Most Common Venue
34	Aachen, North Rhine-Westphalia, Germany	50.775555	6.083611	15	Supermarket	German Restaurant	Park	Café	Bar	Italian Restaurant	...	Gym Pool	Halal Restaurant	Harbor / Marina	Fried Chicken Joint	Franconian Restaurant	Fountain	Financial or Legal Service	Event Space	Exhibit	Factory
41	Bonn, North Rhine-Westphalia, Germany	50.733334	7.100000	15	Italian Restaurant	Ice Cream Shop	Supermarket	Plaza	Greek Restaurant	Café	...	Drugstore	Dumpling Restaurant	Fish Market	Dutch Restaurant	Financial or Legal Service	Food Court	Flea Market	Flower Shop	Gym Pool	Afghan Restaurant
45	Düsseldorf, North Rhine-Westphalia, Germany	51.233334	6.783333	15	Café	Coffee Shop	Italian Restaurant	Vietnamese Restaurant	Park	Cocktail Bar	...	Fried Chicken Joint	Franconian Restaurant	Forest	Gift Shop	Cosmetics Shop	Food Truck	Food Court	Food & Drink Shop	Food	Go Kart Track
46	Stuttgart, Baden-Württemberg, Germany	48.783333	9.183333	15	Bar	Park	Plaza	Italian Restaurant	Cocktail Bar	Ice Cream Shop	...	Football Stadium	Food Truck	Food Court	General Entertainment	Food & Drink Shop	Food	Flea Market	Fish Market	Gift Shop	Financial or Legal Service
54	Cologne, North Rhine-Westphalia, Germany	50.933594	6.961899	15	Park	Café	Plaza	Coffee Shop	Hotel	Bakery	...	Gourmet Shop	Grilled Meat Restaurant	Grocery Store	Gym	Gym Pool	Fried Chicken Joint	Fountain	Harbor / Marina	Fish Market	Factory
55	Dresden, Germany	51.050407	13.737262	15	Plaza	German Restaurant	Bakery	Hotel	Science Museum	Café	...	Furniture / Home Store	Fountain	Forest	Football Stadium	Food Truck	Food Court	Garden Center	Flower Shop	Gas Station	Flea Market
59	Cologne, Germany	50.935173	6.953101	15	Park	Café	Coffee Shop	Plaza	Hotel	French Restaurant	...	Food & Drink Shop	Furniture / Home Store	Food	Flower Shop	Flea Market	Garden	Fish Market	Field	Eastern European Restaurant	Garden Center
67	Hanover, Lower Saxony, Germany	52.373920	9.735603	15	Park	Supermarket	Café	Italian Restaurant	Coffee Shop	Bakery	...	College Stadium	Colombian Restaurant	Comedy Club	Comfort Food Restaurant	Concert Hall	Dive Bar	Discount Store	Trattoria/Osteria	Dog Run	Flea Market
70	Berlin, Germany	52.520008	13.404954	15	Bookstore	Coffee Shop	Bakery	Concert Hall	Park	Indie Movie Theater	...	Forest	Football Stadium	Garden Center	Gas Station	Food & Drink Shop	Food	Flower Shop	Flea Market	Fish Market	Grocery Store
73	Nuremberg, Germany	49.460983	11.061859	15	Park	Supermarket	Coffee Shop	Italian Restaurant	Café	History Museum	...	Food Court	Gym	Gift Shop	Go Kart Track	Golf Course	Grilled Meat Restaurant	Airfield	Grocery Store	African Restaurant	Food Truck
74	Hamburg, Germany	53.551086	9.993682	15	Café	Coffee Shop	Park	Seafood Restaurant	Wine Bar	Plaza	...	Halal Restaurant	Harbor / Marina	Hardware Store	Garden Center	Furniture / Home Store	Frozen Yogurt Shop	Fried Chicken Joint	English Restaurant	Event Space	Factory
79	Wiesbaden, Germany	50.078217	8.239761	15	Café	Park	German Restaurant	Plaza	Supermarket	Falafel Restaurant	...	Eastern European Restaurant	Food	Forest	Food Court	Airfield	Golf Course	Gourmet Shop	American Restaurant	Airport Service	Grilled Meat Restaurant

Place Name	
34	Aachen, North Rhine-Westphalia, Germany
41	Bonn, North Rhine-Westphalia, Germany
45	Düsseldorf, North Rhine-Westphalia, Germany
46	Stuttgart, Baden-Württemberg, Germany
54	Cologne, North Rhine-Westphalia, Germany
55	Dresden, Germany
59	Cologne, Germany
67	Hanover, Lower Saxony, Germany
70	Berlin, Germany
73	Nuremberg, Germany
74	Hamburg, Germany
79	Wiesbaden, Germany

So we end up with a short list of german cities where we can potentially open a new store, since people live there or visit there have similar tendency as in Düsseldorf.

Result

The short list has narrowed down the choices to 12 choices.

The major international cities like Berlin, Hamburg are listed. These cities obviously serve as international connection point and hosts a large number of international institutions, companies, tourists groups, also universities. Our data analysis by clustering confirms its open culture.

Cities like Aachen, Bonn, Cologne are in the same Rheinland area as Düsseldorf. These cities are also commercially active cities which host famous universities in Germany, which accounts for the international elements of the cities.

Dresden, Hannover, Stuttgart and Wiesbaden are all capitals of the respective federal states, which are comparatively economically well performing states. Their appearance on the lists are all comprehensible.

In the contrast the bavarian metropolitian Munich, which is regarded as the wealthiest city, is not listed.

Discussion section

We have considered purely in terms of venues. What other elements could be important to our consideration? for example:

1. GDP per capital
2. the number of foreigners
3. the existing numbers of potential competitors
4. distance to the headquarter in Düsseldorf ... How to incorporate these elements into the data set. For example which kinds of stores should be considered as competitors and which should not be. For example is coffee shop to be thought as competitors?

Does the number of top venues also affect the results? Does the choice of the number of clusters merely narrow down the number of each cluster group or does it affects the structure of the clustering?

To further the investigation, how should we find the right neighborhood, even the street to open our new store?

These questions are all topics one can dig deeper.

Conclusion

Data Analysis really helps in the decision making. Firstly just by running the dataset which is unbiased and objective, we end up with a list which somewhat confirms our conception of the social-economic conception of german cities. We have a data set of around 100 major cities in Germany, but we get to the conclusion that big cities like Berlin and Hamburg are listed at the top of the potential lists.

Secondly, signature cities in the vicinity in the Rhineland as Aachen, Bonn and Cologne share the same culture atmosphere and do host famous international universities in their cities. Their candidacy as top choices is fully advocated. It is also worthwhile to mention that Bonn and Aachen are relatively small cities in germany in comparison to other cities, but due to the huge amount of international students in their univesities they are culturally very embracing. The fact that the data analysis tells this truth just by venue information is astonishing and magical.

Third, capitals of several economic strong federal states, like Hannover, Stuttgart, Wiesbaden, Dresden are also listed as the potentials.

Last but not least, there are also famous cities, which might be considered more traditional, are not particularly suitable for an asian shop, so delivers our analysis this information.

Appendix

One can find the Jupyter notebook with the code in

<https://github.com/Hyperboloiz/battlegermany/blob/main/code.ipynb>

(<https://github.com/Hyperboloiz/battlegermany/blob/main/code.ipynb>)

In []: