CAPSTONE PROJECT

Restaurant opening in Munich

Leonardo Mangold

Introduction

The purpose of this project is to analyze the neighborhoods of Munich in order to determine possible location for opening a restaurant. I intend to analyze relevant data and provide value insights.

Data

1. District data of Munich which we can find at: https://www.muenchen.de/int/en/living/postal-codes.html

2. Geographical coordinates of Munich and each of its neighborhoods

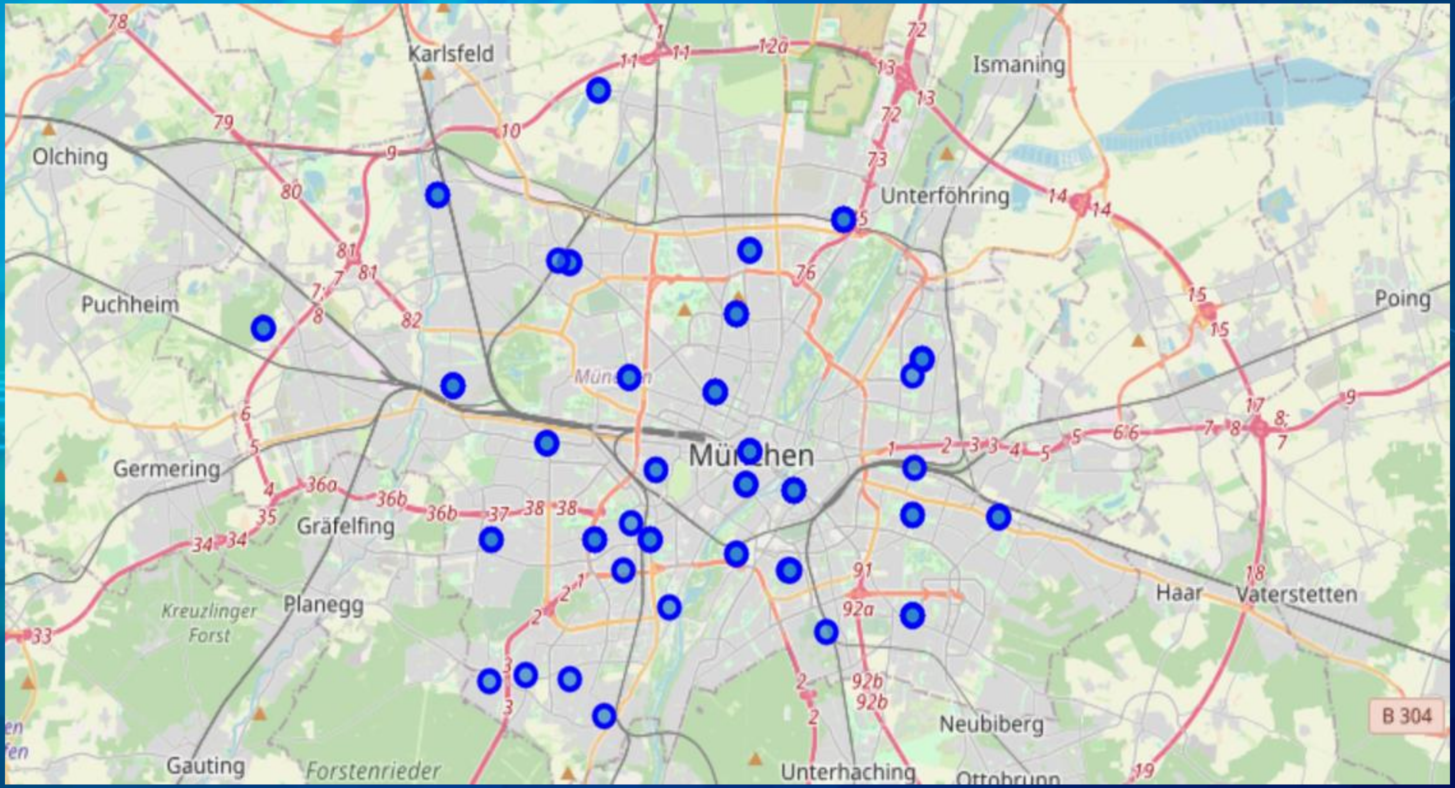
3. Venue data for neighborhoods in Munich

Importing and creating dataset

First we import the dataset. Following, we use the geolocator library to add Latitude and Longitude to the Dataframe.

	District	Postal Code	Latitude	Longitude
0	Allach-Untermenzing	80995	48.195157	11.462973
1	Allach-Untermenzing	80997	48.195157	11.462973
2	Allach-Untermenzing	80999	48.195157	11.462973
3	Allach-Untermenzing	81247	48.195157	11.462973
4	Allach-Untermenzing	81249	48.195157	11.462973

Visualizing Districts on a map



Foursquare venues

We use Foursquare to bring the most common venues of each Postal Code zone.

	District	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Allach-Untermenzing	48.195157	11.462973	Bäckerei Schuhmair	48.197175	11.459016	Bakery
1	Allach-Untermenzing	48.195157	11.462973	Sport Bittl	48.191447	11.466553	Sporting Goods Shop
2	Allach-Untermenzing	48.195157	11.462973	dm-drogerie markt	48.194118	11.465640	Drugstore
3	Allach-Untermenzing	48.195157	11.462973	Sicilia	48.193331	11.459387	Italian Restaurant
4	Allach-Untermenzing	48.195157	11.462973	Lidl	48.194428	11.465612	Supermarket

One hot encoding and mean

In order to determine the venues categories presence in each District, we use one hot encoding and then we calculate the mean.

	District	ATM	Afghan Restaurant	American Restaurant	Arcade	Art Gallery	Art Museum	Asian Restaurant	Athletics & Sports	Austrian Restaurant	Auto Dealership	Automotive Shop	BBQ Joint	Bagel Shop	Bake
0	Allach-Untermenzing	0.0	0.000000	0.0	0.0	0.0	0.00	0.0	0.0	0.0	0.0	0.125	0.0	0.0	0.1250(
1	Altstadt-Lehel	0.0	0.000000	0.0	0.0	0.0	0.01	0.0	0.0	0.0	0.0	0.000	0.0	0.0	0.0000(
2	Au-Haidhausen	0.0	0.026316	0.0	0.0	0.0	0.00	0.0	0.0	0.0	0.0	0.000	0.0	0.0	0.0000(
3	Berg am Laim	0.0	0.000000	0.0	0.0	0.0	0.00	0.0	0.0	0.0	0.0	0.000	0.0	0.0	0.1034
4	Bogenhausen	0.0	0.000000	0.0	0.0	0.0	0.00	0.0	0.0	0.0	0.0	0.000	0.0	0.0	0.1388

10 most common venues category

Finally, we create a function that brings the 10 most common venue category of each District.

	District	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	Allach-Untermenzing	Supermarket	Drugstore	Bakery	Sporting Goods Shop	Automotive Shop	Italian Restaurant	Yoga Studio	Fish Market	Fast Food Restaurant	Farmers Market
1	Altstadt-Lehel	Bavarian Restaurant	Café	Plaza	Hotel	German Restaurant	Restaurant	Coffee Shop	Cocktail Bar	Church	Clothing Store
2	Au-Haidhausen	Italian Restaurant	Concert Hall	Coffee Shop	French Restaurant	Thai Restaurant	Gourmet Shop	Doner Restaurant	Bistro	Movie Theater	Rock Club
3	Berg am Laim	Supermarket	Drugstore	Café	Metro Station	Hotel	Gastropub	Bakery	Cafeteria	Light Rail Station	Eastern European Restaurant
4	Bogenhausen	Bus Stop	Drugstore	Bakery	Italian Restaurant	Greek Restaurant	Park	Bank	Supermarket	Pharmacy	Water Park

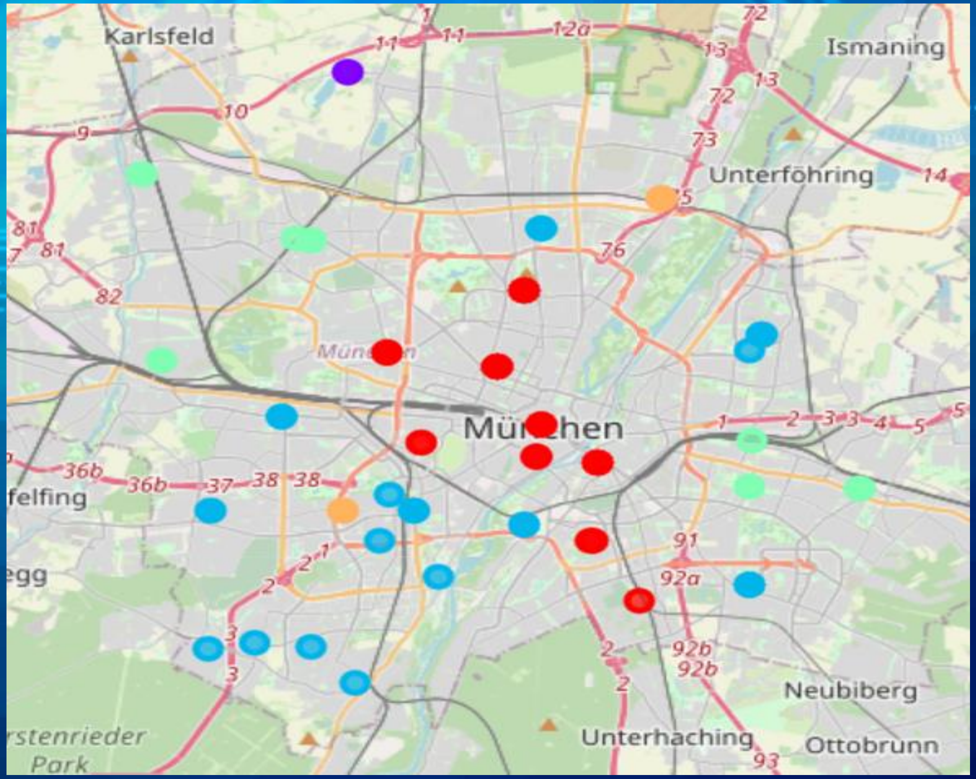
Clustering

Our Dataframe is now ready. We proceed to cluster Munich's Districts into 5 clusters and we add them to the Dataframe.

	Cluster Labels	District	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	3	Allach-Untermenzing	Supermarket	Drugstore	Bakery	Sporting Goods Shop	Automotive Shop	Italian Restaurant	Yoga Studio	Fish Market	Fast Food Restaurant	Farmers Market
1	0	Altstadt-Lehel	Bavarian Restaurant	Café	Plaza	Hotel	German Restaurant	Restaurant	Coffee Shop	Cocktail Bar	Church	Clothing Store
2	0	Au-Haidhausen	Italian Restaurant	Concert Hall	Coffee Shop	French Restaurant	Thai Restaurant	Gourmet Shop	Doner Restaurant	Bistro	Movie Theater	Rock Club
3	3	Berg am Laim	Supermarket	Drugstore	Café	Metro Station	Hotel	Gastropub	Bakery	Cafeteria	Light Rail Station	Eastern European Restaurant
4	2	Bogenhausen	Bus Stop	Drugstore	Bakery	Italian Restaurant	Greek Restaurant	Park	Bank	Supermarket	Pharmacy	Water Park

Clustering visualization

Lets visualize on a map how the Districts are clustered.



Analyzing cluster 0

This cluster is spread across Munich, including its centre and most tourist locations. As a result, there should be a strong demand, but rent would be higher than in other Districts.

```
# Cluster 0 Value Counts
cluster0['1st Most Common Venue'].value_counts()
```

```
Café                     15
Vietnamese Restaurant     8
Bavarian Restaurant       7
Italian Restaurant        7
Bakery                    5
Park                      4
```

Its most common venues include Cafes, different types of restaurants (Italian, Vietnamese, Asian, Stakehouse) as well as Cocktail Bars. On the one hand this Districts presents lots of potential customers but, on the other hand, competence is numerous and diverse. In order to operate in this cluster, we must offer a highly differentiated product in order to attract customers attention and convince them to try our restaurant.

Analyzing cluster 1

Districts in this cluster are away from the centre of Munich and do not present restaurants as most common venues. Therefore, we ought to discard this cluster.

```
# Cluster 1 Value Counts
cluster1['1st Most Common Venue'].value_counts()
```

```
Motorcycle Shop    3
```

Analyzing cluster 2

Also spreading across Munich but excepting the centre, these Districts present lots of supermarkets, German restaurants and bus stops. Looks like residential Districts (lots of supermarkets and bus stops). People in these Districts seem to prefer traditional food (German restaurants) over Fast Food, Italian, Asian, Falafel, Stakehouse, etc. So, it would be an interesting option to consider if our target are Munich's residents and not tourists.

```
# Cluster 2 Value Counts
cluster2['1st Most Common Venue'].value_counts()
```

```
Supermarket          22
German Restaurant    10
Bus Stop              6
```

Analyzing cluster 3

This cluster represents a single District which is far away from the others. Although it has some food offers (Greek, Doner, Fast Food and Falafel restaurants), there aren't many. We can assume these Districts present more accessible prices to consumers, as well as lower costs for companies (rent for example). Might be a good opportunity if we do not have too much capital to invest.

```
# Cluster 3 Value Counts
cluster3['1st Most Common Venue'].value_counts()
```

```
Supermarket    13
Spa             6
Bakery          5
```

Analyzing cluster 4

```
# Cluster 4 Value Counts
cluster4['1st Most Common Venue'].value_counts()
```

```
Fast Food Restaurant    8
Ice Cream Shop          5
```

In this case we have Fast Food as the most common venues. Considering their low prices and how difficult it is to compete with them, we ought to discard this cluster.