**To set up a Shopping Mall in Munich**

**Business problem**

The objective of this project is to analyze and select the best locations in the city of Munich, Germany, to open a new shopping mall. This project is mainly focused on geospatial analysis of the Munich City to understand which would be the best place to open a new mall. Using data science methodology and machine learning techniques like clustering, this project aims to provide solutions to answer the business question: In the part of Munich, if a property developer is looking to open a new shopping mall, where would you recommend that they open it?

**Data Required**

• List of neighbourhoods in Munich.

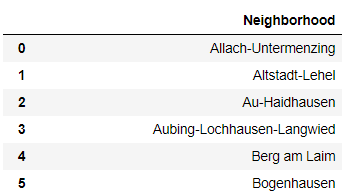
• Latitude and longitude coordinates of those neighbourhoods. This is required in order to plot the map and also to get the venue data

• Venue data, particularly data related to shopping malls. We will use this data to perform clustering on the neighbourhoods.

**Methodology**:

1. **Web Scraping to get the information the Neighbourhoods data**:

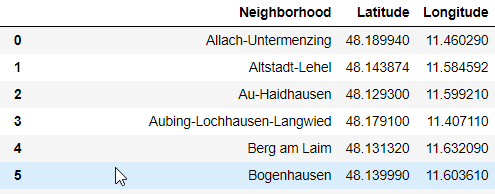
To extract the list of neighbourhoods in Munich used Python requests and beautifulsoup packages.



This data is created after scraping.

2. **Get Latitude and Longitude for the Neighbourhoods**:

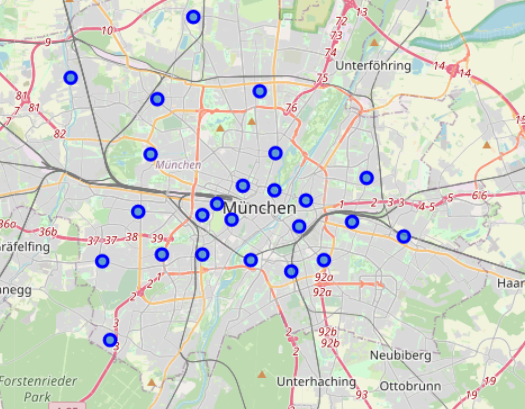
Now we need to get the latitude and longitude of the location. So that we can use that with Foursquare API to find the explore the location. For which we are going to use geocoders package.



This is the data which is obtained, after merging the latitude and longitude from the geocoder with the original DataFrame.

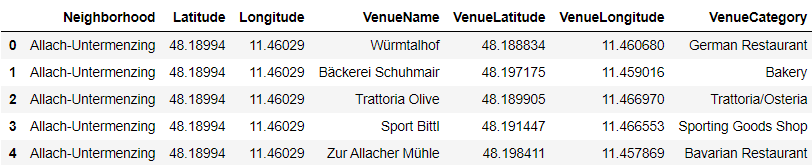
3**. Visualize the neighbourhoods**:

After gathering the data, populate the data and visualize the neighbourhoods in a map using Folium package.



4. **Explore the Neighbourhoods**:

Use Foursquare APIs to explore the neighbourhoods and extract the venues into a new DataFrame



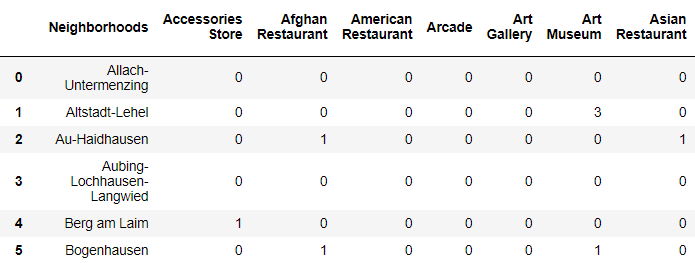
To check out the no of unique categories and the top 50 places



5. **Analyzing each neighbourhood**:

Apply one hot encoding to all the venues, as it is easy to work with binary data than categorical values.

Then group rows by neighbourhood, by taking the sum of the frequency of occurrence of each category.



To get the Number of Shopping Malls in Munich.



Creating a dataframe for Shopping Mall data only



6. **Clustering the neighbourhoods**

Now we need to cluster the neighbourhoods into different clusters. The results will allow us to identify which neighbourhoods have a higher concentration of shopping malls and fewer number of shopping malls. So that it will help us answer the question to which neighbourhoods are most suitable to open new shopping malls.

We set the number of clusters to 2 and run the algorithm. After applying the K-Means clustering algorithm, all the neighbourhoods get segregated and form different clusters.



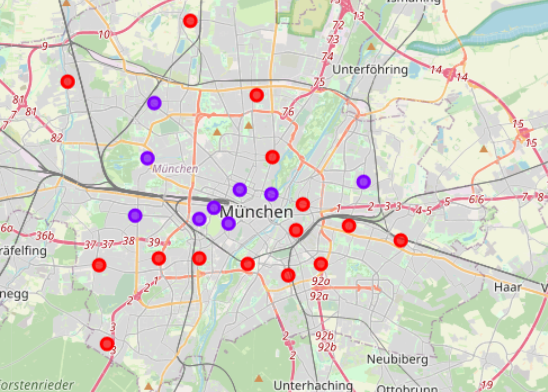
Here the Shopping Mall column represents the number of shopping malls in that particular area and Cluster Labels represents the cluster number (either 0 or 1)



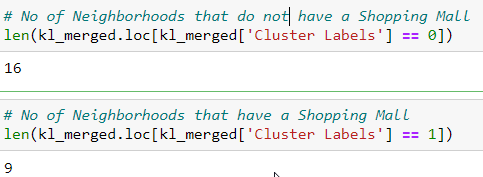
Now here we can clearly see all the cluster numbers in the sorted order that is 0 or 1.

7. **Visualizing the resulting clusters**





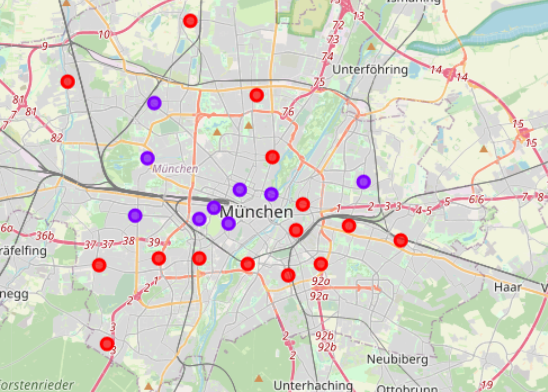
8. **Examine the clusters**



**Results**

There are 16 places in cluster 0 which is the highest among the 2 clusters, and cluster 0 contains all the places which have a shopping mall. Cluster 1 contains 9 places and all of them contain exactly 1 shopping mall.

The results from the K-means clustering show that we can categorize the neighbourhoods into 2 clusters based on the frequency of occurrence for “Shopping Mall”:  
• Cluster 0: Neighbourhoods with no shopping malls   
• Cluster 1: Neighbourhoods with one shopping malls



We visualize the results of the clustering in the map with cluster 0 in red colour and cluster 1 in purple colour

**Conclusions**

The best place to build a shopping mall is in Cluster 0 as it has no shopping malls in that place and as the count of shopping malls in Munich are less, there is less competitions and high potential to open a Shopping Mall. Munich is a economically wealthy city, as the city's unemployment rate is around only four percent, and even lower among youth, there is a great affinity to open a new shopping mall.

This model can not only be used for Shopping mall but also for others business which are yet to be opened in the near future such as cafes, grocery store and many more. This is just the demonstration of one of one such Business.