# Data

Following data will be needed for the project:

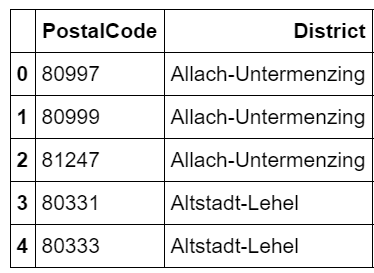
* The list of Munich districts
* Latitude and longitude coordinates of the districts
* Venue data, particularly data on restaurants in Munich and their location

# Sources of data

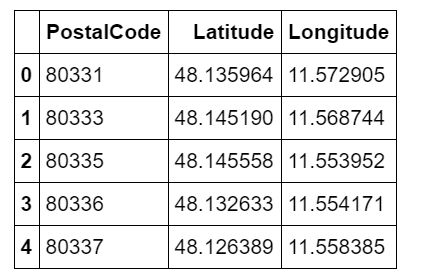
* The list of Munich districts including their postal codes was downloaded from the website <https://www.muenchen.de/int/en/living/postal-codes.html> and converted into a Microsoft Excel spreadsheet.
* Latitude and longitude coordinates of the districts based on their postal codes were downloaded from the website <https://public.opendatasoft.com/explore/dataset/postleitzahlen-deutschland/table/> as a Microsoft Excel spreadsheet.
* Venue data were obtained through Foursquare API.

# Data

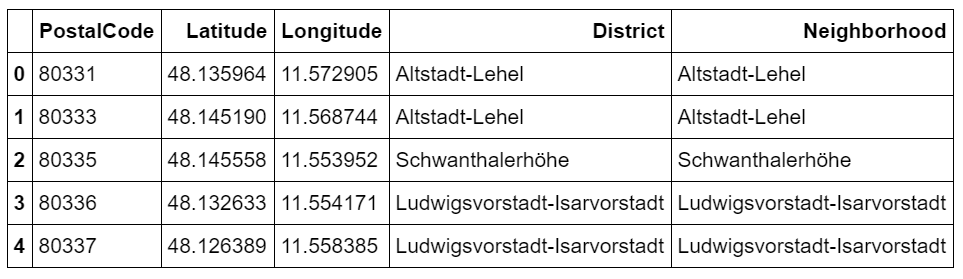
First, the Microsoft Excel Spreadsheet containing the list of Munich districts and their postal codes in Munich and corresponding districts was converted into a dataframe.



The second Microsoft Excel Spreadsheet was converted into a dataframe containing a list of all postal codes in Munich and their coordinates.

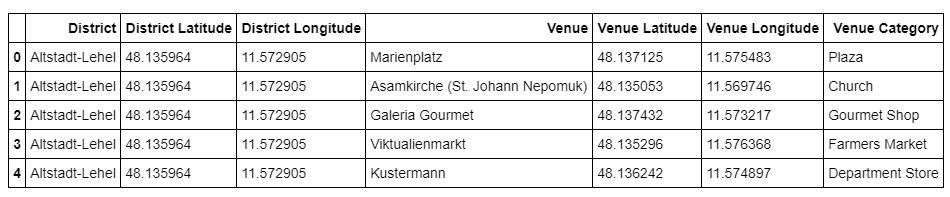


Both dataframes have been merged, resulting in a dataframe “munich\_df” containing coordinates of every district in Munich.



## Obtaining data on venues in Munich

The next step was to obtain data on all Munich venues using Foursquare API. API calls containing coordinates of districts were made to Foursqare. Foursquare returned venue data in a json format. Using a python loop, data on venue name, category and its coordinates were extracted and converted into a pandas dataframe.



The following steps will be to cluster the districts and to visualize the data.