

An aerial photograph of Munich, Germany, showing the city center. The prominent Gothic spire of the Frauenkirche (Cathedral of Our Lady) is the central focus. To its left, the two large green domes of the Frauenkirche's towers are visible. The surrounding area is filled with traditional European architecture, including colorful buildings with red-tiled roofs. The Marienplatz square is visible in the foreground, with its characteristic clock tower. The image is overlaid with a semi-transparent blue geometric design on the right side.

# Opening up new gym in Munich

Own Capstone Project  
(as part of IBM Applied Data Science with Python course)

# How about Munich?

- ▶ Munich (München) - 3rd largest city in Germany (1.5 mil inhabitants) and growing fast
- ▶ Located at the south of Germany in the province of Bavaria close to the Alps mountains
- ▶ Famous for its beer festival Oktoberfest
- ▶ Economically strongest city in Germany - large business potential





# Project Introduction

- ▶ Purpose of this project is to find neighborhoods with a space for a new gym venue
- ▶ Munich has 25 neighborhoods (Stadtbezirke) that vary in size and wealthiness
- ▶ Comparison of neighborhoods
  - ▶ number of gyms vs population
  - ▶ Affluence - average rent per sqm of private estates
- ▶ Goal:
  - ▶ find the neighborhoods with high population and low number of gyms
  - ▶ what kind of gym to offer (simple vs luxurious)



# Data Collection and Source

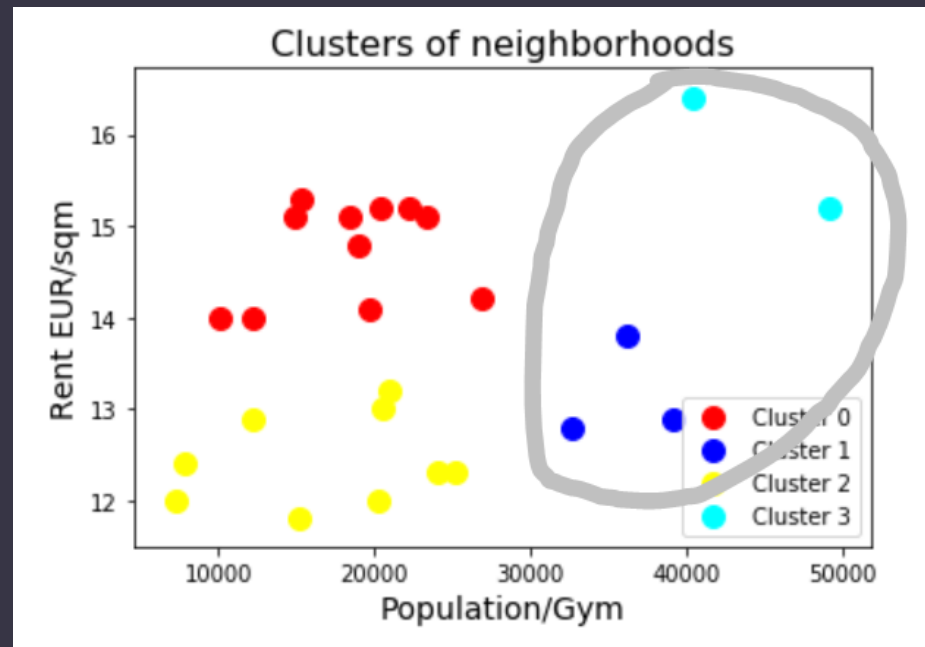
- ▶ 3 data inputs:
  - ▶ Foursquare API
    - ▶ List of gyms and their addresses
    - ▶ Format: json file
  - ▶ Demographical data
    - ▶ number of inhabitants in each neighborhood
    - ▶ source: Munich city council statistical website
    - ▶ Format: MS Excel file
  - ▶ Private property rent prices
    - ▶ Average rent per sqm in each neighborhood
    - ▶ Source: <https://suedbayerische-immobilien.de/Mietpreise-Muenchen-Stadtteile>
    - ▶ Format: table on the webpage

# Methodology

- ▶ Foursquare API - venue info
  - ▶ Collected number of gyms in each area and their addresses
  - ▶ Radius set high to cover large neighborhoods -> duplicated venues, filtered unique venues and checked the corresponding venues using Geopy Nominatim library
- ▶ Demographic file
  - ▶ Selected population 15 - 65 years old (customer base)
  - ▶ Calculated population share for additional gym in each area - „Potential Customers“
- ▶ Rental price data
  - ▶ Table extracted from website using Beautiful Soup, applied name conversion
- ▶ Merged all three columns into one dataframe

# K-Means Clustering

- ▶ Final data frame segmented into 4 clusters using K-Means algorithm
- ▶ Inputs - population per gym and the average rent in each neighborhood
- ▶ Clusters plotted into scatter plot
- ▶ Clusters 1 and 3 - our target
  - ▶ neighborhoods with large population and no gyms
  - ▶ ranging over a variety of rent prices - cheaper as well as affluent areas



# Optimal Neighborhoods for New Gym

- ▶ Cluster 1 - medium to lower rent areas, welcome simple gym with average membership fees
  - ▶ Obergiesing - Fasangarten
  - ▶ Untergiesing - Harlaching
  - ▶ Hadern
- ▶ Cluster 3 - high rent areas offering upscale venues, ideal for luxury gym with a lot of additional facilities
  - ▶ Ludwigsvorstadt - Isarvorstadt
  - ▶ Schwabing - West

Area Name	Population 15 - 65 Yrs	Number of Gyms	Potential Customers	Rent EUR/sqm	Cluster Label
Ludwigsvorstadt - Isarvorstadt	40440.0	0.0	40440	16.4	3
Schwabing - West	49129.0	0.0	49129	15.2	3
Obergiesing - Fasangarten	39182.0	0.0	39182	12.9	1
Untergiesing - Harlaching	36244.0	0.0	36244	13.8	1
Hadern	32602.0	0.0	32602	12.8	1



Thanks for your attention!