

IBM Data Science Professional Certificate

Applied Data Science Capstone Project

Gunjan S Makhloga

'SHOPs AND STOREs' SHOPPING IN MUNICH

DATA

DATA DESCRIPTION

Data collection plays an important role in deciding the nature of data analysis and the insights that could be drawn from it. As required to solve the business problem, we used Munich and its district dataset from publicly available source detailed in next section. The following data were pulled from various data sources:

- Munich districts and their postal codes

	District	PostalCode
0	Allach-Untermenzing	80995, 80997, 80999, 81247, 81249
1	Altstadt-Lehel	80331, 80333, 80335, 80336, 80469, 80538, 80539
2	Au-Haidhausen	81541, 81543, 81667, 81669, 81671, 81675, 81677
3	Aubing-Lochhausen-Langwied	81243, 81245, 81249
4	Berg am Laim	81671, 81673, 81735, 81825
5	Bogenhausen	81675, 81677, 81679, 81925, 81927, 81929
6	Feldmoching-Hasenberg	80933, 80935, 80995
7	Hadern	80689, 81375, 81377
8	Laim	80686, 80687, 80689
9	Ludwigsvorstadt-Isarvorstadt	80335, 80336, 80337, 80469

- Geographical coordinates in the form of latitudes and longitudes for each postal code

	District	PostalCode	Latitude	Longitude
0	Allach-Untermenzing	80995	48.211665	11.525658
0	Allach-Untermenzing	80995	48.211665	11.525658
1	Altstadt-Lehel	80331	48.136064	11.573450
0	Allach-Untermenzing	80995	48.211665	11.525658
1	Altstadt-Lehel	80331	48.136064	11.573450

- 100 venues that were within 500 meters radius for each postal code
- Venues and their category

DATA SOURCES

- Munich district data and their postal codes were scraped from web page <https://www.muenchen.de/int/en/living/postal-codes.html> using Python GET request and BeautifulSoup packages.
- The latitudes and longitudes coordinate of Munich and each of its district postal code were obtained by using Python Geocoder package.
- RESTful API calls were made to Foursquare API to retrieve the information about the venues for each postal code. It also provided the categories of the venue data. To the interest of this project, we used only 'Shops' and 'Stores' categories to solve the business problem.