|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Data name | Data type | # features | # samples | source |
| NO2\_2019 | Time series | 5 | 8760 | BLU |
| NO2\_2020\_01 | Time series | 5 | 744 | BLU |
| NO2\_2020\_02 | Time series | 5 | 696 | BLU |
| NO2\_2020\_03 | Time series | 5 | 744 | BLU |
| NO2\_2020\_04 | Time series | 5 | 720 | BLU |
| NO2\_2020\_05 | Time series | 5 | 744 | BLU |
| NO2\_2020\_06 | Time series | 5 | 531 | BLU |
| PM2x5\_2019 | Time series | 4 | 8760 | BLU |
| PM2x5\_2020\_01 | Time series | 4 | 744 | BLU |
| PM2x5\_2020\_02 | Time series | 4 | 696 | BLU |
| PM2x5\_2020\_03 | Time series | 4 | 744 | BLU |
| PM2x5\_2020\_04 | Time series | 4 | 720 | BLU |
| PM2x5\_2020\_05 | Time series | 4 | 744 | BLU |
| PM2x5\_2020\_06 | Time series | 4 | 531 | BLU |
| CO\_2019 | Time series | 2 | 8760 | BLU |
| CO\_2020\_01 | Time series | 2 | 744 | BLU |
| CO\_2020\_02 | Time series | 2 | 696 | BLU |
| CO\_2020\_03 | Time series | 2 | 744 | BLU |
| CO\_2020\_04 | Time series | 2 | 720 | BLU |
| CO\_2020\_05 | Time series | 2 | 744 | BLU |
| CO\_2020\_06 | Time series | 2 | 531 | BLU |

BLU = Bayerisches Landesamt für Umwelt. Link: <https://www.lfu.bayern.de/luft/immissionsmessungen/messwertarchiv/index.htm>

**Motivation:**

In order to give an environment friendly recommandation for the tourists in Munich and encourage more people to visit parcs and spend more time in the nature it was important to investigate the quality of air in the city after the measures taken by the government. Due to the lack of information concerning the CO2 emissions and as the Bavarian State Office for the Environment provides up-to-date data sets on other several gases, it was possible to analyze multiple data sets of NO2 (Nitrogen Dioxide), CO (Carbon Monoxide) and Particulate Matter (PM2.5) thanks to 5 monitoring stations in Munich located in Allach, Johanneskirchen, Landshuter Allee, Stachus und Lothstraße.

**Data preprocessing:**

1. Reading the Data Sets :

* Start exctracting the values from the third row through ‘header=2‘ and choosing the columns that contain only the stations located in Munich.
* dataCO\_Apr\_hourly = pd.read\_excel('CO\_2020\_04.xls',header=2,usecols =['Zeitpunkt','München/Landshuter Allee','München/Stachus'] )

1. Understanding the data’s shape and type:

* print(dataCO\_Apr\_hourly.shape)
* print(dataCO\_Apr\_hourly.dtypes)

1. Finding the missing values and replacing them with the mean value of their column as they are of a negligeable amount through the function REPLACE

* dataCO\_Apr\_hourly['München/Stachus'].replace(0, dataCO\_Apr\_hourly['München/Stachus'].mean(), inplace=True)

1. Changing the name of the column ‘Zeitpunkt‘ into ‘DATE‘ through the function RENAME

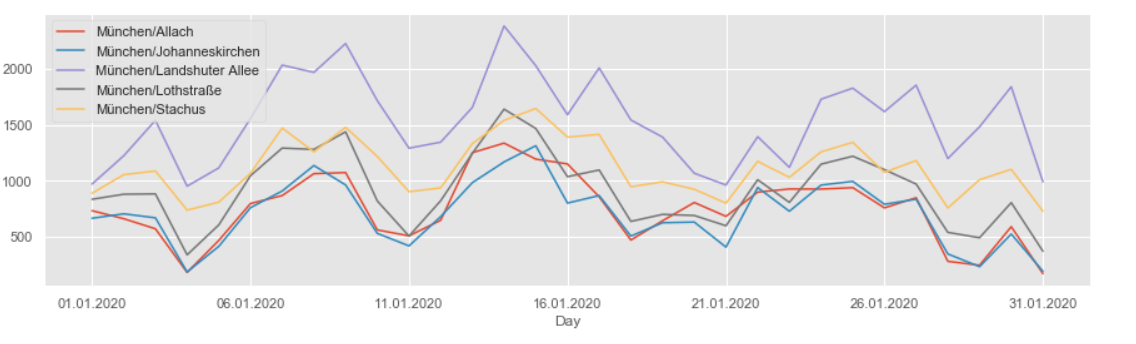
* dataCO\_Apr\_hourly= dataCO\_Apr\_hourly.rename(columns={'Zeitpunkt': 'DATE'})

1. Transforming the hourly data into daily data:

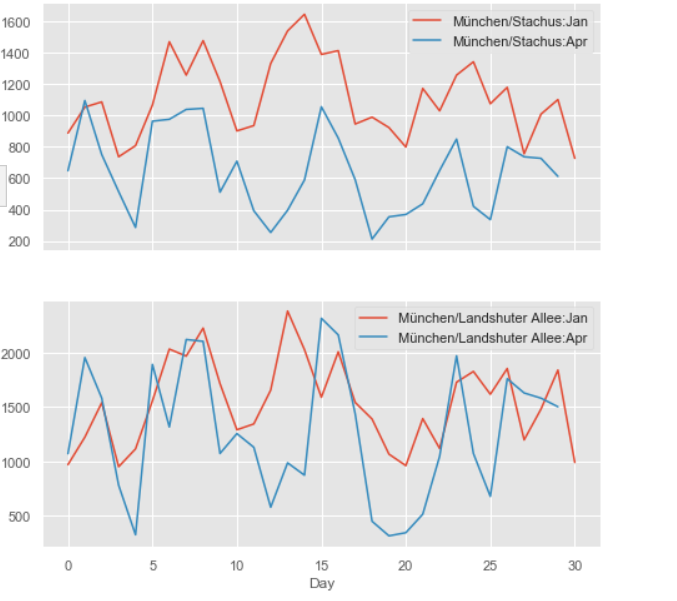
* through the functions TO\_DATETIME which sets the column DATE as a datetime , SET\_INDEX which sets the datetime as an index which will allow us to RESAMPLE the data according to Days instead of Hours
* dataPm2x5\_2019\_hourly['DATE']=pd.to\_datetime(dataPm2x5\_2019\_hourly['DATE'], infer\_datetime\_format=True)
* dataPm2x5\_2019\_hourly = dataPm2x5\_2019\_hourly.set\_index(['DATE'])
* print(dataPm2x5\_2019\_hourly)
* dataPm2x5\_2019\_daily=dataPm2x5\_2019\_hourly.resample('D').sum()
* through splitting the column of DATE into two columns DAY and HOUR and then grouping the values by the DAY and summing them for each day
* dataCO\_Jan\_hourly[['Day','Hour']] = dataCO\_Jan\_hourly['DATE'].astype(str).str.split(expand=True)
* dataCO\_Jan\_daily = dataCO\_Jan\_hourly.groupby(['Day']).sum()

**Data Analysis:**

As an example in the plot below, we can see the NO2 emissions through the month of January before the lockdown. Landshuter Allee and Stachus are the biggest sources of NO2 emissions in Munich.



As we extracted the data for other months as well, it was suitable to compare the air quality in different districts between January (before the lockdown) and April (after the lockdown).



As a next step to see how efficient the lockdown was at reducing the emission of those gases we plotted the values of last and this year before the lockdown and after the lockdown in the district of Stachus.

