

Air Quality Analysis

Exercise 1

- Create a new Python script.
- Read in the datasets **airquality** and **sites**.
- Merge the two data sets.
- Save the merged data as a csv file, to be used also by the other team members.

Exercise 2

- Create a new Python script.
- Read in the merged data set from exercise 1.
- Display the 10 worst sites in terms of no2 emissions on January 1 2020.
- Display the 10 best sites in terms of their average o3 emissions.
- Display the 10 days with the highest no2 emissions for the **site_name** Kiel-Theodor-Heuss-Ring.

Exercise 3

- Create a new Python script or Jupyter Notebook, and read in the data from exercise 1.
- Calculate minimum, mean, and maximum emissions for no2.
- Calculate mean, number of non-missing observations, and the sum for the variable no2.
- Calculate for the variable no2 the mean and the sum, and for the variable elevation just the mean (because the sum makes no sense here).
- Calculate the minimum, mean, and maximum emissions for each emissions type per **site_area** and **site_type**.

Exercise 4

- Create a new Python script or Jupyter Notebook, and read in the data from exercise 1.
- Which are the different site areas that exist in the data?
- What is the number of unique site types in the data?
- What is the proportion of observations belong to the different site areas?
- Identify all sites that have the substring 'Kiel' in the site name
- Create a new column that contains the number of characters of the site names?

Exercise 5

- Create a new Python script or Jupyter Notebook, and read in the data from exercise 1.
- Freely explore the data