
Exercise - Data

Exercises in Data Modeling (190.021), WS2024

¹dm@ai-lab.science, January 13, 2025

This is a cheat sheet for the exercise on data analysis.

Since the relevant infos on what the notebook is about can be found in the notebook itself no further explanations will be made here. This sheet's main use is for you to know what strings you can type in for what variables.

1 Infos on the dataset

For more details about the used dataset, visit the following link: Urban Air Quality and Health Impact Dataset on Kaggle

2 What to generally fill in for 'None'

The cities which you can use for the string variables **city1** and **city2** are:

- Phoenix
- San Jose
- San Antonio
- Los Angeles
- San Diego
- New York City
- Chicago
- Philadelphia
- Dallas
- Houston

The columns which you can use for the string variables **col1** and **col2** are:

- datetimeEpoch
- tempmax
- tempmin

- temp
- feelslikemax
- feelslikemin
- feelslike
- dew
- humidity
- snow
- snowdepth
- windgust
- windspeed
- winddir
- pressure
- cloudcover
- visibility
- solarradiation
- solarenergy
- uvindex
- severerisk
- sunriseEpoch
- sunsetEpoch
- City
- Temp_Range
- Heat_Index
- Severity_Score
- Month
- Is_Weekend
- Health_Risk_Score

The remaining variables are to be filled in with the following datatypes.

- **threshold**: A positive integer or float value.
- **random_state**: An integer.
- **test_size**: A float between 0.0 and 1.0.

If you did not fill out the variables the program will throw an Exception. The output will look similar to picture 1. Follow the instructions given and the program will work.

```

city1 is still None. Please enter a valid city name.

-----
Exception                                     Traceback (most recent call last)
Cell In[6], line 5
    3 if city1 is None:
    4     print("city1 is still None. Please enter a valid city name.")
----> 5     raise Exception("Execution stopped because city1 is not defined.")
    6 if city2 is None:
    7     print("city2 is still None. Please enter a valid city name.")

Exception: Execution stopped because city1 is not defined.

```

Figure 1: Example of an Exception

3 What to fill in for the moodle test

Sometimes it is necessary to use the following variables to answer the questions in the moodle test, sometimes it is not necessary to use these exact values. However, these values are quite useful in illustrating how the presented concepts work. Note that, to properly and fully understand the contents, it is still advantageous to try some own variables.

Chapter a)

- `city1 = 'Chicago'`
- `city2 = 'Pheonix'`
- `col1 = 'tempmin'`
- `col2 = 'solarradiation'`
- `threshold = 3.0`

Chapter b)

- `col1 = 'uvindex'`
- `col2 = 'dew'`

Chapter c)

- `test_size = 0.5`
- `random_state = 42`
- `col1 = 'Temp_Range'`
- `col2 = 'solarradiation'`
- `threshold = 3.0`