University of Oldenburg
Python Programming in Energy Science, WiSe 2023-2024
Dr. Lukas Vollmer, lukas.vollmer@iwes.fraunhofer.de

Assignment 6

Allowed library imports: `pandas`, `numpy`, `matplotlib`

E1. Populating pandas dataframe from numpy arrays

(a) Create a 2×24 ndarray with random integer values in the range [0,20).

(b) Create a pandas DataFrame `df` with two columns named `Data1` and `Data2`.

- directly from the array
- from a dictionary created from the array
- from two pd.Series objects created from the individual array columns

(c) Insert a new column called `DataProd` which is the product of the other two columns.

E2. Exporting and importing dataframe

(a) Change the data type of the columns to float.

(b) Export the dataframe as csv-file with two decimal places.

(c) Open the csv-file in an external editor. Manually add a column `Error-Log`. Add an entry `Error` for each time step if `Data2 > 10`. Leave the other rows blank.

(d) Load the new csv-file with pandas. Make sure to recover the index as index.

(e) Visualize the dataframe by using print(df). What happenend with the blank rows? Which datatype is your new column?

E3. Nans, Grouping and modifying data

(a) Replace all nan values in the `Error-Log` column by `Okay`.

(b) Use the group functionality to group the data into `Okay` and `Error` and calculate the mean values of both `Data`-columns when grouped.

(c) Set the index in `df` to the square product of the current index

(d) Replace the rows in the Data1, Data2 columns, that have an `Error` entry in the `Error-Log` column by `np.nan` (only use one line for the command)

E4. Time series data

(a) Load the data file `timeseries_la_haute_borne_2017.csv` provided in exercise A5 with pandas. You will have to convert the index with pd.datetime(utc=True) to get a consistent datetime index.

(b) Calculate the wind direction and wind speed and add them in two additional columns

(c) How many NaN-values exist per column? In which months do the most NaN values occur?

(d) Fill the NaN-values of the wind speed with the values of the previous time step. By how many percentage points does the mean wind speed change?