

Exercise class 7

Introduction to Programming
and Numerical Analysis

Class 3

Annasofie Marckstrøm Olesen
Spring 2023

UNIVERSITY OF COPENHAGEN



Intro to data analysis in Python and problem set 3

Working with data in Python

The package Pandas is key to working with data in Python - Pandas syntax can be a little particular.

Learning Pandas can be a bit tedious and very syntax heavy. Important that you are willing to experiment a little and it is ok to look at answers. Focus on:

- Referencing and slicing - particularly the `.loc`-operator.
- Copies vs. views

Also make sure you understand the difference between making a copy of a Dataframe and changing it in place. The latter is done by providing the keyword argument `inplace=True`.

Problem set 3

Takes you through the basics of loading and cleaning data from Statistics Denmark.

Tasks teach you to load, rename or delete variables, reshape dataset, select subsets of the data etc. Most are relatively short, so make sure you understand the syntax for each operation.

Task 10 may be a little bit tricky, because error codes are caught by the try-statement. Try to run the lines individually, if you get stuck.

In the problem, make sure you follow the pre-defined structure. If you've done it correctly, you should be able to call the `plot_timeseries()` function on your data.

Note: Bug in task 10

There are a few bugs in the pre-written try:-block to task 10, causing you to get errors even if your code is written correctly:

- The code refers to a variable named `con`. However, the variable in your dataset is at this point named `cons`. Make sure to correct either the code or the name of your variable.
- At a point, `dt['inc'].values` and `dt['con'].values()` are called - but the parentheses are a mistake, it should just be `dt['inc'].values()` and `dt['con'].values` (`.values` is an attribute, not a method).

Plan for today

Plan for today is:

- Now-16.00: Work on Tasks in problem set 3
- 16.00-16.15: Break
- 16.15-16.25: Go through solutions to Tasks
- 16.25-16.50: Work on Problems in problem set 3
- 16.50-17.00: Go through solution to Problems

You are of course also welcome to ask questions related to the inaugural project.

Next time

Video lectures

- Combining data sets
- Fetching data
- Split-apply-combine

Exercise:

- A few words on data project
- Problem set 4: data analysis