



UNIVERSITY OF  
COPENHAGEN

# Sixth exercise class

Class 1 & 5

Introduction to numerical programming and analysis

---

Jonathan Wenzel Pedersen

Spring 2024

1. Plan for today

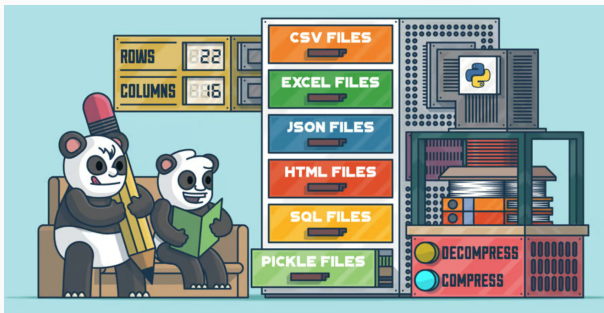
2. Hints

## **Plan for today**

---

# What are we doing today?

Today we will be working with Problem set 3, which is all about pandas, a extremely useful package in python. Today you will do a number of small problems to get you familiar with pandas



## Hints



# Dataframe, tips

Q1: New column can be created using `df['New variable'] = 'values'`

Q2: Same as question 1, remember pandas allows mathematical operations across rows by simply using the operation (for example \*)

Q3: Use `describe()`.

# Indexing, tips

Pandas allows us to pass in a True/False statement and returns only the rows that are True in that statement and we can also change these values using this method. It works with the codes:

```
df['True/False statement'] = 'new value' or  
df.loc['True/False statement', 'column names'] = 'new  
value'. Column names can be : if its all columns. The True/False  
statement can for example be df['Column 2'] == 0.5
```

Q4: Use True/False like above

Q5: Same you just need two (Think how you can use logical operators for this)

Q6: Consider how to do it for only specific column (which code should you use)

Q7: Same as question 6.

## Dropping and renaming, tips

Q8: The command `.drop()`, be AWARE that `.drop()` uses index or column values. The True/False statement you used earlier returns Series i.e. not a index or column, be aware of this (try to think about if and a possible solution - search google or ask me if you get stuck)

Q9: Consider using the `.rename()` command.



## Functions, tips

Q10: You need to make the 4 assets equal each other i.e. create 3 functions that lead to the same result. **Note! code in `dt['assets3']` has wrong code it needs to be `.values` no parenthesis**

## Cleaning, tips

Q11: Use what you have used in previous questions, part d is that if you observe the dataset every second year is missing replace the missing with the previous year, you can use the pandas code `.ffill()`