Class 7, Problem Set 4

Introduction to Programming and Numerical Analysis

Plan for today

1. Inaugural project
2. Data project
3. Working on PS4
    - Concepts
    - Split-apply-combine

Inuagural project

- Overall **solid** projects
- Remember that markdown is your friend and proof-reading is too!
- Non-passed projects should be handed in along with the data project

Working in Jupyter Lab

IMPORTANT: when your are done and ready to hand-in:

Restart kernel and run all cells

Data project

- From the start bear in mind that you should tell a good story
    - Nice documentation
    - Nice visuals
    - Nice markdown
    - Nice (sub-) conlusion(s) ... and so on
- Minimum requirements:
    - Import data from an online source of your own choosing
    - Present the data visually
    - Apply some method(s) from descriptive economics

Remember...

UNIVERSITY OF
COPENHAGEN

that the exam is a portfolio!

Using an API, interactive figures, maybe doing some regression/ML? → Awesome data project
→ Higher chance of passing the exam

PS4

- Get familiar with Jeppes dst-module
- The concepts are the same as last time!
- Split-apply-combine

UNIVERSITY OF
COPENHAGEN

Split-apply-combine

It is basic data analysis formalized:

- Split data into sub-groups
- Apply some function for all groups
- Combine with your data

Split-apply-combine example

Say that you income data for each municipality for each year retrieved using Jeppe's module and you want to see how income is fluctuating around the mean

- Group by municipality
- Apply a function that calculates the mean income across years e.g. a lambda function or just do it the easy way by column operation
- Add this information to your dataset

Complementary to lectures

PS4

Let's go!