

PROJECT 1: DATA ANALYSIS PROJECT

- **Vision:** *Programming is more than writing code.* The ultimate goal of the projects in this course is that you learn to structure a programming project from start to finish. Your answers should of course be correct, but your code should also be well-structured and documented.
- **Objectives:** In your data analysis project, you should show that you can
 1. Download online data using an API
 2. Apply data cleaning and data structuring methods
 3. Apply data analysis methods
 4. Make the figures nice to look at and easy to understand
 5. Learn new functionalities in a known package on your own
 6. Structure a code project
 7. Document code
 8. Present results in text form and in figures
- **Structure:** Your data analysis project should consist of
 1. A single self-contained notebook (.ipynb) presenting the analysis
 2. Documented Python module files (.py) (if any)

Example of structure: See this [repository](#) and
ProgEcon-lectures/06_YourFirstProject/01_Structure
- **Hand-in:** Upload to GitHub in your repository
`github.com/NumEconCopenhagen/projects-2025-GROUPNAME/dataproject/`
- **Deadline:** See [Calendar](#).
- **Feedback:** Your TA will provide feedback on your project.
- **Exam:** Your data analysis project will be a part of your exam portfolio. You can incorporate feedback before handing in the final version.
This first version does not need to be perfect!

1 Aggregate inflation in Denmark

1.1 Basics

Using table PRIS113 from Denmark Statistics produce figures with, respectively,

1. The consumer price index (CPI), P_t , indexed to 100 on average in 2020
2. The month-to-month inflation rate, $\pi_t = P_t/P_{t-1} - 1$
3. The 12-month inflation rate, $\pi_t^{12} = P_t/P_{t-12} - 1$

When did the post-pandemic inflation surge end in Denmark?

Hint I: You can use the pandas method `.pct_change`.

Hint II: You can use the function `Y = pd.to_datetime(X, format='%YM%m')` to create a date-variable, and then e.g. get the year as `Y.dt.year`

1.2 Instantaneous inflation

The *instantaneous inflation rate* is defined as

$$\pi_t^{12,\alpha} = \left(\prod_{k=0}^{11} (1 + \pi_{t-k})^{\kappa(k,\alpha)} \right) - 1$$
$$\kappa(k,\alpha) = \frac{(T-k)^\alpha}{\sum (T-k)^\alpha} T$$

When $\alpha = 0$ this is just the 12-month inflation rate. For higher α more weight is placed on inflation in the recent months.

1. Plot $\kappa(k,\alpha)$ for $k \in \{0, 1, 2, \dots, 11\}$ and $\alpha \in \{0, 1, 2, 3\}$
2. Compute $\pi_t^{12,\alpha}$ for $\alpha \in \{0, 1, 2, 3\}$ (e.g. using the `.rolling(12).apply` method)
3. Plot $\pi_t^{12,\alpha}$ for $\alpha \in \{0, 1, 2, 3\}$ since 2019

Can this be used to say something new about when the post-pandemic inflation surge ended in Denmark compared to your analysis above?

1.3 Core inflation

Using table PRIS111 from Denmark Statistics produce a figure with 12-month inflation rates since 2019 for the following series:

- VAREGR = 000000: Consumer price index, total"
- VAREGR = 151000: ... excl. energy
- VAREGR = 141000: ... excl. energy and unprocessed food (core inflation).

What does this tell us about the most likely drivers of the post-pandemic inflation surge in Denmark?

1.4 Disaggregated inflation in Denmark

Using table PRIS113 from Denmark Statistics do the following:

1. Create a list of all 4-digit product categories (e.g. 01.1.1.1 Rice)
Hint: You can start from the following code

```
levels = DstApi('PRIS111').variable_levels('VAREGR', language='en' )  
for row in levels.itertuples(): print(row.id, row.text)
```
2. Compute the 12-month inflation rate for the 4-digit product categories using `.groupby` and compute the 25th, 50th and 75th percentile for each month in the data using `.agg`. Plot and describe what you see.
3. Plot the histogram of the percent change in prices for 4-digit product categories from August 2020 to August 2025. Describe what you see.
4. Illustrate the top 10 and bottom 10 product categories.

2 International Comparison

2.1 CPI vs HICP

Compare the CPI from Denmark Statistics with the Harmonized Index for Consumer Prices (HICP) from FRED code CP0000DKM086NEST (All Items). Are the two indexes comparable?

2.2 Across countries

Using FRED do the following for Denmark, Austria, the Euro Area and the United States since 2019:

1. Compare HICP (All Items) across countries
2. Compute and compare the 12-month HICP inflation rate across countries
3. Print the min., max., and mean 12-month inflation rate year-by-year and across countries

Summarize the cross-country differences in the path of inflation since 2019.

3 Extension

Extend the above inflation analysis with some new data of your choice and 1-3 new figures illustrating an economic insight.

Note: *This is intended as an open question, and we encourage you to be creative.*