

Effective Programming Practices for Economists

Reproducible Research

Paths handling in projects

Janoś Gabler and Hans-Martin von Gaudecker

Recap: Path goals

- Start paths relative to the root folder of the project
- Only make assumptions about directory structure inside the project
- Define the path in a way that is portable across operating systems

We showed you how to do this in the pathlib screencast, but in the project templates there is a simpler way

How to use it

assume your project slug is `my_project``

```
from my_project.config import SRC, BLD

def task_clean_data(
    raw_data=SRC / "original_data" / "data.csv",
    produces=BLD / "data" / "clean_data.pkl",
):
    ...
```

- ``SRC``, ``BLD`` and some other paths are defined inside ``config.py``
- Can be imported from anywhere in the project
- Only works inside the project environment

How it is implemented

in config.py

```
from pathlib import Path

SRC = Path(__file__).parent.resolve()
BLD = SRC.joinpath("../", "..", "bld").resolve()

TEST_DIR = SRC.joinpath("../", "..", "tests").resolve()
PAPER_DIR = SRC.joinpath("../", "..", "paper").resolve()
```

- You can define more shorthands if you need them
- Makes it easy to rename entire folders in one central place