Effective Programming Practices for Economists

Data management with pandas

Functional data management: Fundamental rules

Janoś Gabler and Hans-Martin von Gaudecker

Example

```python
>>> def clean(df):
...     out = pd.DataFrame()
...     out["a"] = convert_to_ordered_cat(df["Q001"]
...     out["b"] = convert_to_ordered_cat(df["Q002"]
...     out["c"] = _clean_c(df["Q003"], out["b"])
...     return out
```

Rules

- Start with an empty DataFrame for the clean data

- Touch every variable just once