

Effective Programming Practices for Economists

Data management with pandas

Functional data management

Janoš Gabler and Hans-Martin von Gaudecker

Example

consider this hypothetical survey about a programming course

```
>>> raw = pd.read_csv("survey.csv")
>>> raw
```

| Q001 | Q002 | Q003 |
|---------------------|----------------|--------|
| 0 strongly disagree | agree | python |
| 1 strongly agree | strongly agree | Python |
| 2 -77 | disagree | R |
| 3 agree | -77 | Python |
| 4 -99 | -99 | Python |
| 5 nan | strongly agree | Python |
| 6 neutral | strongly agree | Python |
| 7 disagree | agree | python |
| 8 strongly disagree | -99 | PYTHON |
| 9 -77 | -99 | Ypthon |

From the metadata you know

- Q001: I am a coding genius
- Q001: I learned a lot
- Q003: What is your favourite language
- -77 not readable
- -99 no reply