

Tidy Data

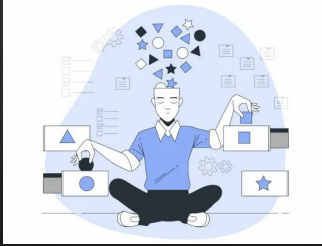
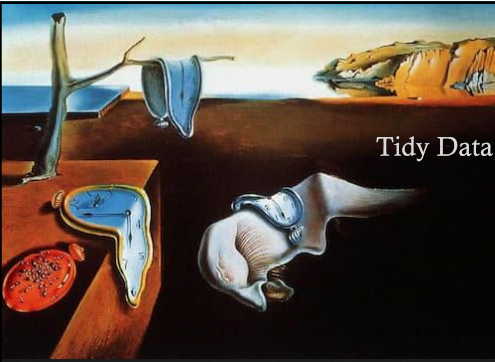
Working with tables

Prequel to the "Machine Learning Course"

Juanna Pastor

How do I organize my data?

Handling data with Python

Tidy Data


SAVE TIME

Journal of Statistical Software
November 2007, Volume 21, Issue 12. <http://www.jstatsoft.org/>

Reshaping Data with the reshape Package

Hadley Wickham


Iowa State University



Tidy Data

"Happy families are all alike; every unhappy family is unhappy in its way"

Leo Tolstoy



Messy Data

- ◆ Values in headers
- ◆ Multiple variables in column
- ◆ Variable in row and column
- ◆ Multiple types observations
- ◆ Same unit in multiple tables

Our course

Table handling with pandas

Data types

Data operations

And more...

