

## COMP 333 — Week 4 Tidy Data

### Tidy Data

Tidy Data describes well-structured data arrays

that follow three rules for structuring observations and their features:

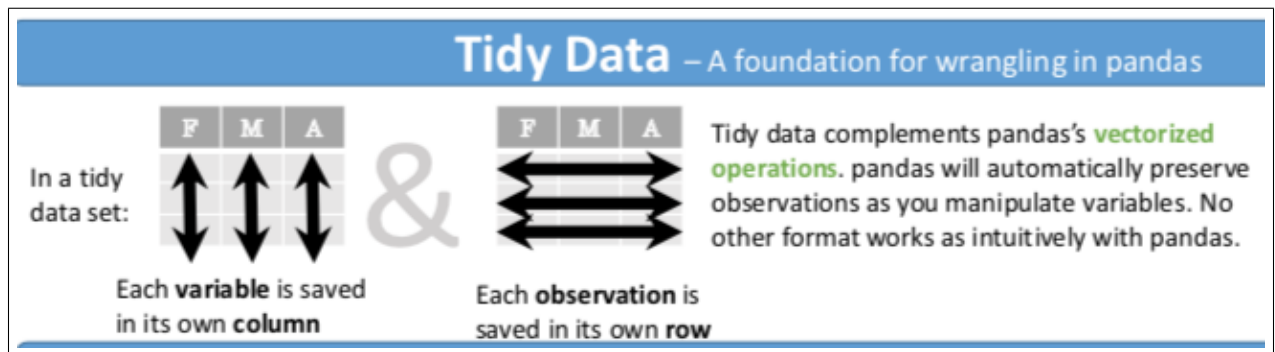
1. Each variable forms a column.
2. Each observation forms a row.
3. Each type of observational unit forms a table.

The concept of tidy data came from the R community.

They name some R libraries as the “*tidyverse*”.

See the original paper <https://vita.had.co.nz/papers/tidy-data.pdf>

pandas operations support the transformation of data into dataframes that are tidy data.



Tidy Data should be your aim when structuring data in Data Wrangling.