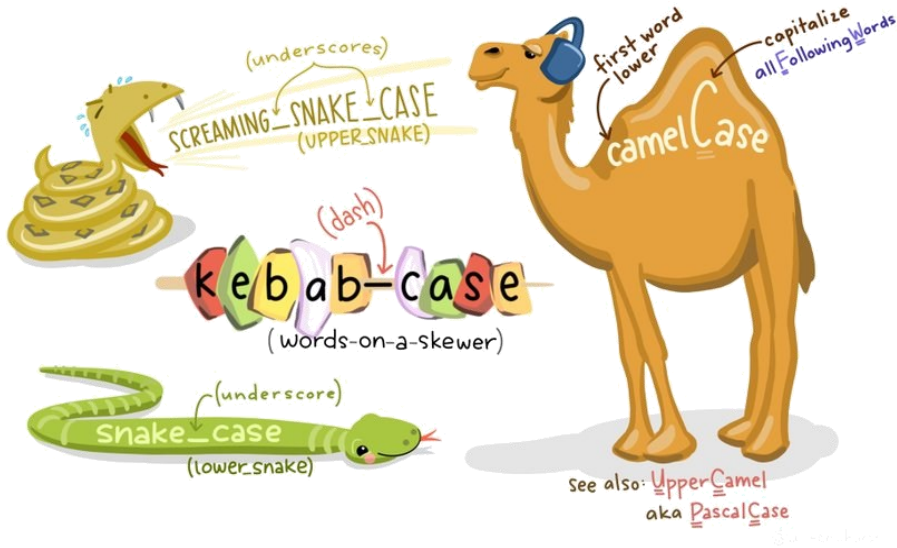


# Cheat Sheet

## How to structure data?

### File naming



Artwork by [@allison\\_horst](#) licensed under CC-BY 4.0

#### Do:

- ✓ Human and machine readable
- ✓ Content-specific and descriptive
- ✓ Keep naming consistent
- ✓ Date files
- ✓ Pad numbers with leading zeros (`_v001`)
- ✓ Document your conventions

#### Don't:

- ✗ No special characters (Spaces, `{ } [ ] < > ( ) * % # ' , " ? ! & @ $ ~ . ä`)
- ✗ Cryptic names
- ✗ Too long names

Date      No spaces      Reference to project, experiment, ...      If needed: File owner      If needed: Version

2 0 2 3 - 1 1 - 1 4 \_ H C N 4 \_ H E K 2 9 3 \_ I V 6 0 s \_ F N - 0 0 1 . c s v

Based on artwork by fdm.nrw licensed under CC-BY 4.0

## Documentation

#### Should include:

- Author information and contact points
- Project description
  - Goals
  - Hypothesis
- Definition of acronyms and abbreviations
- Information on how the data was created (Methods, Time and Place)
- Limitations
- Data-specific information
  - Data structure (file overview) and relation
  - Taken measures to clean data
  - Explanation of variables, labels and code
  - Differences between dataset version
- Information on access and reuse (license)
- Acknowledgments

