

Principles of Data Management

- **Well documented data are easier to understand.**
 - Share your data. This allows:
 - results of existing analyses to be reproduced and improved upon;
 - data to be combined in meta-analyses to reach general conclusions;
 - new approaches to be applied to the data and new questions asked using it; and
 - approaches to inquiry that could not be considered without broad scale data sharing.
 - Data management plans are required by many funding agencies today.
 - Provide metadata.
 - Provide an unprocessed form of the data.
 - Establish and use a simple and consistent data and folder structure across your projects.
- **Properly formatted data are easier to use in a variety of software.**
 - Use standard data formats.
 - Use standard file formats.
 - Use standard table formats.
 - Use standard formats within cells.
 - Use good null values.
 - Blank v. 0 v. "NA" or "NULL"
 - Make it easy to combine your data with other datasets.
 - As a corollary, keep your code and your data separate.
 - Perform basic quality control, including basic sanity checks:
 - If a column should contain numeric values, check for non-numeric values in the data.
 - Check that empty cells actually represent missing data, and not mistakes in data entry, and indicate that they are empty using the appropriate null values.
 - Check for consistency in unit of measurement, data type (e.g., numeric, character), naming scheme (e.g., taxonomy, location), etc.
- **Data that are shared in established repositories with open licenses is easier for others to find and use.**
 - Use an established repository.
 - Use an established and open license.
 - Using well established licenses (or waivers) in order to clearly communicate the rights and responsibilities of both the people providing the data and the people using it.
 - Using the most open license (or waiver) possible, because even minor restrictions on data use can have unintended consequences for the reuse of the data (Schofield et al. 2009, Poisot et al. 2013).
 - Make it clear how to cite your data appropriately (README, INSTALL, CITATION, LICENSE files).

Adapted and expanded by Neal Davis from Ethan P. White, *et al.* Nine simple ways to make it easier to (re)use your data. *Ideas in Ecology and Evolution* 6(2): 1–10, 2013.

Data License Suggestions

- [GNU General Public License](#) (GPL)
- [MIT license](#)
- [BSD license](#)
- [Creative Commons](#) CC0, CC-BY, CC-BY-ND, CC-BY-NC, CC-BY-ND-SA, CC-BY-NC-SA

GitHub Features

- Git/Mercurial Repository Hosting
- GitHub Pages
- Wiki
- Issue Tracking
- Pull Requests Collaboration Model

git Quick Reference

init	create a new repository.
status	see what you are in the middle of doing.
add	manage the index file.
commit	advance the current branch.
log	see what happened.
reset	undo changes
revert	undo a commit
checkout	(with parameters) undo changes.
clone	prime local repo
pull	keep up-to-date with upstream
push	send changes upstream

Markdown Quick Reference

# Header 1	I have more [^1] to say up here.
## Header 2	
### Header 3	[^1]: To say down here.
#### Header 4	
##### Header 5	- Item
##### Header 6	- Item
Emphasize _emphasize_	1. Item
Strong __Strong__	2. Item
A [link](http://example.com "Title").	`This is code`
Some text with [a link][1] and another [link][2].	```python for i in range(x): print i ... ---
[1]: http://example.com/ "Title" [2]: http://example.org/ "Title"	
Logo: ![Alt](/wp.png "Title")	<blockquote>All HTML works too.</blockquote>
Linked logo: [![alt text](/wp-smaller.png)] (http://wordpress.com/ "Title")	Adapted from WordPress documentation and John Gruber's Daring Fireball .