Reproducible research

Andreas Bjerre-Nielsen

Git: Version control

Why version control?

> *Your closest collaborator is you six months ago, but you don't reply to emails.*

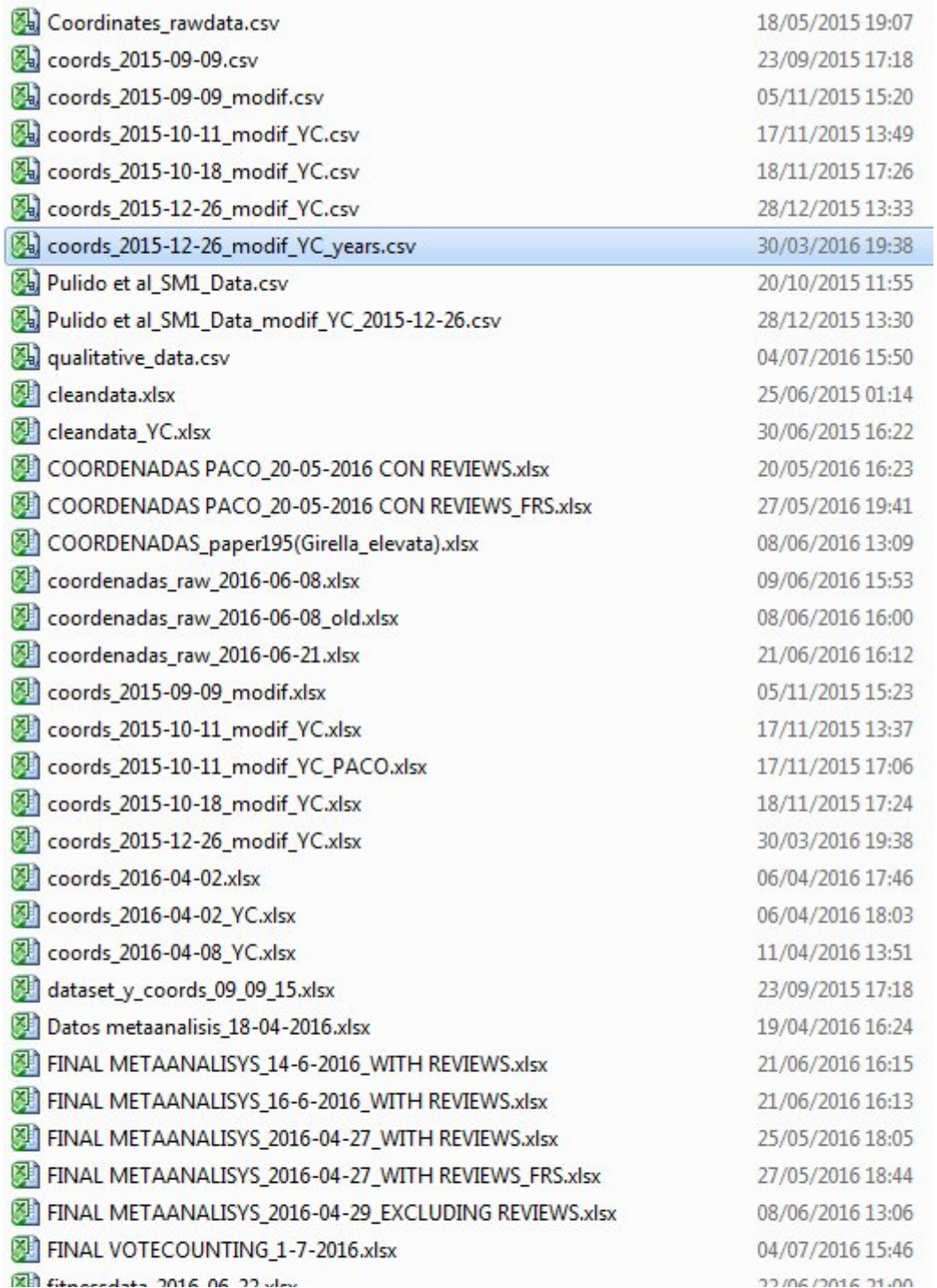
Detailed log of all changes.

Easy to revert back to previous versions (remembers forever).

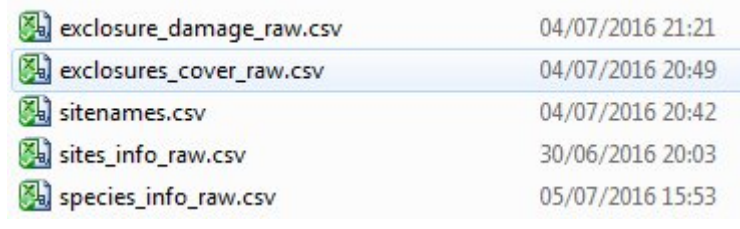
Clear attribution of work (who contributed what).

Other differences from DropBox/Google Drive etc.? Some files are shared, some not.

.. no version control


Coordinates_rawdata.csv 18/05/2015 19:07
coords_2015-09-09.csv 23/09/2015 17:18
coords_2015-09-09_modif.csv 05/11/2015 15:20
coords_2015-10-11_modif_YC.csv 17/11/2015 13:49
coords_2015-10-18_modif_YC.csv 18/11/2015 17:26
coords_2015-12-26_modif_YC.csv 28/12/2015 13:33
coords_2015-12-26_modif_YC_years.csv 30/03/2016 19:38
Pulido et al_SM1_Data.csv 20/10/2015 11:55
Pulido et al_SM1_Data_modif_YC_2015-12-26.csv 28/12/2015 13:30
qualitative_data.csv 04/07/2016 15:50
cleandata.xlsx 25/06/2015 01:14
cleandata_YC.xlsx 30/06/2015 16:22
COORDENADAS PACO_20-05-2016 CON REVIEWS.xlsx 20/05/2016 16:23
COORDENADAS PACO_20-05-2016 CON REVIEWS_FRS.xlsx 27/05/2016 19:41
COORDENADAS_paper195(Girella_elevata).xlsx 08/06/2016 13:09
coordenadas_raw_2016-06-08.xlsx 09/06/2016 15:53
coordenadas_raw_2016-06-08_old.xlsx 08/06/2016 16:00
coordenadas_raw_2016-06-21.xlsx 21/06/2016 16:12
coords_2015-09-09_modif.xlsx 05/11/2015 15:23
coords_2015-10-11_modif_YC.xlsx 17/11/2015 13:37
coords_2015-10-11_modif_YC_PACO.xlsx 17/11/2015 17:06
coords_2015-10-18_modif_YC.xlsx 18/11/2015 17:24
coords_2015-12-26_modif_YC.xlsx 30/03/2016 19:38
coords_2016-04-02.xlsx 06/04/2016 17:46
coords_2016-04-02_YC.xlsx 06/04/2016 18:03
coords_2016-04-08_YC.xlsx 11/04/2016 13:51
dataset_y_coords_09_09_15.xlsx 23/09/2015 17:18
Datos metaanalisis_18-04-2016.xlsx 19/04/2016 16:24
FINAL METAANALISYS_14-6-2016_WITH REVIEWS.xlsx 21/06/2016 16:15
FINAL METAANALISYS_16-6-2016_WITH REVIEWS.xlsx 21/06/2016 16:13
FINAL METAANALISYS_2016-04-27_WITH REVIEWS.xlsx 25/05/2016 18:05
FINAL METAANALISYS_2016-04-27_WITH REVIEWS_FRS.xlsx 27/05/2016 18:44
FINAL METAANALISYS_2016-04-29_EXCLUDING REVIEWS.xlsx 08/06/2016 13:06
FINAL VOTECOUNTING_1-7-2016.xlsx 04/07/2016 15:46

version control



What is version control

what is `git`? Git is an open source command line program for version control.

what is `github` / `gitlab`? Companies/web services that hosts Git repositories and enables 'social coding'

What is GitHub for Mac/Windows? A GUI for Git.

Makes it easier to use.

Ultimately just does command line Git.

Getting started

Key terms (local):

- Repository (repo): a directory where Git looks for changes
- Initialize (init): have Git begin watching a directory
- add: stage a file so that Git starts watching it
- master: the main branch. By convention this should be the most stable version.
- push: commit changes to a remotely hosted repository
- pull: merge changes from a remotely hosted repository

Getting started (2)

About providers:

- `GUI GitHub` push and pull are combined into sync
- `gitlab.com` provides free private repositories

Example

Pushing a change to the SDS repo

Example (2)

Cloning the SDS repo

Markdown

Why use Markdown?

Easy to learn and use.

Focus on **content**, rather than **coding** and debugging **errors**.

It's flexible. Markdown was created to simplify HTML, but with the right tools, your Markdown files can easily be converted to many different formats!

What is Markdown?

Markdown is a particular type of markup language.

Markup languages are designed produce documents from plain text.

Some of you may be familiar with *LaTeX*. This is another (less human friendly) markup language for creating pdf documents.

LaTeX gives you much greater control, but it is restricted to pdf and has a **much** greater learning curve.

markdown was created for the web (you know it if you use Github, Stackoverflow, etc.)

Example

Suppose we want to create a nested list

- fruits
  - apples
    - macintosh
  - pears
  - peaches
- vegetables
  - chard

Latex

```
\begin{itemize}
\item fruits
\begin{itemize}
\item apples
  \begin{itemize}
  \item macintosh
  \end{itemize}
\item pears
\item peaches
\end{itemize}
\item vegetables
  \begin{itemize}
    \item chard
  \end{itemize}
\end{itemize}
```

HTML

```
<ul>
        <li>fruits
        <ul>
            <li>apples
            <ul>
                <li>macintosh</li>
            </ul></li>
            <li>pears</li>
            <li>peaches</li>
        </ul></li>
        <li>vegetables
        <ul>
            <li>brocolli</li>
        </ul></li>
    </ul>
```

Markdown

```
* fruits
    - apples
        - macintosh
    - pears
    - peaches
* vegetables
    - chard
```

Jupyter and Markdown

We can use markdown within Jupyter. In command mode pressing `m` will convert your cell to a markdown cell, y will convert it back to code.

TRY THIS!

Markdown commands

- A heading can be created with # and for smaller sizes ##, ### ..
- A link can is `[LINK_TEXT](URL)`
    - A link can also be picture no link text is provided
- Italicized is `*[ITAL_TEXT]*`
- Bold is `**[BOLD_TEXT]**`

See more in Markdown tutorial.