# Social Data Science

## Data Science Tools

Sebastian Barfort

August 12, 2016

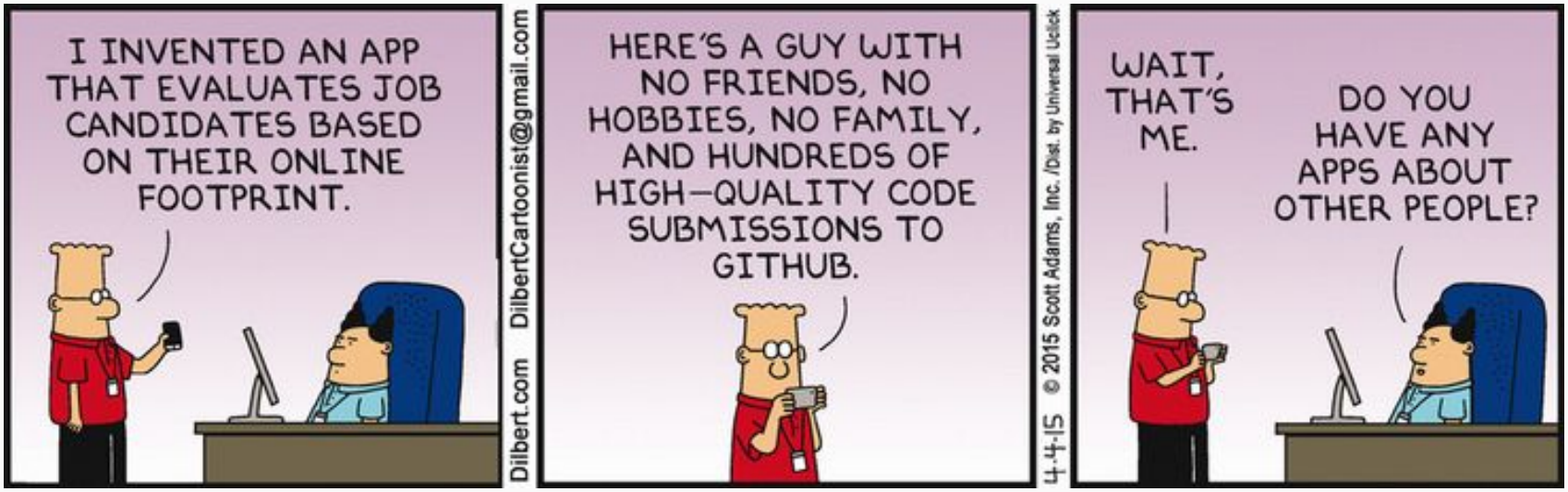
University of Copenhagen
Department of Economics

# TODAY

Git & Github

markdown & rmarkdown

Shiny


I INVENTED AN APP THAT EVALUATES JOB CANDIDATES BASED ON THEIR ONLINE FOOTPRINT.
Dilbert.com DilbertCartoonist@gmail.com
HERE'S A GUY WITH NO FRIENDS, NO HOBBIES, NO FAMILY, AND HUNDREDS OF HIGH-QUALITY CODE SUBMISSIONS TO GITHUB.
4-4-15 © 2015 Scott Adams, Inc. /Dist. by Universal Uclick
WAIT, THAT'S ME.
DO YOU HAVE ANY APPS ABOUT OTHER PEOPLE?

# Git & Github

# WHY VERSION CONTROL?

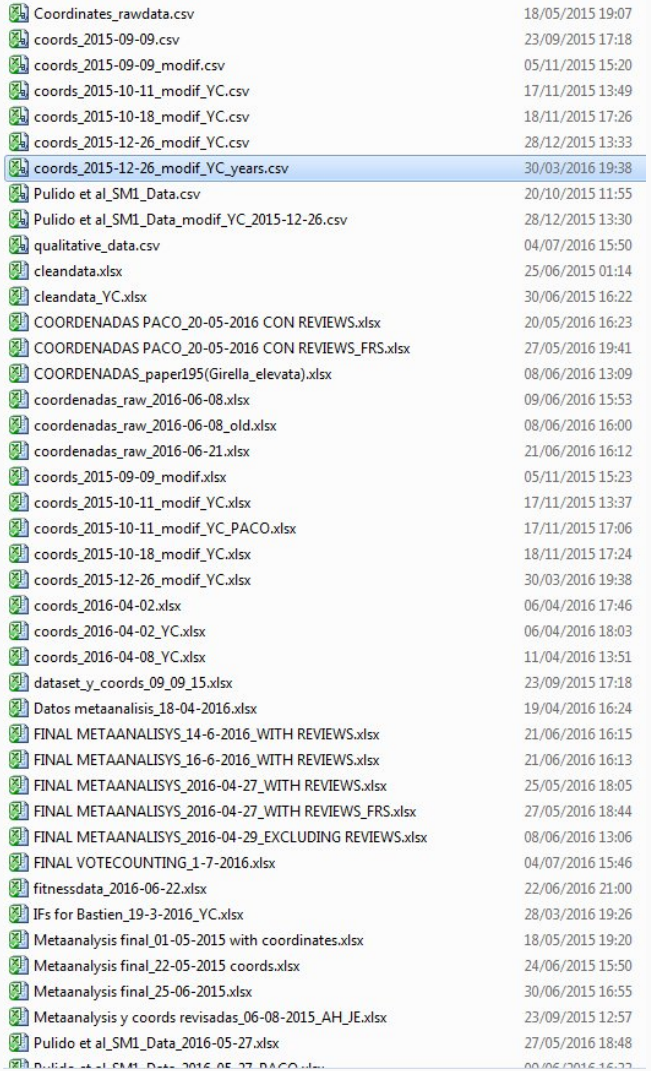
> *Your closest collaborator is you six months ago, but you don't reply to emails.*

Detailed log of all changes.

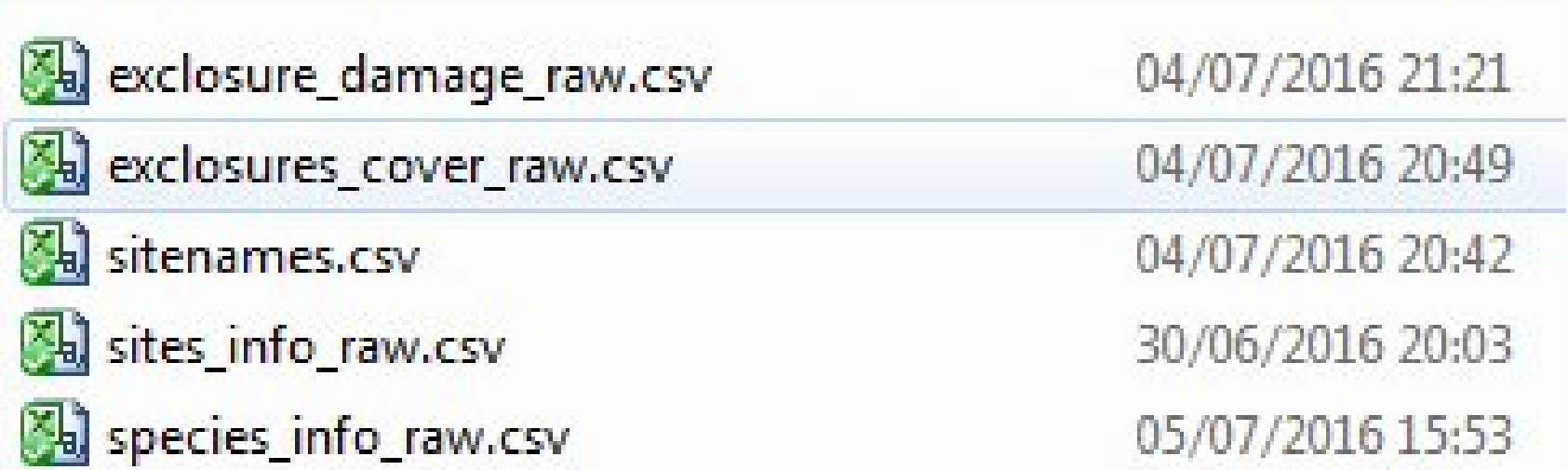
Easy to revert back to previous versions.

Clear attribution of work (who contributed what).

# Without version control



## With version control



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

**what is `github`?** A company/web service 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.

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

In GUI GitHub push and pull are combined into sync

# markdown & rmarkdown

# 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

```
\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}
```

```
<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>
```

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

# WHAT IS rmarkdown?

`rmarkdown` is a new package from RStudio that merges a lot of technologies in an exciting way.

In fact, in RStudio now, you integrate R code into your Markdown document and render as HTML/pdf/Word in the click of a button!

https://vimeo.com/178485416

# Shiny

https://www.rstudio.com/products/shiny/shiny-user-showcase/