Creating Reproducible Reports using R Markdown

Symposium: Using RStudio for Visualization and Analysis of Weed Science Experiments

Maxwel Coura Oliveira, PhD

Department of Agronomy
University of Wisconsin-Madison

December 2019

Outline

- What is literate programming?
- Why is it useful?
- How to use R Markdown to create reproducible reports
  - Text formatting
  - Code chunks
  - Output formats

Literate programming

human readable **text**

+

machine readable **code**

=

Reproducible **document**

Programs as works of literature

- Developed by Donald Knuth, Stanford University
- **Traditional**: telling a computer what to do
- **Novel**: telling a human what you want the computer to do
- Improves documentation and the programs themselves

Literate programming in research

- Tailor reports to an audience

- Repeatable and assures reproducibility

- Works well with version control
- Works well with languages used in research:

R Markdown

- Weave narrative text and code
  - R
  - Python
  - Stata
  - SAS
- Produce documents in many formats
- Reproducible

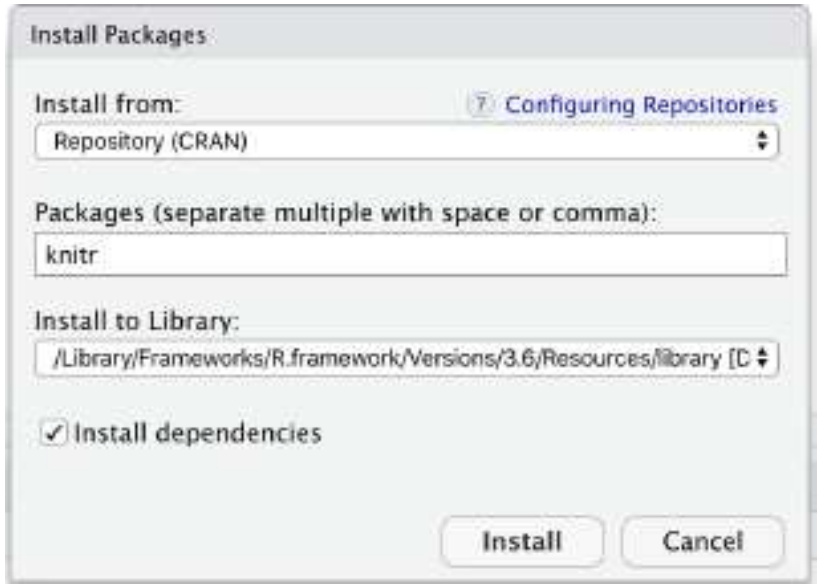
Setup

- Tools > install packages

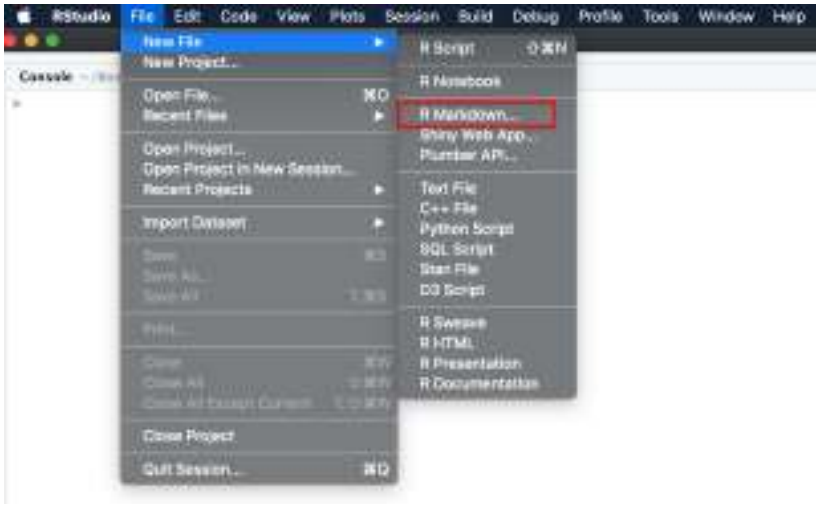
```
install.packages("markdown")
```

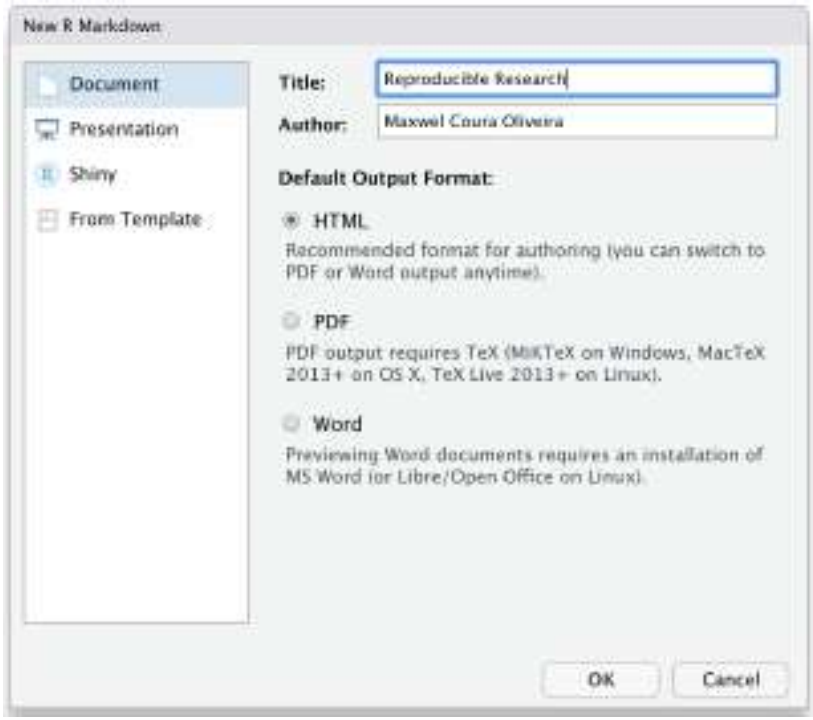
- Install TeX if you want create PDFs

- Open **R** project

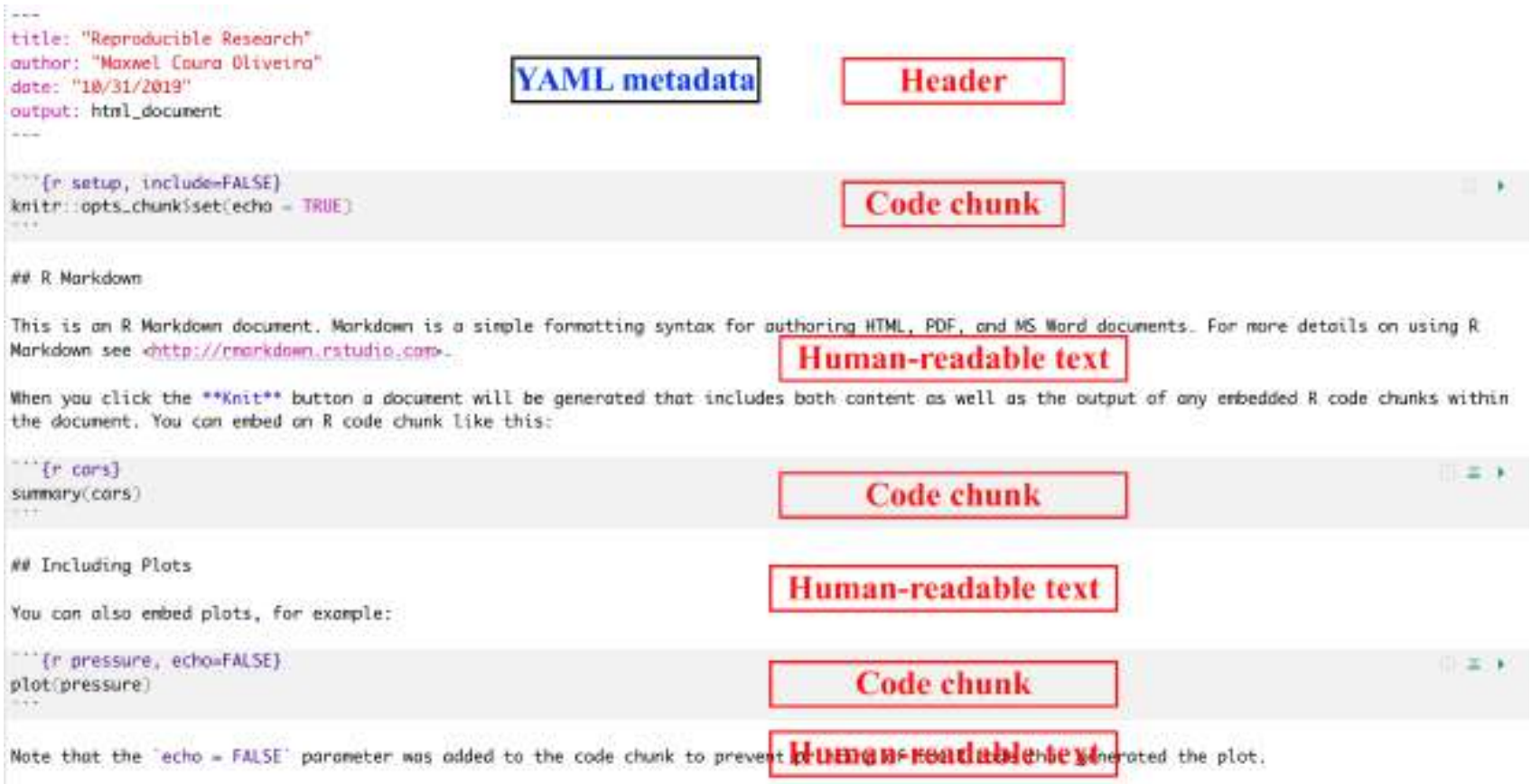


Open/Create a markdown document



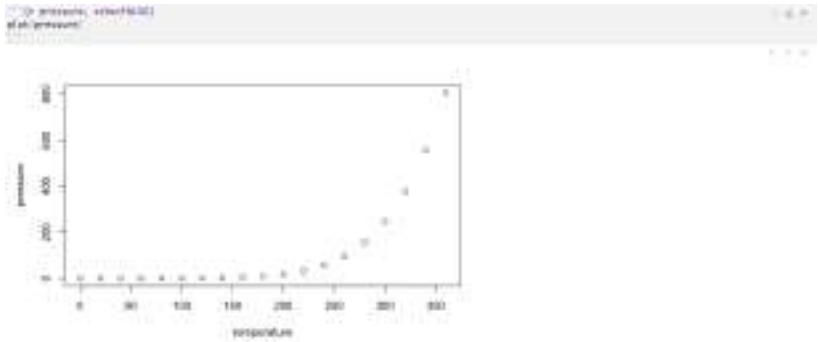


Sections of R Markdown



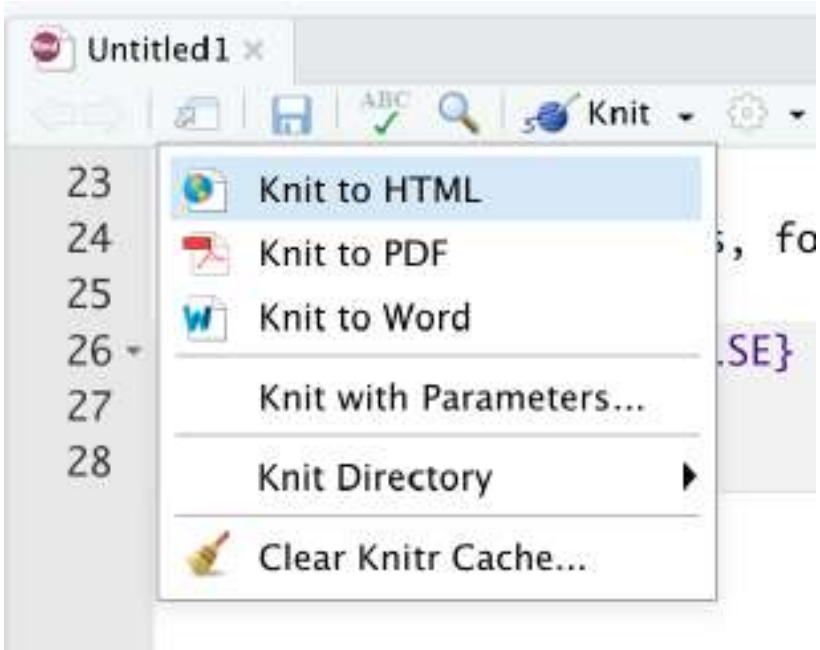
Notebook interface

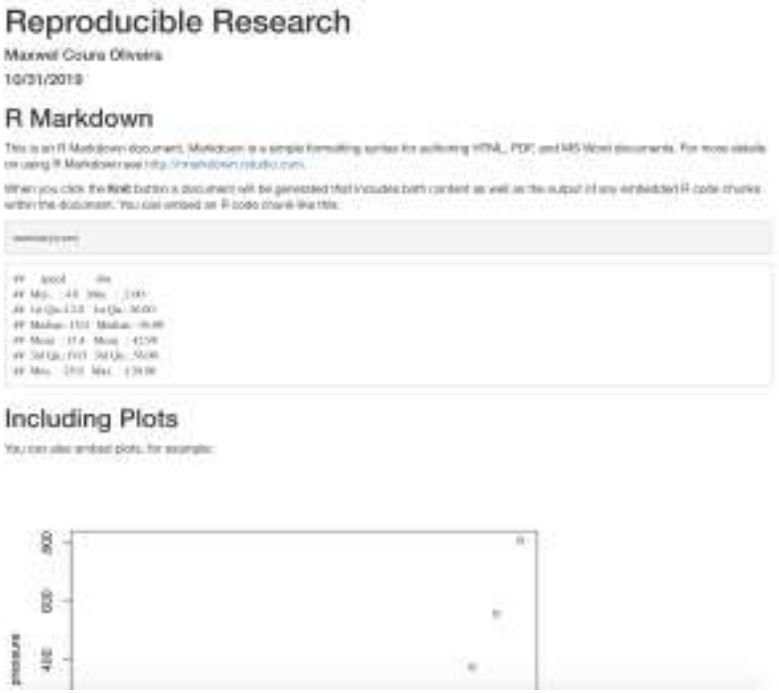
- Press green triangle on a code chunk to run the code and show the output



Knit: from text to document

render(input = "name.Rmd",
output = "html_document")
OR

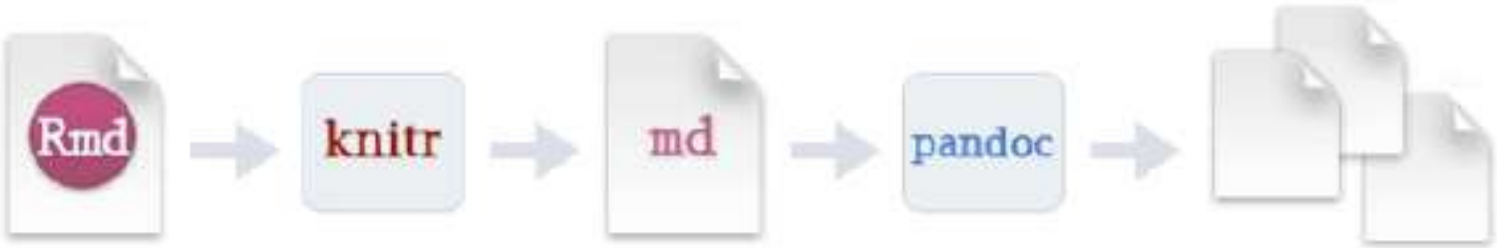




How it works

- knitr package converts the **R Markdown** document to markdown

- Pandoc converts the markdown document to the final

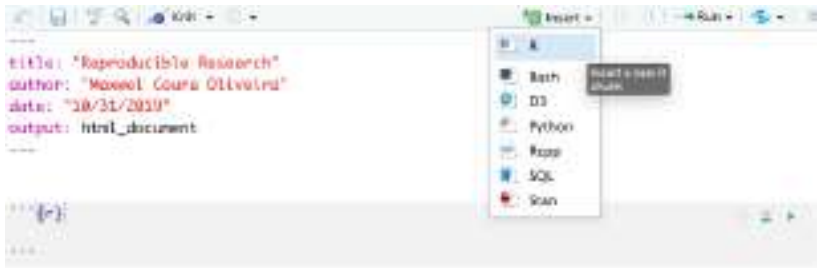


Human readable text

- Outline first, code
- Use markdown rich formatting syntax to
- Cheat sheet: Link

Adding code chunks

- Chunk delimiters
  - beginning: ```{r name}
- end: ```
- Shortkey: Ctrl+Alt+i
- Editor toolbar



Setup code block

```
#load the libraries we need
library(tidyverse)
library(knitr)
library(lattice) #this is for the barley dataset
```

```
data <- barley
```

Code chunk options

- include = FALSE
  - Hide code and results from document
  - results can still be used later
- echo = FALSE
  - hides code, keeps result
- message = FALSE
  - hides messages
- warning = FALSE
  - hides warnings
- fig.cap = " "
  - adds figure captions

Setting options for all code chunks

- Set default options for all code chunks
- knitr::opts_chunk$set() function
- Put it in your first codeblock
- Must be after you load the rmarkdown package

Tables

- Tables look like console output by default
- kable() function makes it pretty

```
nd <- barley %>%
  group_by(site, year) %>%
  summarize(mean = mean(yield))
```

```
kable(nd, digits=1)
```

site	year	mean
Grand Rapids	1932	20.8
Grand Rapids	1931	29.1
Duluth	1932	25.7
Duluth	1931	30.3
University Farm	1932	29.5
University Farm	1931	35.8
Morris	1932	41.5
Morris	1931	29.3
Crookston	1932	31.2
Crookston	1931	43.7
Waseca	1932	41.9
Waseca	1931	54.3

Inline code

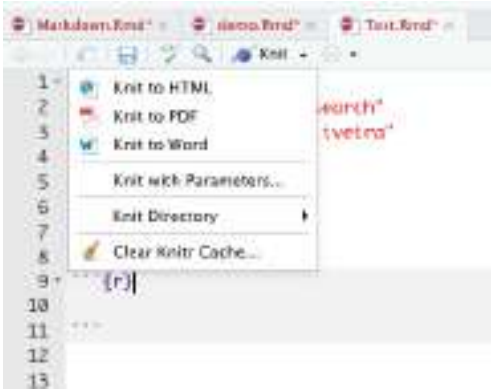
- Syntax `r`
- Include in test for responsive text
- Allows you to write one report and run it over changing datasets
- Example: mean

```
460
461 The overall yield mean is `r mean(data$yield)`
462
```

The overall yield mean is 34.4205553

Change output formats

- select output from knit menu



- Specify output in header

```
---
title: "Demo"
author: "Maxwel Coura Oliveira"
date: "11/1/2019"
output: html_document
---
```

- with render

```
```{r}
library(rmarkdown)
render("demo.Rmd", output_format="html_document")
```
```

Manuscripts

```
library(rticles)
```

Package 'rticles'

November 12, 2019

Type Package

Title Article Formats for R Markdown

Version 0.12

Description A suite of custom R Markdown formats and templates for authoring journal articles and conference submissions.

License GPL-3

Imports utils, rmarkdown, knitr, yaml, tinytex, xfun

SystemRequirements GNU make

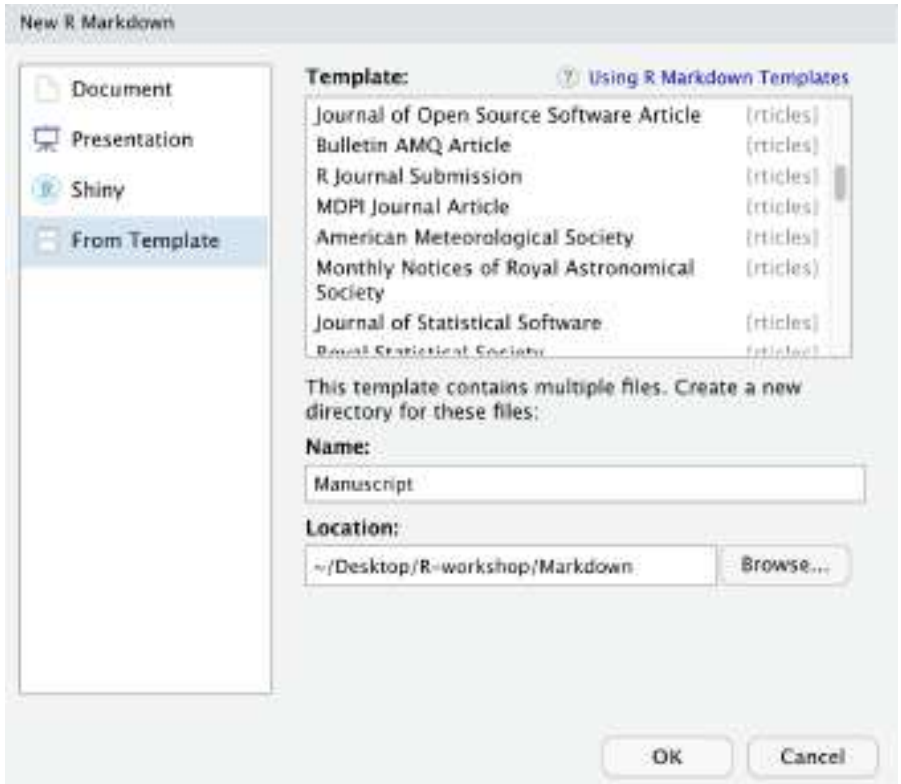
URL https://github.com/rstudio/rticles

Opening a template

- File > New File > R Markdown > From Template > **Template:**

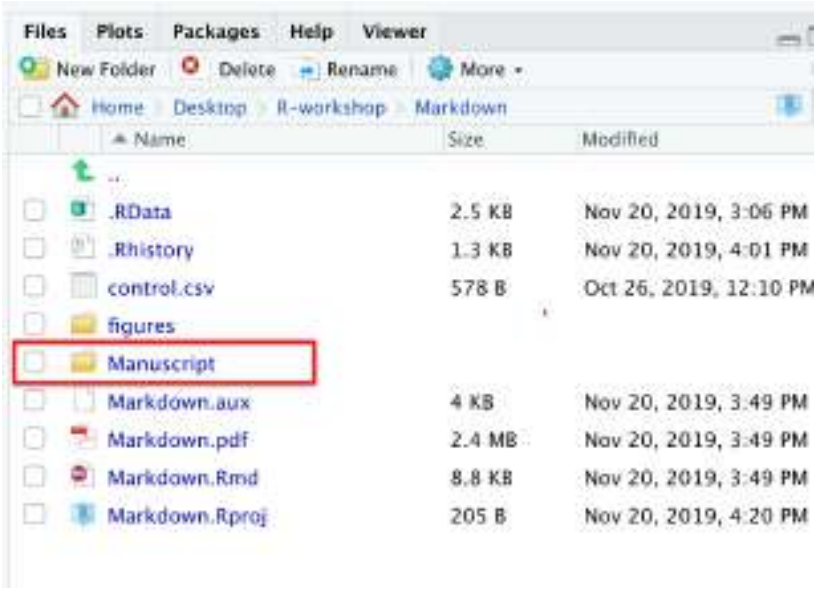
- Choose *PLOS Journal Article*

- Change name to Manuscript

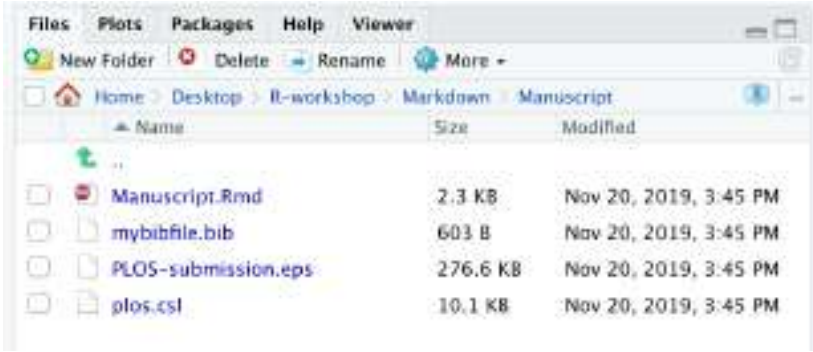


New folders

- It creates a new folder name Manuscript



- Manuscript > Rmd, bib, PLOS-submission.eps and plos.csl



PloS one template

- Rmarkdown template: the manuscript!

```yaml
---
title: Title of submission to PLOS journal
author:
  - name: Alice Anonymous
    email: alice@example.com
    affiliation: Some Institute of Technology
    corresponding: alice@example.com
  - name: Bob Security
    email: bob@example.com
    affiliation: Another University
address:
  - code: Some Institute of Technology
    address: Department, Street, City, State, Zip
  - code: Another University
    address: Department, Street, City, State, Zip
abstract: |
  Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eget porta erat. Morbi consectetur est vel gravida pretium. Suspendisse ut dui eu ante cursus gravida non sed sem. Nullam sapien tellus, commodo id velit id, eleifend volutpat quam. Phasellus mauris velit, dapibus finibus elementum vel, pulvinar non tellus. Nunc pellentesque pretium diam, quis maximus dolor faucibus id. Nunc convallis sodales ante, ut ullamcorper est egestas vitae. Nam sit amet enim ultrices, ultrices elit pulvinar, volutpat risus.

author_summary: |
  Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur eget porta erat. Morbi consectetur est vel gravida pretium. Suspendisse ut dui eu ante cursus gravida non sed sem. Nullam sapien tellus, commodo id velit id, eleifend volutpat quam. Phasellus mauris velit, dapibus finibus elementum vel, pulvinar non tellus. Nunc pellentesque pretium diam, quis maximus dolor faucibus id. Nunc convallis sodales ante, ut ullamcorper est egestas vitae. Nam sit amet enim ultrices, ultrices elit pulvinar, volutpat risus.

bibliography: mybibfile.bib
output: rticles::plos_article
csl: plos.csl
---

_Text based on plos sample manuscript, see [http://journals.plos.org/ploscompbiol/s/latex](http://journals.plos.org/ploscompbiol/s/latex)_
```

YAML metadata

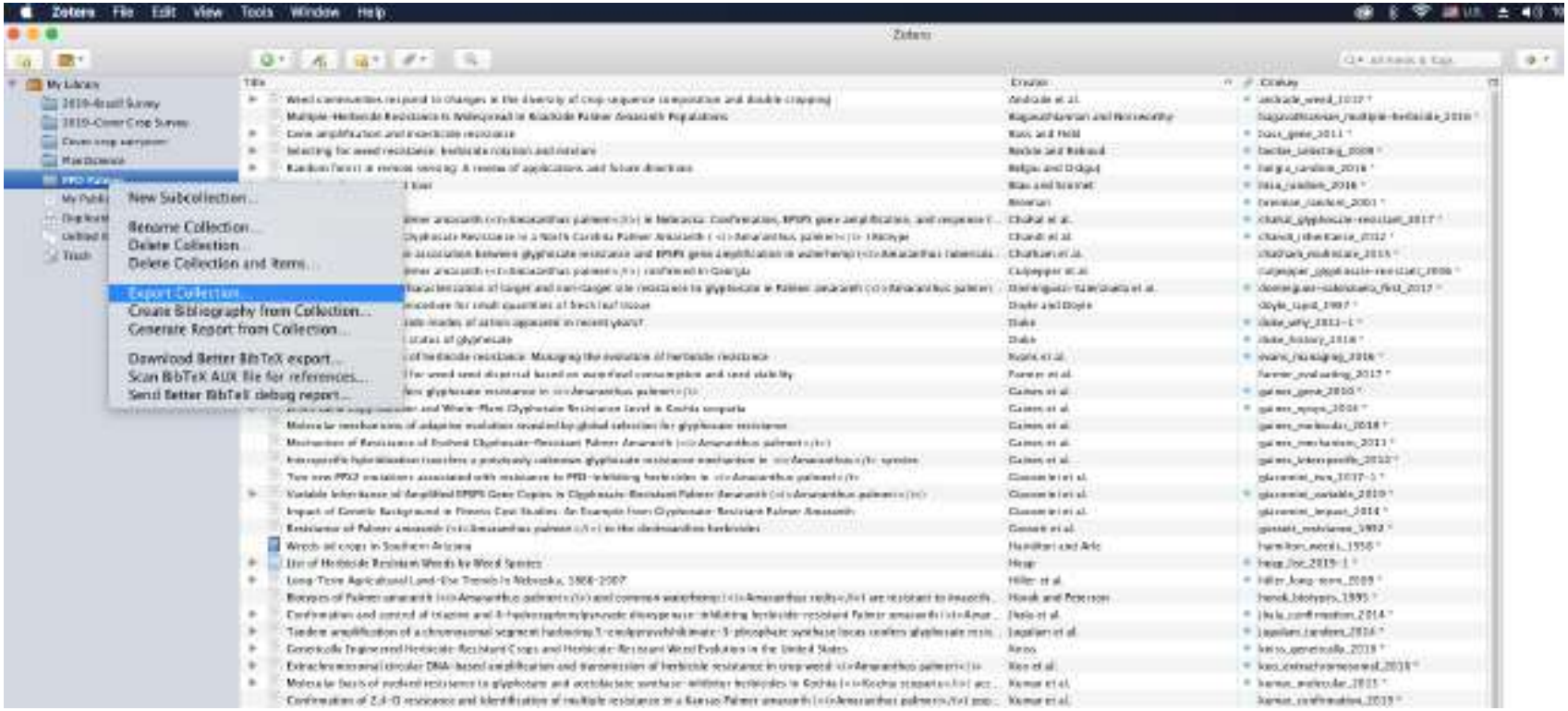
Bibliography

- The word **BibTeX** stands for a tool and a file format which are used to describe and process lists of references, mostly in conjunction with LaTeX documents.

- Zotero is a free, easy-to-use tool to help you collect, organize, cite, and share research.

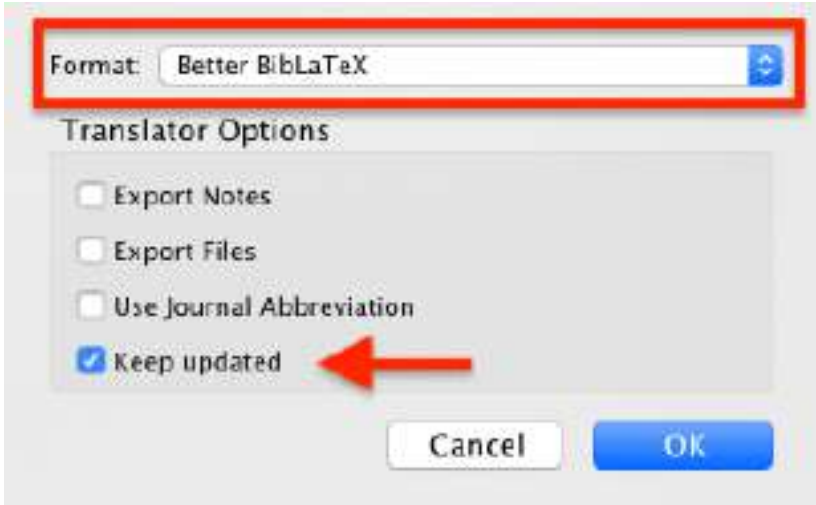
Zotero

- Right click in your citation folder

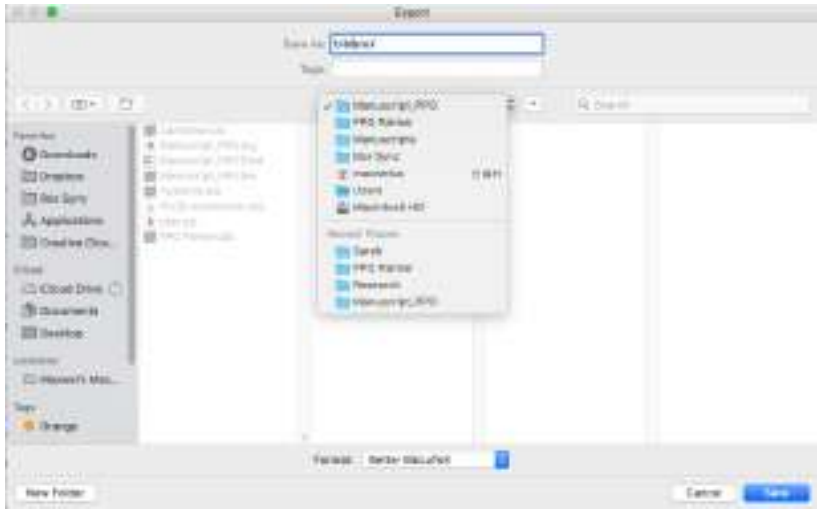


Saving the bibliography

- Save as Better BibLaTeX

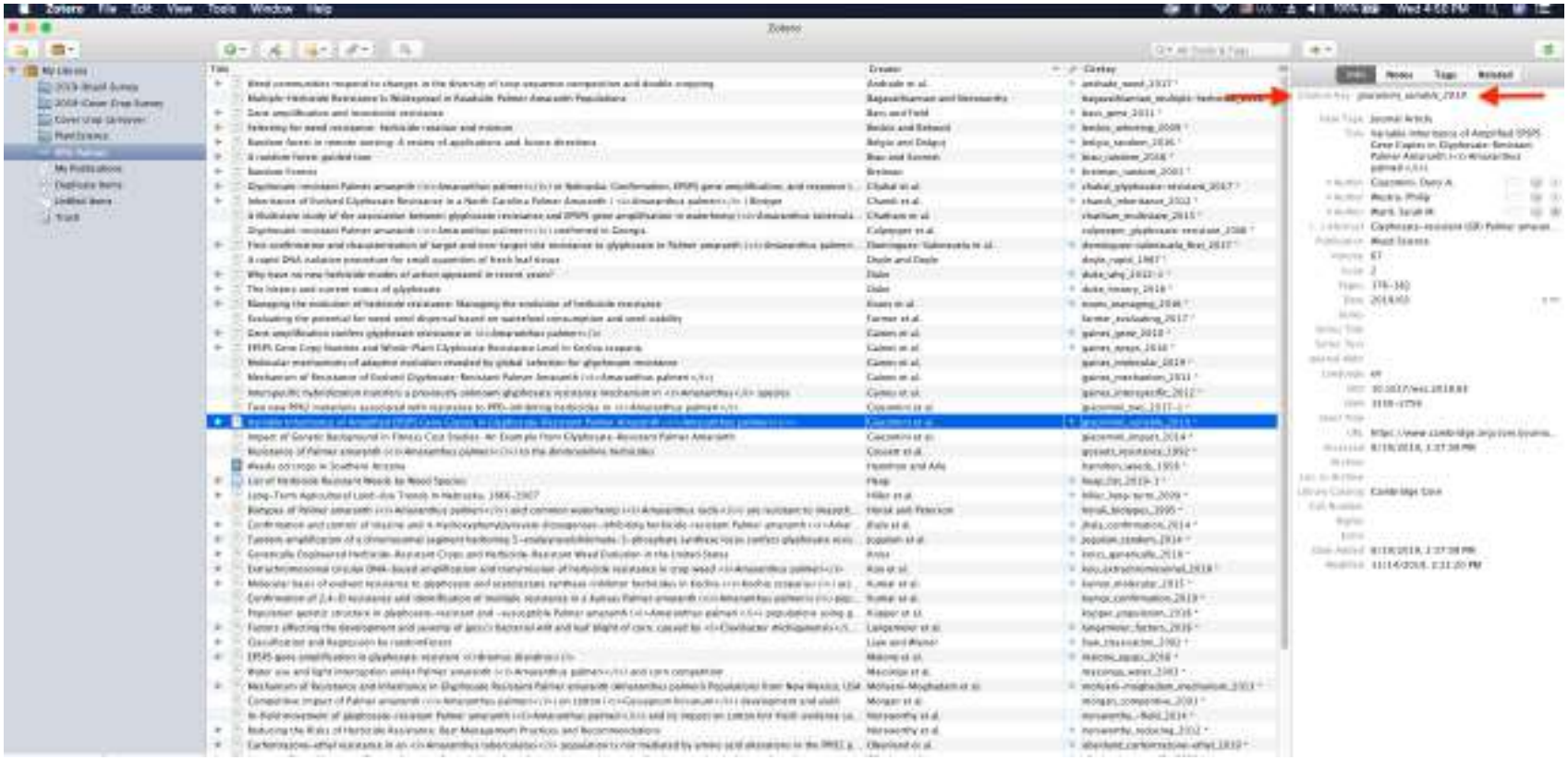


- Save your bib file in the same folder of your manuscript



Citation key

- Copy and paste citation key as: [@giacomini_variable_2019]



Summary

- Literate programming makes reproducible research more machine readable
- R markdown documents facilitate literate programming in RStudio
- R markdown has 3 sections
  - Header: determines output and adds parameters
  - Markdown Text: can include inline code
  - Code chunks: can be customized to mute code or output
- Manuscript templates

Need help?

- Email: max.oliveira@wisc.edu
- R markdown cheatsheet: Link
- R markdown tutorial: Link
- Table generator: Link