Audio Connection and Chat (Zoom)

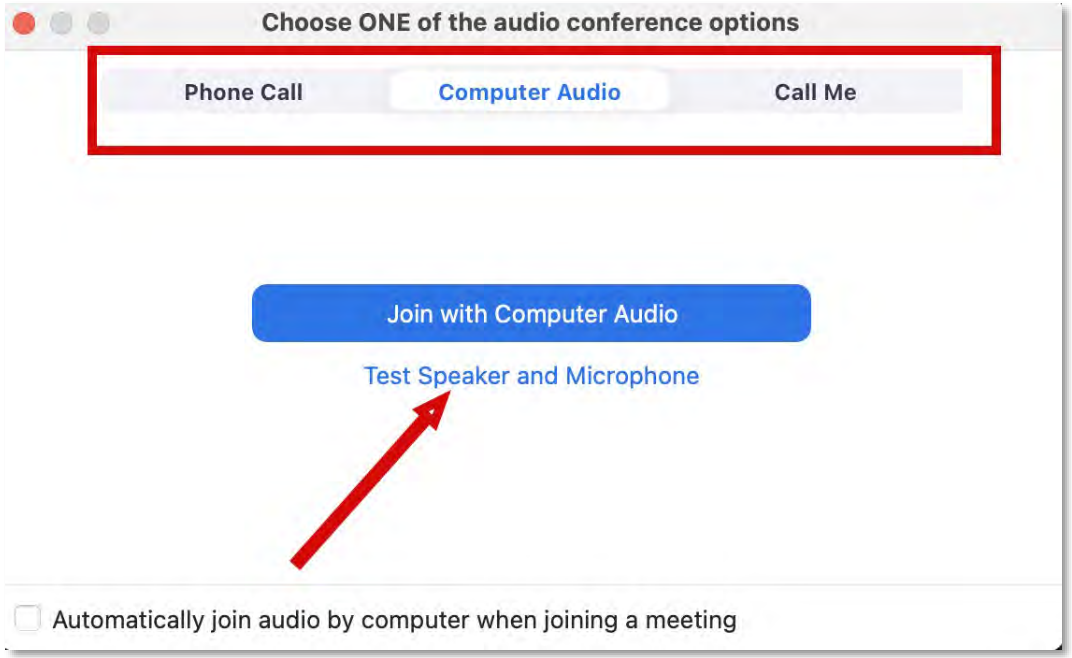
You will not hear any sound until the webinar starts.

Connect Audio

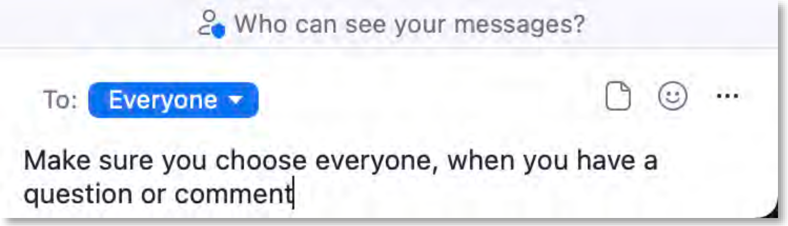
1. When you join Zoom, the ***Join Audio*** preferences box pops-up (Phone Call, Computer Audio, or Call Me)
2. Choose an option that works best for you
3. Join using that option
4. Use Test Speakers and Microphone option to optimize your webinar experience

Chat

Please send your chat to ***Everyone*** to make sure the monitor sees your question


Choose ONE of the audio conference options

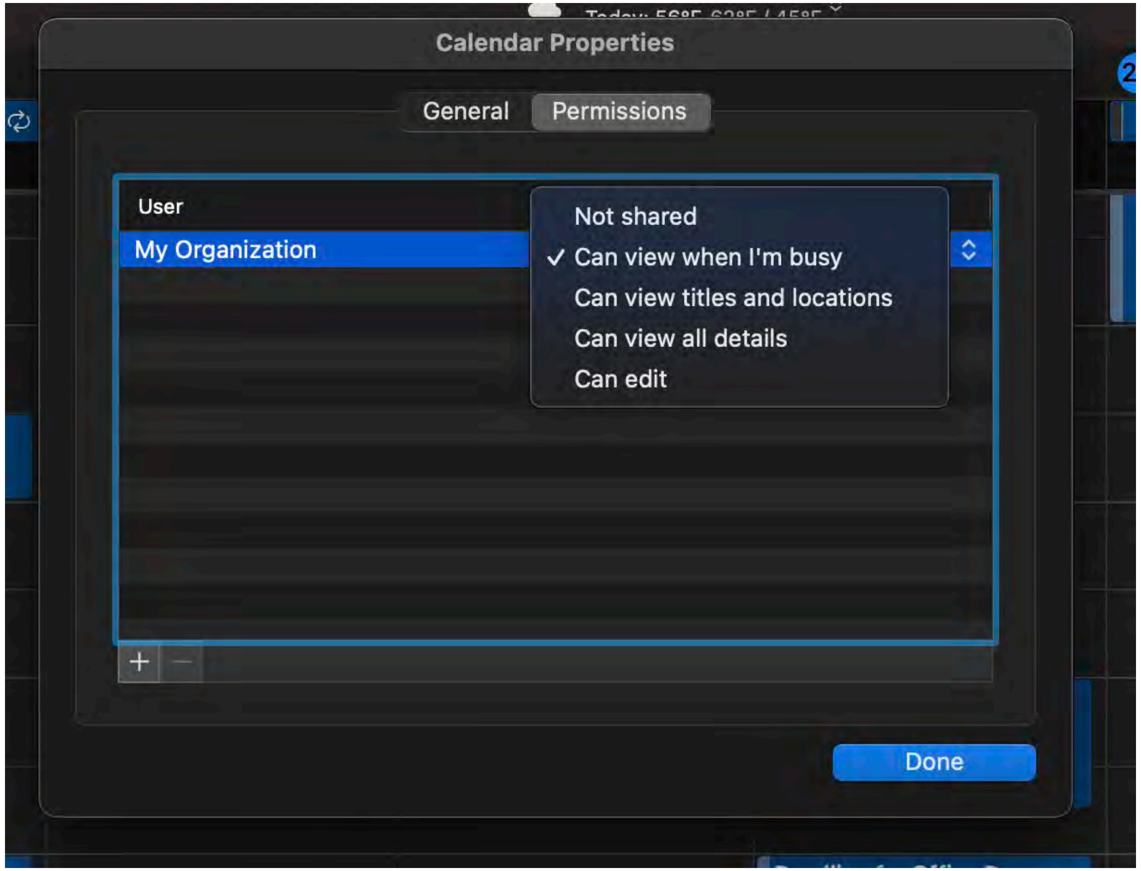
□ Automatically join audio by computer when joining a meeting


Who can see your messages?

To: Everyone

Make sure you choose everyone, when you have a question or comment

Resources from PowerPoint



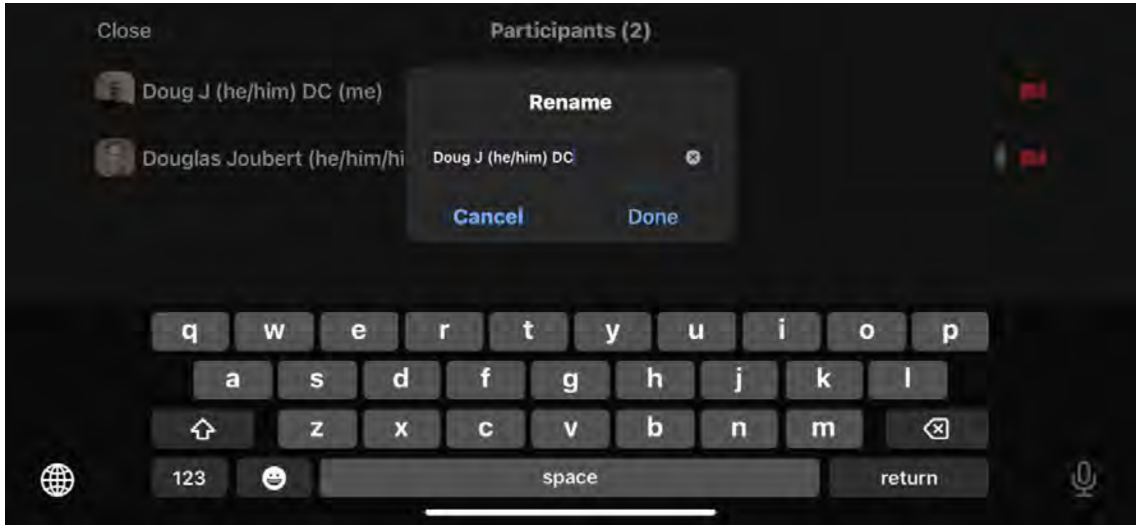
Calendar Properties

General | Permissions

User

My Organization

Not shared
✓ Can view when I'm busy
Can view titles and locations
Can view all details
Can edit



Participants (2)

Doug J (he/him) DC (me)
Douglas Joubert (he/him/hi

Rename

Doug J (he/him) DC

Please rename yourself, so we can:

- Send you the student version of the PowerPoint
- Send your training certificate
- Add you to our list-serve

National Institutes of Health
Office of Management

Project Management and Reproducibility In RStudio

Doug Joubert

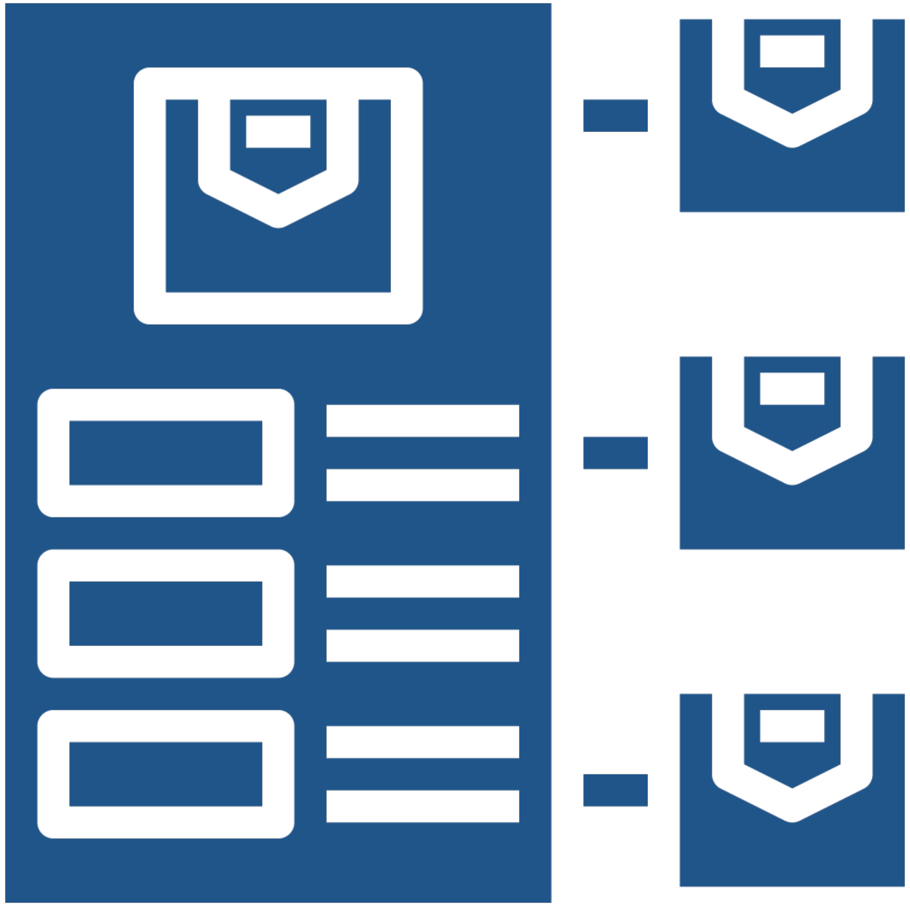
2023-02-28

NIH Library | nihlibrary.nih.gov

Office of Research Services
National Institutes of Health
U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

Class Description

- Focuses on data and project management using R and Rstudio
- Some familiarity or experience in R and RStudio is recommended but not required



Class Objectives

- Define scientific reproducibility
- Discuss best practices for organizing data in an RStudio project
- Discuss the importance of using a data dictionary and read me files
- Ensure that their data is machine readable

Configuration for Exercises

R and RStudio

- R is a programming language that is especially powerful for data exploration, visualization
- RStudio is an integrated development environment (IDE) that makes using R easier
- R and RStudio are two separate pieces of software
- **Must install R before you install RStudio**

R and RStudio: Windows

1. Download R from the CRAN website
2. Run the .exe file that was just downloaded

R and RStudio: Windows

1. Go to the RStudio download page
2. Under Installers select RStudio x.yy.zzz - Windows Vista/7/8/10 (where x, y, and z represent version numbers)
3. Double click the file to install it

R and RStudio: Mac

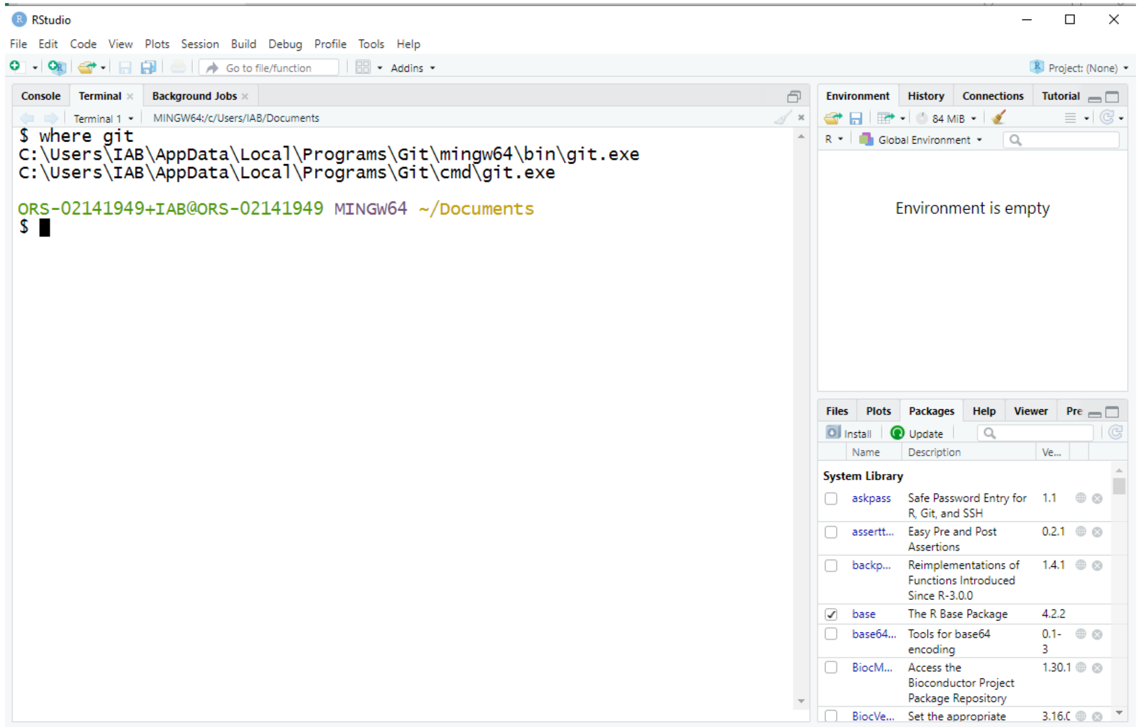
1. Download R from the CRAN website
2. Select the .pkg file for the latest R version
3. Double click on the downloaded file to install R
4. It is also a good idea to install XQuartz (needed by some packages)

R and RStudio: Mac

1. Go to the RStudio download page
2. Under Installers select RStudio x.yy.zzz - Mac OS X 10.6+ (64-bit) (where x, y, and z represent version numbers)
3. Double click the file to install RStudio

Git: Windows

- Open Power Shell on Windows or Terminal in RStudio
- type: `where git`



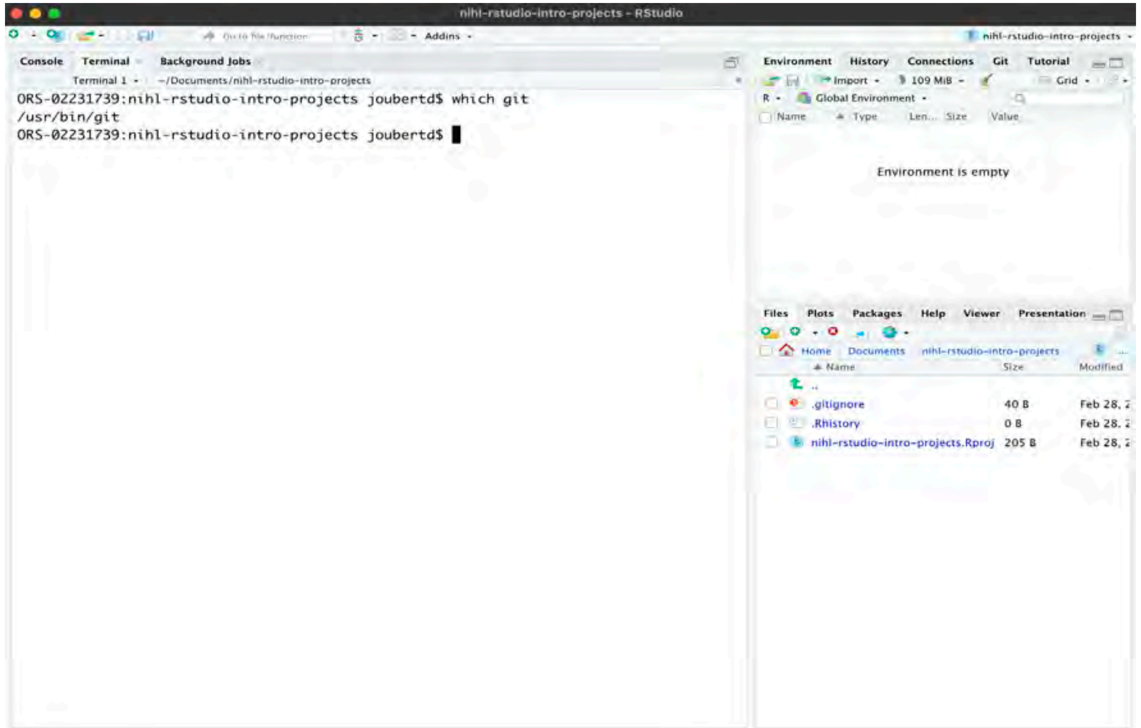
```
$ where git
C:\Users\IAB\AppData\Local\Programs\Git\mingw64\bin\git.exe
C:\Users\IAB\AppData\Local\Programs\Git\cmd\git.exe

ORS-02141949+IAB@ORS-02141949 MINGW64 ~/Documents
$ 
```

System Library

	Name	Description	Ve...
□	askpass	Safe Password Entry for R, Git, and SSH	1.1
□	assertt...	Easy Pre and Post Assertions	0.2.1
□	backp...	Reimplementations of Functions Introduced Since R-3.0.0	1.4.1
☑	base	The R Base Package	4.2.2
□	base64...	Tools for base64 encoding	0.1-3
□	BiocM...	Access the Bioconductor Project Package Repository	1.30.1
□	BiocVe...	Set the appropriate	3.16.(

Git: Mac

- Open Power Shell on Windows or Terminal in RStudio
- type: `which git`

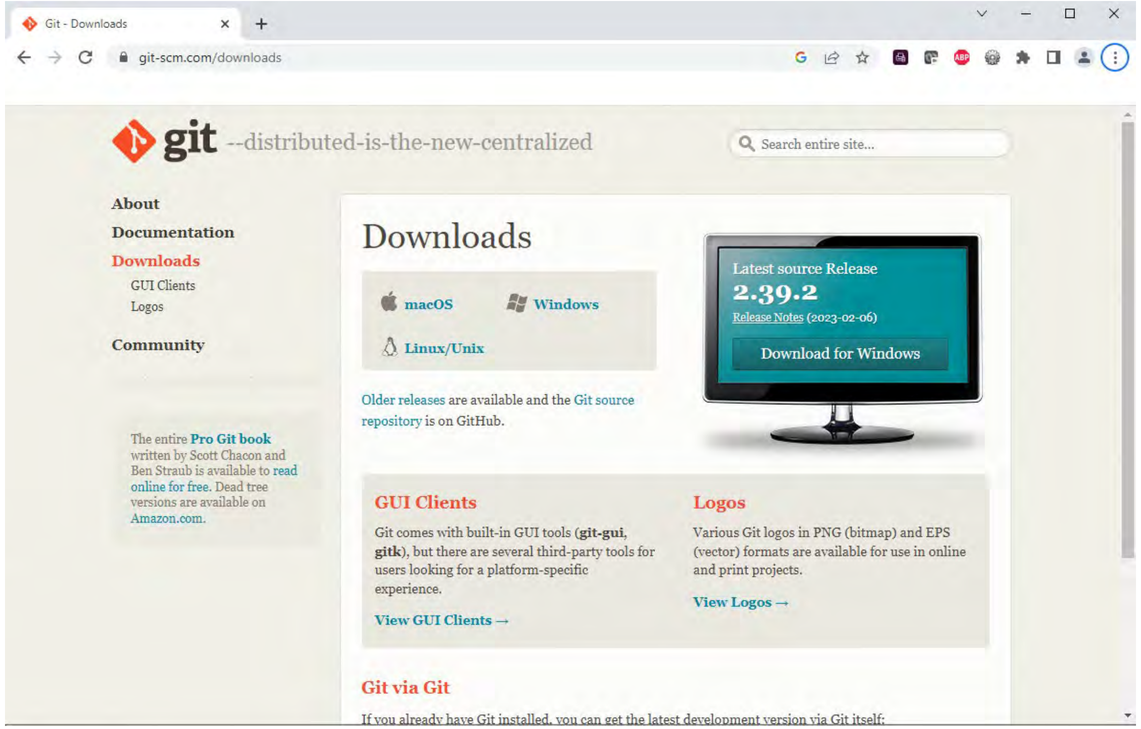


```
ORS-02231739:nihl-rstudio-intro-projects joubertd$ which git
/usr/bin/git
ORS-02231739:nihl-rstudio-intro-projects joubertd$ █
```

Name	Size	Modified
..		
.gitignore	40 B	Feb 28, 2
.Rhistory	0 B	Feb 28, 2
nihl-rstudio-intro-projects.Rproj	205 B	Feb 28, 2

Git: Installing

- Git installation instructions to install Git, if not on your computer



git --distributed-is-the-new-centralized

Search entire site...

About
Documentation
Downloads
GUI Clients
Logos
Community

The entire Pro Git book written by Scott Chacon and Ben Straub is available to read online for free. Dead tree versions are available on Amazon.com.

Downloads

macOS
Windows
Linux/Unix

Older releases are available and the Git source repository is on GitHub.

Latest source Release
2.39.2
Release Notes (2023-02-06)
Download for Windows

GUI Clients

Git comes with built-in GUI tools (**git-gui**, **gitk**), but there are several third-party tools for users looking for a platform-specific experience.

View GUI Clients →

Logos

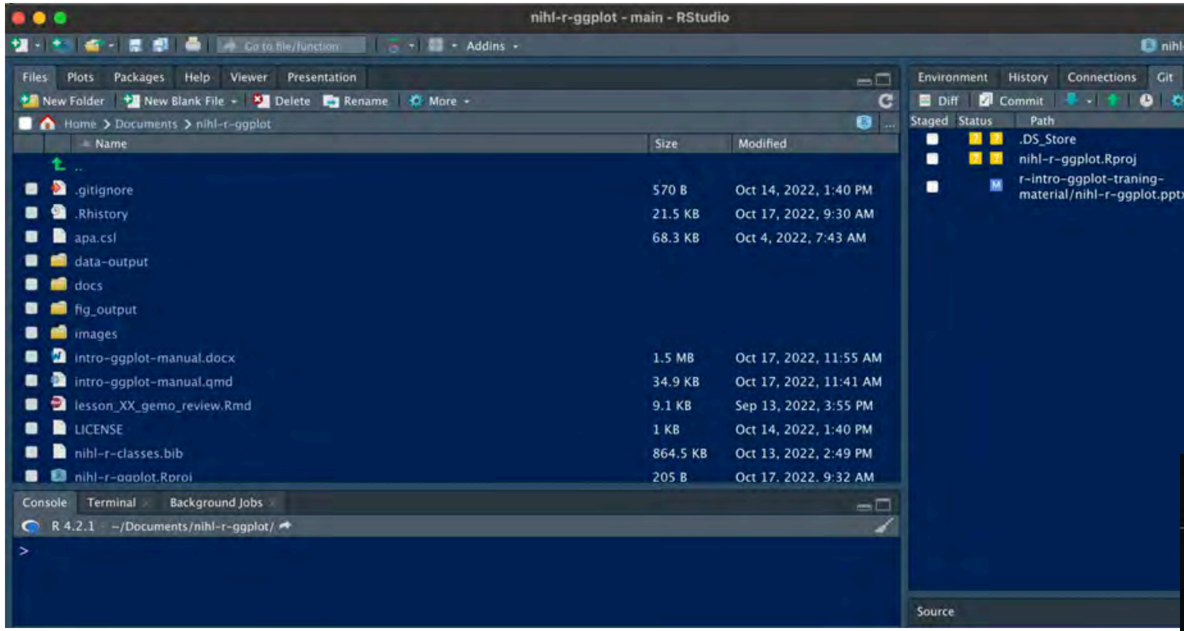
Various Git logos in PNG (bitmap) and EPS (vector) formats are available for use in online and print projects.

View Logos →

Git via Git

If you already have Git installed, you can get the latest development version via Git itself:

Project Configuration



nihl-r-ggplot - main - RStudio

Home > Documents > nihl-r-ggplot

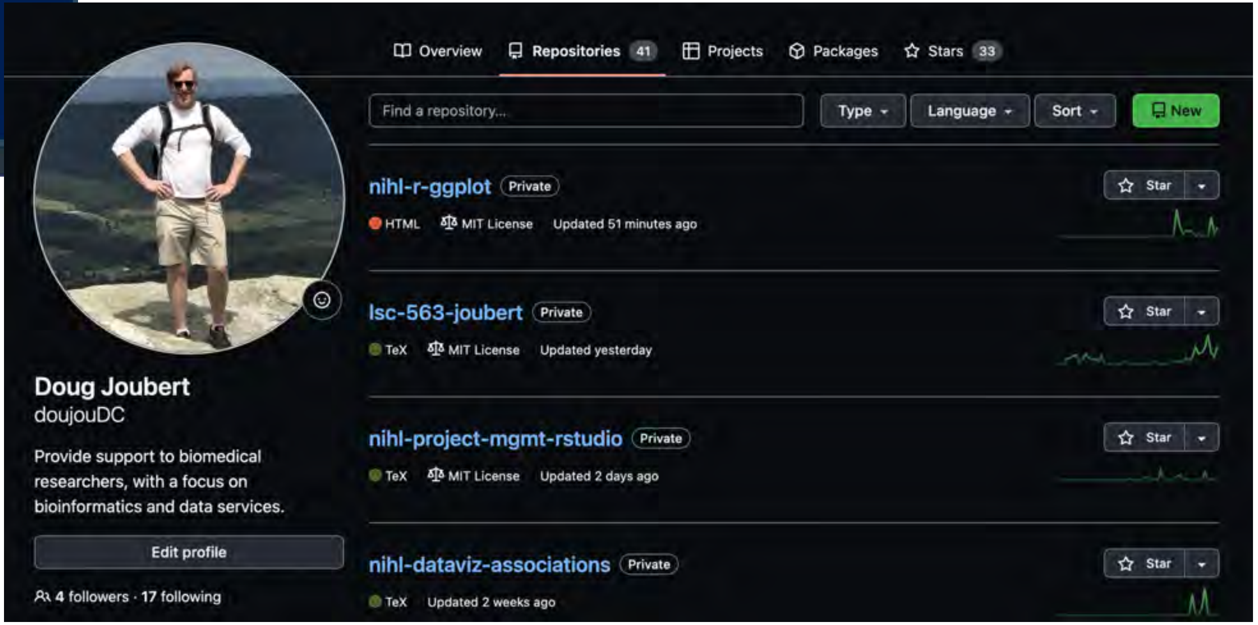
Name	Size	Modified
..		
.gitignore	570 B	Oct 14, 2022, 1:40 PM
.Rhistory	21.5 KB	Oct 17, 2022, 9:30 AM
apa.csl	68.3 KB	Oct 4, 2022, 7:43 AM
data-output		
docs		
fig_output		
images		
intro-ggplot-manual.docx	1.5 MB	Oct 17, 2022, 11:55 AM
intro-ggplot-manual.qmd	34.9 KB	Oct 17, 2022, 11:41 AM
lesson_XX_gemo_review.Rmd	9.1 KB	Sep 13, 2022, 3:55 PM
LICENSE	1 KB	Oct 14, 2022, 1:40 PM
nihl-r-classes.bib	864.5 KB	Oct 13, 2022, 2:49 PM
nihl-r-ggplot.Rproj	205 B	Oct 17, 2022, 9:32 AM

Staged	Status	Path
□	? ?	.DS_Store
□	? ?	nihl-r-ggplot.Rproj
□	M	r-intro-ggplot-traning-material/nihl-r-ggplot.pptx

```
R 4.2.1 · ~/Documents/nihl-r-ggplot/
>
```

- Follows RStudio project framework
- Code in .qmd file
- Wilson, G., Bryan, J., Cranston, K., Kitzes, J., Nederbragt, L., & Teal, T. K. (2017). Good enough practices in scientific computing. PLOS Computational Biology, 13(6), e1005510-e1005510.

- All class repos sync'd to GitHub
- Includes:
  - Class data
  - Folder structure
  - Training material
  - Exercises



Doug Joubert
doujouDC

Provide support to biomedical researchers, with a focus on bioinformatics and data services.

4 followers · 17 following

nihl-r-ggplot Private
HTML · MIT License · Updated 51 minutes ago

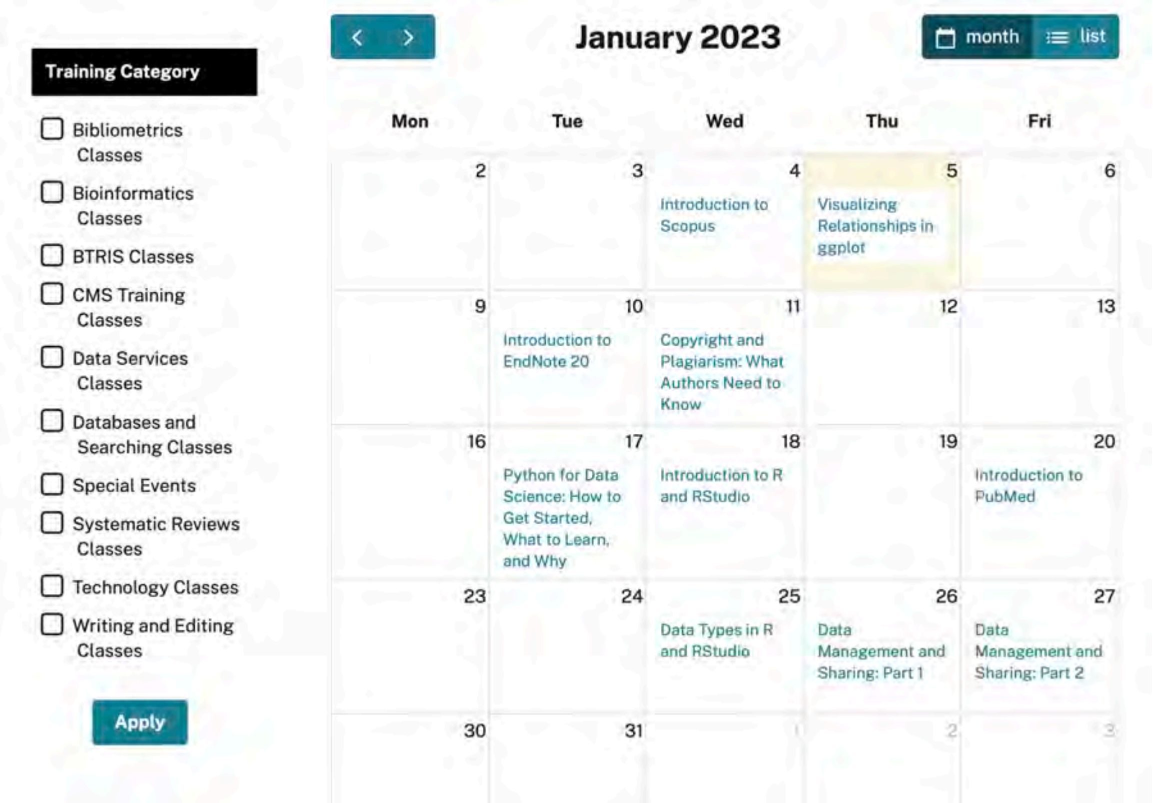
lsc-563-joubert Private
TeX · MIT License · Updated yesterday

nihl-project-mgmt-rstudio Private
TeX · MIT License · Updated 2 days ago

nihl-dataviz-associations Private
TeX · Updated 2 weeks ago

NIH Library R Classes: Basics

- Introduction to R and RStudio
- Introduction to R Data Types
- Data Wrangling in R
- Introduction to Project Management in RStudio (A)
- Reproducibility in RStudio: Basic Markdown
- Introduction to Data Visualization in R: ggplot (A)



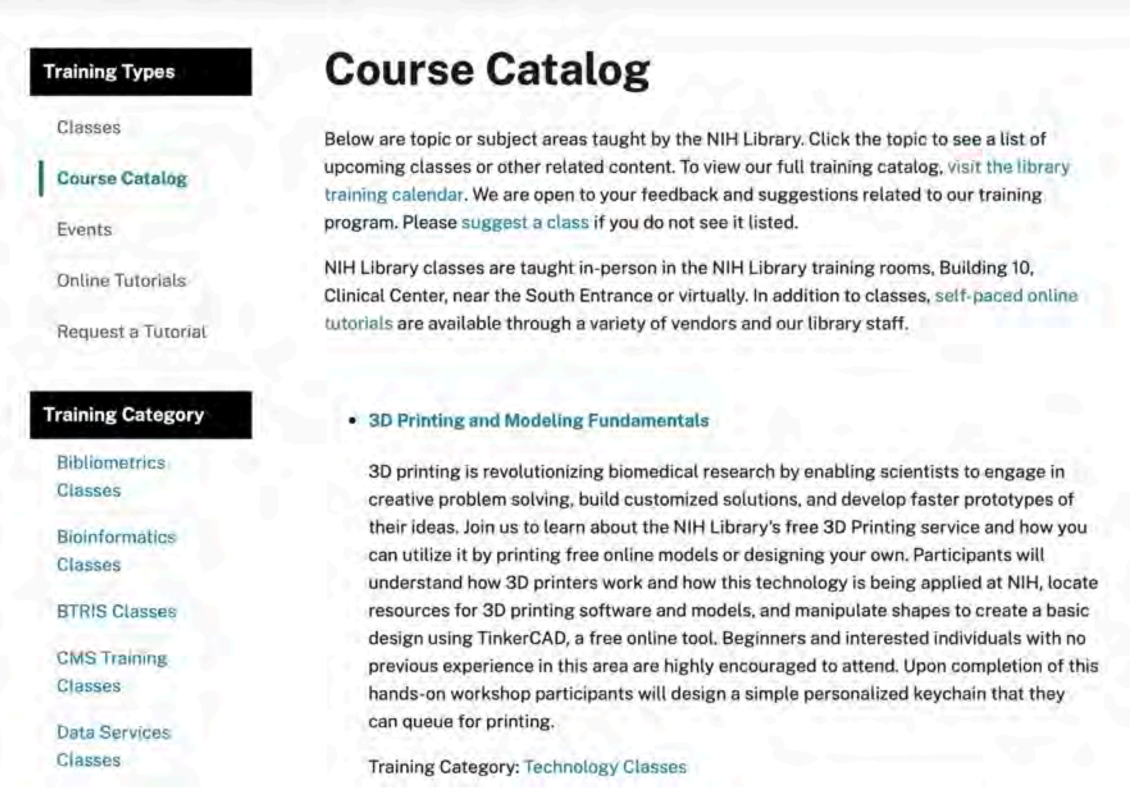
□ Bibliometrics Classes
□ Bioinformatics Classes
□ BTRIS Classes
□ CMS Training Classes
□ Data Services Classes
□ Databases and Searching Classes
□ Special Events
□ Systematic Reviews Classes
□ Technology Classes
□ Writing and Editing Classes

January 2023

Mon	Tue	Wed	Thu	Fri
2	3	4 Introduction to Scopus	5 Visualizing Relationships in ggplot	6
9	10 Introduction to EndNote 20	11 Copyright and Plagiarism: What Authors Need to Know	12	13
16	17 Python for Data Science: How to Get Started, What to Learn, and Why	18 Introduction to R and RStudio	19	20 Introduction to PubMed
23	24	25 Data Types in R and RStudio	26 Data Management and Sharing: Part 1	27 Data Management and Sharing: Part 2
30	31	1	2	3

NIH Library R Classes: Advanced

- Reproducibility in RStudio: Advanced Markdown
- Working with Git in RStudio
- Introduction to Data Visualization in R: Customization in ggplot



Training Types

- Classes
- Course Catalog
- Events
- Online Tutorials
- Request a Tutorial

Training Category

- Bibliometrics Classes
- Bioinformatics Classes
- BTRIS Classes
- CMS Training Classes
- Data Services Classes

Course Catalog

Below are topic or subject areas taught by the NIH Library. Click the topic to see a list of upcoming classes or other related content. To view our full training catalog, visit the library training calendar. We are open to your feedback and suggestions related to our training program. Please suggest a class if you do not see it listed.

NIH Library classes are taught in-person in the NIH Library training rooms, Building 10, Clinical Center, near the South Entrance or virtually. In addition to classes, self-paced online tutorials are available through a variety of vendors and our library staff.

- **3D Printing and Modeling Fundamentals**

3D printing is revolutionizing biomedical research by enabling scientists to engage in creative problem solving, build customized solutions, and develop faster prototypes of their ideas. Join us to learn about the NIH Library's free 3D Printing service and how you can utilize it by printing free online models or designing your own. Participants will understand how 3D printers work and how this technology is being applied at NIH, locate resources for 3D printing software and models, and manipulate shapes to create a basic design using TinkerCAD, a free online tool. Beginners and interested individuals with no previous experience in this area are highly encouraged to attend. Upon completion of this hands-on workshop participants will design a simple personalized keychain that they can queue for printing.

Training Category: Technology Classes

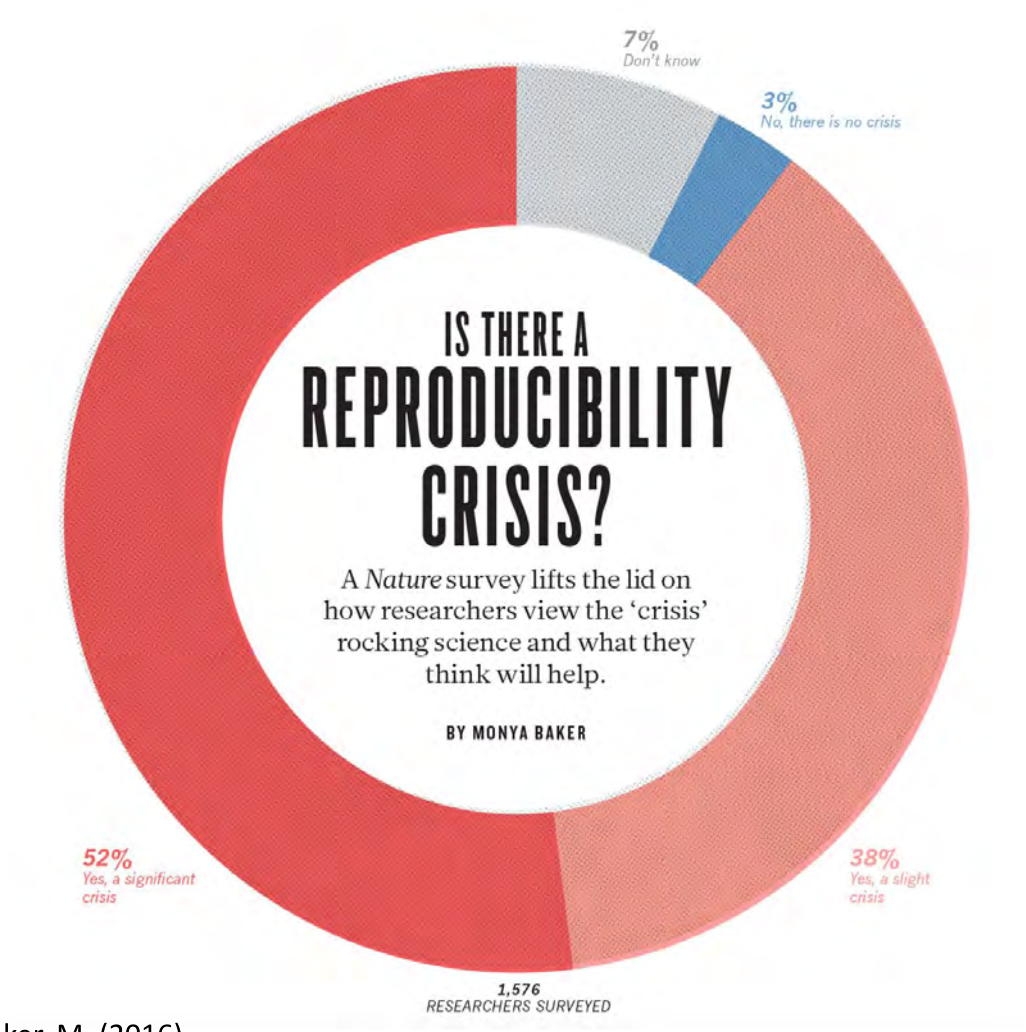
Scientific Reproducibility

According to the U.S. National Science Foundation (NSF) subcommittee on replicability in science (2015):

Science should routinely evaluate the reproducibility of findings that enjoy a prominent role in the published literature. To make reproduction possible, efficient, and informative, **researchers should sufficiently document the details of the procedures used to collect data**, to **convert observations into analyzable data**, and to **perform data analysis**.

Bollen, K., et al (2015)

Reproducibility Crisis?



IS THERE A REPRODUCIBILITY CRISIS?

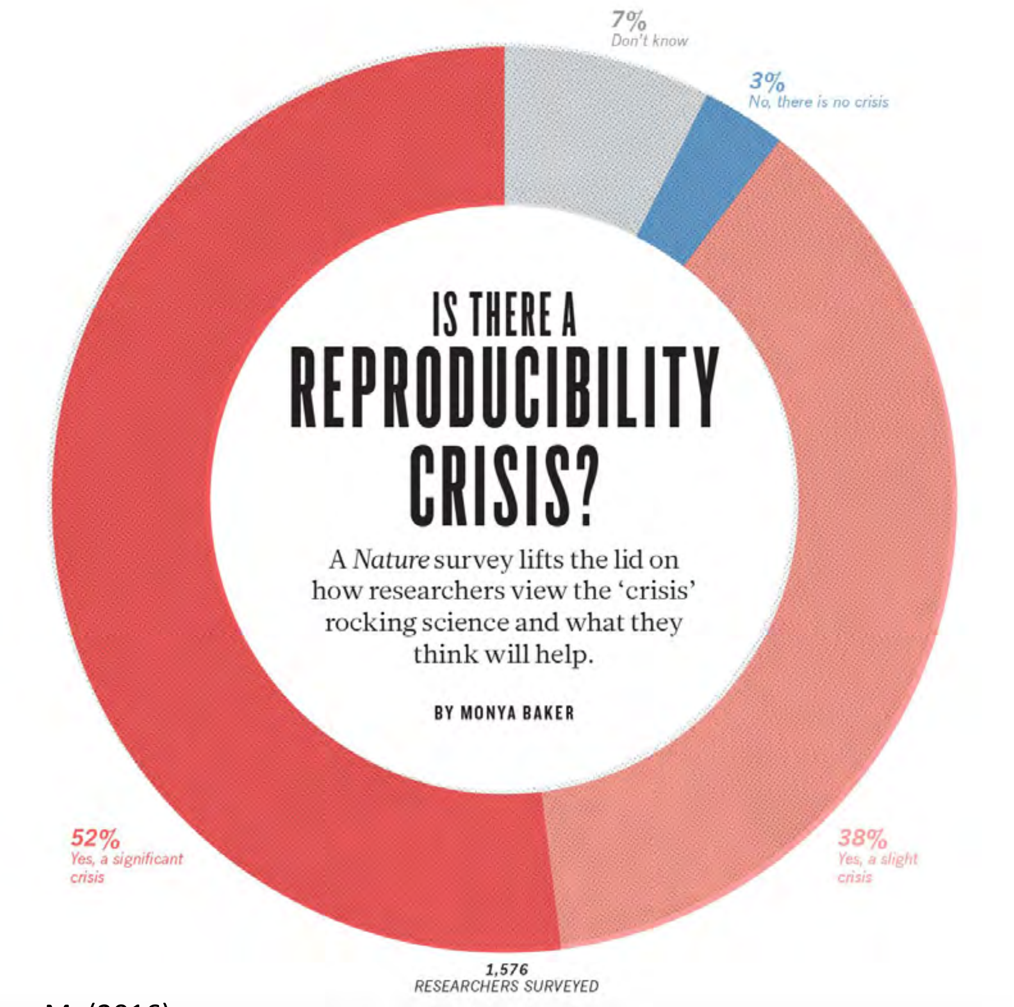
A *Nature* survey lifts the lid on how researchers view the 'crisis' rocking science and what they think will help.

BY MONYA BAKER

Baker, M. (2016).

- 1,576 researchers took online questionnaire
- More than 70% of researchers have tried and failed to reproduce another scientist's experiments
- More than half have failed to reproduce their own experiments
- Specific factors in handout

Reproducibility Crisis?



IS THERE A REPRODUCIBILITY CRISIS?

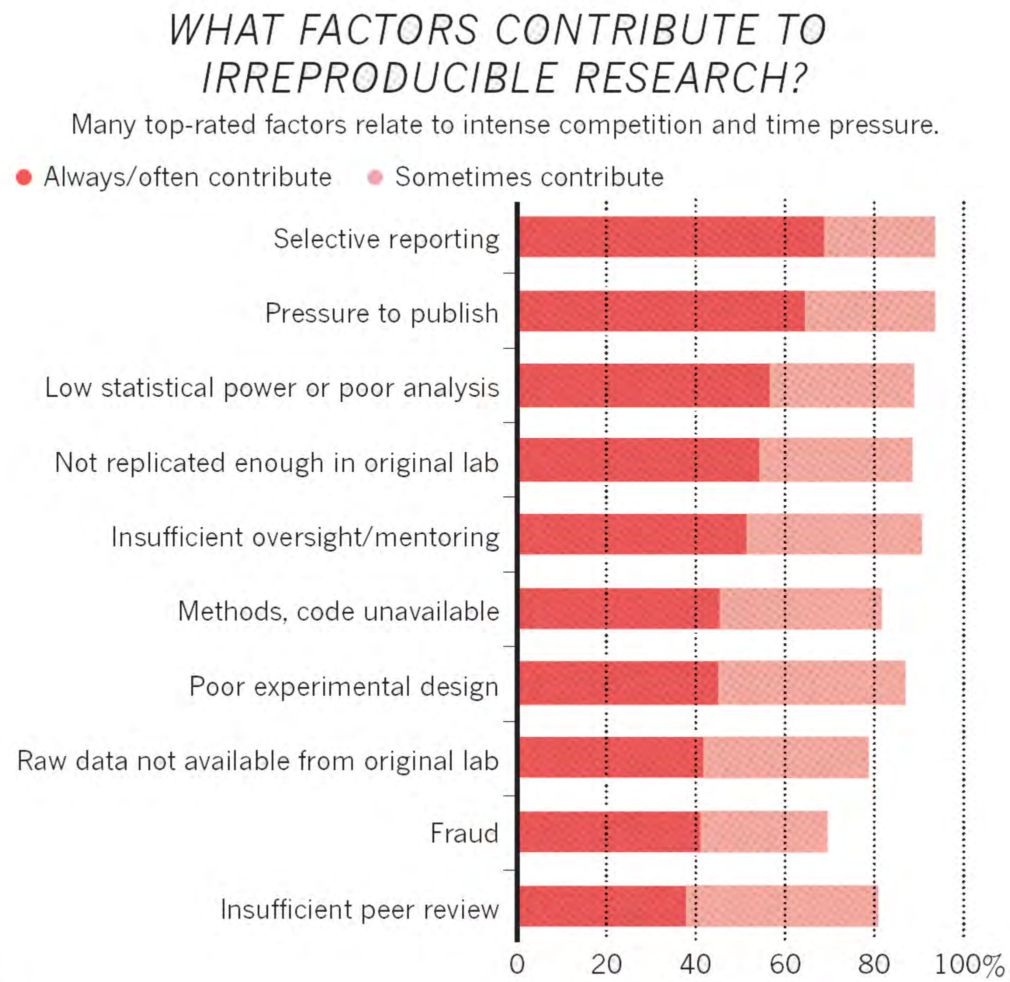
A *Nature* survey lifts the lid on how researchers view the 'crisis' rocking science and what they think will help.

BY MONYA BAKER

Baker, M. (2016).

- Sometimes-contradictory attitudes towards reproducibility:
  - > 52% agree that there is a significant crisis
  - < 31% think that failure to reproduce published results means results are wrong
  - Most still trust the published literature

Problems in Reproducibility



WHAT FACTORS CONTRIBUTE TO IRREPRODUCIBLE RESEARCH?

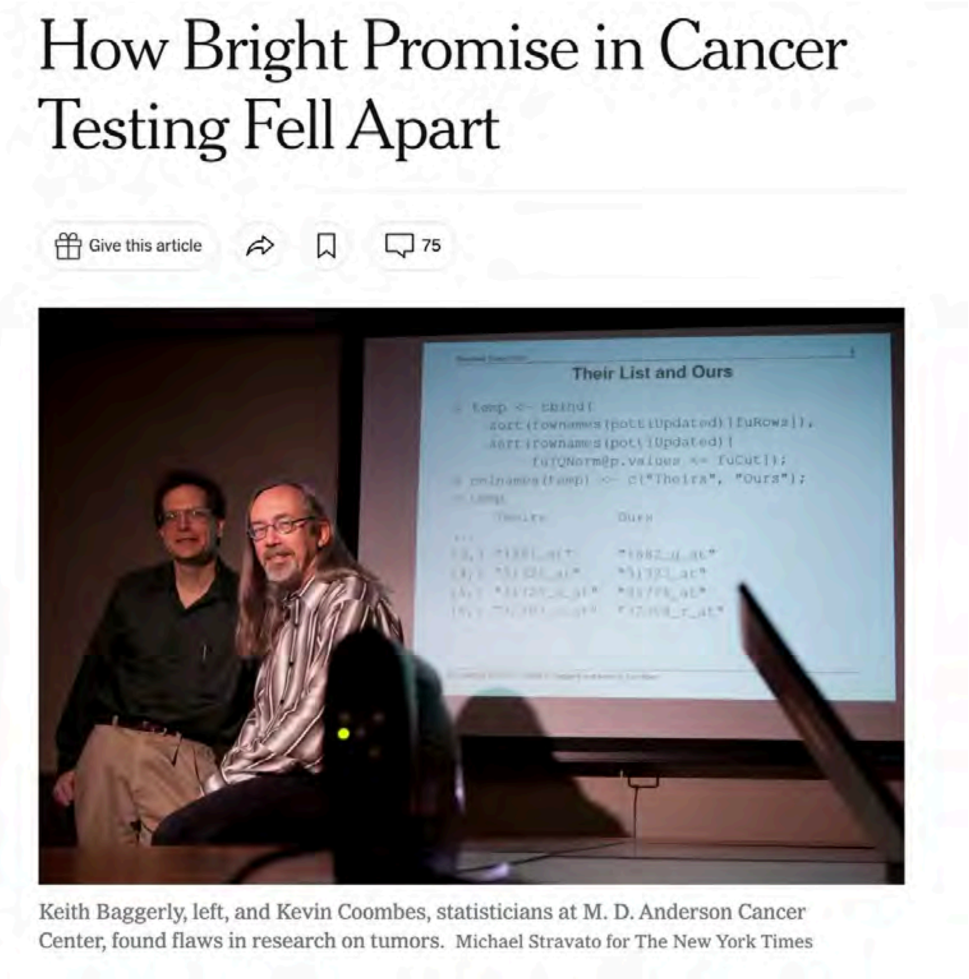
Many top-rated factors relate to intense competition and time pressure.

Baker, M. (2016).

- > 60% said that two factors were problems:
  - Pressure to publish
  - Selective reporting
- > 50% pointed to:
  - Insufficient replication in the lab
  - Poor oversight
  - Low statistical power

Embracing Reproducibility Practices

- 5 selfish reasons to use reproducibility practices:
  - Helps to avoid data loss and disaster
  - Makes it easier to write papers
  - Helps reviewers see it your way
  - Enables continuity of your work
  - Helps to build your reputation



How Bright Promise in Cancer Testing Fell Apart

Give this article 75

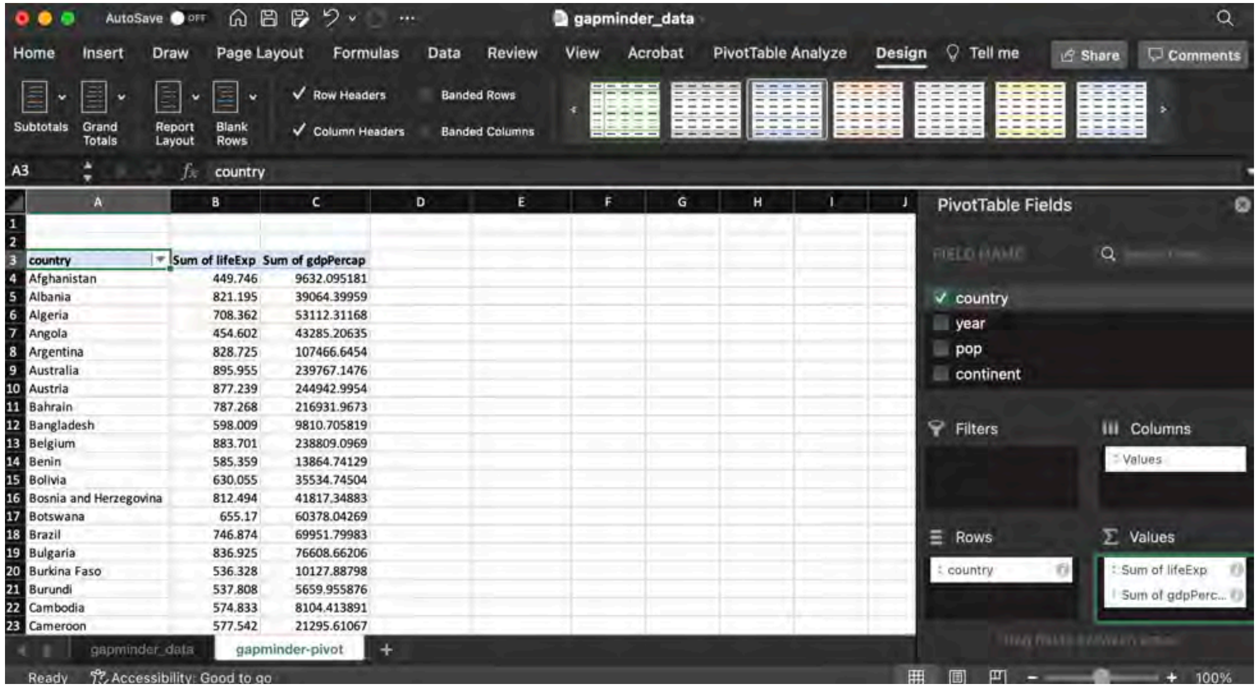
Keith Baggerly, left, and Kevin Coombes, statisticians at M. D. Anderson Cancer Center, found flaws in research on tumors. Michael Stravato for The New York Times

Markowetz, F. (2015).

Image: NY Times

Why Learn R? – It is More Powerful

- Why should I learn R when Pivot Tables do the same thing?
- For example, Figure 4 is showing a Pivot Table for gapminder data
  - Mean
  - Median
  - Min value
  - Max value



✓ Row Headers
✓ Column Headers

A3 *fx* country

country	Sum of lifeExp	Sum of gdpPercap
Afghanistan	449.746	9632.095181
Albania	821.195	39064.39959
Algeria	708.362	53112.31168
Angola	454.602	43285.20635
Argentina	828.725	107466.6454
Australia	895.955	239767.1476
Austria	877.239	244942.9954
Bahrain	787.268	216931.9673
Bangladesh	598.009	9810.705819
Belgium	883.701	238809.0969
Benin	585.359	13864.74129
Bolivia	630.055	35534.74504
Bosnia and Herzegovina	812.494	41817.34883
Botswana	655.17	60378.04269
Brazil	746.874	69951.79983
Bulgaria	836.925	76608.66206
Burkina Faso	536.328	10127.88798
Burundi	537.808	5659.955876
Cambodia	574.833	8104.413891
Cameroon	577.542	21295.61067

PivotTable Fields

☑ country
☐ year
☐ pop
☐ continent

Filters

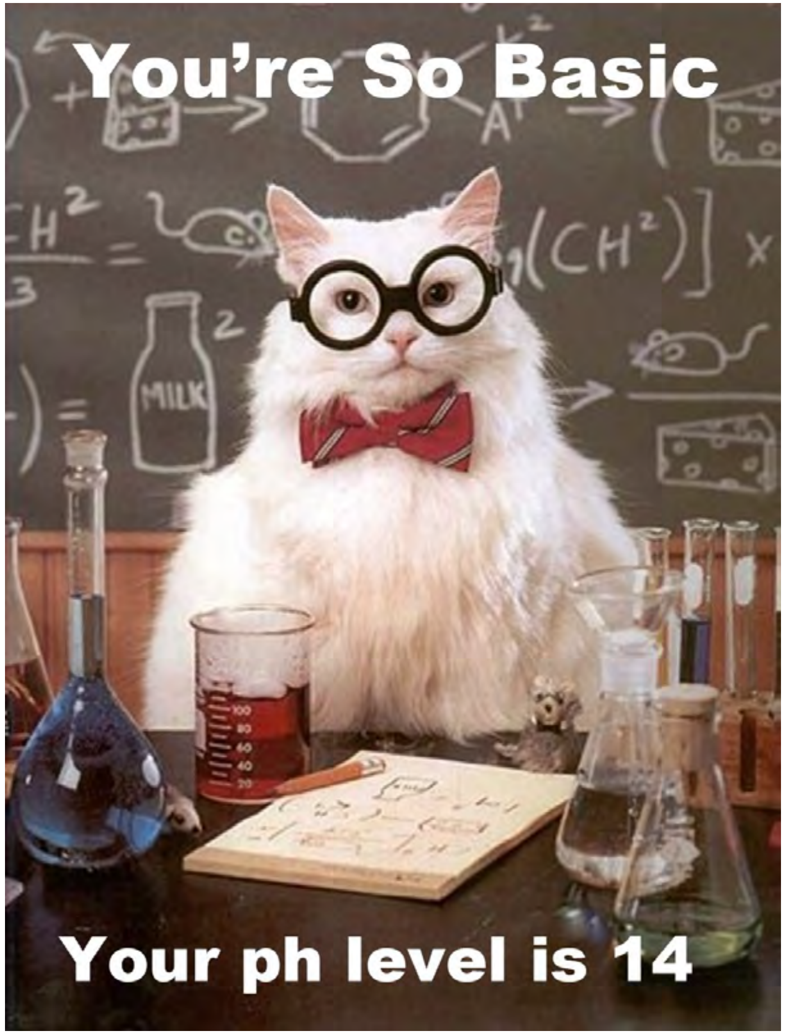
Columns: Values

Rows: country

Values: Sum of lifeExp, Sum of gdpPerc...

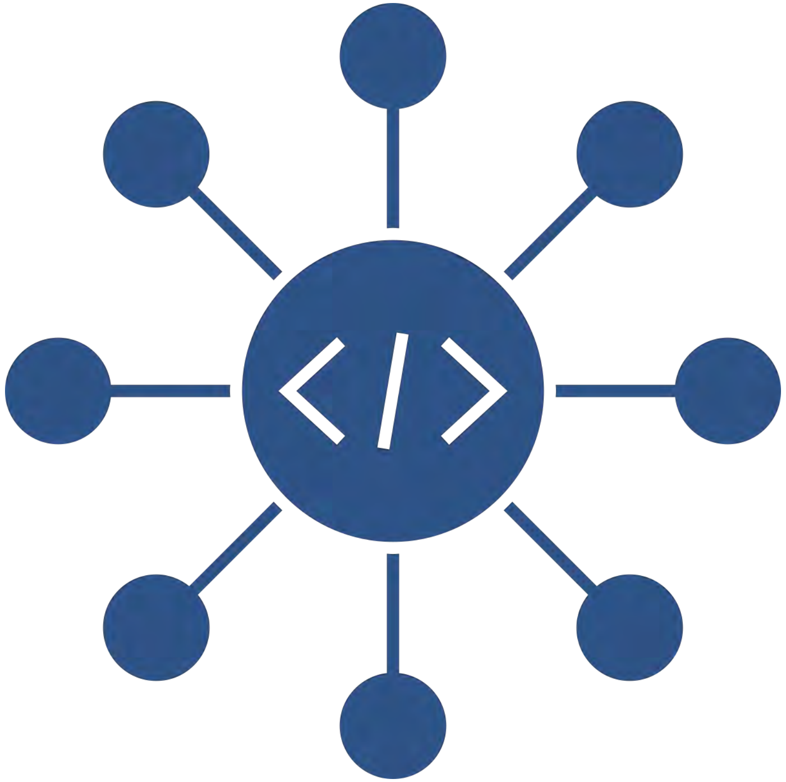
Why Learn R? – It is More Powerful

- With R, using 2 lines of code
  - Group data by country and run one regression model
  - Display statistics like the coefficients or the R2 value
  - I agree with Chemistry Cat, Excel is so basic



Why Learn R? – Free and Open Source

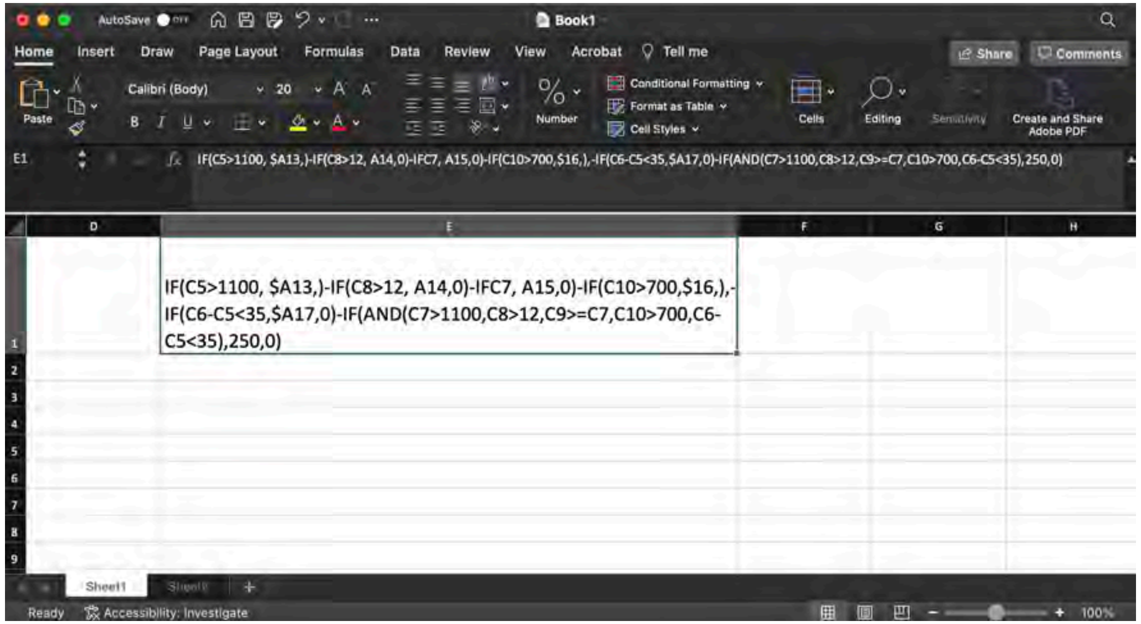
- This gets at a more philosophical reason for using R
- Not everyone has access to Office 365
- Tableau or Stata skills are great, but products are very expensive
- If sharing analysis results in Excel, not everyone will be able to open that file



(Heiss, 2022)

Why Learn R? – Reproducibility

- Only way this would be reproducible is if you write down all the steps for:
  - Every menu you clicked
  - Every cell you clicked on to add a formula or changed the formatting
  - Have you ever seen an Excel spreadsheet with an accompanying set of instructions?



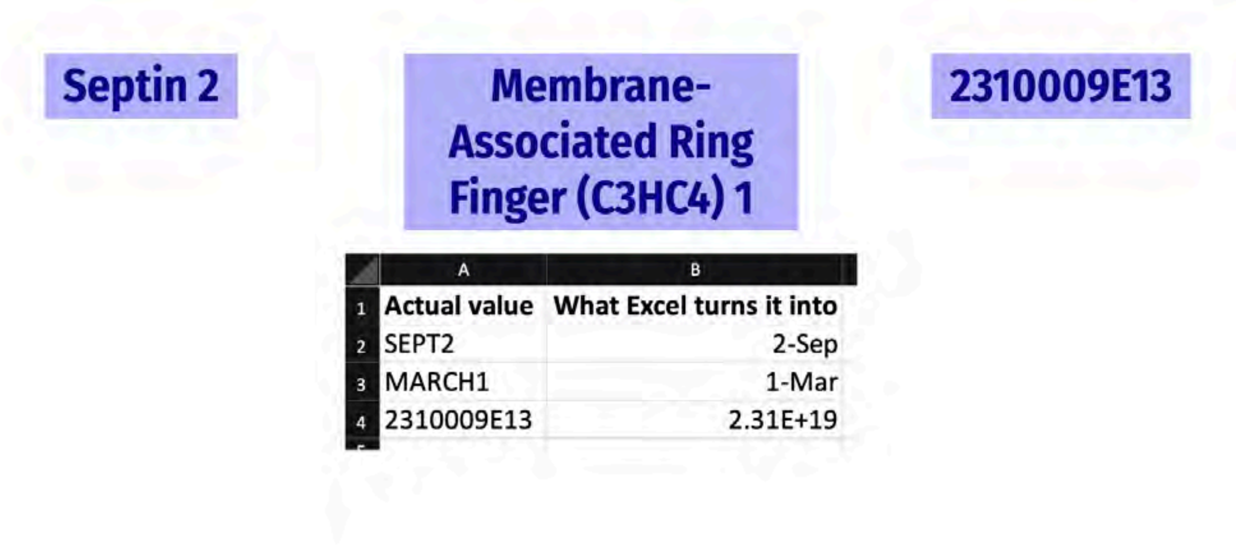
```
E1  IF(C5>1100, $A13,)-IF(C8>12, A14,0)-IFC7, A15,0)-IF(C10>700,$16,),-IF(C6-C5<35,$A17,0)-IF(AND(C7>1100,C8>12,C9>=C7,C10>700,C6-C5<35),250,0)
```

```
IF(C5>1100, $A13,)-IF(C8>12, A14,0)-IFC7, A15,0)-IF(C10>700,$16,),-
IF(C6-C5<35,$A17,0)-IF(AND(C7>1100,C8>12,C9>=C7,C10>700,C6-
C5<35),250,0)
```

(Heiss, 2022)

Why Learn R? – Reproducibility

- A 2016 study found gene name errors in 20% of the papers that they reviewed
- Web of Science search using the terms “Reproducible statistical analyses” OR “Reproducibility” resulted in over 7,000 published papers, in the last 10 years



Septin 2

Membrane-Associated Ring Finger (C3HC4) 1

2310009E13

	A	B
1	Actual value	What Excel turns it into
2	SEPT2	2-Sep
3	MARCH1	1-Mar
4	2310009E13	2.31E+19

(Ziemann, Eren, and EL-Osta)

General Recommendation

- **Don’t Touch the Raw Data**: no analysis on the original data, or, if you do, then explain what you did to the data.
- **Self-documenting and Reproducible Code**: consider writing your reports or papers in markdown. Markdown combines text with code.
- **Use Open Formats**: open formats as much as possible. That means sharing your data in csv or tab-delimited format.

(Heiss, 2022)

RStudio Project Demo

Outline for Demo

- Creating a project in RStudio
- Brief discussion about “best practices”

Best Practices for Managing Projects

Using RStudio for Project Management

- RStudio is an integrated development environment (IDE) for R and Python:
  - Free and open-source
  - Designed to make it easy to write and reuse code
  - Convenient to view and interact with the objects stored in your environment
  - Collaboration and Publishing Tools
  - Documents using R Markdown
- We have a separate class that focuses on Intro to R

Files and folders

- Using the poll, list some problems that you have encountered when dealing with files and folders

Files and folders

- Cannot find your files on your computer (or your cloud storage)
- Multiple versions of files with names such as "finaldraft_4.txt”
- Path issues when trying to run code
- Reviewers or colleagues cannot re-run your code/analyses

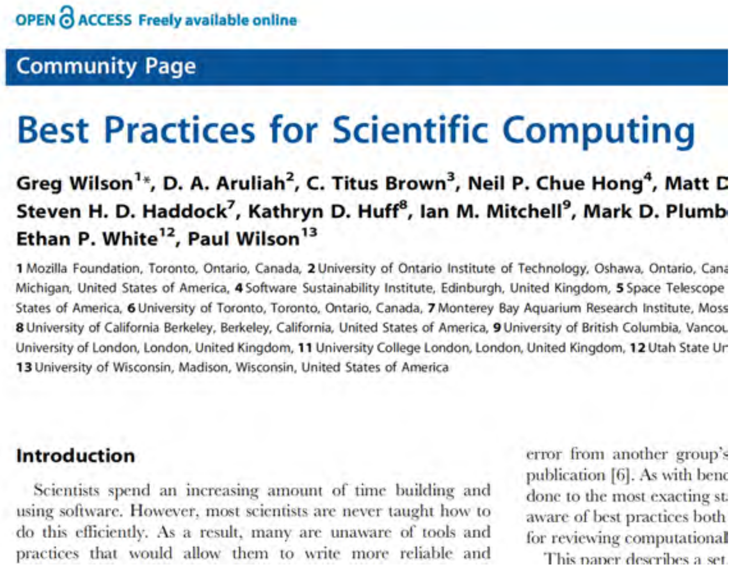
Storage and Sharing Issues

- Using the poll, list some problems that you have encountered when dealing with storage or sharing

Storage and Sharing Issues

- Files are only saved to your computer
- Collaborators don't share the files needed
- Files are shared via email attachments
- Difficult to know if you have the latest version of documents

Practice Good File Organization



OPEN ACCESS Freely available online

Community Page

Best Practices for Scientific Computing

**Greg Wilson[1]*, D. A. Aruliah[2], C. Titus Brown[3], Neil P. Chue Hong[4], Matt D
Steven H. D. Haddock[7], Kathryn D. Huff[8], Ian M. Mitchell[9], Mark D. Plumb
Ethan P. White[12], Paul Wilson[13]**

1 Mozilla Foundation, Toronto, Ontario, Canada, **2** University of Ontario Institute of Technology, Oshawa, Ontario, Cana
Michigan, United States of America, **4** Software Sustainability Institute, Edinburgh, United Kingdom, **5** Space Telescope
States of America, **6** University of Toronto, Toronto, Ontario, Canada, **7** Monterey Bay Aquarium Research Institute, Moss
8 University of California Berkeley, Berkeley, California, United States of America, **9** University of British Columbia, Vancou
University of London, London, United Kingdom, **11** University College London, London, United Kingdom, **12** Utah State Ur
13 University of Wisconsin, Madison, Wisconsin, United States of America

Introduction

Scientists spend an increasing amount of time building and using software. However, most scientists are never taught how to do this efficiently. As a result, many are unaware of tools and practices that would allow them to write more reliable and

error from another group's publication [6]. As with benc
done to the most exacting st
aware of best practices both
for reviewing computational
This paper describes a set

F1000Research F1000

COMMENTARY

REVISED Software Carpentry: lessons learned review: 3 approved]

Greg Wilson

Software Carpentry Foundation, Austin, TX, USA

v2 **First published:** 19 Feb 2014, **3**:62
https://doi.org/10.12688/f1000research.3-62.v1
Latest published: 28 Jan 2016, **3**:62
https://doi.org/10.12688/f1000research.3-62.v2

Open P

Approv

version

Abstract
Since its start in 1998, Software Carpentry has evolved from a week-

Good Enough Practices for Scientific Computing

Greg Wilson[1], Jennifer Bryan[2], Karen Cranston[3], Justin Kitzes[4], Lex Nederbragt[5], Tracy K. Teal[6]

1) Software Carpentry Foundation / gvwilson@software-carpentry.org

2) University of British Columbia / jenny@stat.ubc.ca

3) Duke University / karen.cranston@duke.edu

4) University of California, Berkeley / jkitzes@berkeley.edu

5) University of Oslo / lex.nederbragt@ibv.uio.no

6) Data Carpentry / tkteal@datacarpentry.org

> **A PDF of this paper can be found at http://arxiv.org/abs/1609.00037. We would appreciate feedback of all kinds: please email comments to the authors.**

Abstract

We present a set of computing tools and techniques that every researcher can and should consider adopting. These recommendations synthesize inspiration from our own work, from the experiences of the thousands of people who hav taken part in Software Carpentry and Data Carpentry workshops over the past six years, and from a variety of other guides. Our recommendations are aimed specifically at people who are new to research computing.

Author Summary

Computers are now essential in all branches of science, but most researchers are never taught the equivalent of basic

Practice Good File Organization

- Good Enough Practices for Scientific Computing gives the following recommendations for project organization:
  - Each project in its own directory, which is named after the project
  - Text documents associated with the project in the doc directory
  - Raw data and metadata in the data directory (raw-data)
  - Files generated during cleanup and analysis in a results directory

Wilson., et al, 2017

Practice Good File Organization

- Good Enough Practices for Scientific Computing gives the following recommendations for project organization:
  - Project's scripts and programs in the src directory
  - Programs brought in from elsewhere or compiled locally in the bin directory
  - Name all files to reflect their content or function

Wilson., et al, 2017

Practice Good File Organization

- Additional filles to include:
  - README file, to communicate important information about your project
  - LICENSE file, so that others are free to use, change, and distribute the software
  - CITATION.cff file, to let others know how you would like them to cite your work
- Student version of the PowerPoint has more resources to explore

Wilson., et al, 2017

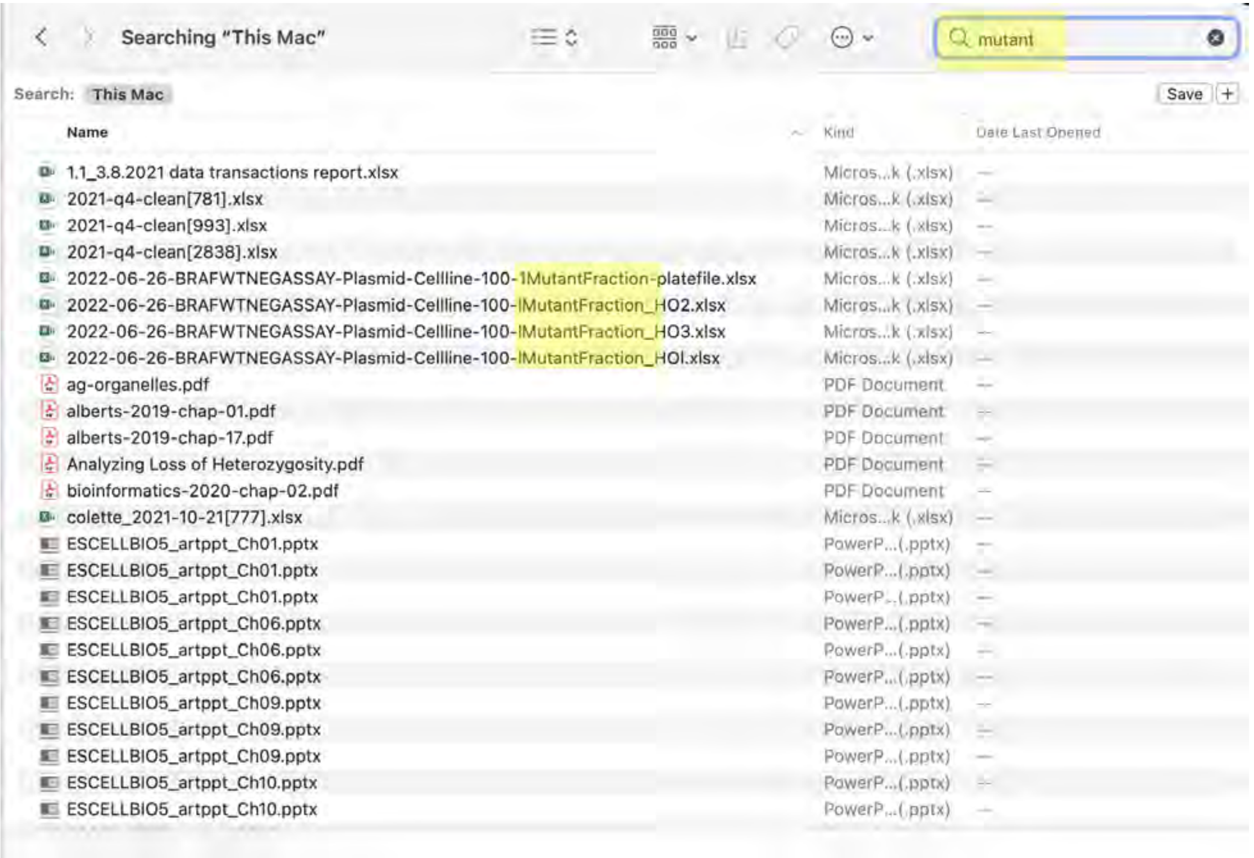
Naming Files

1. Machine readable
2. Human readable
3. Plays well with default ordering

Wilson., et al, 2017

Naming Files: Why This is Important

- Globbing: using wildcard characters to request or evaluate sets of files with the same partial names or sets of character
- Regular expression friendly
- “Findability”
- Global name changes



Name	Kind	Date Last Opened
1.1_3.8.2021 data transactions report.xlsx	Micros...k (.xlsx)	--
2021-q4-clean[781].xlsx	Micros...k (.xlsx)	--
2021-q4-clean[993].xlsx	Micros...k (.xlsx)	--
2021-q4-clean[2838].xlsx	Micros...k (.xlsx)	--
2022-06-26-BRAFWTNEGASSAY-Plasmid-Cellline-100-1MutantFraction-platefile.xlsx	Micros...k (.xlsx)	--
2022-06-26-BRAFWTNEGASSAY-Plasmid-Cellline-100-lMutantFraction_HO2.xlsx	Micros...k (.xlsx)	--
2022-06-26-BRAFWTNEGASSAY-Plasmid-Cellline-100-lMutantFraction_HO3.xlsx	Micros...k (.xlsx)	--
2022-06-26-BRAFWTNEGASSAY-Plasmid-Cellline-100-lMutantFraction_HOl.xlsx	Micros...k (.xlsx)	--
ag-organelles.pdf	PDF Document	--
alberts-2019-chap-01.pdf	PDF Document	--
alberts-2019-chap-17.pdf	PDF Document	--
Analyzing Loss of Heterozygosity.pdf	PDF Document	--
bioinformatics-2020-chap-02.pdf	PDF Document	--
colette_2021-10-21[777].xlsx	Micros...k (.xlsx)	--
ESCELLBIO5_artppt_Ch01.pptx	PowerP...(.pptx)	--
ESCELLBIO5_artppt_Ch01.pptx	PowerP...(.pptx)	--
ESCELLBIO5_artppt_Ch01.pptx	PowerP...(.pptx)	--
ESCELLBIO5_artppt_Ch06.pptx	PowerP...(.pptx)	--
ESCELLBIO5_artppt_Ch06.pptx	PowerP...(.pptx)	--
ESCELLBIO5_artppt_Ch06.pptx	PowerP...(.pptx)	--
ESCELLBIO5_artppt_Ch09.pptx	PowerP...(.pptx)	--
ESCELLBIO5_artppt_Ch09.pptx	PowerP...(.pptx)	--
ESCELLBIO5_artppt_Ch09.pptx	PowerP...(.pptx)	--
ESCELLBIO5_artppt_Ch10.pptx	PowerP...(.pptx)	--
ESCELLBIO5_artppt_Ch10.pptx	PowerP...(.pptx)	--

Naming Files

- **Machine-readable**
  - No spaces, unsupported punctuation, accented characters, or case-sensitive file names
  - Deliberate use of delimiters (i.e. for splitting file names)
  - Consistently use the same delimiters: data-analyses-fig1.R as an example
- **Human-readable**
  - Name contains brief description of content: i.e. anova-analyses-control.R

Ordering Files: Chronological

- With **chronological ordering**, file name starts with date:
  - 2022-02-26-BRAFWTNEGASSAY-FFPEDNA-CRC-1-41-AO2.csv
  - 2022-02-26-BRAFWTNEGASSAY-FFPEDNA-CRC-1-41-AO3.csv
  - 2022-02-26-BRAFWTNEGASSAY-FFPEDNA-CRC-1-41-AO4.csv

Consider using ISO 8601 date standard

Ordering Files: Logical

- With **logical ordering**, the filename starts with a number or keyword/number combo.
  - 01-marshall-data.r *see code directory*
  - 02-pre-dea-filtering.r *see code directory*
  - 03-explore-dea-limma-voom.r
  - 04-exploe-dea-results.r helper
  - 01-load-counts.rmd
  - helper02-load-exp-des.r
  - helper03-extract-and-tidy.r

Ordering Files: Logical

- As illustrated on the previous slide, left-pad your numbers to facilitate sorting. If you do not do this, your data sorts like this…which is really sad
  - 01-marshall-data.r *see code directory*
  - 04-exploe-dea-results.r
  - 2-explore-dea-limma-voom.r
  - 3-helper-extract-and-tidy.r
  - helper01-load-counts.rmd

Adapted from https://datacarpentry.org/rr-organization1/01-file-naming/index.html. For more tips on file naming, check: The Dos and Don’ts of File Naming.

Resources

- Configuration for exercises
- Introduction to Reproducible Publications with RStudio
- R for Reproducible Scientific Analysis

Resources

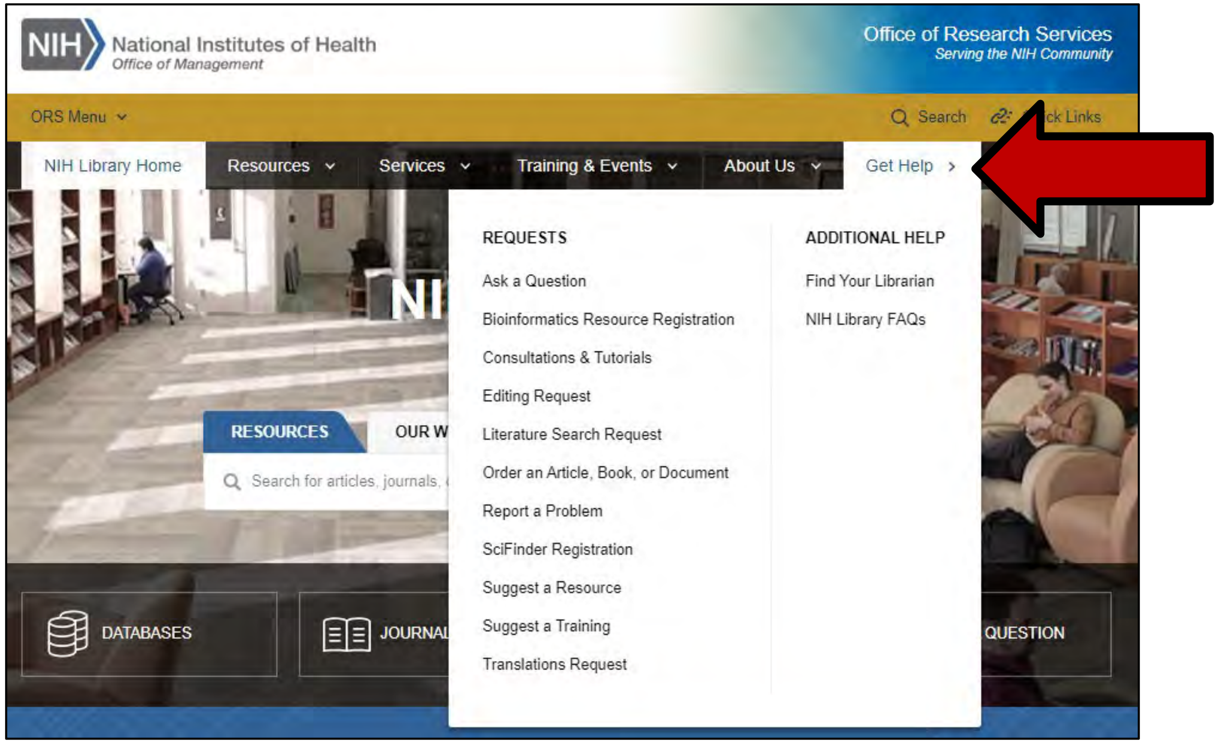
- Reproducible Research Data and Project Management in R
- Using Projects in Rstudio
- email me for a copy: douglas.joubert@nih.gov

We Can Help

- Classes on a variety of data-related topics, including:
  - Data management
  - Data visualization
  - Data analysis
  - R and RStudio
- Computers which offers a suite of tools for data analysis, processing, and visualization

Contact Us for Ongoing Support

NIH Library
Office of Research Services
Serving the NIH Community



Doug Joubert
Bioinformatics Support Program
301-827-3829
douglas.joubert@nih.gov

NIH Library Help Desk
(301) 496-1080

- **Ask a Question:** https://www.nihlibrary.nih.gov/get-help/ask-question
- **Request a Tutorial:** https://www.nihlibrary.nih.gov/get-help/consultations-tutorials
- **Sign up for Additional Classes:** https://www.nihlibrary.nih.gov/training/calendar

DEPARTMENT OF HEALTH & HUMAN SERVICES · USA

NIH National Institutes of Health
Office of Management