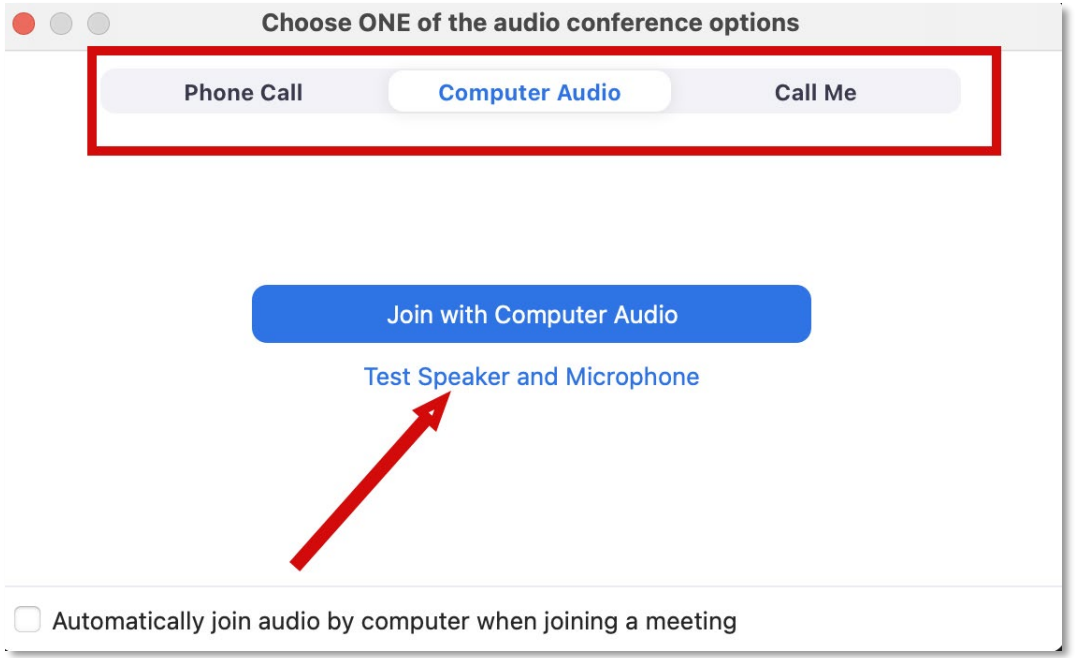
Audio Connection and Chat (Zoom)

You will not hear any sound until the webinar starts.

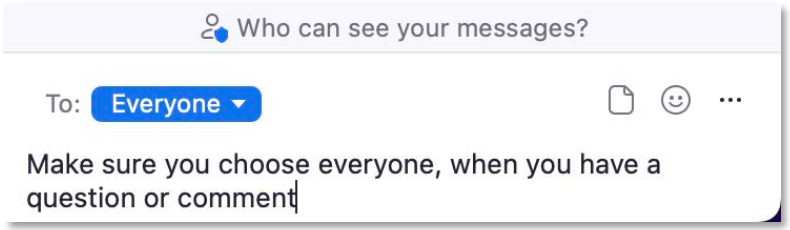
Connect Audio

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

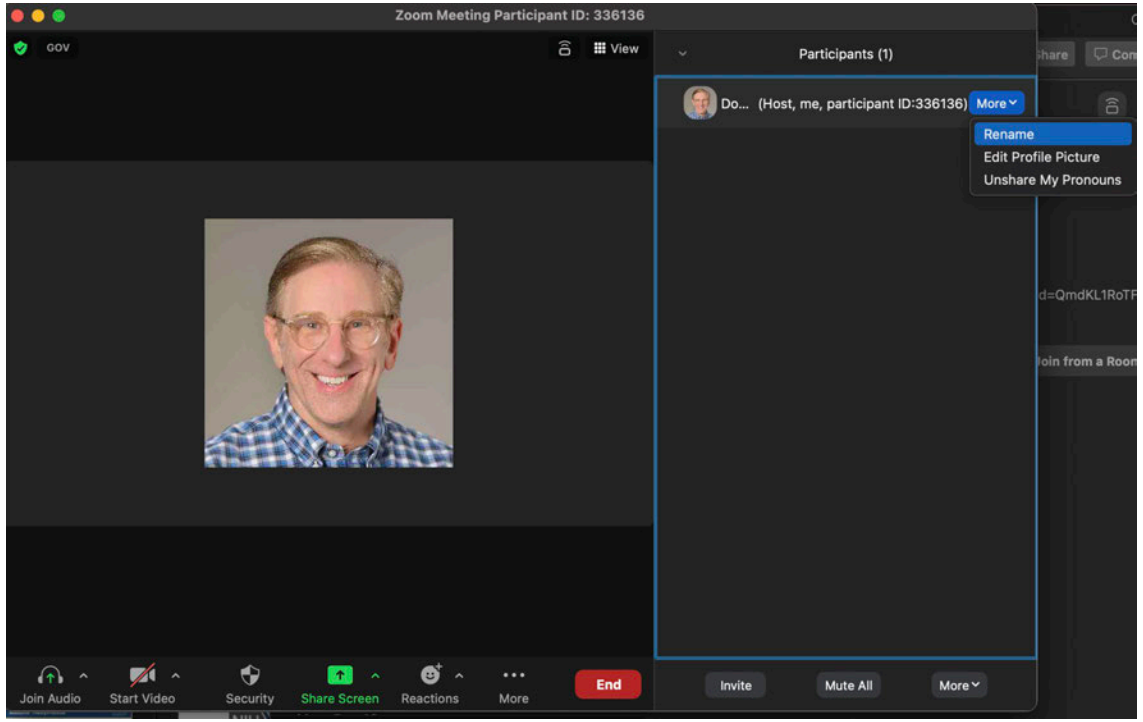


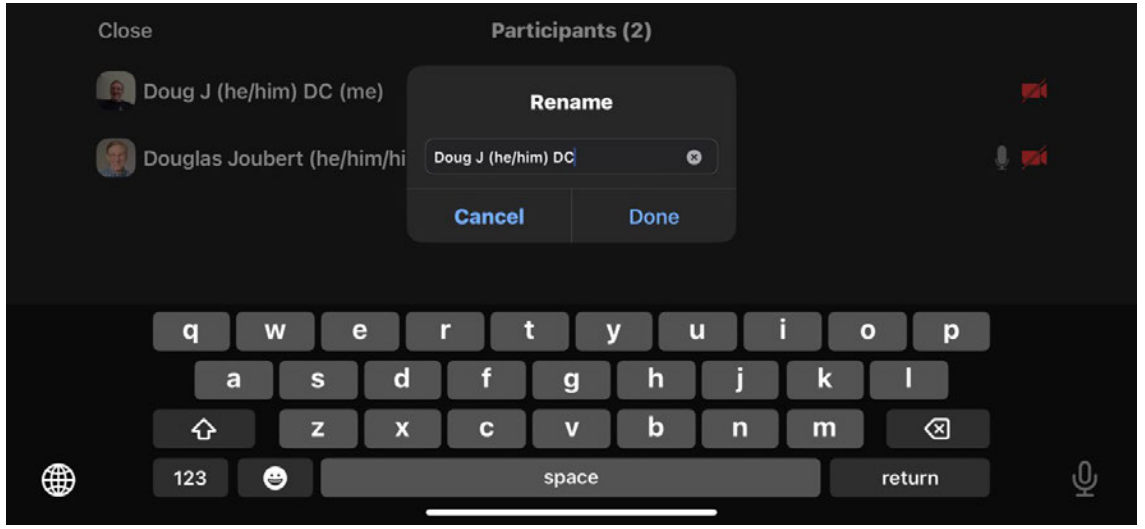
Chat

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



Rename Yourself via Participants List





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

Reproducibility in RStudio: Advanced Markdown

Doug Joubert
2023-03-09

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

Class Objectives

- Upon completion of this class students should be able to:
  - Download and install Zotero
  - Link RStudio to Zotero
  - Create a bibliography and link it to a markdown document
  - Insert citations using RStudio Visual Interface, and via the command line
  - Download and link a CSL file which specifies the formatting to use when generating the citations and bibliography

Testing Your Knowledge

Testing Your Knowledge

- Line 12 add the following link to the NSF site (see chat)

Testing Your Knowledge

- Line 12 add the following link to the NSF site (see chat)
- Lines 44-47 need to be bullet items

Testing Your Knowledge

- Line 12 add the following link to the NSF site (see chat)
- Lines 44-47 need to be bullet items
- Line 72, Add URL to text (see chat)

Testing Your Knowledge

- Line 12 add the following link to the NSF site (see chat)
- Lines 44-47 need to be bullet items
- Line 72, Add URL to text (see chat)
- Lines 74-80 need to be numbered list

Testing Your Knowledge

- Line 12 add the following link to the NSF site (see chat)
- Lines 44-47 need to be bullet items
- Line 72, Add URL to text (see chat)
- Lines 74-80 need to be numbered list
- Line 105, text needs to be formatted as code: "data-analyses-fig1.R"

Testing Your Knowledge

- Line 12 add the following link to the NSF site (see chat)
- Lines 44-47 need to be bullet items
- Line 72, Add URL to text (see chat)
- Lines 74-80 need to be numbered list
- Line 105, text needs to be formatted as code: "data-analyses-fig1.R”
- Line 113, "chronological ordering" needs to be bold

Figure Options

Figure Options

- Several ways to control the default width and height of figures
- Quarto sets a default width and height for figures appropriate to the target output format
- The table is displaying the defaults (expressed in inches)

Format	Default
Default	7 x 5
HTML Slides	9.5 x 6.5
HTML Slides (reveal.js)	9 x 5
PDF	5.5 x 3.5
PDF Slides (Beamer)	10 x 7
PowerPoint	7.5 x 5.5
MS Word, ODT, RTF	5 x 4
EPUB	5 x 4
Hugo	8 x 5

Figure Options: Header

- Set the default sizes using the `fig-width` and `fig-height` options in YAML header

```yaml
---
title: "My Document"
format:
  html:
    fig-width: 8
    fig-height: 6
  pdf:
    fig-width: 7
    fig-height: 5
---
```

Figure Options: Caption and Alt Text

- Specify the caption and alt text for figures generated from code using the `fig-cap` and `fig-alt` options

```
{r Penguins Version 2}
#| fig-width: 5
#| fig-height: 3
#| fig-cap: "Penguin length and depth"
#| fig-alt: "Penguin data from palmerpenguins package"

penguin_plot <- penguins %>%
ggplot(aes(x = bill_length_mm,
                    y = bill_depth_mm,
                    col = island))
penguin_plot +
  geom_point() +
  labs(x = "Bill Length",
       y = "Bill Depth")
```

Bibliography Files and Citation Formats

Bibliography Options

- Quarto uses Pandoc to automatically generate citations and a bibliography. To use this capability, you will need:
  - A Quarto document formatted with citations (see Citation Markdown).
  - A bibliographic data source, for example a `BibTeX (.bib)` file.
  - Optionally, a CSL file which specifies the formatting to use when generating the citations and bibliography

Bibliography Files: Supported Options

- The bibliography may have any of these formats

Format	File extension
BibLaTeX	.bib
BibTeX	.bibtex
Copac	.copac
CSL JSON	.json
CSL YAML	.yaml
EndNote	.enl
EndNote XML	.xml
ISI	.wos
MEDLINE	.medline
MODS	.mods
RIS	.ris

`.bib` can be used with both `BibTeX` and `BibLaTeX` files. I will be using citations in the `BibTeX` format.

Citation Syntax

- Quarto uses the standard Pandoc for citations
- Citations go inside square brackets, separated by semicolons
- Each citation must have a key, composed of '@' + the citation identifier from the database

```
[@citation]
```

```
Introduces reproducibility [see @doe99,
pp. 33-35; also @smith04, chap. 1].

Introduces reproducibility [@smith04;
@doe99].
```

You can also write in-text citations

```
@doe99, is the first research to highlight
this problem.

@smith04 [p. 34] says nothing on this
topic.
```

Citation Metadata

- Minimal template for a chapter or section in a book with authors

```
@inbook{citekey,
  author    = "",
  title     = "",
  chapter   = "",
  publisher = "",
  year      = ""
}
```

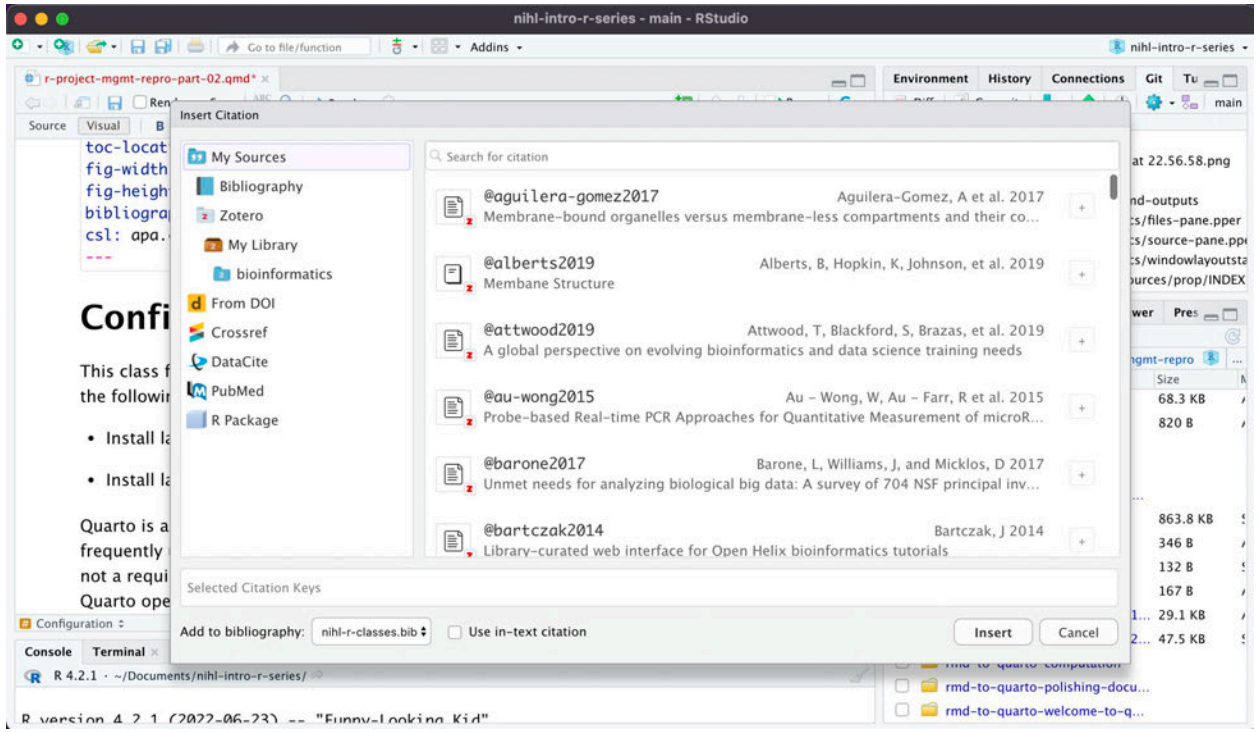
- Pandoc Citations
- BibTeX website

BibTeX Export Insert Copy

```
@inbook{RN1527,
  author = {Wilke, C.},
  title = {Multipanel Figures},
  booktitle = {Fundamentals of data visualization: a primer on making informative and
compelling figures},
  publisher = {O'Reilly Media},
  address = {Sebastopol, CA},
  edition = {First edition.},
  chapter = {21},
  pages = {255-266},
  keywords = {Information visualization.
Visual analytics.
Visualization Data processing.
Data Visualization.},
  ISBN = {9781492031086
1492031089},
  year = {2019},
  type = {Book Section}
}
```

Supported Citation Sources

- Citations can be inserted from a variety of sources:
  - Bibliography
  - DOI (Document Object Identifier) references
  - Searches of Crossref, DataCite, or PubMed.
  - Zotero personal or group libraries
    - Zotero must be installed locally to detect your library



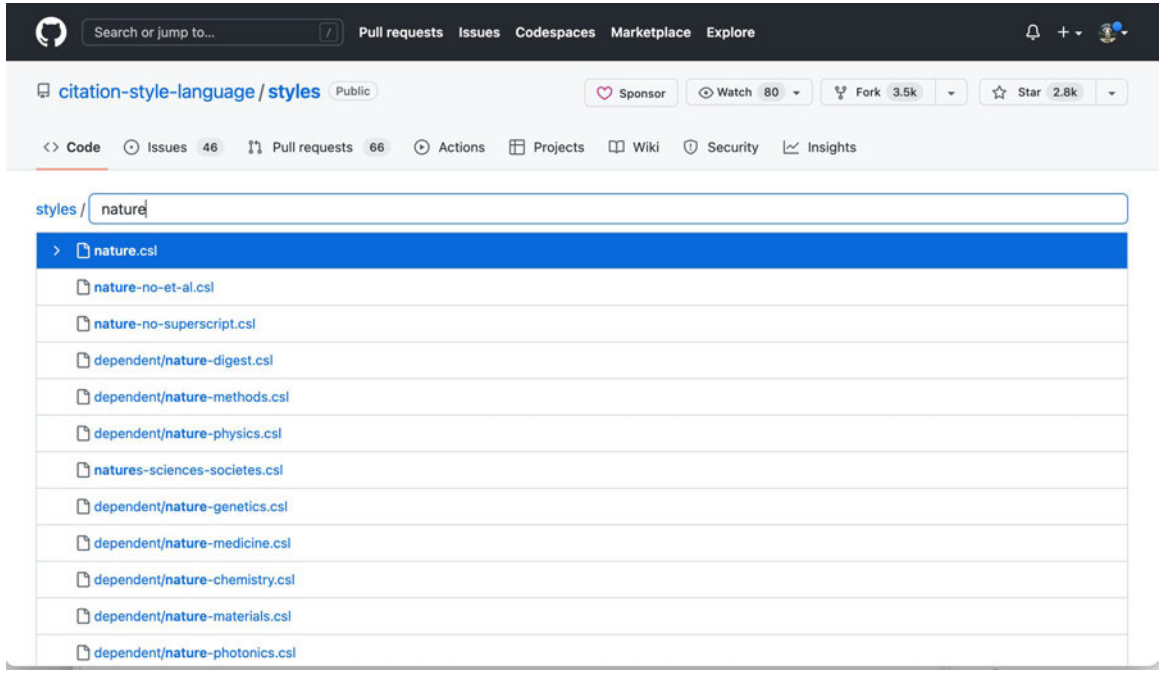
Citation Style

- By default, Pandoc will use the Chicago Manual of Style format
- You can specify a custom formatting using CSL (Citation Style Language)
- You must provide a path to a CSL file using the `csl` metadata field in your document, for example:

```
---
title: "My Document"
bibliography: references.bib
csl: nature.csl
---
```

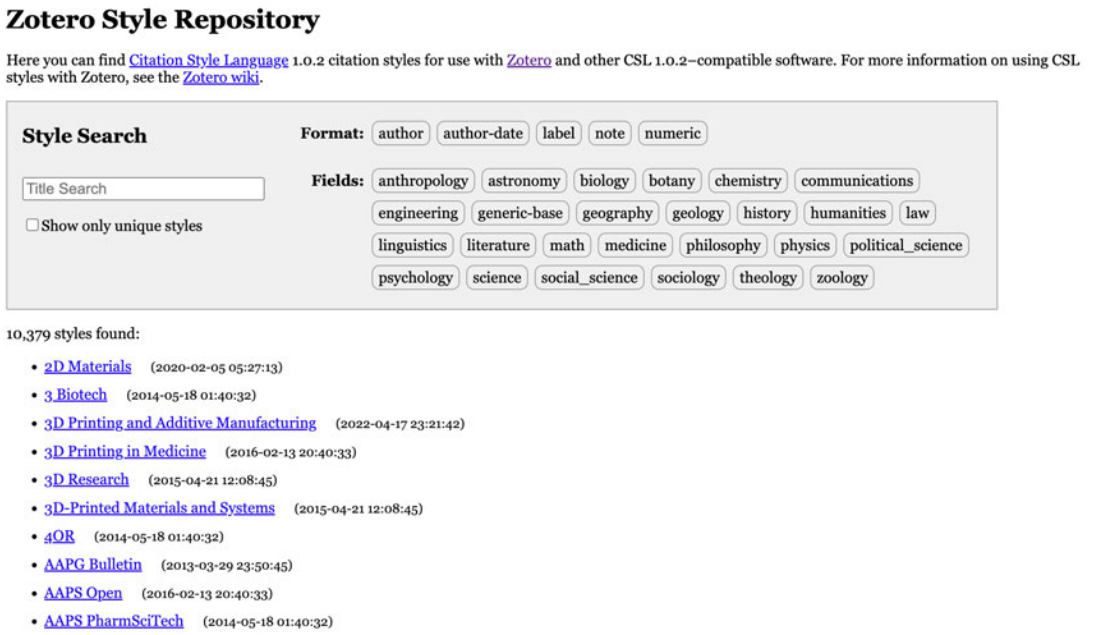
Citation Style: CSL Project

- CSL Project has a list of more than 8,500 Creative Commons CSL definitions
- Search the GitHub CSL style repository by file name
- Press “t” to activate GitHub’s File Finder and start typing the name of the style
- Download a style by clicking on the style’s file name, then click the “Raw” button
- Can also clone or copy repo



Citation Style: Zotero Style Repository

- Zotero also maintains a [style repository]
- Zotero repository is for adding a particular style to Zotero desktop (if Zotero is open)
- Zotero repository will download the file (if Zotero is not open)



Zotero Style Repository

Here you can find Citation Style Language 1.0.2 citation styles for use with Zotero and other CSL 1.0.2–compatible software. For more information on using CSL styles with Zotero, see the Zotero wiki.

Style Search

Title Search

Show only unique styles

Format: author | author-date | label | note | numeric

Fields: anthropology | astronomy | biology | botany | chemistry | communications | engineering | generic-base | geography | geology | history | humanities | law | linguistics | literature | math | medicine | philosophy | physics | political_science | psychology | science | social_science | sociology | theology | zoology

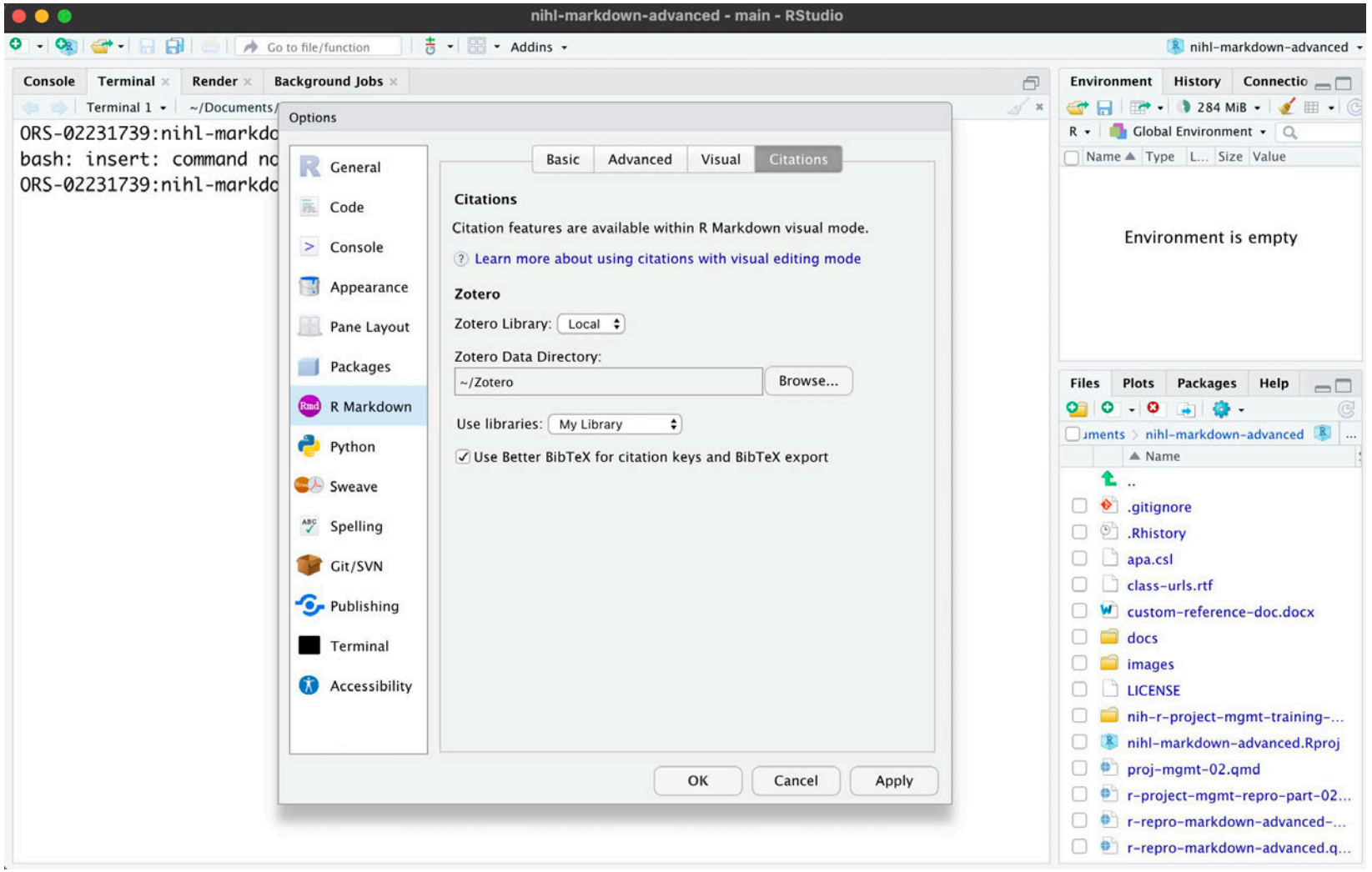
10,379 styles found:

- 2D Materials (2020-02-05 05:27:13)
- 3 Biotech (2014-05-18 01:40:32)
- 3D Printing and Additive Manufacturing (2022-04-17 23:21:42)
- 3D Printing in Medicine (2016-02-13 20:40:33)
- 3D Research (2015-04-21 12:08:45)
- 3D-Printed Materials and Systems (2015-04-21 12:08:45)
- 4OR (2014-05-18 01:40:32)
- AAPG Bulletin (2013-03-29 23:50:45)
- AAPS Open (2016-02-13 20:40:33)
- AAPS PharmSciTech (2014-05-18 01:40:32)

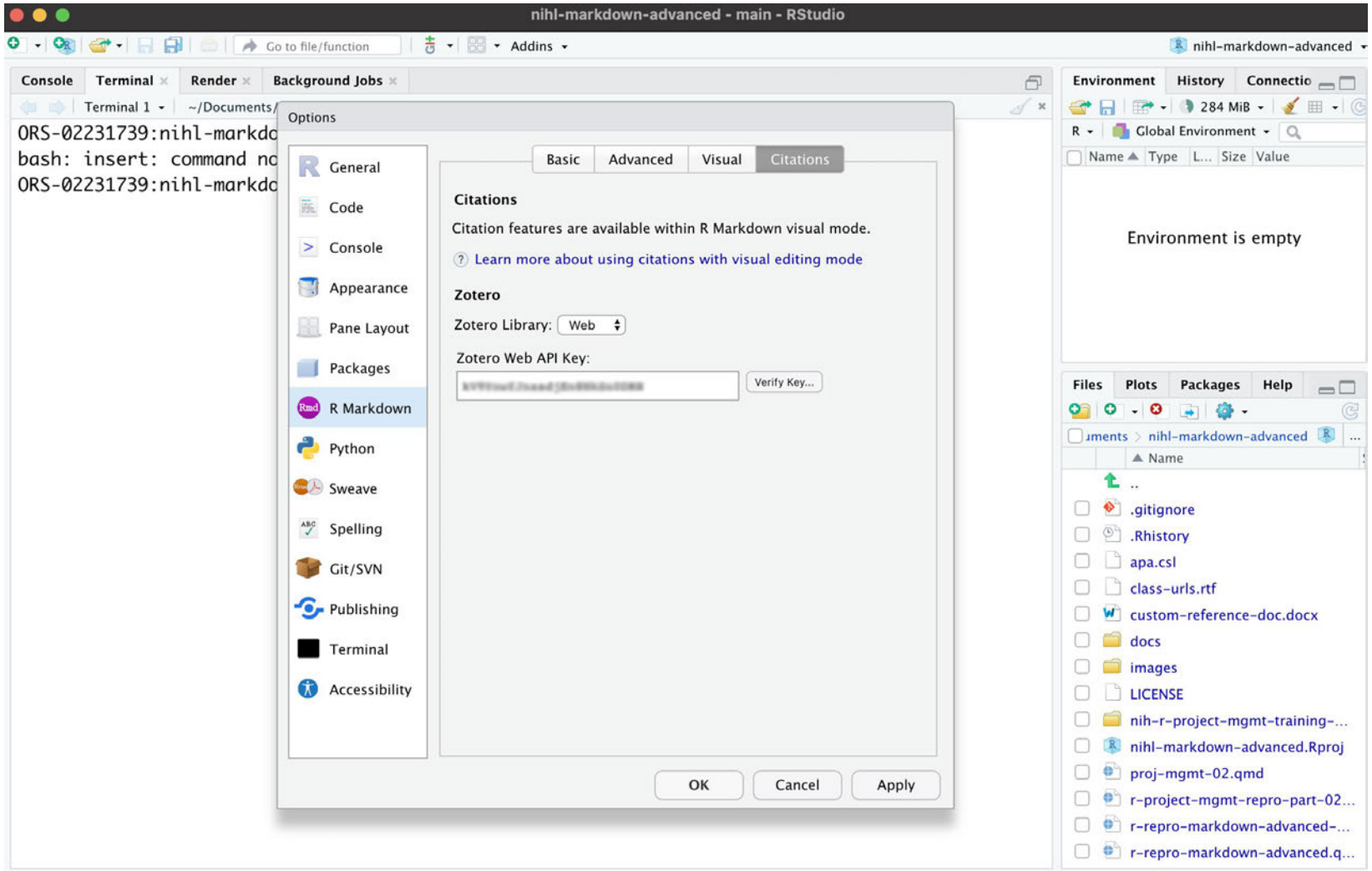
Linking to Citation Manger

- Zotero is a free and open-source reference manager
- Zotero must be installed locally to use it
- RStudio will automatically link to Zotero
- Zotero settings are accessed from *Options>R Markdown>Citations*
- There are two library options:
  - Local option pulls from your computer
  - Web option requires that you generate an API Key

Zotero: Local Option



Zotero: Web Option

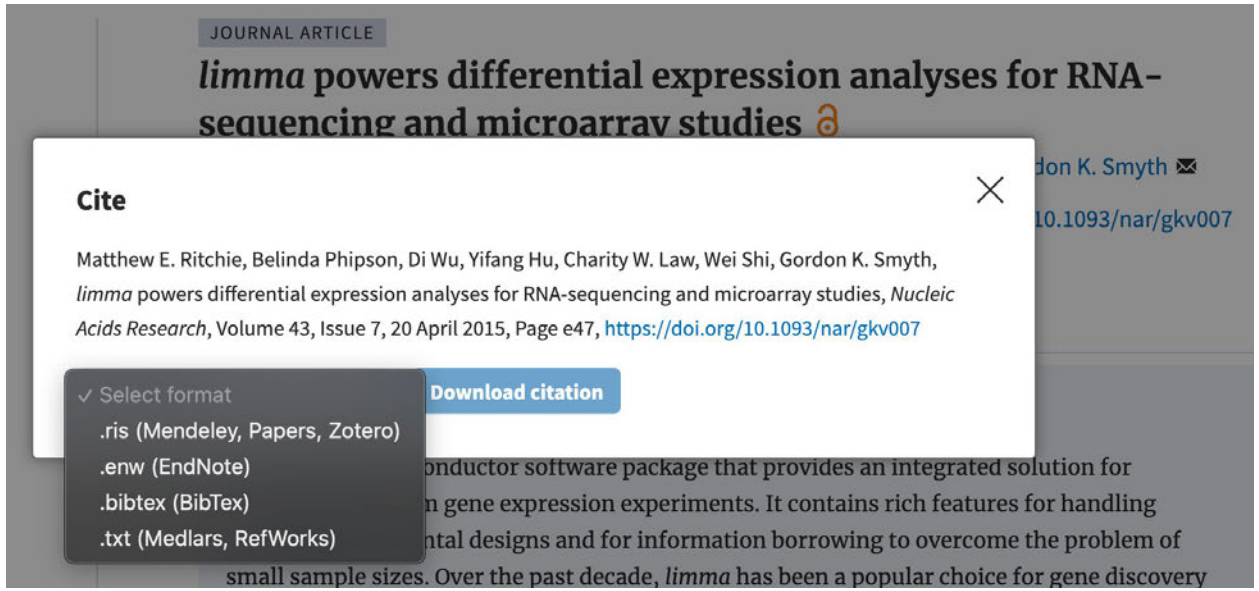


Creating Your Reference List

- Need a list of references saved to a .bib file before you can insert citations into your markdown document
- File with the .bib file extension is a specially formatted text file that lists references
- Ways to create your bib file.
  - Manually
  - Export citations using Endnote or Zotero (citation managers)
  - Use lookup feature to search for publications by DOI (Digital Object Identifier), Crossref, DataCite, or PubMed ID

Bib Files

- Most citation and reference management tools allow you to export citations as .bib BibteX files
- Some scientific databases and journals also allow you to export citations as .bib file

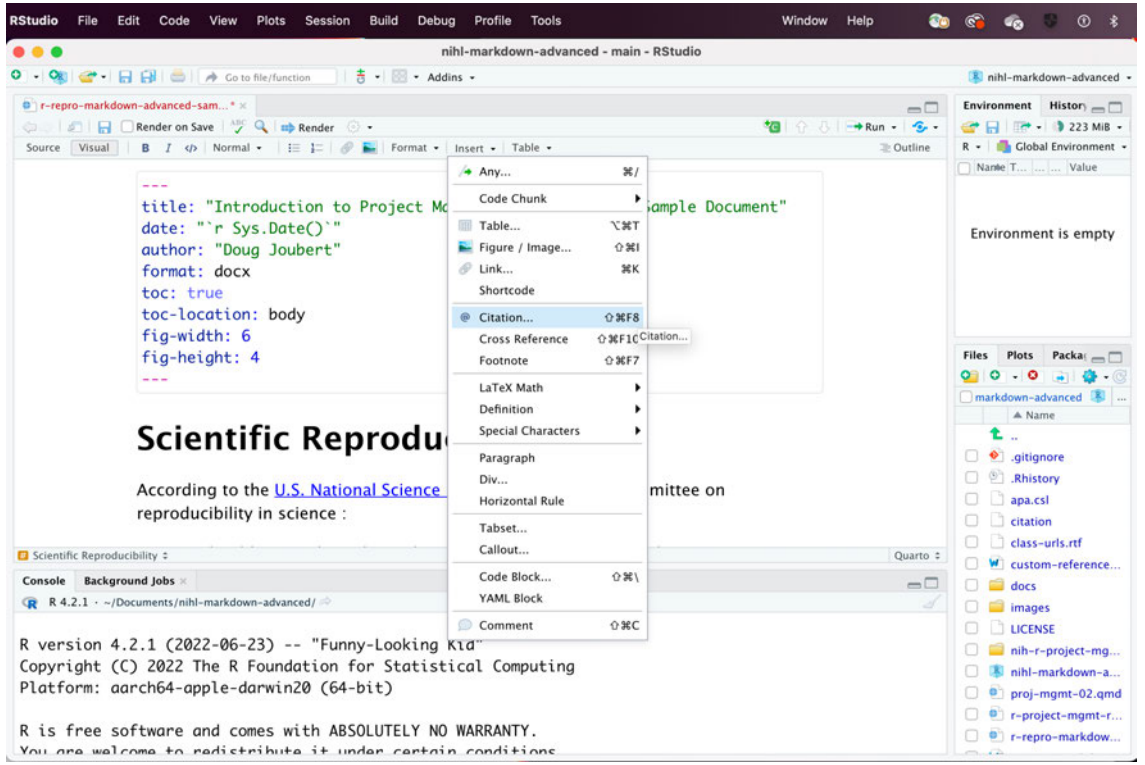


Inserting Citations

Inserting Citations

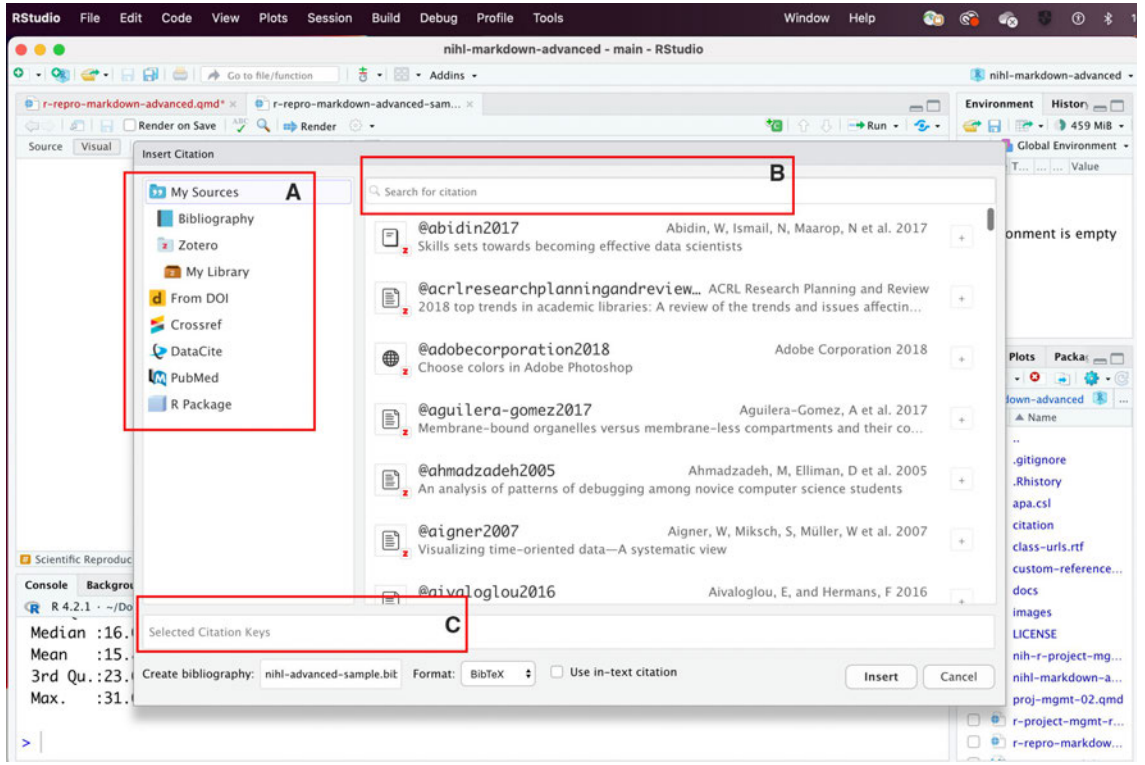
- Using the *Insert -> Citation* command
- Use markdown syntax directly:

```
[@cite] or @cite
```



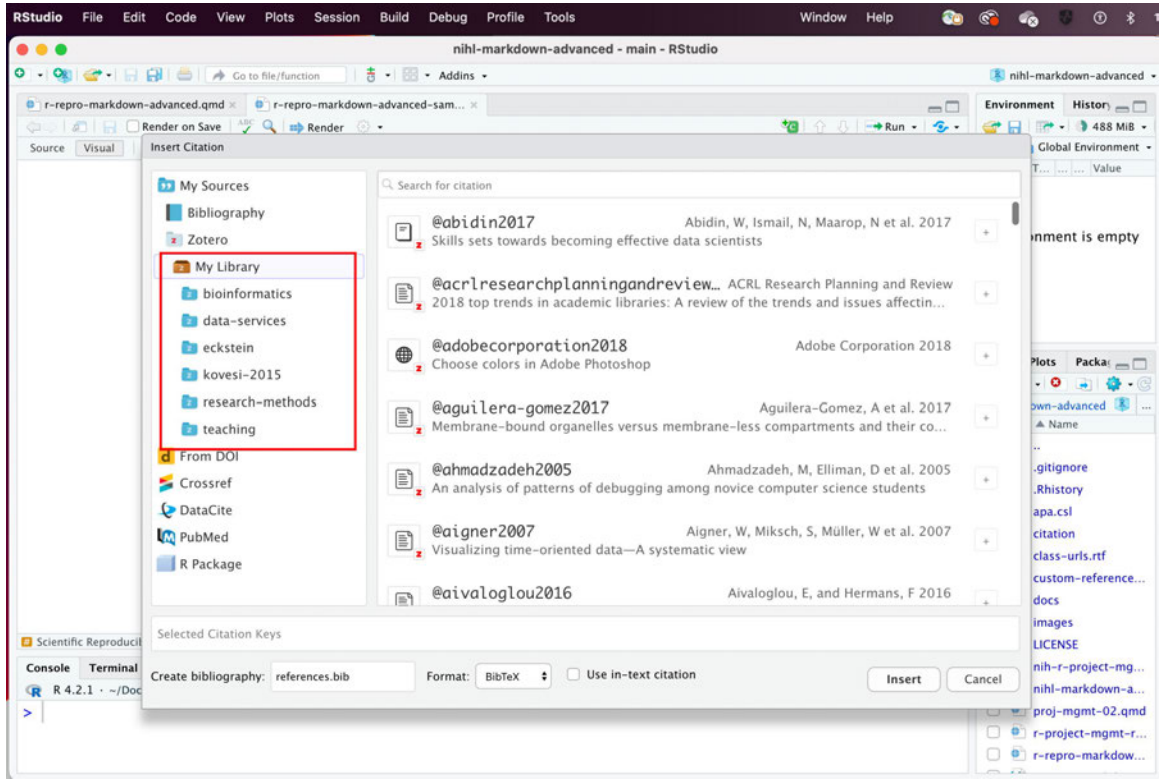
Citation Search Options

- Insert Citation options:
  - Your document bibliography
  - Zotero personal or group libraries
  - DOI (Document Object Identifier) references
  - Searches of Crossref, DataCite, or PubMed
- Zotero citations work the Better BibTeX plugin to generate citation IDs
  - Enabled via Citation Options in Better BibTeX



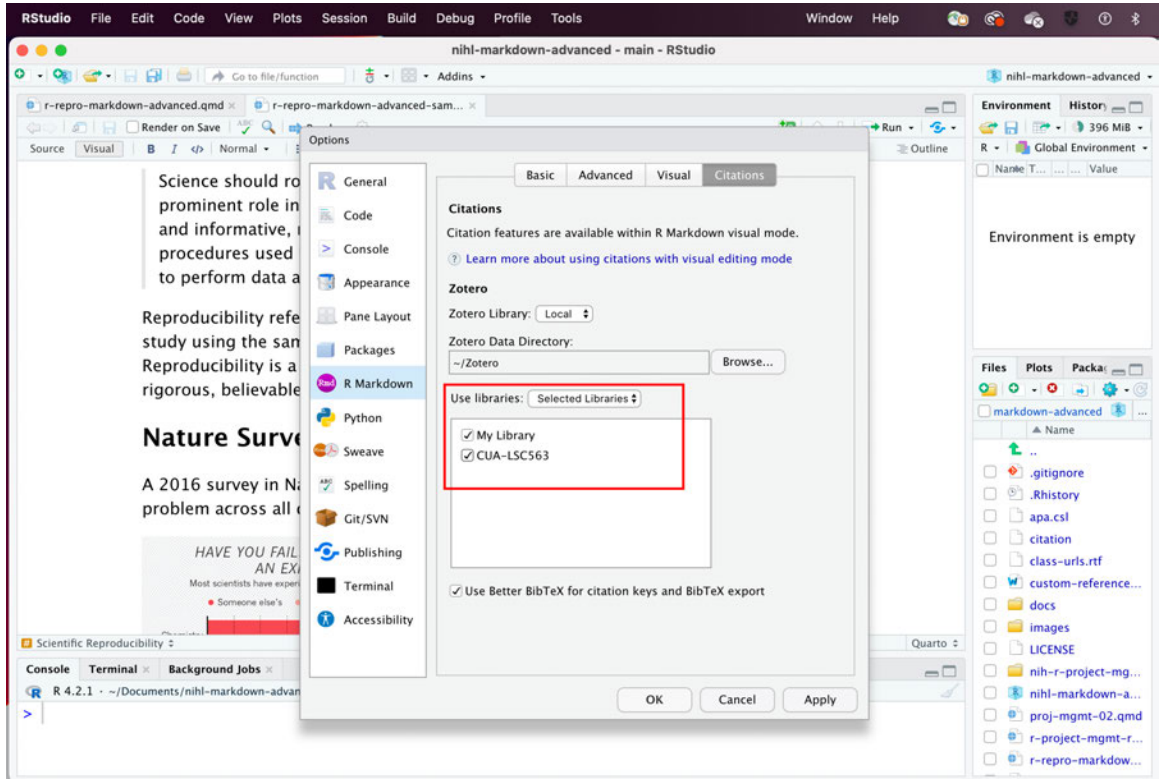
Zotero Library Options

- With Zotero, you can also insert citations directly from your Zotero libraries
- If you have Zotero installed locally My Library will be available
- The groups listed in this image are part of my Zotero library



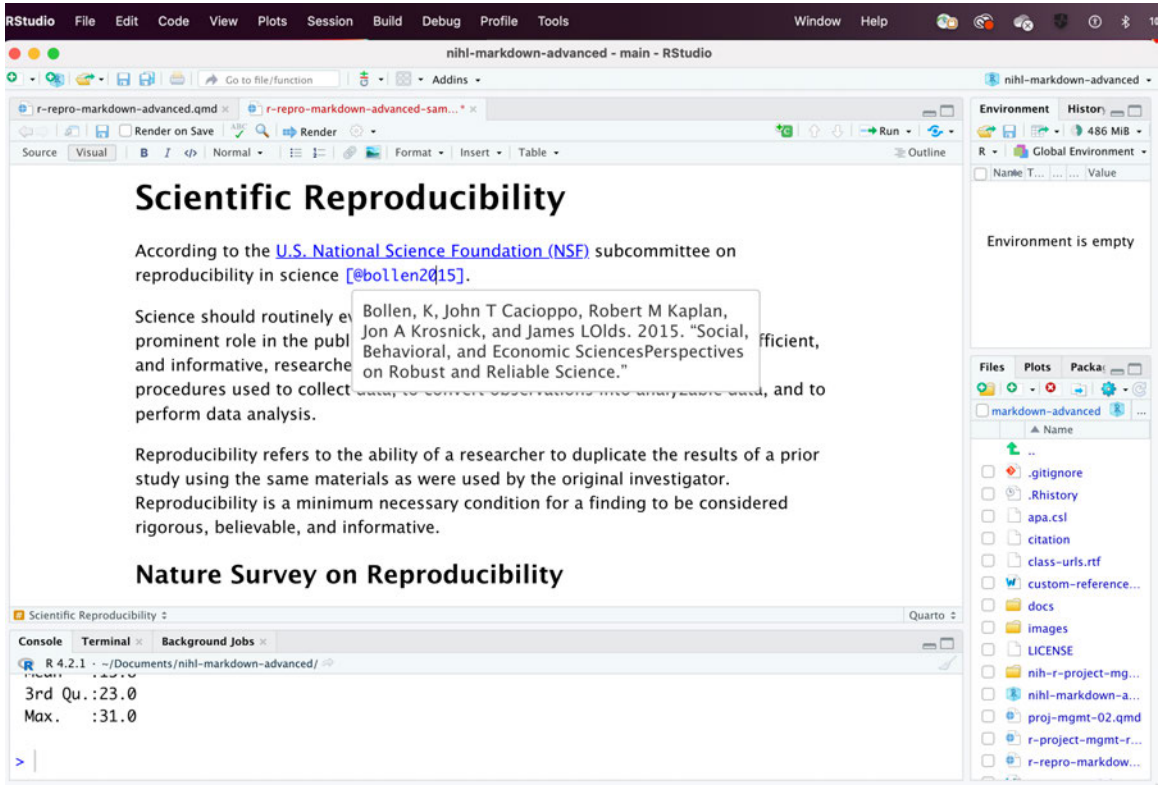
Zotero Library Options: Groups

- Zotero Groups provide a powerful way to share collections
- By default, Zotero Group Libraries are not included in the Insert Citation dialog or citation completions
- Groups can be added in *Global Options>R Markdown*



Citation Preview

- Available after you insert a citation
- Place the cursor over citation to see a preview, with a link to the source if one is available



Bibliography Generation

Creating a Bibliography

- Pandoc will automatically generate a list of works cited
- It will be placed in a div with the id refs if one exists:

```
### References

::: {#refs}
:::
```

- If no div is found, bib will be placed at the end of the document

Figure 21: Head length versus body mass for 123 blue jays. Data source: Keith Tarvin, Oberlin College

The birds' sex is indicated by color, and the birds' skull size by symbol size. Head-length measurements include the length of the bill while skull-size measurements do not. Head length and skull size tend to be correlated, but there are some birds with unusually long or short bills given their skull size.

Tip: Guidance on how to only display certain legends. You can change the position of the legend using this resource and this resource.

Bubble charts have the disadvantage that they show the same types of variables, quantitative variables, with two different types of scales, position and size. As we have already learned, this makes it difficult to visually ascertain the strengths of associations between the various variables. Moreover, differences between data values encoded as bubble size are harder to perceive than differences between data values encoded as position.

Like with Proportional Area Charts, the sizes of the circles need to be drawn based on the circle's area, not its radius or diameter. Not only will the size of the circles change exponentially, but this will lead to misinterpretations by the human visual system.

References

Cleveland, W. S., McGill, M. E., & McGill, R. (1988). The shape parameter of a two-variable graph. *Journal of the American Statistical Association*, *83*(402), 289–300. https://doi.org/10.2307/2288843

Heiss, A. (2021). Relationships. In *Data Visualization*. https://datavizs21.classes.andrewheiss.com/content/07-content/.

Lane, D. M. (2007a). Describing bivariate data. In D. M. Lane & R. Guerra (Eds.), *Introduction to statistics* (pp. 164–184).

Lane, D. M. (Ed.). (2007b). Regression. In *Introduction to Statistics*.

Ribecca, S. (2016). Bubble Chart. In *The Data Visualisation Catalogue*. https://datavizcatalogue.com/methods/bubble_chart.html.

Sarikaya, A., & Gleicher, M. (2018). Scatterplots: Tasks, data, and designs. *IEEE Trans Vis Comput Graph*, *24*(1), 402–412. https://doi.org/10.1109/TVCG.2017.2744184

Triola, M. F., & Lossi, L. (2018). *Exploring data with tables and graphs* (13th edition., pp. 40–79). Pearson.

Wickham, H., & Sievert, C. (2016). Toolbox. In *Ggplot2 : Elegant Graphics for Data Analysis* (1st ed.). Springer International Publishing AG.

Including Uncited Items

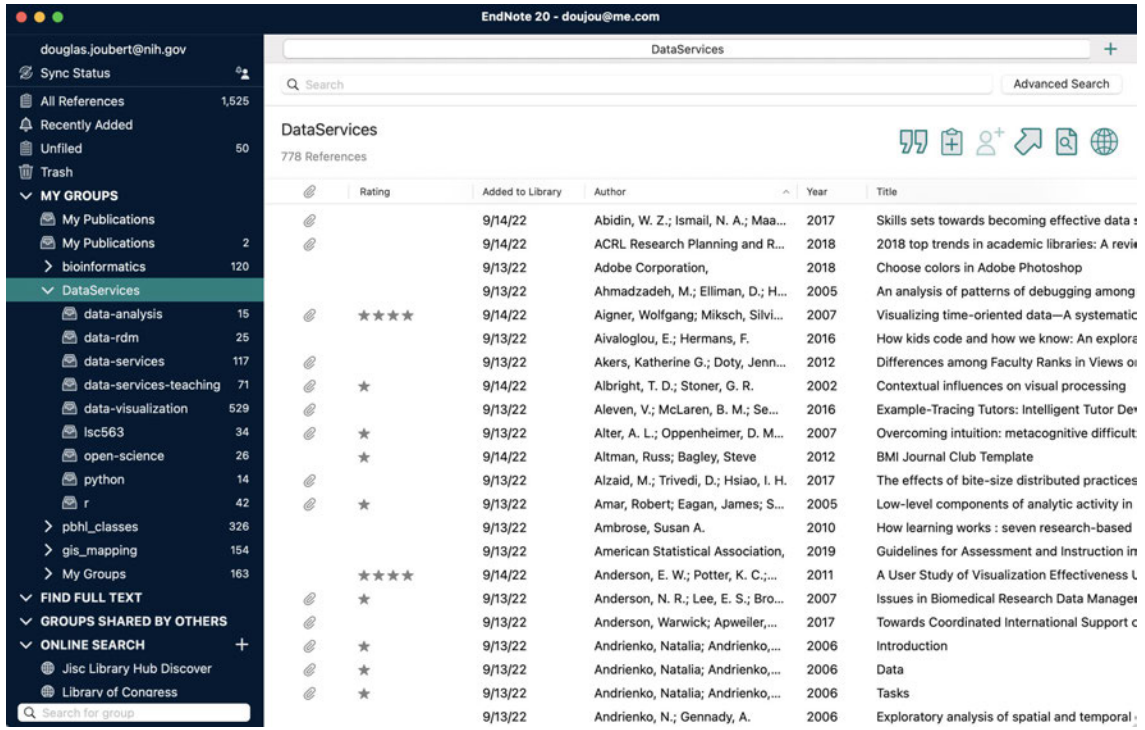
- If you want to include items in the bibliography without citing them in the body text:

```
---
nocite: |
  @wilkeVisualizingAssociationsTwo2019,
@sarikayaScatterplotsTasksData2018
---

@wilsonGoodEnoughPractices2017
```

Exporting References from Endnote

- PowerPoint has information on exporting references from Endnote
- Two ways to export references:
  - Copy and paste into a text document
  - Using the Endnote Export feature
- NIH Library maintains an Endnote [subject guide]



Please Fill Out Our Training Survey

- Helps me become a better teacher
- Helps me identify training gap
- Gives you an opportunity to suggest training

Survey Link

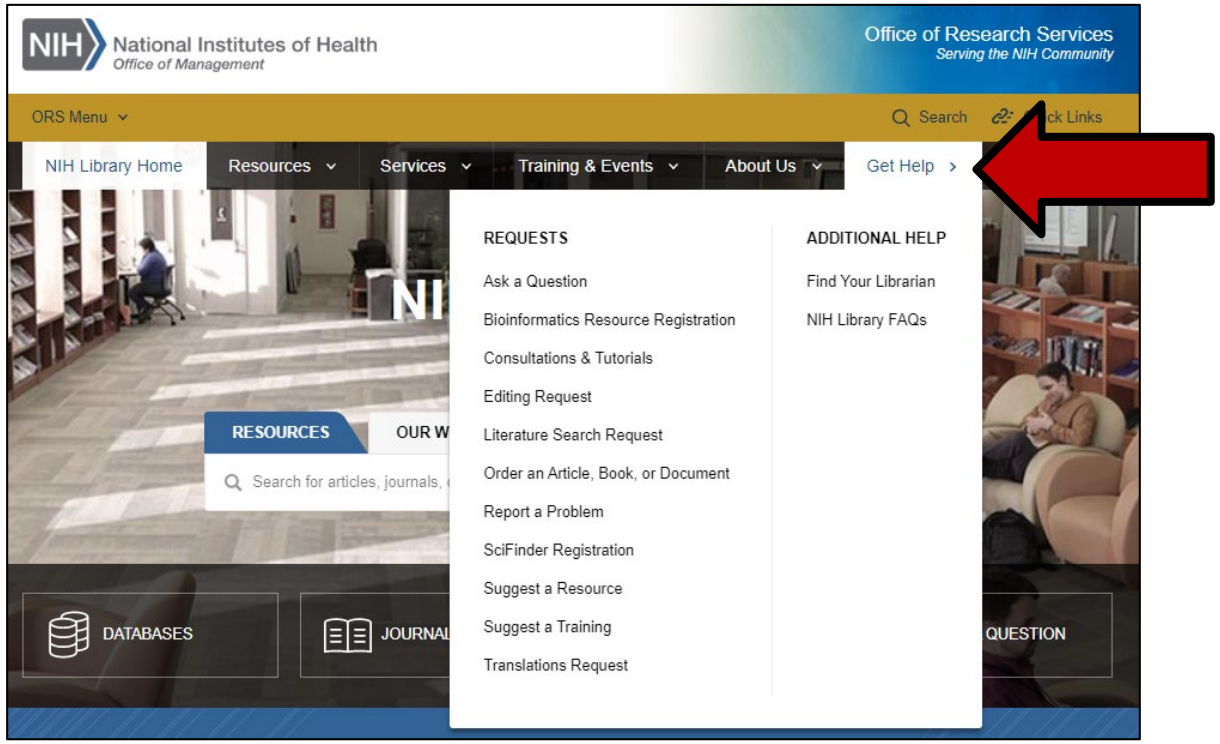
We Can Help





- 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



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

Questions & Comments

National Institutes of Health
Office of Management