

# Intro to R

R Studio & R Markdown

Cultivate Learning Innovation Lab Workshop | July 13, 2020 | Monday

[황보 민] Min Hwangbo

PhD Candidate in Learning Sciences & Human Development  
Graduate Certificate Recipient in Demographic Methods

# Learning Agenda

- Why do social scientists use R?
- Why do I prefer using R Studio & R Markdown (RMD)?
- How to use RMD

---

# [Intro] Inspiration

- Land

- *“The University of Washington & Cultivate Learning acknowledges that it sits on Indigenous Land, which touches the shared waters of all tribes and bands within the Duwamish, Suquamish, Tulalip, and Muckleshoot Tribes.”*

- People

- [Aimée Dechter](#) | Affiliate Assistant Professor & Former Research Coordinator at [Center for Studies in Demography & Ecology \(CSDE\)](#)
- [Chuck] [Charles C. Lanfear](#) | PhD Candidate & R Guru | [2020 Distinguished Teaching Award Recipient](#)
- [Jose] [Jose Hernandez](#) | Data Science Fellow & Research Staff @ eScience Institute
- [Liz] [Elizabeth Sanders](#) | Associate Professor & Quantitative Researcher @ College of Education
- [本橋智光] [Motohashi, Tomomitsu](#). (2018). Maeshoritaizen data bunseki no tame no SQL/R/Python jissen technique ([데이터 전처리 대전](#). 2019).
- [Nicolas] [Nicolas Pröllochs](#) | Tenure-track Professor of Data Science in University of Giessen & Social Network Analysis / Text Mining Expert
- [이근열] Keun Yeol, Lee. | Professor in Busan National University & Qualitative/Dialect/Linguistics Researcher

# Learning Agenda

- Why do social scientists use R?
- Why do I prefer using R Studio & R Markdown (RMD)?
- How to use RMD

---

# [Intro] R?

- A language and environment for **statistical computing** and **graphics**
- Available as a **free** software
- Tool for **reproducible** research practices
- Has a **learning community** (i.e. [Stack OverFlow](#))
- **Easier** to learn compared to SQL or Python
- **Great skill for your current/future personal & career pathway!**

# [Intro] R Studio?

- “Front-end” aka: Integrated Development Environment (IDE) for R
- It functions as a library catalog
  - Organize your code, output, and plots
- Automation
  - Auto-complete code and highlight syntax
- See your result right away!
  - Help view data and objects
- **Enable easy integration of R code into documents!!!**
  - HTML
  - PDF
  - Slides
  - Packages ([LaTeX](#) & [Bookdown](#)) for dissertation?!!

# [Prep] Preparation

- I'm ready! Let's use R Studio!
  - Wait... before you jump in....
- File naming convention
- Folder structure
- Download [R Studio](#)

# [Prep] File Naming Convention (FNC)

- Framework/System for **naming** your **files** in a way that describes what they contain and how they relate to other **files**
- **FileName\_ProjectAcronym\_DateCreatedinMMDDYY**
- **i.e. Workshop1\_INL\_071320**
  - IntroRWkshop: Intro to R workshop file
  - INL: Innovation Lab
  - 071320: July 13, 2020
- **Project Acronyms**
  - Professional Learning & Coaching: **PLC**
    - formerly Coaching & PD: **CPD**
  - Circle Time Magazine: **CTM**
  - Meaningful Makeover: **MM**

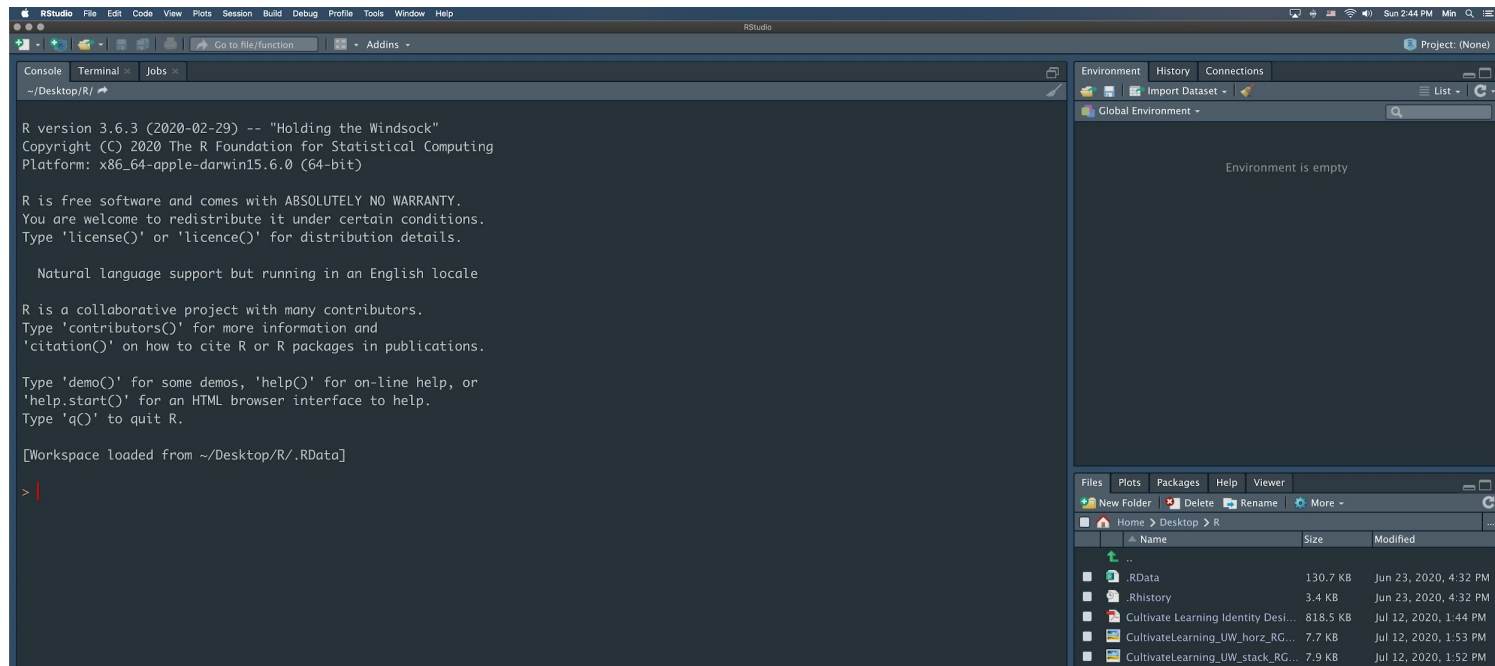


# [Prep] Folder Structure

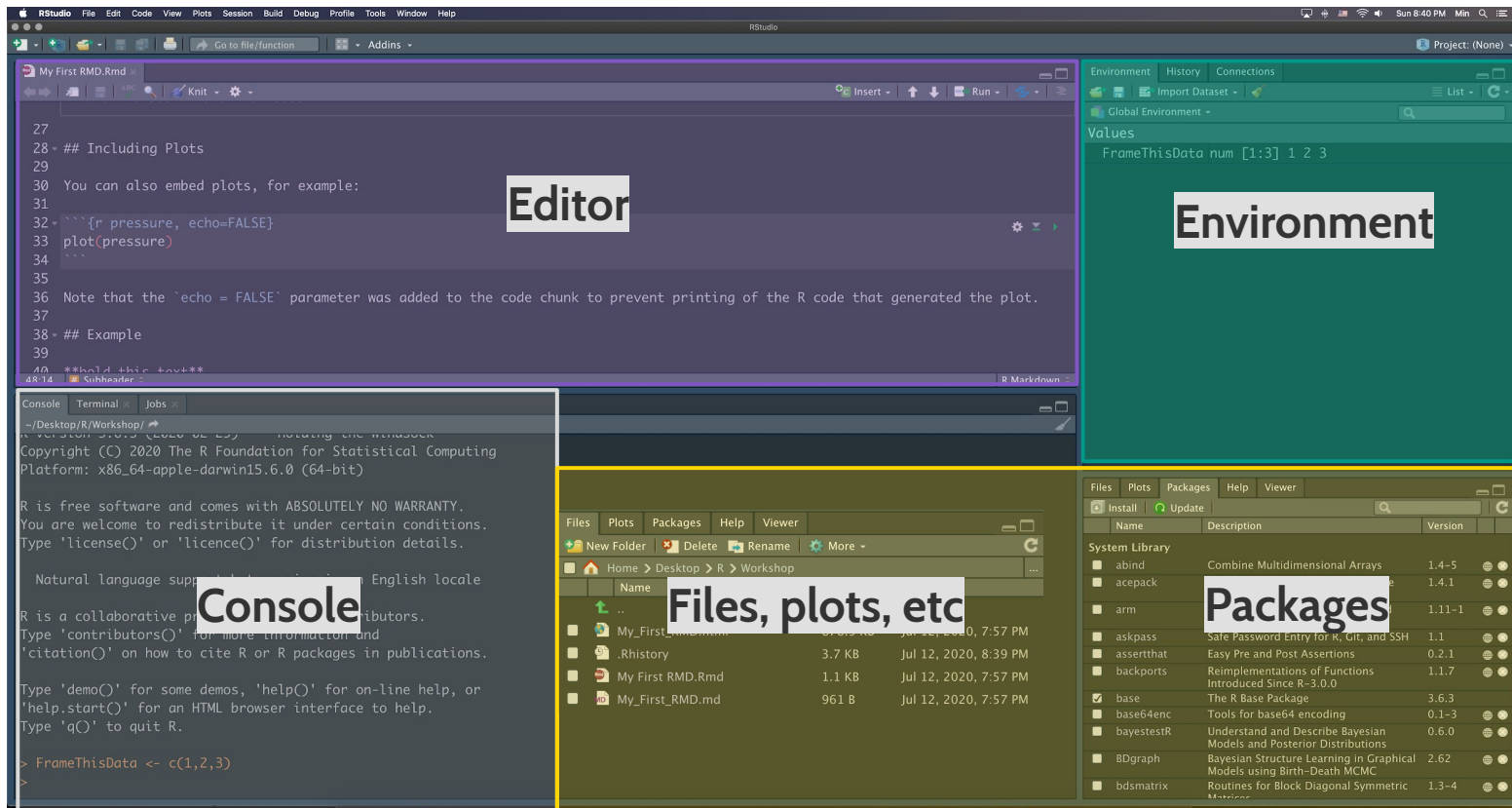
- Cultivate Learning protocol on C-Dash (Cloud server for archiving research relevant data)
  - Data
  - Output
  - Source
  - Documents
- Current Project
  - Create your own R folder on your desktop
    - R
      - Week 1: Intro to R
        - Data (where raw data would live)
        - Output (where you would save your RMD file)

**Questions?**

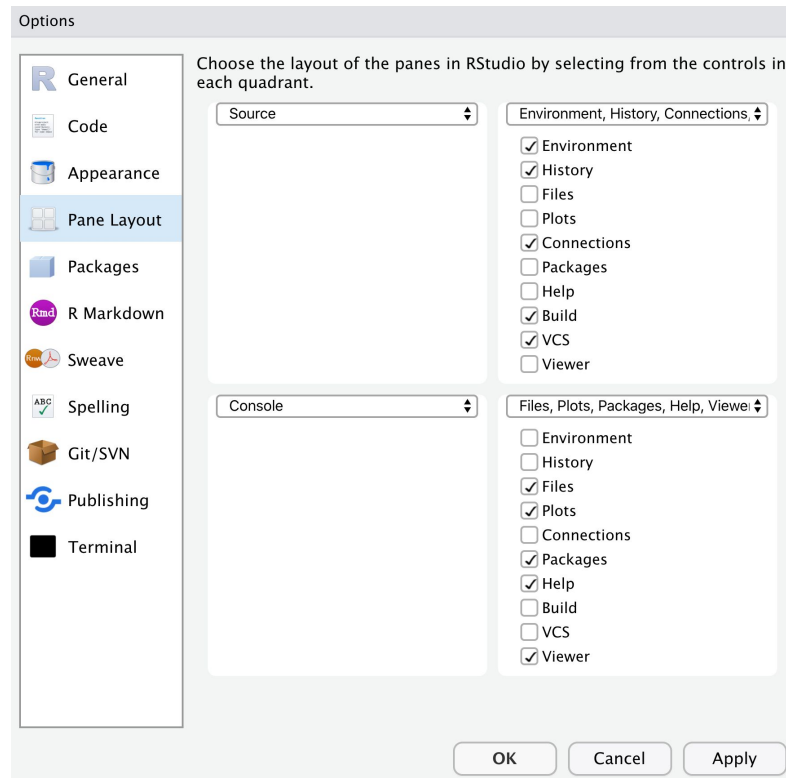
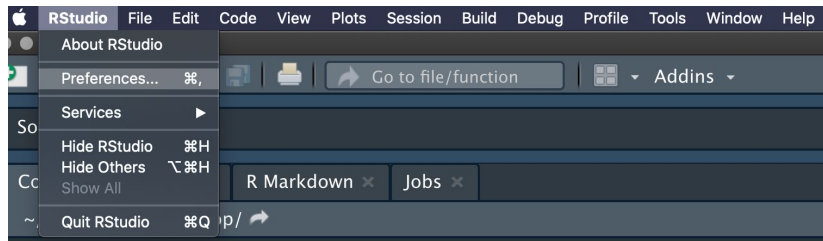
# [R Studio] Integrated Development Environment



# [R Studio] IDE Feature



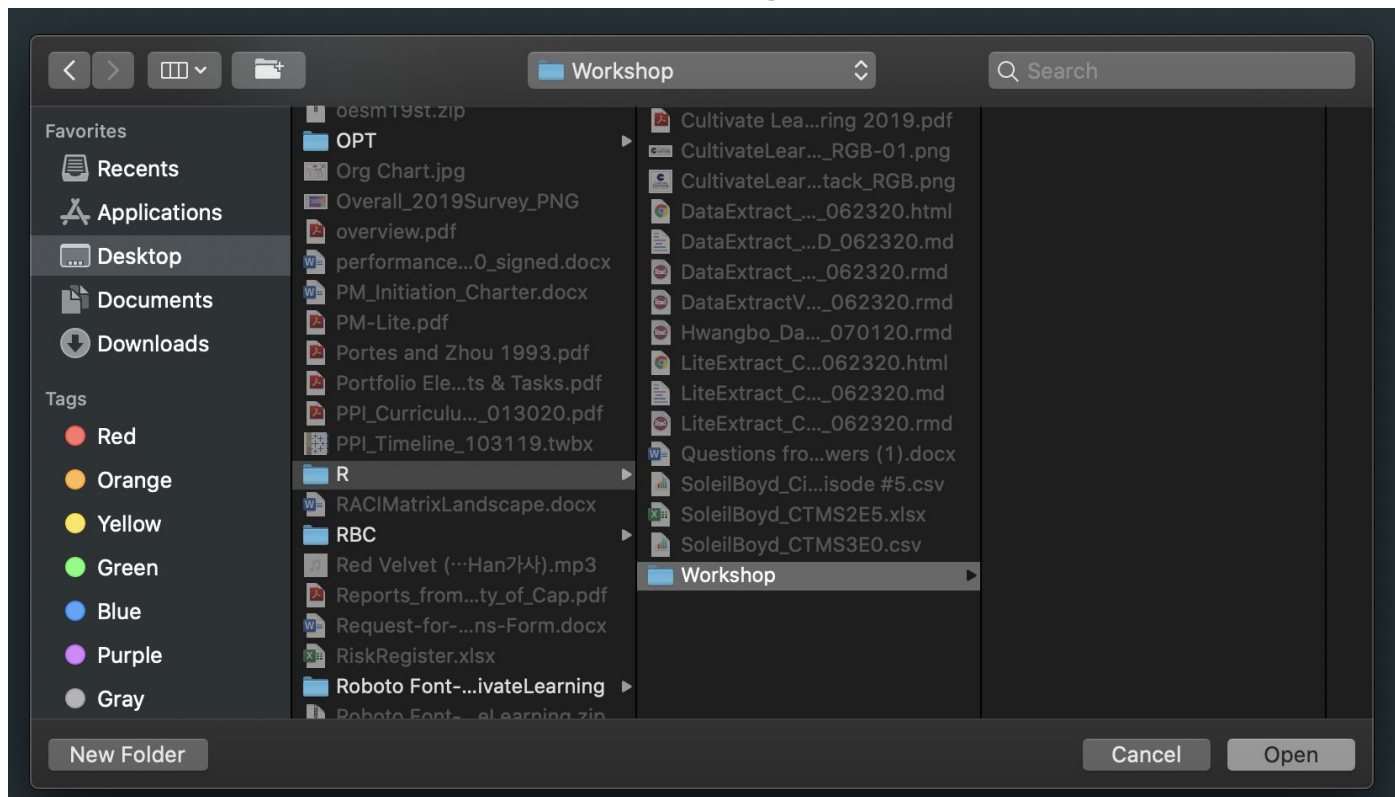
# [R Studio] Changing IDE layout



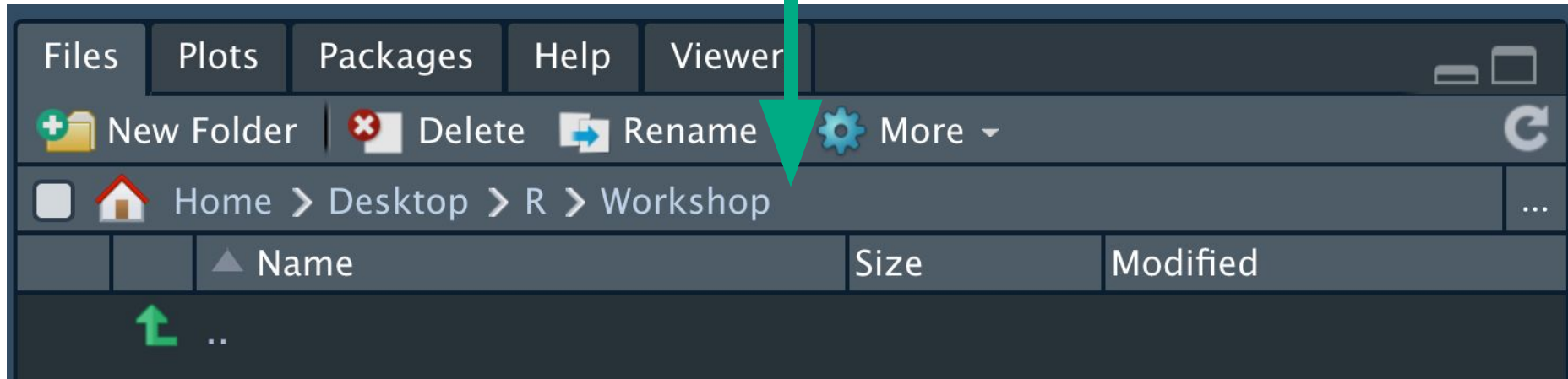
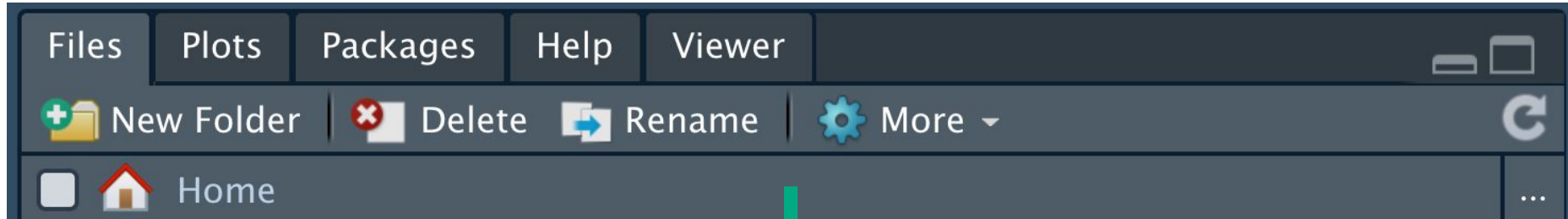
# [R Studio] Setting Directory aka *Building Foundation*



# [R Studio] Setting Directory: Option 1

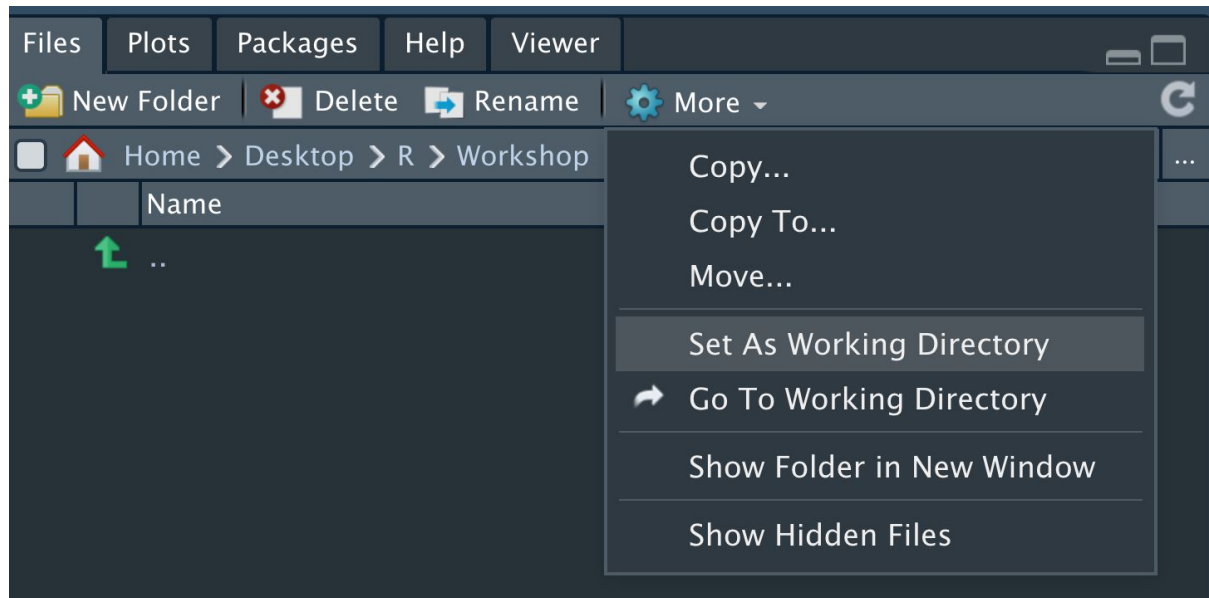


# [R Studio] Setting Directory: Option 1



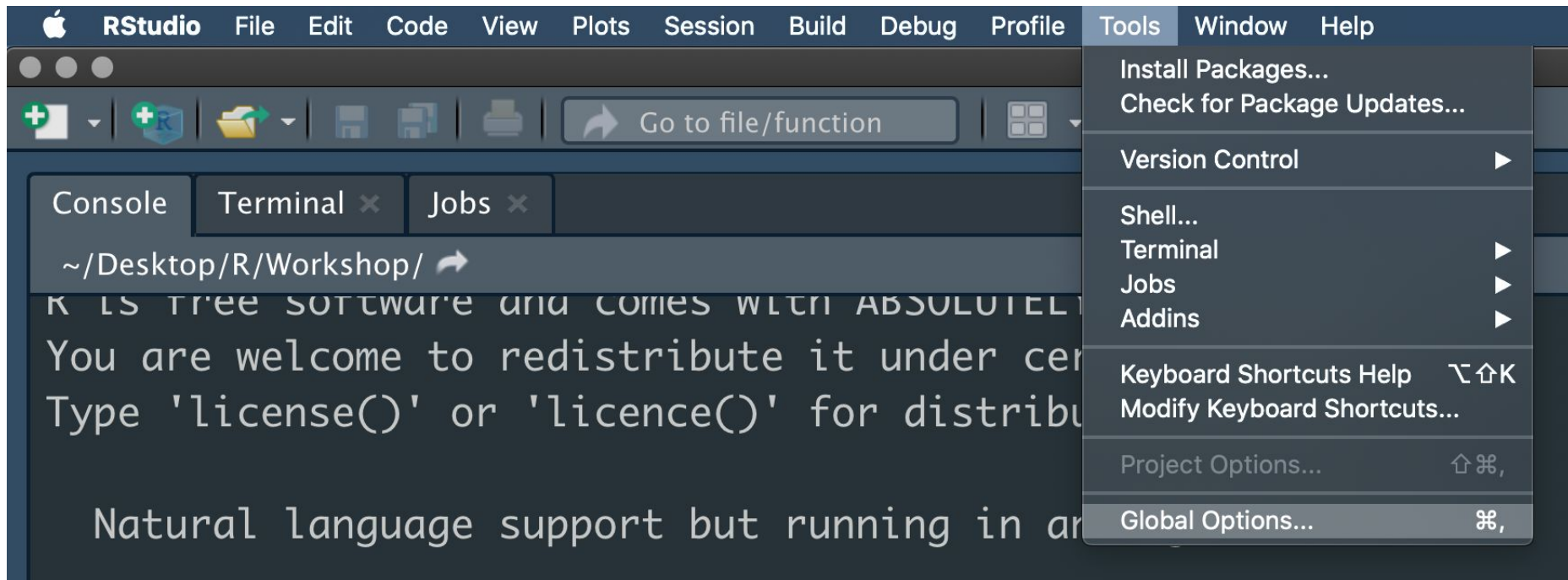


# [R Studio] Setting Directory: Option 1

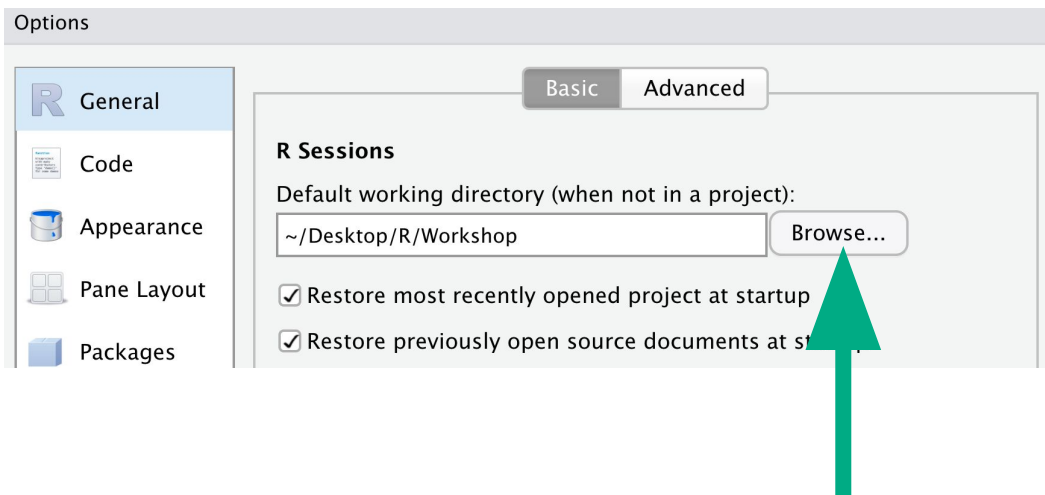


```
> setwd("~/Desktop/R/Workshop")
```

# [R Studio] Setting Directory: Option 2



# [R Studio] Setting Directory: Option 2



**Questions?**

**10 min Break**

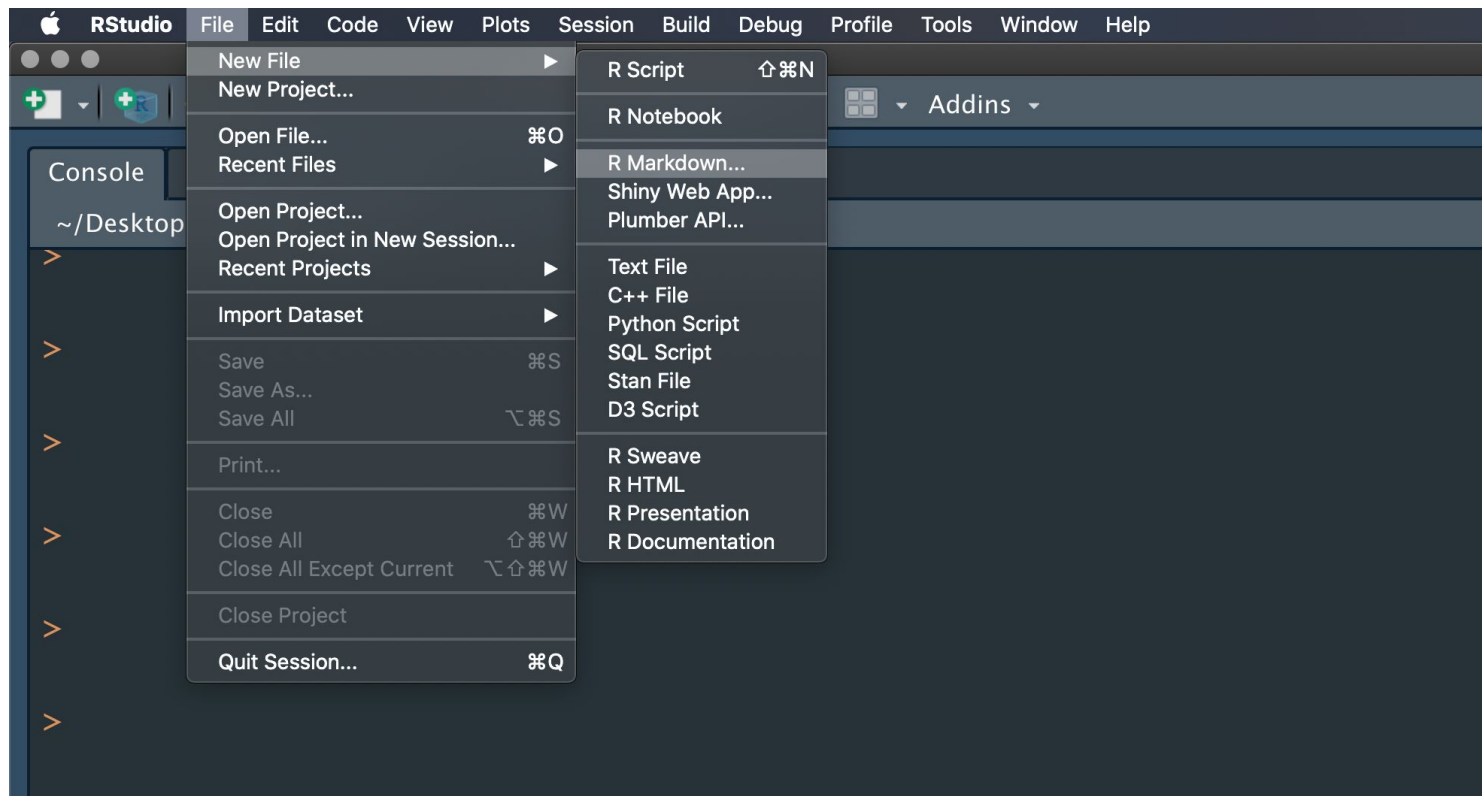
**Coffee, Tea &  
Snack Time**

# Learning Agenda

- Why do social scientists use R?
- Why do I prefer using R Studio & R Markdown (RMD)?
- How to use RMD


---


# [RMD] R Markdown





# [RMD] R Markdown

New R Markdown

 Document

 Presentation

 Shiny

 From Template

**Title:**

**Author:**

**Default Output Format:**

☒ **HTML**  
Recommended format for authoring (you can switch to PDF or Word output anytime).

☐ **PDF**  
PDF output requires TeX (MiKTeX on Windows, MacTeX 2013+ on OS X, TeX Live 2013+ on Linux).

☐ **Word**  
Previewing Word documents requires an installation of MS Word (or Libre/Open Office on Linux).

OK

Cancel



# [RMD] R Markdown

```
1 ---
2 title: "My First RMD"
3 author: "Min Hwangbo"
4 date: "7/12/2020"
5 output: html_document
6 ---
7
8 ```{r setup, include=FALSE}
9 knitr::opts_chunk$set(echo = TRUE)
10 ```
11
12 ## R Markdown
13
14 This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more
15 details on using R Markdown see <http://rmarkdown.rstudio.com>.
16
17 When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded
18 R code chunks within the document. You can embed an R code chunk like this:
19
20 ```{r cars}
21 summary(cars)
22 ```
23
24 ## Including Plots
25
26 You can also embed plots, for example:
27
28 ```{r pressure, echo=FALSE}
29 plot(pressure)
30 ```
31
32 Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.
```

← Headers

← Chunks

← Body

# [RMD] R Markdown: Syntax (Body)

## Syntax

**\*\*bold this text\*\***

*\*italicize this text\**

**# Header**

**## Subheader**

**> Block quote**

## Output

**bold this text**

*italicize this text*

**Header**

**Subheader**

Block quote

Header

Subheader

# [RMD] Headers

```
1 ---
2 title: "My First RMD"
3 author: "Min Hwangbo"
4 date: "7/12/2020"
5 output: html_document
6 ---
```

- The header of a RMD file: Written in YAML code block
- **Customizable** (i.e. Table of Contents: Min's favorite go-to header)

```
1 ---
2 title: "LiteDataExtract_CTM_062320"
3 author: "Min Hwangbo"
4 date: "6/30/2020"
5 output:
6   html_document:
7     preserve_yaml: true
8     toc: true
9     toc_float: true
10    keep_md: true
11 published: false
12 ---
```

# [RMD] Headers w/ HTML example

```
5 output: html_document
6 ---
```



## My First RMD

Min Hwangbo

7/12/2020

### R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```



```
5 output:
6   html_document:
7     preserve_yaml: true
8   toc: true
9   toc_float: true
10  keep_md: true
11 published: false
12 ---
```

### R Markdown

Including Plots

### R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed      dist
##  Min.   : 4.0   Min.   : 2.00
##  1st Qu.:12.0   1st Qu.: 26.00
```

# [RMD] Chunks aka Code Block

- Code is sandwiched between sets of three backticks ``` and `{r}`.
- Input

```
```{r}  
summary(cars)  
```
```

- Output:

```
24 ```{r}  
25 summary(cars)  
26 ```  
  
      speed      dist  
Min.   : 4.0    Min.   : 2.00  
1st Qu.:12.0    1st Qu.: 26.00  
Median :15.0    Median : 36.00  
Mean   :15.4    Mean   : 42.98  
3rd Qu.:19.0    3rd Qu.: 56.00  
Max.   :25.0    Max.   :120.00
```

# [RMD] R Markdown Chunk Options

```
24 ```{r, echo = F, eval = F, include = F, results = 'hide', cache = T}  
25 summary(cars)  
26 ```
```

- `echo=FALSE`: Keeps R code from being shown in the document
- `eval=FALSE`: Shows R code in the document without running it
- `include=FALSE`: Hides all output but still runs code (good for `setup` chunks where you load packages!)
- `results='hide'`: Hides R's (non-plot) output from the document
- `cache=TRUE`: Saves results of running that chunk so if it takes a while, you won't have to re-run it each time you re-knit the document

# [RMD] R Markdown Chunk Options: Naming

```
24 ▾ ```{r 1st_Cars, echo = F, eval = F, include = F, results = 'hide', cache = T}  
25 summary(cars)  
26 ```
```

| speed        | dist           |
|--------------|----------------|
| Min. : 4.0   | Min. : 2.00    |
| 1st Qu.:12.0 | 1st Qu.: 26.00 |
| Median :15.0 | Median : 36.00 |
| Mean :15.4   | Mean : 42.98   |
| 3rd Qu.:19.0 | 3rd Qu.: 56.00 |
| Max. :25.0   | Max. :120.00   |

```
27  
28 ▾ ## Including Plots
```

```
29  
30 You can also embed plots, for example:
```

```
31  
32 ▾ ```{r pressure, echo=FALSE}  
33 plot(pressure)  
34 ```
```

```
35  
36 Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.  
37
```

# [RMD] R Markdown Cheatsheet

# R Markdown :: CHEAT SHEET

## What is R Markdown?

The diagram illustrates the process of generating dynamic documents from R Markdown files. It starts with a blue folder icon labeled '.Rmd files - An R Markdown (.Rmd) file'. Below it, a document icon shows a code editor window with a red 'Run' button. To the right, a paragraph explains that an Rmd file is a record of research containing code and narration for reproducibility. Below this, another document icon shows a presentation slide with a 'Reproducible Research' title and a 'button' label. A paragraph explains how clicking a button or typing a command can rerun the code and export results as a finished report. At the bottom, a document icon shows various output formats like HTML, PDF, Word, and RTF. A paragraph explains that dynamic documents allow users to choose the format for their finished report.

**.Rmd files - An R Markdown**

(.Rmd) file is a record of your research. It contains the code that a scientist needs to reproduce your work along with the narration that a reader needs to understand your work.

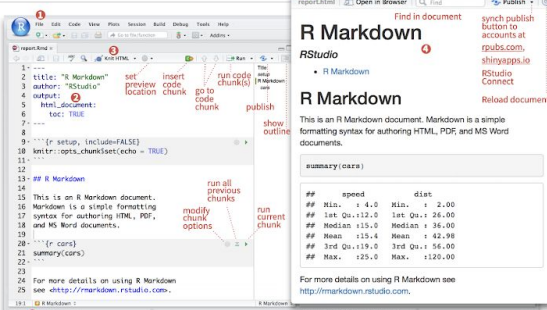
**Reproducible Research** - At the click of a button, or the type of a command, you can rerun the code in an R Markdown file to reproduce your work and export the results as a finished report.

**Dynamic Documents** - You can choose to export the finished report in a variety of formats, including html, pdf, MS Word, or RTF documents; html or pdf based slides. Notebooks, and more.

## Workflow



- 1 **Open a new .Rmd file** at File ► New File ► R Markdown. Use the wizard that opens to pre-populate the file with a template
- 2 **Write document** by editing template
- 3 **Knit document to create report**; use knit button or `render()` to knit
- 4 **Preview Output** in IDE window
- 5 **Publish** (optional) to web server
- 6 **Examine build log** in R Markdown console
- 7 **Use output file** that is saved along side .Rmd



## render

Use `markdown::render()` to render/knit at cmd line. Important args:

|                               |   |                    |                                       |   |                                 |
|-------------------------------|---|--------------------|---------------------------------------|---|---------------------------------|
| <b>input</b> - file to render | <b>output_options</b> - List of render options (as in YAML) | <b>output_file</b> | <b>params</b> - list of params to use | <b>envir</b> - environment to evaluate code chunks in | <b>encoding</b> - of input file |
| <b>output_format</b>          |   | <b>output_dir</b>  |                                       |   |                                 |

### Embed code with knitr syntax

**INLINE CODE**  
Insert with ``r <code>``. Results appear as text without code.  
Built with ``r getRversion()`` Built with 3.2.3

**CODE CHUNKS**  
One or more lines surrounded with ````(r)` and `````. Place chunk options within curly braces, after `r`. Insert with . `getRversion()`

**GLOBAL OPTIONS**  
Set with `knitr::opts_chunk$set()`, e.g.  
````{r include=FALSE}  
knitr::opts\_chunk\$set(echo = TRUE)  
````

## IMPORTANT CHUNK OPTIONS

- cache** - cache results for future knits (default = FALSE)
- cache.path** - directory to save cached results in (default = "cache/")
- child** - file(s) to knit and then include (default = NULL)
- collapse** - collapse all output into single block (default = FALSE)
- comment** - prefix for each line of results (default = "##")

- dependson** - chunk dependencies for caching (default = NULL)
- echo** - Display code in output document (default = TRUE)
- engine** - code language used in chunk (default = 'R')
- error** - Display error messages in doc (TRUE) or stop render when errors occur (FALSE) (default = FALSE)
- eval** - Run code in chunk (default = TRUE)

- fig.align** - 'left', 'right', or 'center' (default = 'default')
- fig.cap** - figure caption as character string (default = NULL)
- fig.height**, **fig.width** - Dimensions of plots in inches
- highlight** - highlight source code (default = TRUE)
- include** - Include chunk in doc after running

**message** - display code messages in document (default = TRUE)  
**results** (default = 'markup')  
 'asis' - passthrough results  
 'hide' - do not display results  
 'hold' - put all results below all code  
**tidy** - tidy code for display (default = FALSE)  
**warning** - display code warnings in document (default = TRUE)

Options not listed above: R.options, aniopts, autodep, background, cache.comments, cache.lazy, cache.rebuild, cache.vars, dev, dev.args, dpi, engine.opts, engine.path, fig.asp, fig.env, fig.fig, fig.keep, fig.lib, fig.path, fig.pos, fig.process, fig.retina, fig.scap, fig.show, fig.showtext, fig.subcap, interval, out.extra, out.height, out.width, prompt, purl, ref.label, render.size, split, tidy.opts

## .rmd Structure

**YAML Header**  
Optional section of render (e.g. pandoc) options written as key:value pairs (YAML).

Between lines of ---

**Text**  
Narration formatted with markdown, mixed with

### Code Chunks

Chunks of embedded code. Each chunk:

Begins with ````{r}`  
 ends with `````  
 R Markdown will run the code and append the results to the doc.  
 It will use the location of the `.Rmd` file as the **working directory**

## Parameters

Parameterize your documents to reuse with new inputs (e.g., data, values, etc.)

1. **Add parameters** - Create and set parameters in the header as sub-values of params
2. **Call parameters** - Call parameter values in code as params<name>
3. **Set parameters** - Set values with Knit with parameters or the params argument of render():  

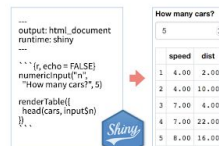
```
render("doc.Rmd", params = list(n = 1,  
d = as.Date("2015-01-01")))
```



## Interactive Documents

Turn your report into an interactive Shiny document in 4 steps

1. Add runtime: shiny to the YAML header.
2. Call Shiny input functions to embed input objects.
3. Call Shiny render functions to embed reactive output.
4. Render with `rmarkdown::run` or click Run Document in RStudio IDE.



Embed a complete app into your document with  
`shiny::shinyAppDir()`

**Publish on RStudio Connect**, to share R Markdown documents securely, schedule automatic updates, and interact with parameters in real time.  
[www.rstudio.com/products/connect/](https://www.rstudio.com/products/connect/)



# Summary

- Anyone can learn R
- Organizational skill -> Programming skill
- No need to reinvent the wheel: *Use codes that are pre-built!*

# Next Steps

1. Create your own RMD file, knit it, then upload it on our Shared drive
2. Next session will be on Data Frame & Data Reshape

**Thank you!**

감사합니다!