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)

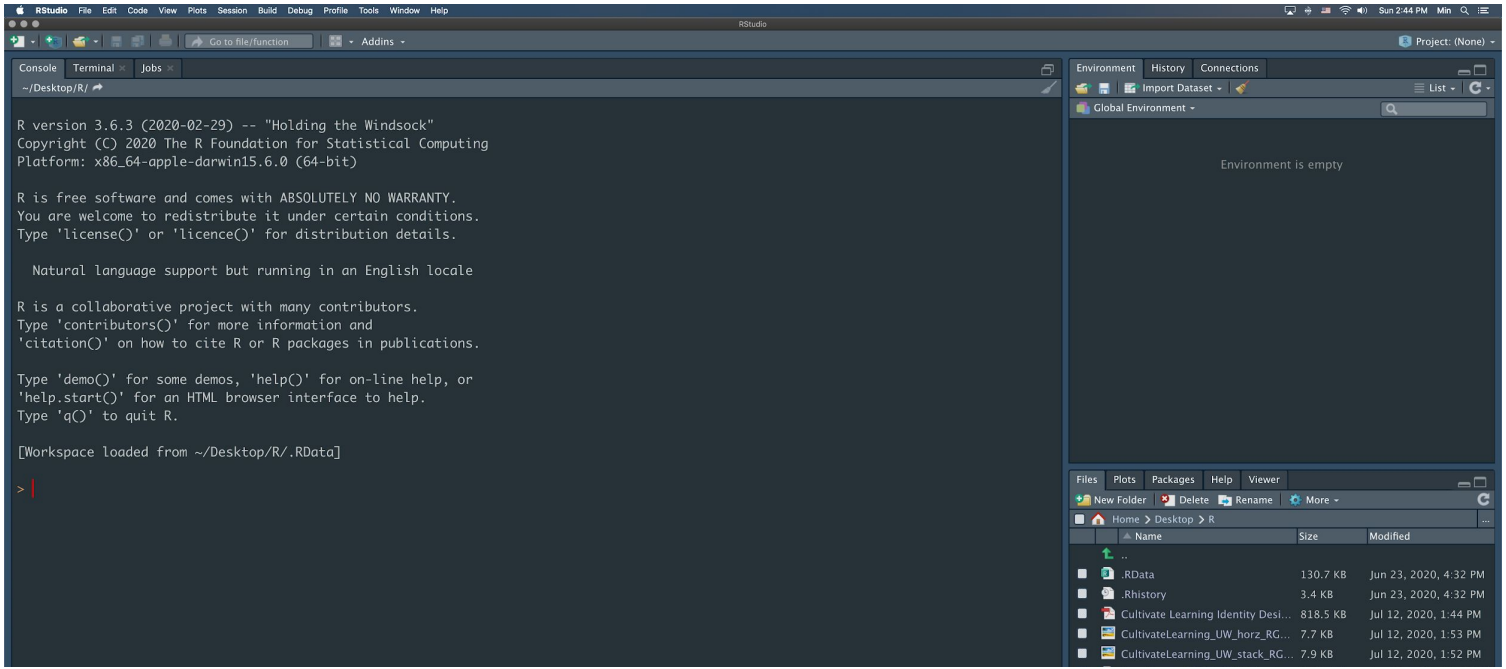
- Framework/System for **naming** your **files** in a way that describes what they contain and how they relate to other **files**
- **FileName_ProjectAcronym_DateCreatedinMMDDYY**
- **i.e. Workshop1_INL_071320**
  - IntroRWkshop: Intro to R workshop file
  - INL: Innovation Lab
  - 071320: July 13, 2020
- Project Acronyms
  - Professional Learning & Coaching: **PLC**
    - formerly Coaching & PD: **CPD**
  - Circle Time Magazine: **CTM**
  - Meaningful Makeover: **MM**

[Prep] Folder Structure

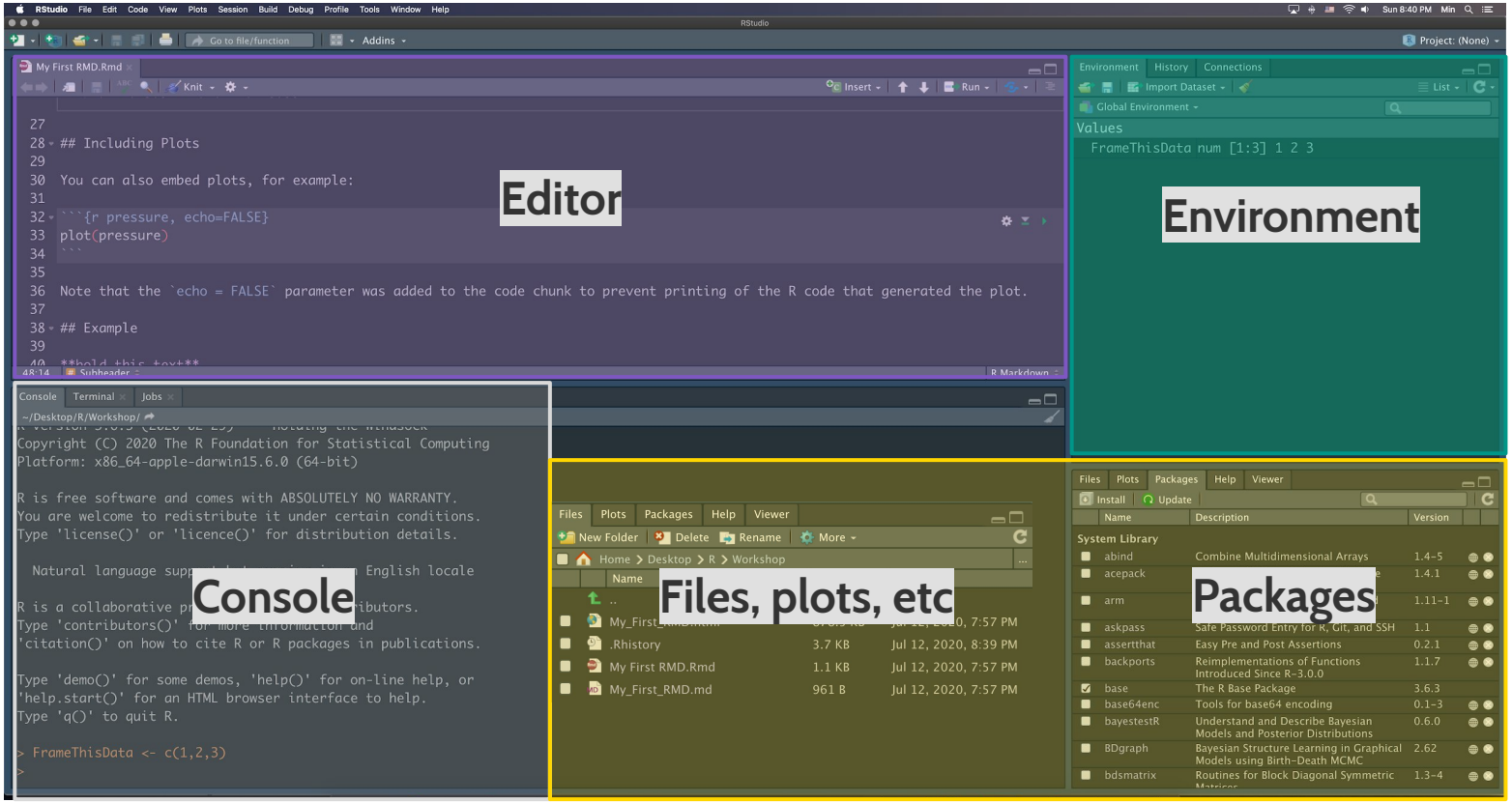
- Cultivate Learning protocol on C-Dash (Cloud server for archiving research relevant data)
  - Data
  - Output
  - Source
  - Documents
- Current Project
  - Create your own R folder on your desktop
    - R
      - Week 1: Intro to R
        - Data (where raw data would live)
        - Output (where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

Questions?

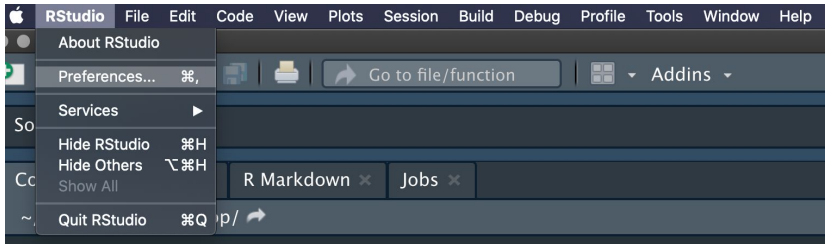
[R Studio] Integrated Development Environment

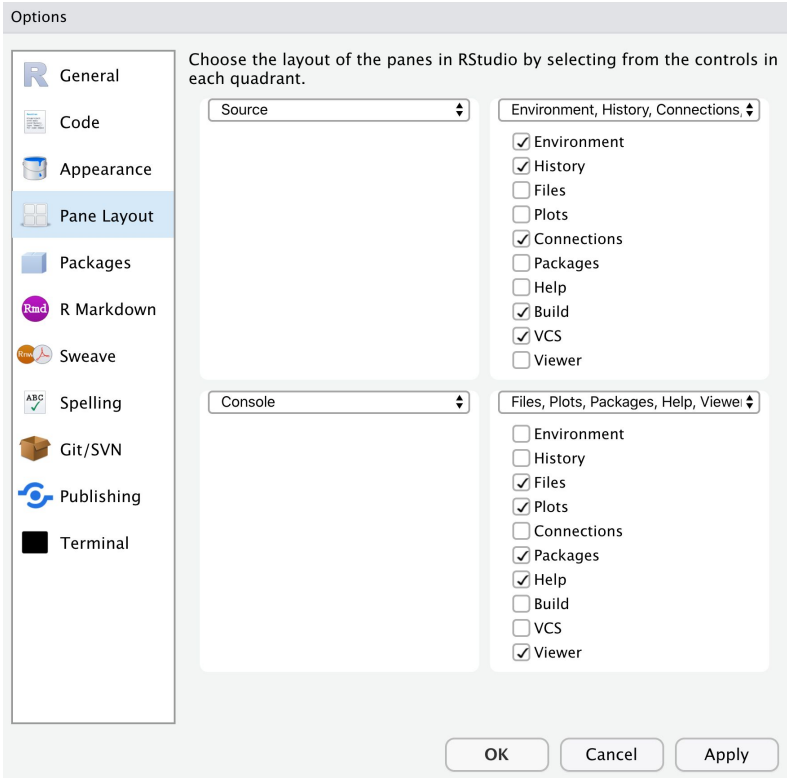


[R Studio] IDE Feature

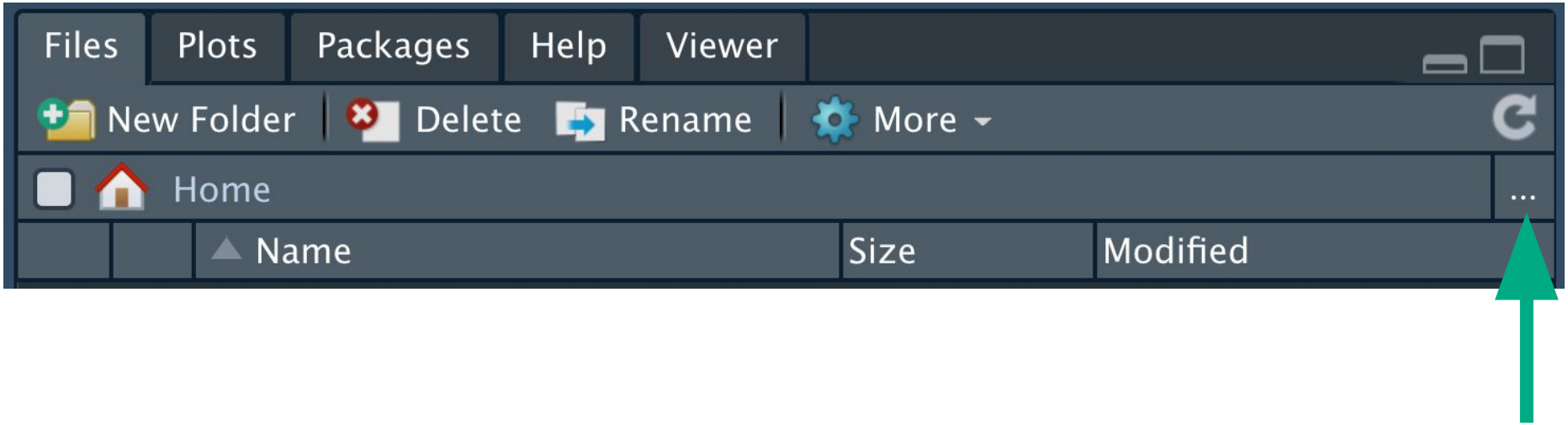


[R Studio] Changing IDE layout

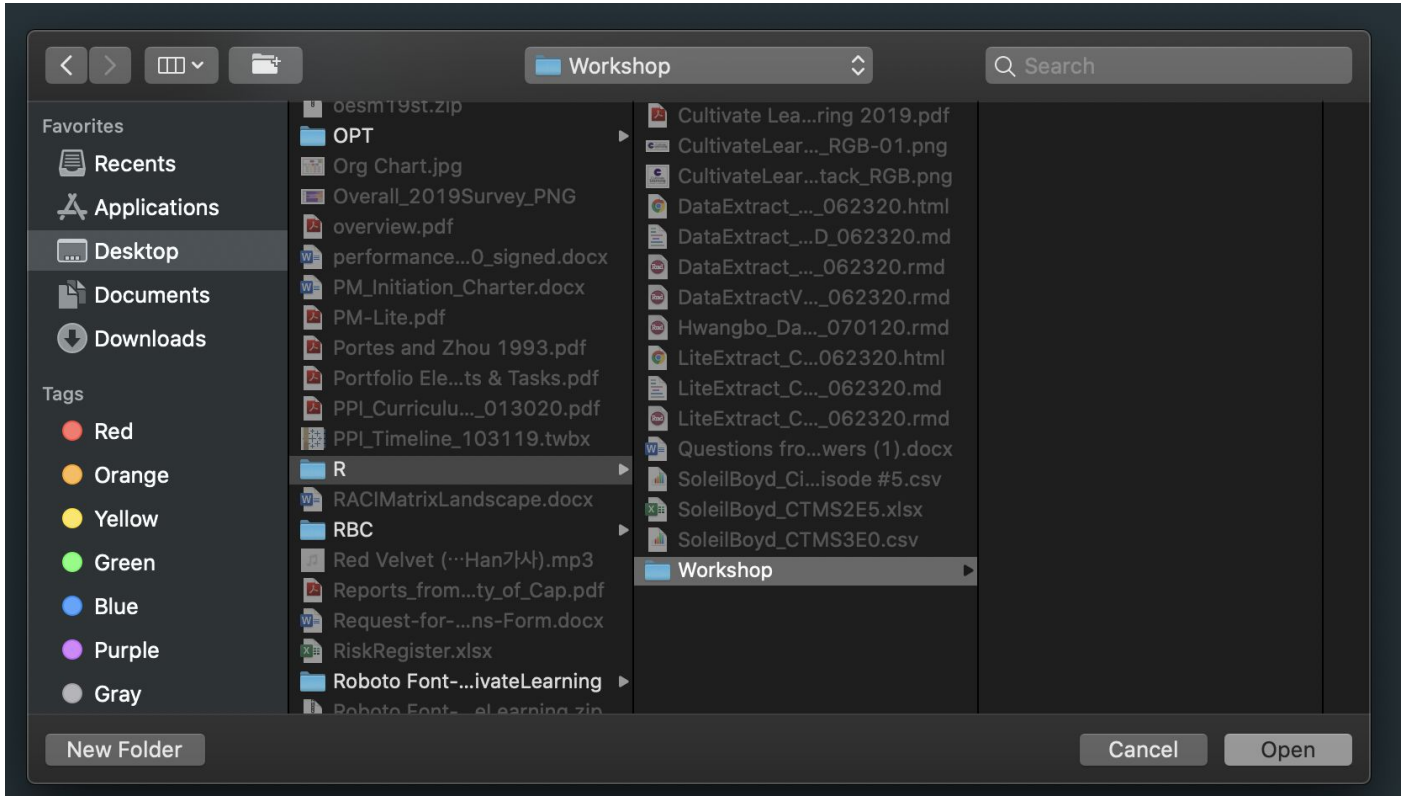




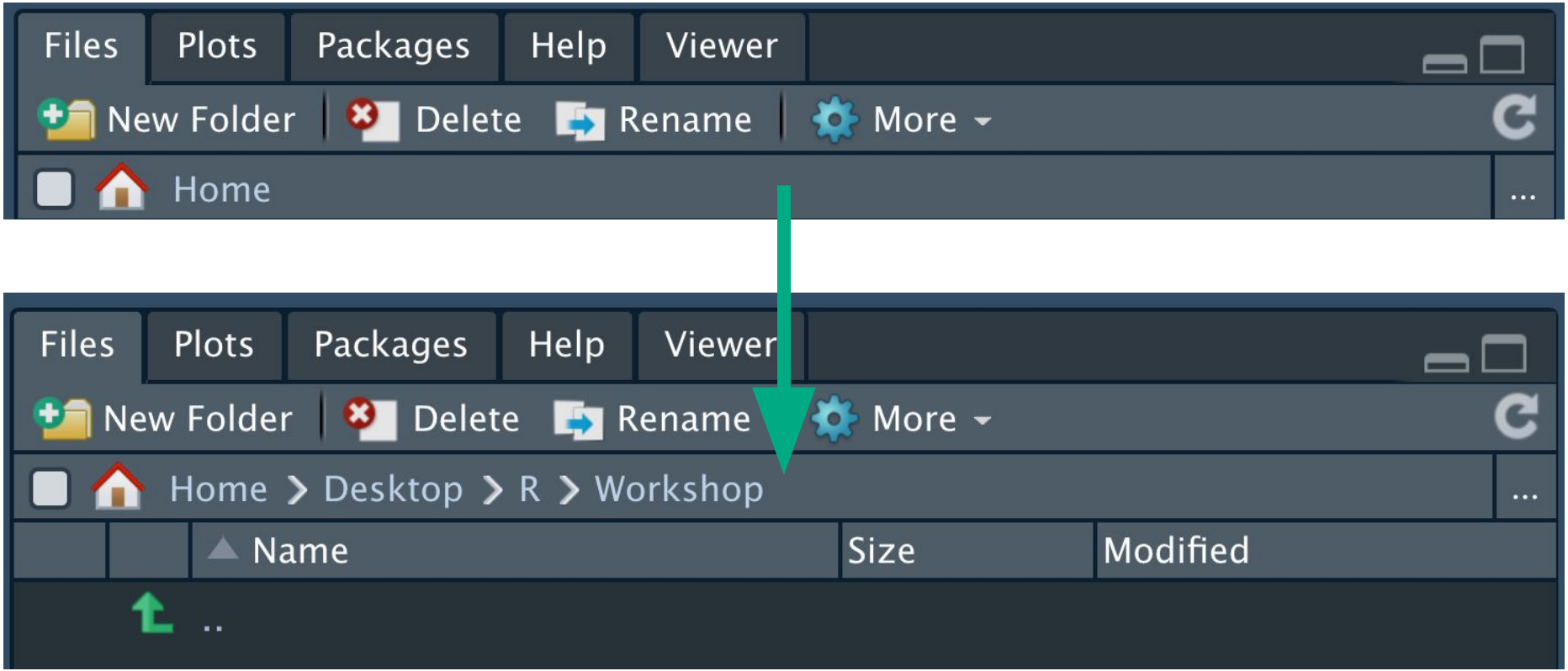
[R Studio] Setting Directory aka *Building Foundation*



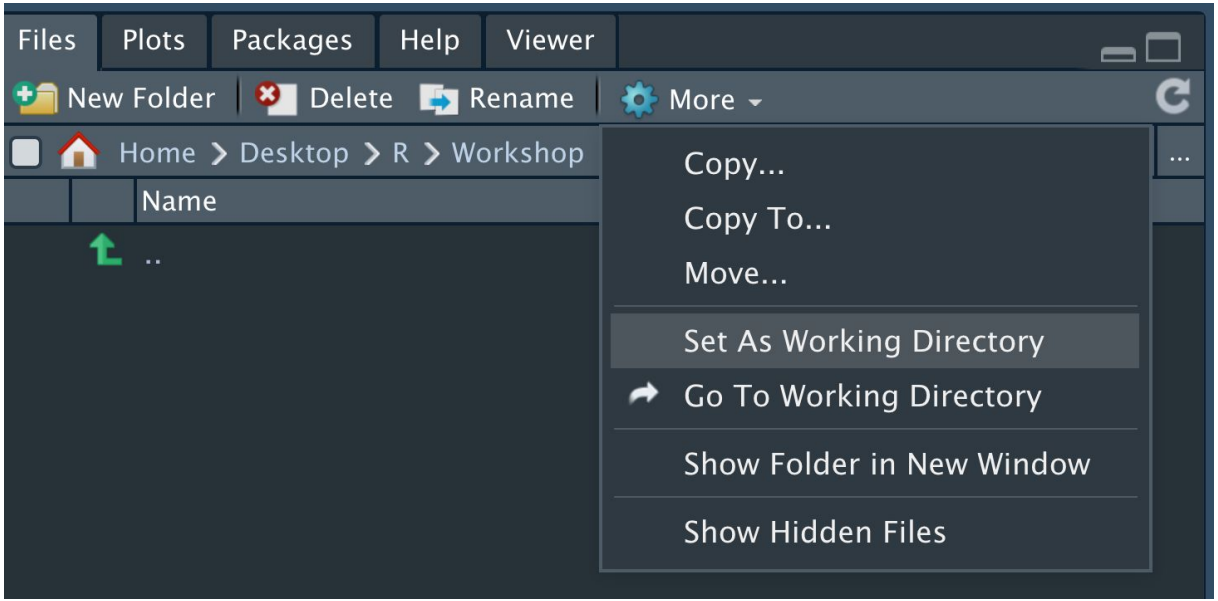
[R Studio] Setting Directory: Option 1



[R Studio] Setting Directory: Option 1

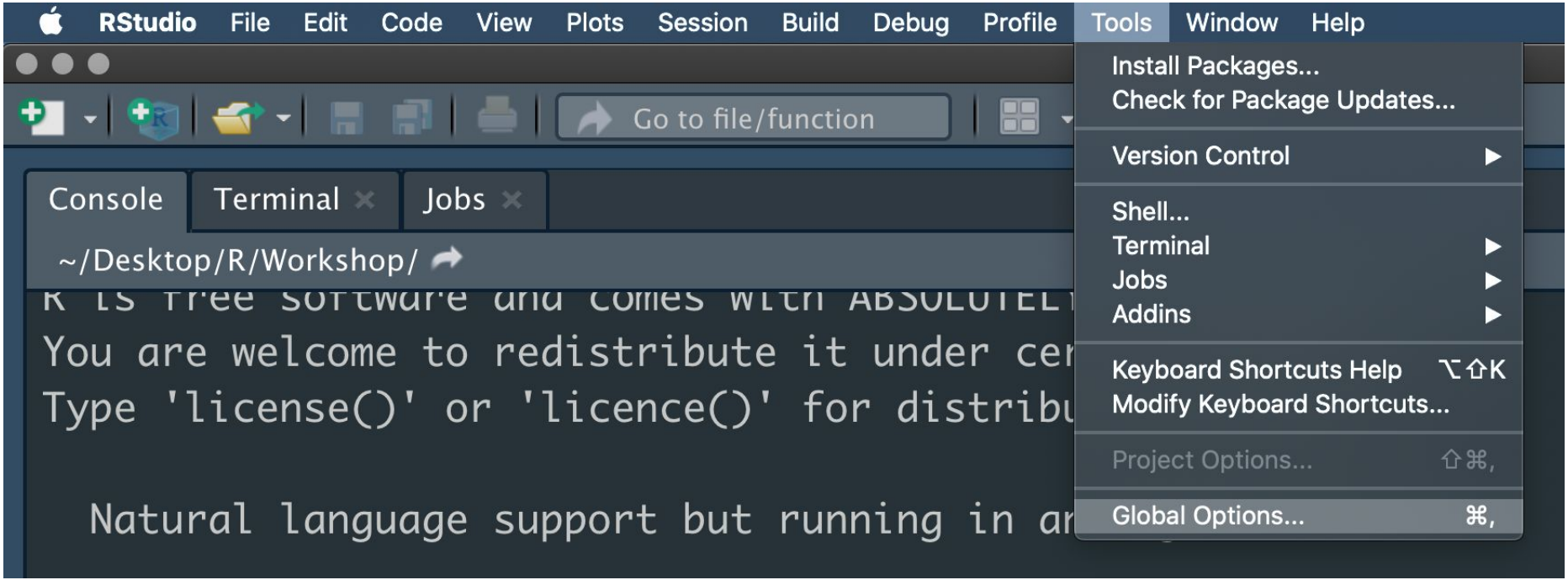


[R Studio] Setting Directory: Option 1

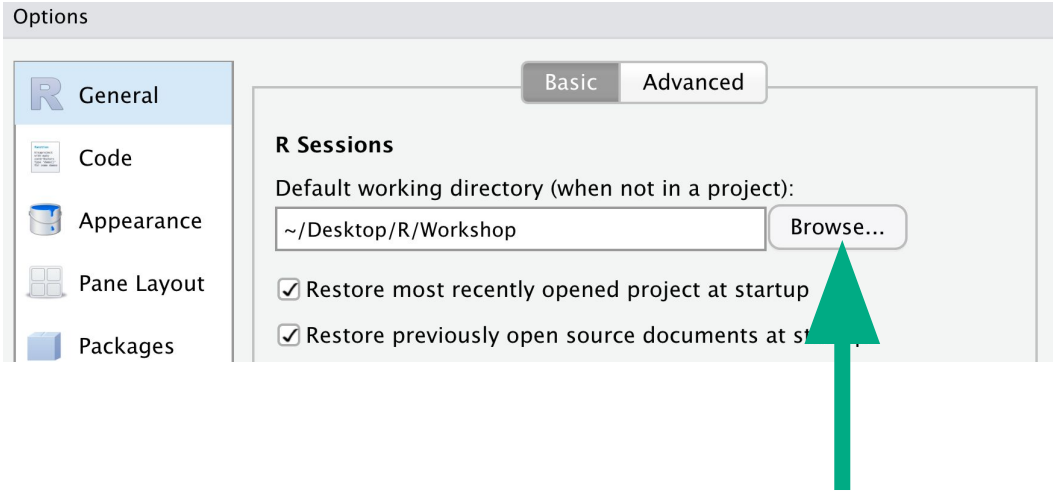



```
> setwd("~/Desktop/R/Workshop")
```

[R Studio] Setting Directory: Option 2



[R Studio] Setting Directory: Option 2



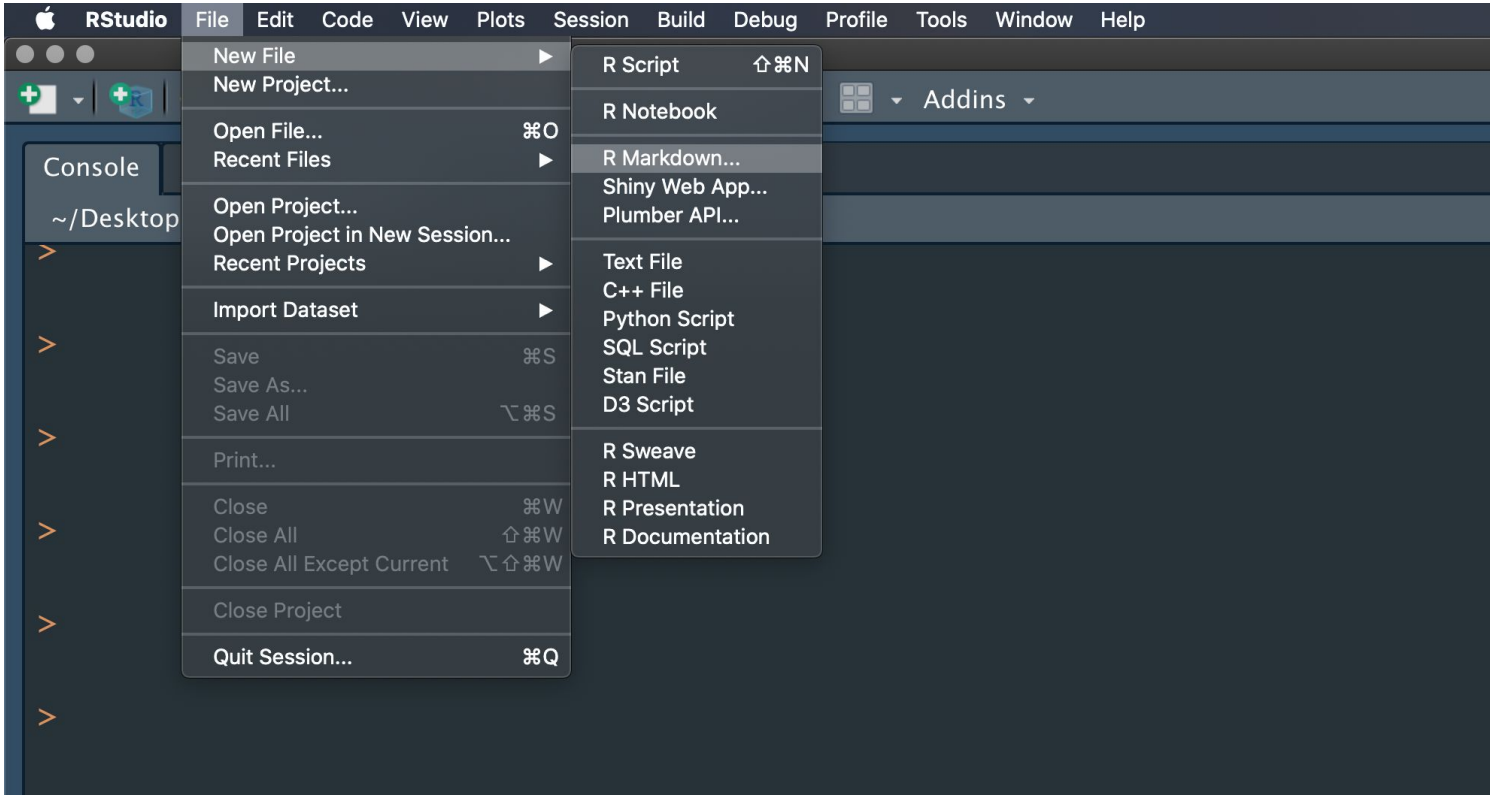
Questions?

10 min Break
Coffee, Tea & Snack Time

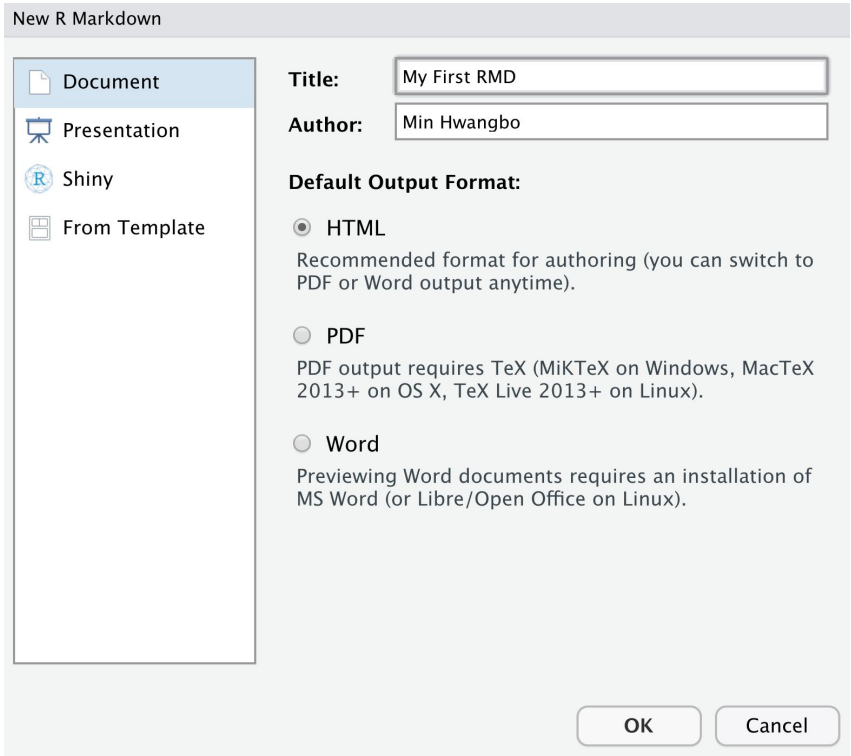
Learning Agenda

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

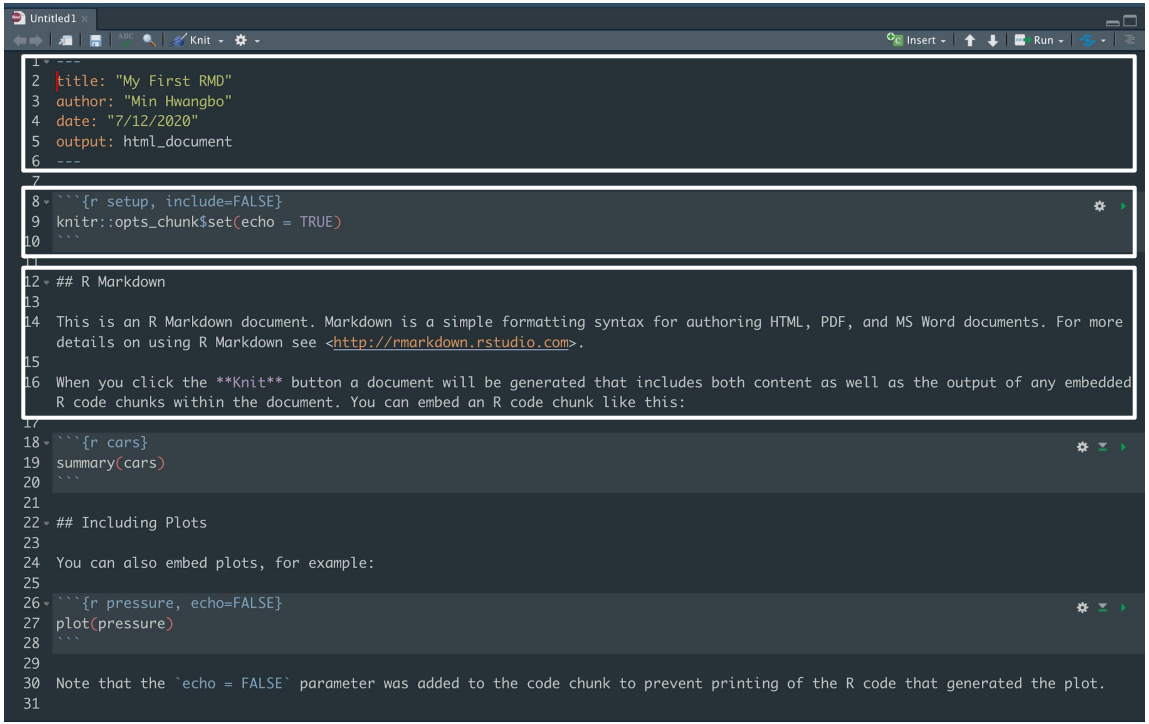
[RMD] R Markdown



[RMD] R Markdown



[RMD] R Markdown



Headers

Chunks

Body

[RMD] R Markdown: Syntax (Body)

Syntax

****bold this text****

italicize this text

 # Header

Subheader

> Block quote

Output

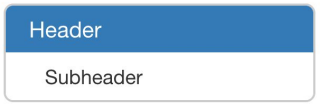
bold this text

italicize this text

Header

Subheader

> Block quote



[RMD] Headers

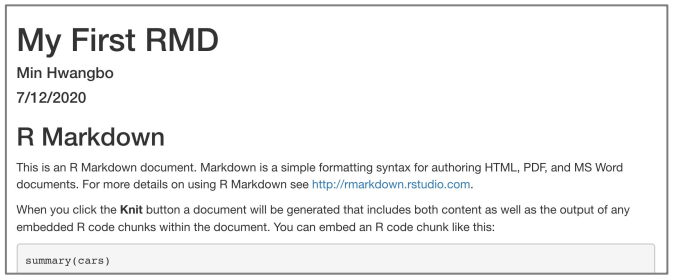
```
---
title: "My First RMD"
author: "Min Hwangbo"
date: "7/12/2020"
output: html_document
---
```

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

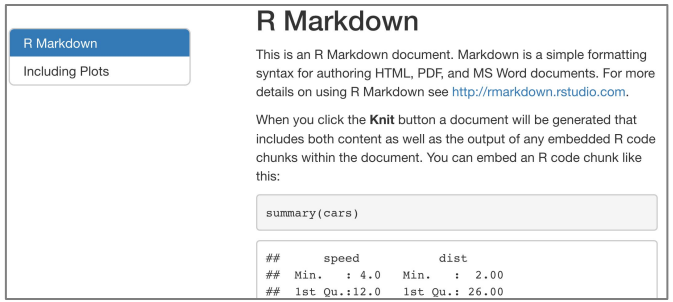
```
---
title: "LiteDataExtract_CTM_062320"
author: "Min Hwangbo"
date: "6/30/2020"
output:
  html_document:
    preserve_yaml: true
    toc: true
    toc_float: true
    keep_md: true
published: false
---
```

[RMD] Headers w/ HTML example

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




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

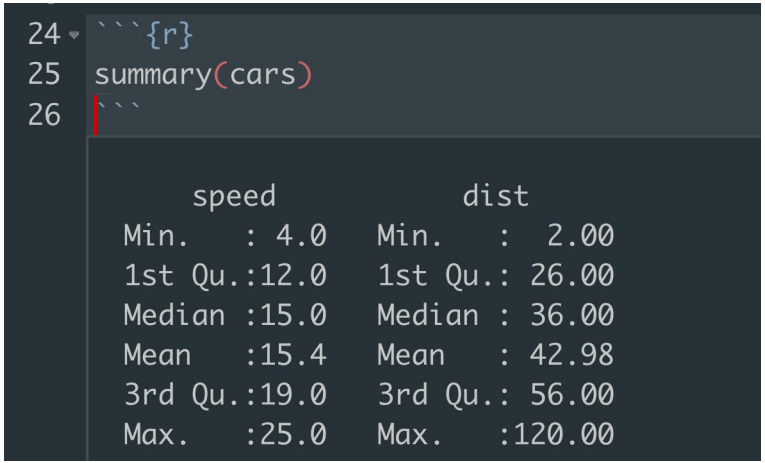


[RMD] Chunks aka Code Block

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

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

- Output:



[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

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

```
     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
```

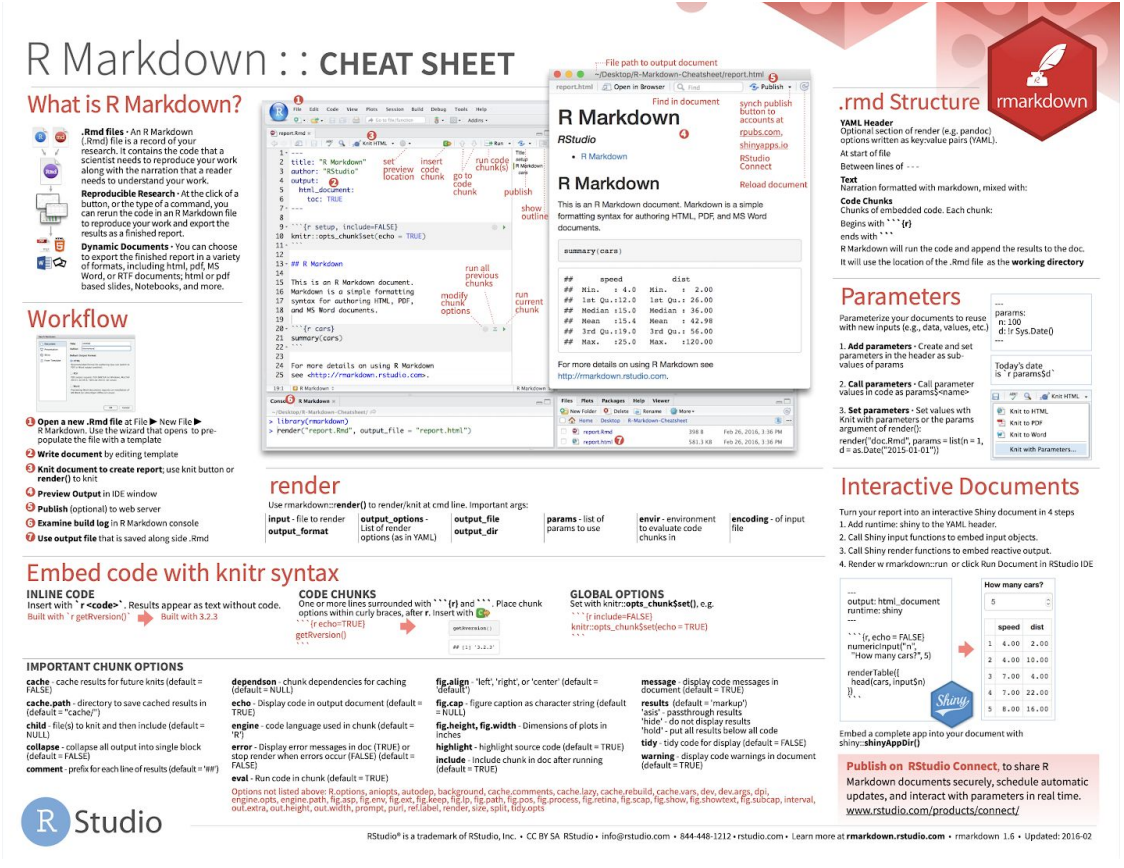
```
## Including Plots

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

````
```{r pressure, echo=FALSE}
plot(pressure)
```
````

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

[RMD] R Markdown Cheatsheet



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!