Lecture 2: R Markdown, Version Control with Git(Hub), and Other Productivity Tools

James Sears*
AFRE 891 SS 24
Michigan State University

*Parts of these slides are adapted from "Advanced Data Analytics" by Nick Hagerty and "Data Science for Economists" by Grant McDermott.

Table of Contents

Prologue

Prologue

Before we dive in, let's double check that we all have

- ☑ Installed **R**.
- ☑ Installed **RStudio**.
- ☑ Signed up for an account on **Github**
- ☑ Installed **Git** and **Github Desktop**
- ☑ Log into your Github account on Github Desktop

R Markdown

R Markdown

Before we dive into version control, let's chat about **R Markdown**.

R Markdown is a document type that allows for integration of R code and output into a Markdown document.

Resources:

- Website: **rmarkdown.rstudio.com**
- **R Markdown Cheatsheet**
- Book: **R Markdown: The Definitive Guide** (Yihui Xie, JJ Allaire, and Garrett Grolemund)

R Markdown

Before we dive into version control, let's chat about **R Markdown**.

R Markdown is a document type that allows for integration of R code and output into a Markdown document.

Other points:

- We'll be completing assignments using R Markdown.
- FWIW, my lecture slides and notes are all written in R Markdown too. (E.g. This slide deck is built using the **xaringan** package with the metropolis theme.)

R Markdown: Getting Started

☑ Installed R.

☑ Installed RStudio.

☐ Add the `rmarkdown` package

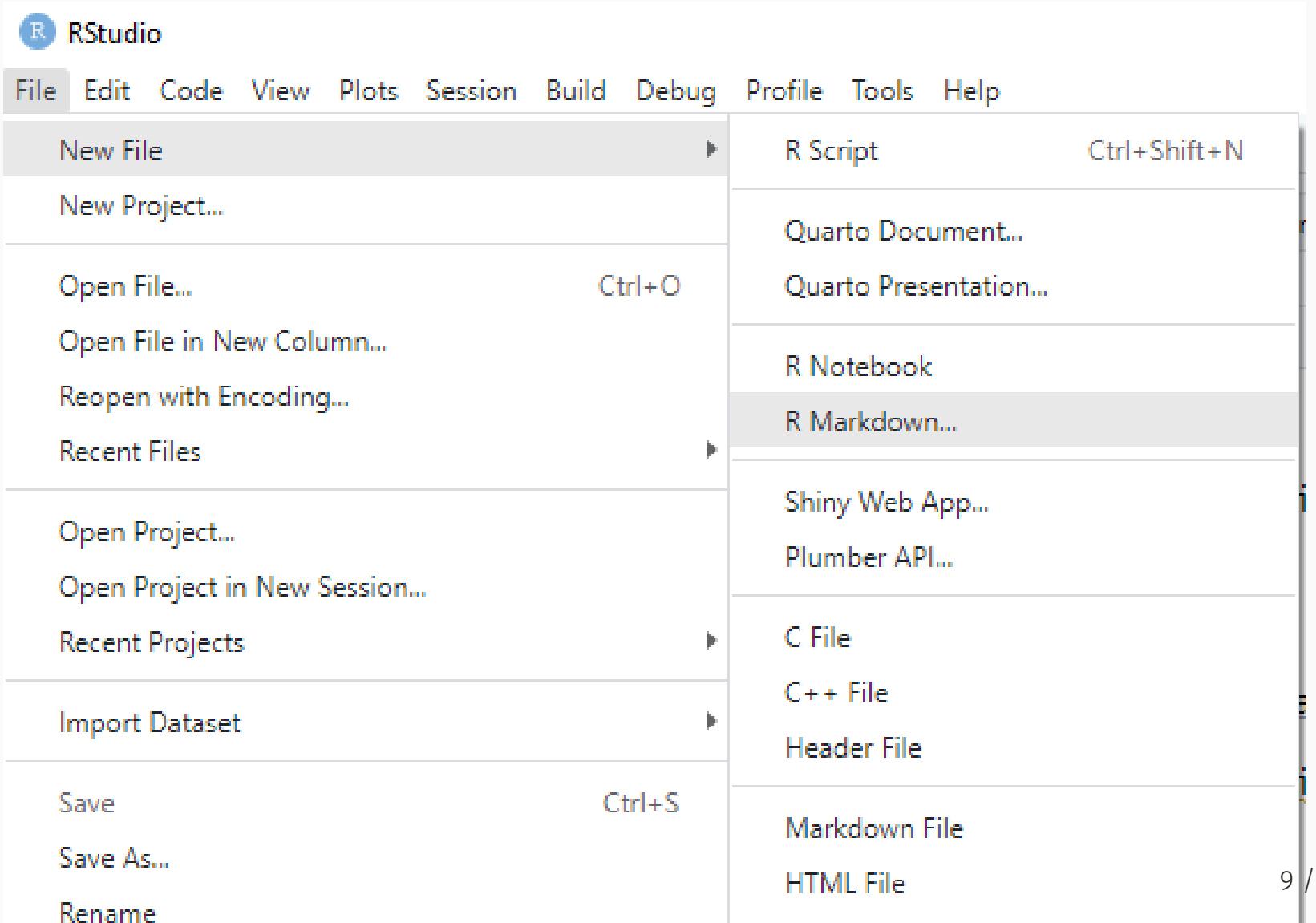
```
install.packages("rmarkdown")
```

☐ Install LaTeX

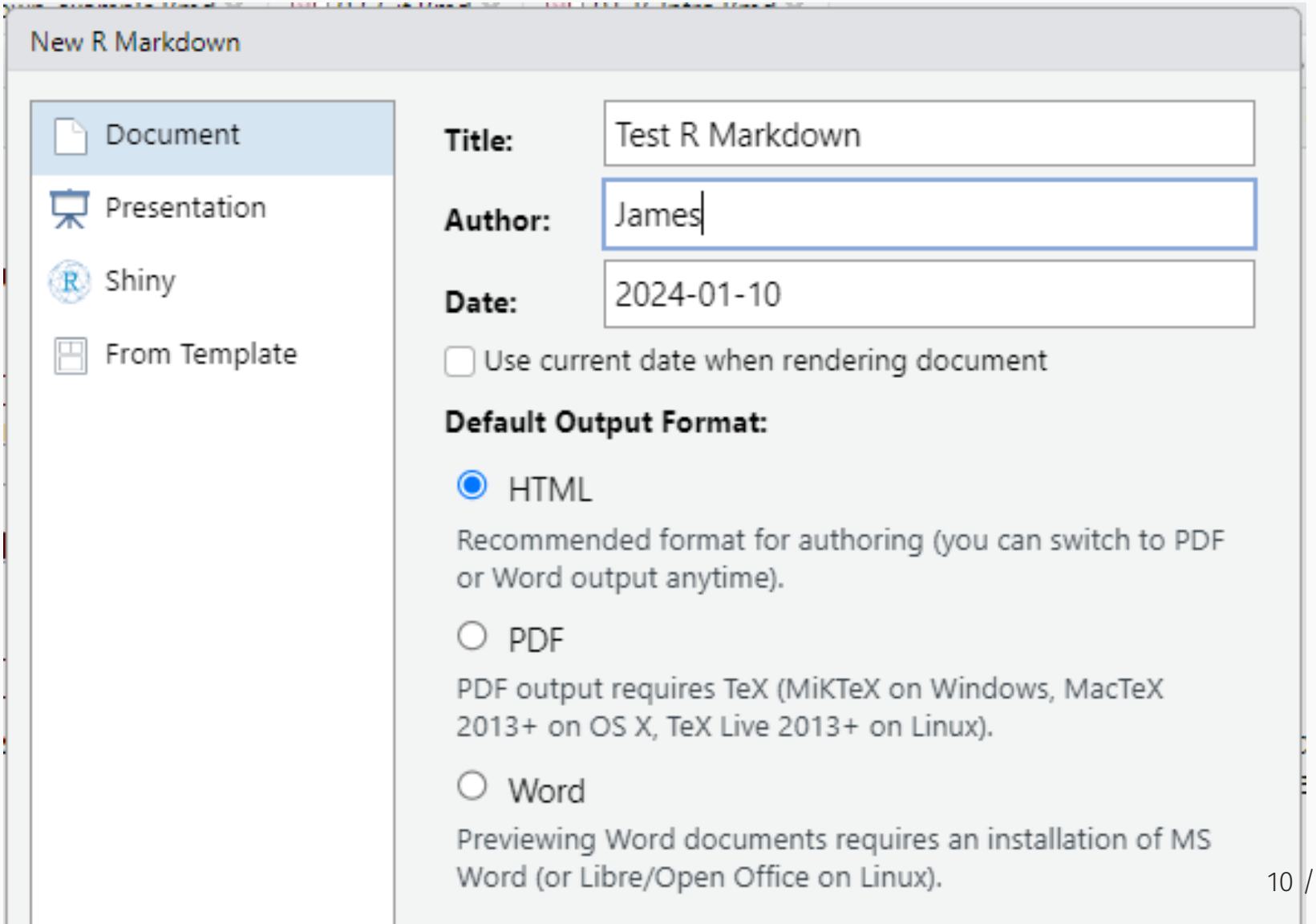
- If just for this, can use **TinyTex**

```
# Install only if you don't have LaTeX already
install.packages("tinytex")
tinytex::install_tinytex()
```

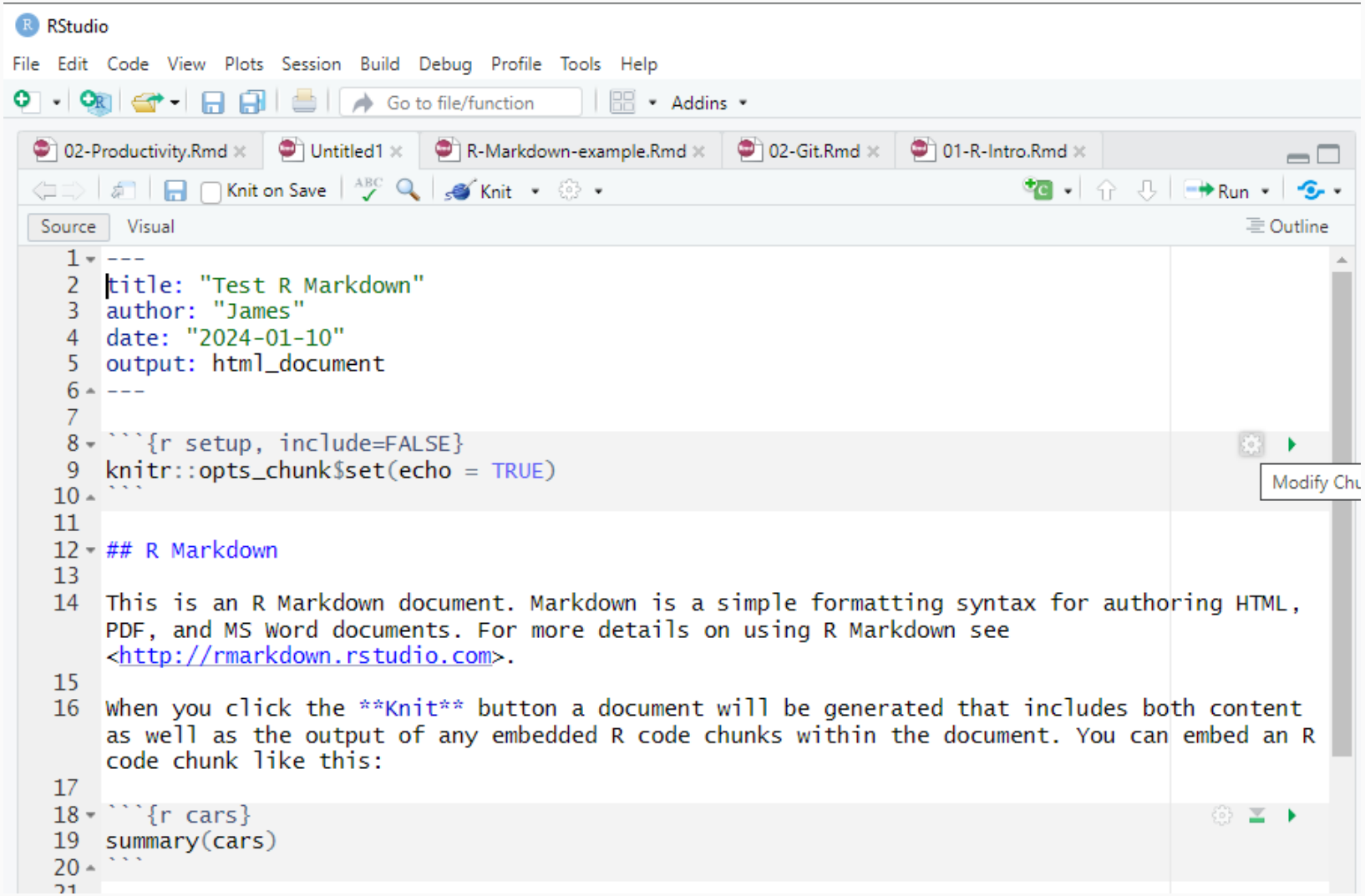
R Markdown: Creating a New .Rmd File



R Markdown: Creating a New .Rmd File



R Markdown: Creating a New .Rmd File



R Markdown Components

R Markdown combines

1. **Markdown:** lightweight markup language
2. **LaTeX:** typesetting for math
3. **R:** include code and generate output

Let's do some practice: **open a new .Rmd file** and try adding content as we go

Markdown

Markdown allows for formatting text in a lightweight way

I highly recommend the handy **Markdown Guide** for more details

Markdown: Heading

Headings emphasize text and add chunks to your script

Largest headingwith one leading # (slide title above)

Second Largest (##)

Third Largest (###)

Getting Smaller... (####)

Normal Text for comparison

Markdown: Text Format

Bold text with **your text**

Italicize with *single asterisks*

Add `code text` with grave accents (the back tick symbol)

- `
- The shift output of the tilde key `~` on keyboard

End a line with two spaces to start a new paragraph

- or leave a line space between sentences

Can also start a new line with backslash (\)

Markdown: Text Format

Add superscripts^{2} with ^carets^

Add ~~strikethroughs~~ with ~tildes~

Add a line break (horizontal rule)

---

with ***

Markdown: Text Format

Draw **tables** using | and -

Col A	Col B	Col C
This	is	a
Table		wow

Markdown: Lists

Add an **ordered list** with **1.**

1. First Item
2. Second Item
3. No need to change the number - keep using 1. It will automatically update.

Add an **unordered list** with *** or -**

- A thing
- Another related thing
    - Indent to nest
        1. Can mix ordered and unordered

Markdown: Inputs

Add a link with []()

- [text label](URL)
- Add direct link with <link> https://www.markdownguide.org

Add an image with

- ![alt text](URL)

practice by adding `images/smile.png`:

R Markdown: LaTeX

Another advantage of R Markdown is that it integrates LaTeX functionality for typesetting math.

Add an **inline equation** with TeX

$$Var(X) = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n} \qquad Y_{it} = \beta_0 + \beta_1 X_{it} + \epsilon_{it}$$

Add multiple rows of LaTeX with

$$

LaTeX lines here

$$

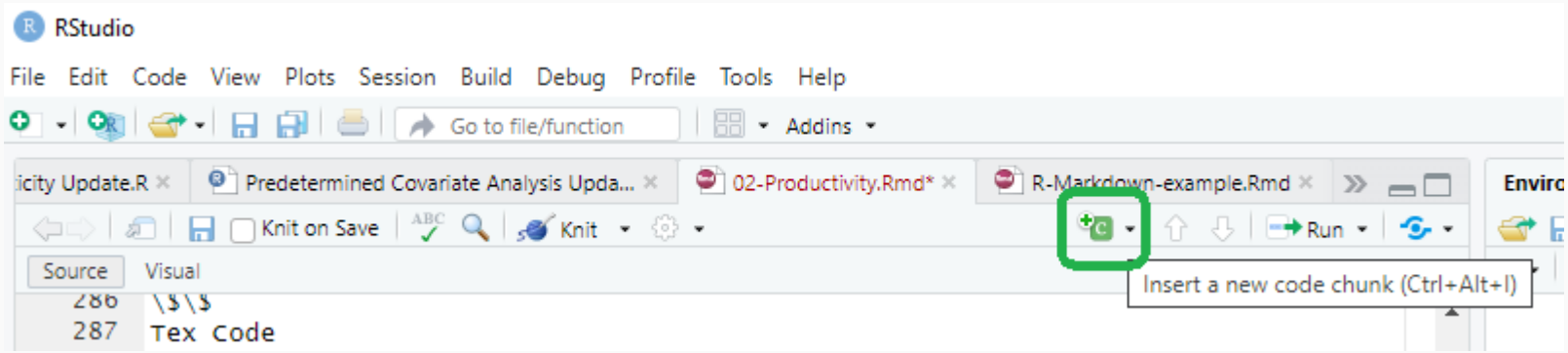
Use the **standard LaTeX commands** for symbols/characters

R Markdown: R Code

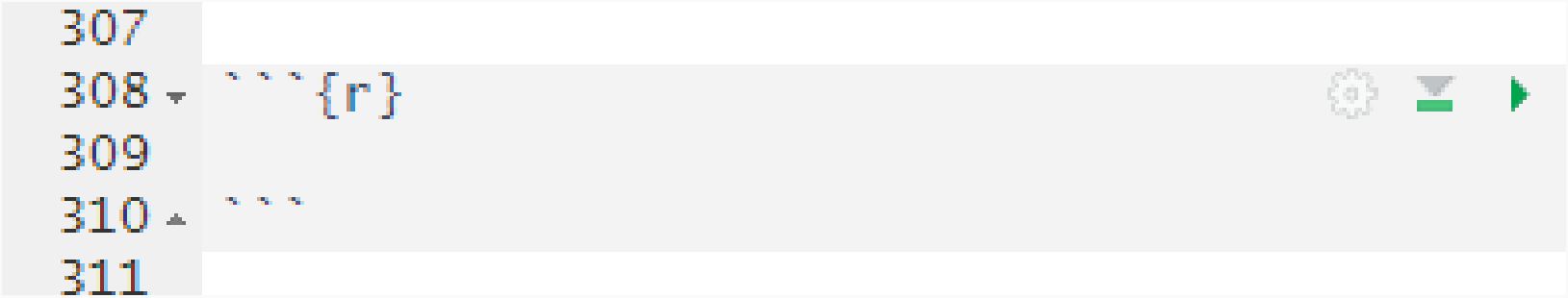
R code is primarily executed with **code chunks**

Add a chunk with

- `Cmd + Option + I (Ctrl + Alt + I on PC)`
- The `Insert` button in the UI
- Manually type



R Markdown: R Code Chunks



Code chunks allow us to add as many lines of code as we want

- Output will appear underneath after executing the full chunk
- Can customize whether it runs, how output is displayed
- Can run manually
  - Line by line with `Cmd/Ctrl + Enter`
  - Entire chunk with `Run Entire Chunk` button

R Markdown: R Code Chunk Options

You can **add chunk options** in brackets after `r` and separated by commas.

Some commonly-used options include:

- **Chunk label** (`ex_chunk`)
- `include = FALSE` will run the chunk but hide it from the final document
- `eval = FALSE` will display code without evaluating it
- `results = 'hide'` runs code but hides output from the final document

````
52
53 ```{r sum, echo = FALSE, warning = FALSE}
54 2+2
55
56 ```
````

R Markdown: R Code Chunk Options

You can **add chunk options** after `r` and separated by commas.

Some commonly-used options include:

- `echo = FALSE` runs the code but hides the chunk from the final document
- `error = FALSE` (`warning: FALSE`) will hide error (warning) messages generated by the code

- LOTS of options for output figures: figure size (fig.width, fig.height, fig.dim), output document scale (out.width, out.height), alignment (fig.align), caption (fig.cap)

  Learn more **about chunk options here**

R Markdown: R Code

You can call R objects from earlier chunks **inline** with

`` `r` ``

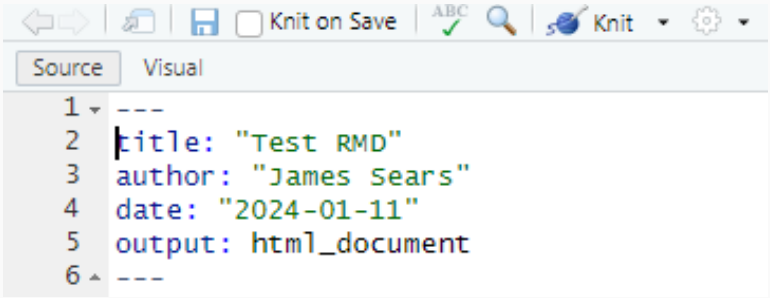
```
four = 2+2
```

This can output in line with text: 2 + 2 = 4

R Markdown File Organization

1. Header

RStudio automatically builds the R Markdown file from a template, which begins with a **header**



- Title
- Author
- Date
- Output Format
  - Main options[1]: HTML (`html_document`), PDF (`pdf_document`), LaTeX (`latex_document`), or Word (`word_document`)

1: See **CH 3 of "R Markdown: The Definitive Guide" for more on how to customize output formats**

2. R Setup

By default, RStudio adds a **setup** code chunk next.

````
```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```
````

- Can set global options
- Useful as your preamble
- For **R Notebooks**, this will automatically be run and is the only place where you can change your working directory

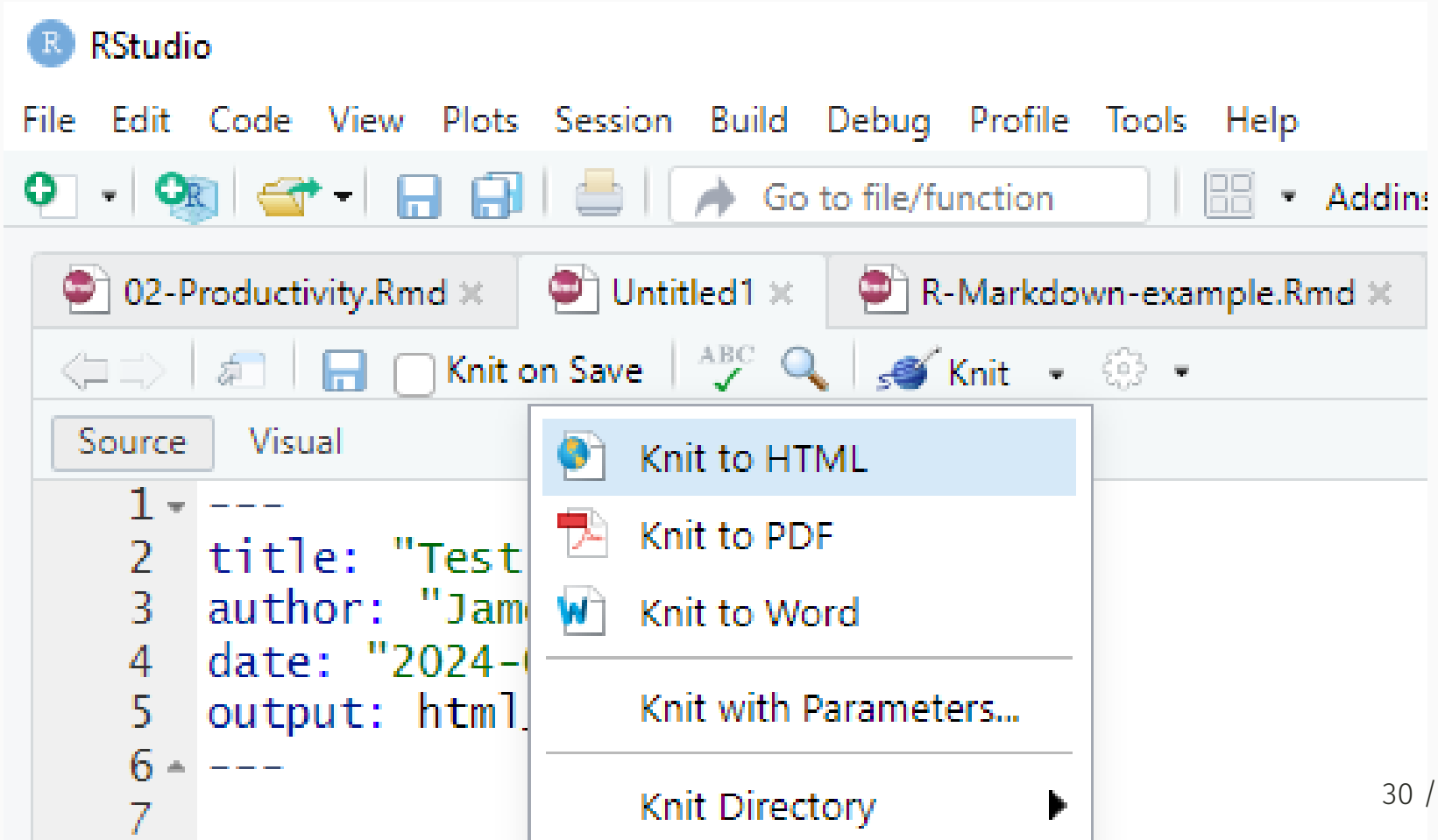
3. Contents

From here on you can build the report/notebook as needed for the task.

- Add any writing and outside graphics or **bibTeX citations**
- Add code chunks to carry out desired analysis
- Employ sections and formatting to structure the document as desired

Compiling/Knitting

When you are ready to compile your final document, use the `Knit` button or `Ctrl/Cmd + Shift + K`



R Markdown: Knit to Compile Output

Markdown Practice!

Markdown Practice

1. Create a new R Markdown file named "R-Markdown-Ex.RMD"
2. In the setup chunk, load the `dslabs` and `tidyverse` packages
    - Use the `data()` function to read in the `divorce_margarine` dataset
3. Add a header labeled "Correlation vs. Causation" and a text explanation below for why we often want to differentiate between the two
4. Add a code chunk with the label `plot`
    - Type the following code:

```
ggplot(divorce_margarine) +
geom_point(aes(x = margarine_consumption_per_capita,
           y = divorce_rate_maine)) +
labs(title = "Relationship between Margarine Consumption and
          Divorce Rates in Maine",
 subtitle = "2000-2009",
 x = "Margarine Consumption per Capita",
y = "Divorce Rate")`
```

5. Knit and save a PDF/HTML copy of the file to the "output" folder