Intro to RStudio

Angela He
Stanford Software and Services for Data Science

What are R and RStudio?

- R is a programming language
- We will be using RStudio with R, which is a user interface for working with R

Why R and RStudio?

- Compared to Stata:
    - Free and open-source
        - Anyone can open and edit files you share
            - Vs. Stata: you need Stata to open and edit Stata files
        - Lots of cool packages that people create and share regularly
    - R can hold multiple datasets at a time
    - Steeper learning curve but used more widely than Stata and will help you learn other programming languages
- Compared to Python:
    - More gradual learning curve, good first introduction to programming
    - Easier to use for data analysis

Set-up Steps

1. Create a new project
2. Create a markdown file
3. Install/load the packages you need (you can add more later too)
4. Read in your dataset(s)
5. Now you can begin your data cleaning and analysis!

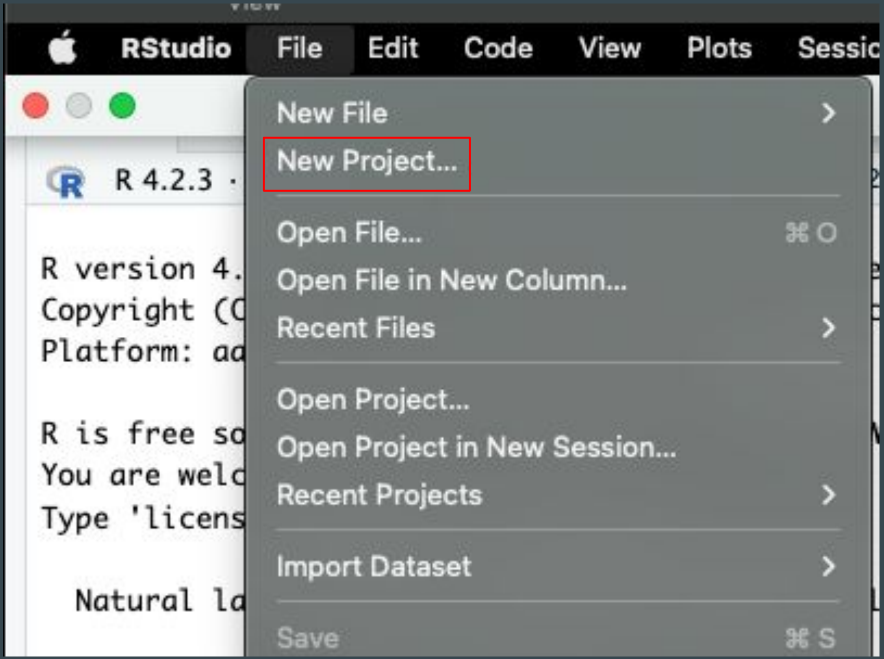
1. Create a new project

•••

Why use a project file?

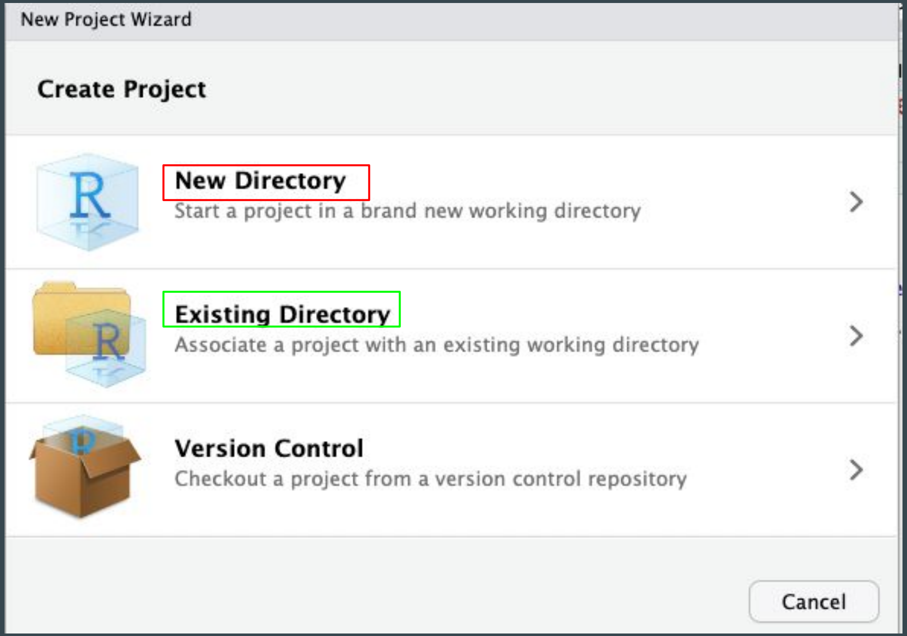
- All of your work associated with the project (e.g. data, scripts) are stored in the same portable folder
- You won’t need to worry about configuring your working directory no matter which computer you are on (helps with reproducibility too!)

File > New Project



```
R version 4.
Copyright (C
Platform: aa

R is free so
You are welc
Type 'licens

  Natural la
```

- You can choose New Directory or Existing Directory
- Choose New Directory if you want to create a new folder to put your project in
- Choose Existing Directory if you have an existing folder you want to put your project in



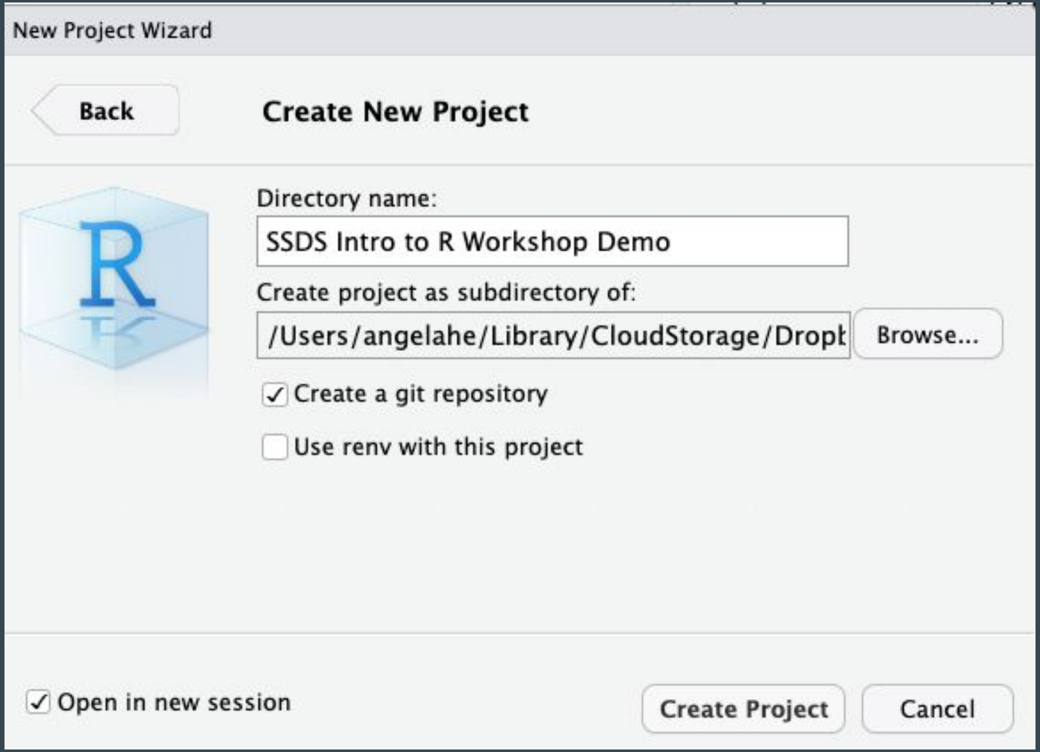
New Project Wizard

Create Project

New Directory
Start a project in a brand new working directory

Existing Directory
Associate a project with an existing working directory

Version Control
Checkout a project from a version control repository



Directory name: SSDS Intro to R Workshop Demo

Create project as subdirectory of: /Users/angelahe/Library/CloudStorage/Dropb

☑ Create a git repository

☐ Use renv with this project

☑ Open in new session

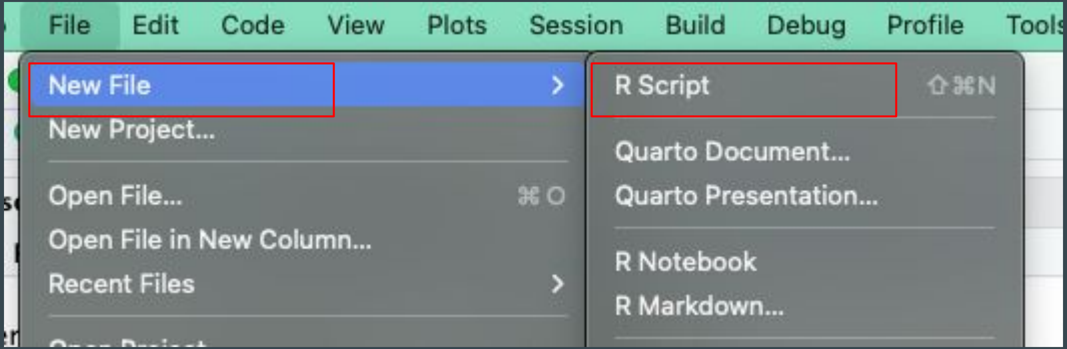
2. Create a new markdown file

•••

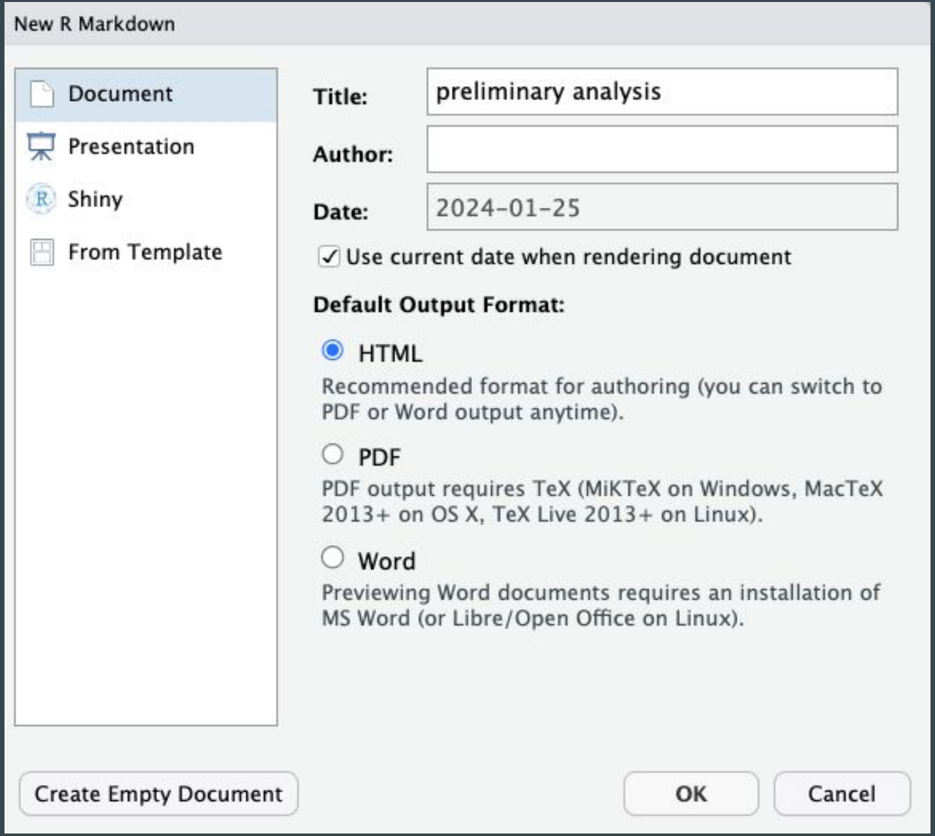
Why use a project file?

- All of your work associated with the project (e.g. data, scripts) are stored in the same portable folder
- You won't need to worry about configuring your working directory no matter which computer you are on (helps with reproducibility too!)

File > New File > R Script

- 

You can add a title and author to your markdown file



Title: preliminary analysis

Author:

Date: 2024-01-25

☑ Use current date when rendering document

Default Output Format:

● HTML

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

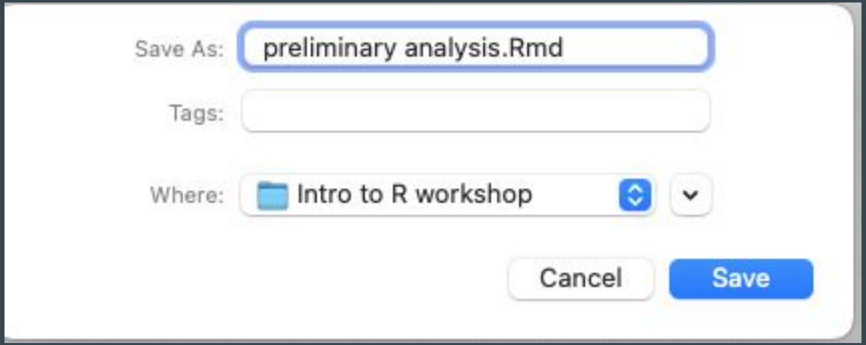
○ PDF

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).

After the new markdown file is created, make sure you press ctrl + s to save the file to your project folder

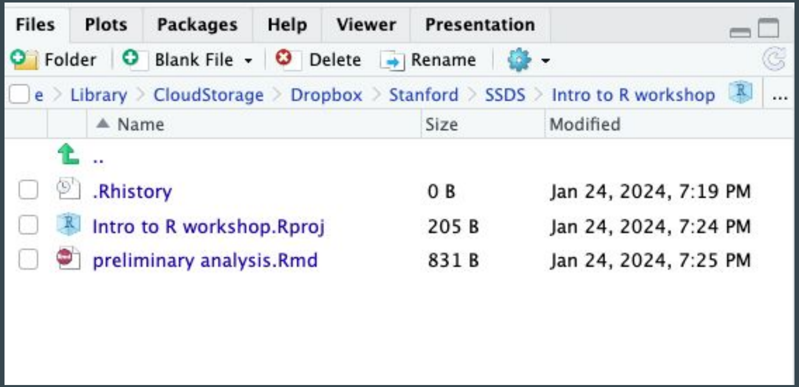


Save As: preliminary analysis.Rmd

Tags:

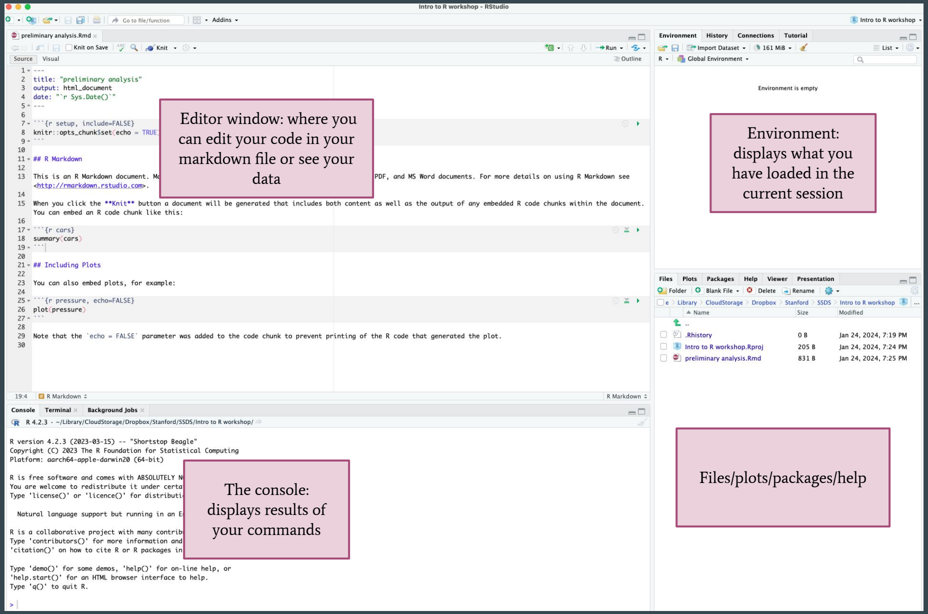
Where: Intro to R workshop

You should now see your markdown file in the bottom right quadrant


Files
Plots
Packages
Help
Viewer
Presentation
Folder
Blank File
Delete
Rename
e > Library > CloudStorage > Dropbox > Stanford > SSDS > Intro to R workshop
...

Name	Size	Modified
..		
.Rhistory	0 B	Jan 24, 2024, 7:19 PM
Intro to R workshop.Rproj	205 B	Jan 24, 2024, 7:24 PM
preliminary analysis.Rmd	831 B	Jan 24, 2024, 7:25 PM

The four quadrants in RStudio



~~~
---
title: "preliminary analysis"
output: html_document
date: "`r Sys.Date()`"
---

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

R Markdown

This is an R Markdown document. Mo   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:

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

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.
```

Editor window: where you can edit your code in your markdown file or see your data

Environment: displays what you have loaded in the current session

	Name	Size	Modified
	..		
□	.Rhistory	0 B	Jan 24, 2024, 7:19 PM
□	Intro to R workshop.Rproj	205 B	Jan 24, 2024, 7:24 PM
□	preliminary analysis.Rmd	831 B	Jan 24, 2024, 7:25 PM

```
R version 4.2.3 (2023-03-15) -- "Shortstop Beagle"
Copyright (C) 2023 The R Foundation for Statistical Computing
Platform: aarch64-apple-darwin20 (64-bit)

R is free software and comes with ABSOLUTELY N
You are welcome to redistribute it under certa
Type 'license()' or 'licence()' for distributi

  Natural language support but running in an E

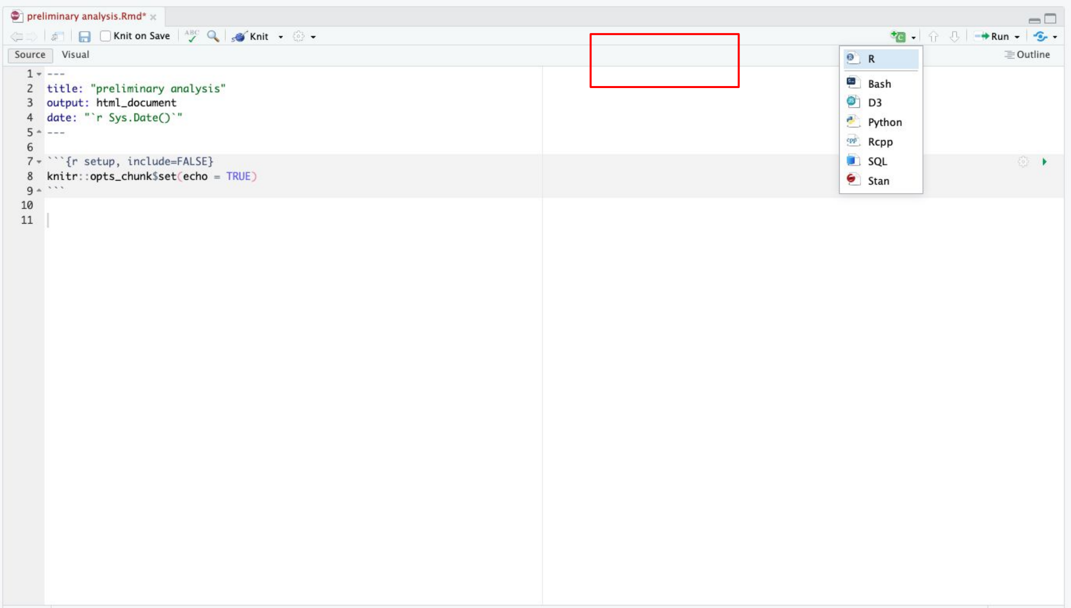
R is a collaborative project with many contrib
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

>
```

The console: displays results of your commands

Files/plots/packages/help
```

# How to write code in your markdown file

- You can organize your code into “chunks”
- To create a new chunk:



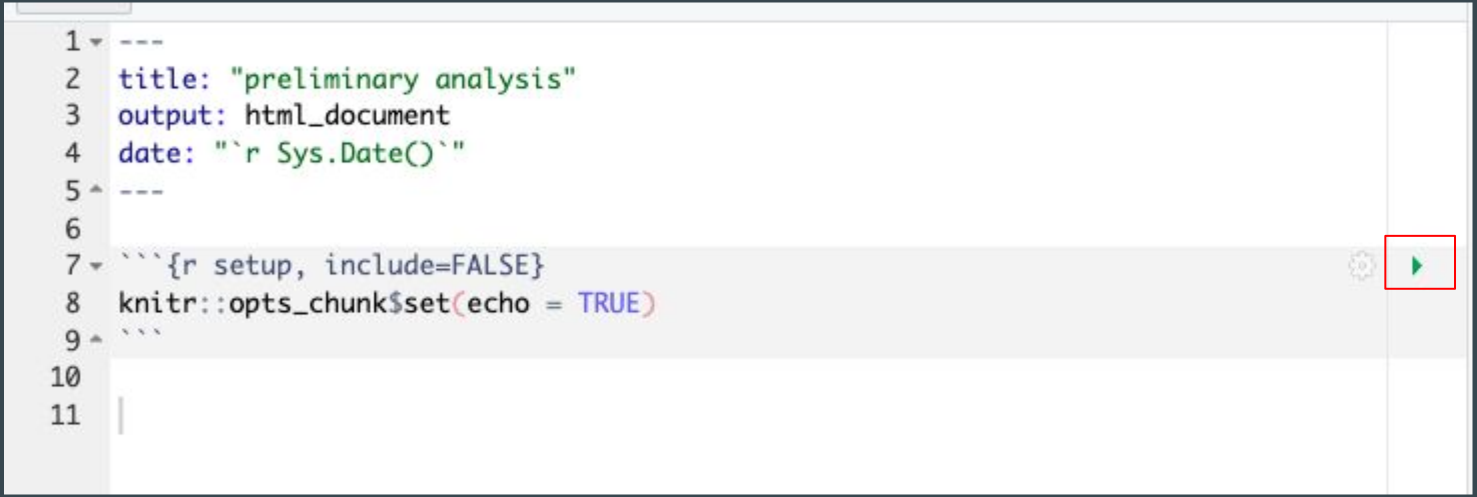
~~~
---
title: "preliminary analysis"
output: html_document
date: "`r Sys.Date()`"
---

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


~~~

How to run code

- Press the green play button

  OR

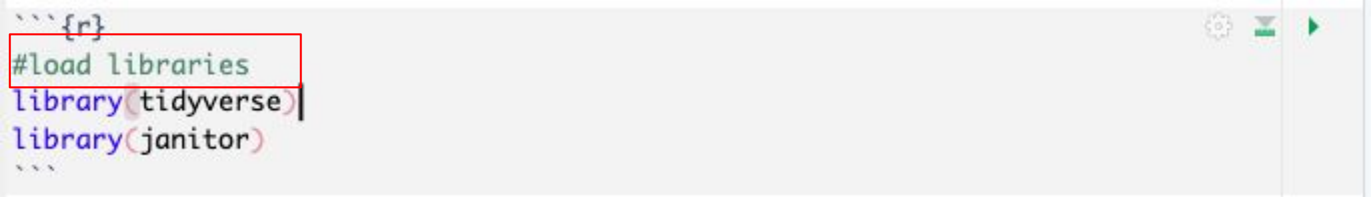
- Highlight code and “ctrl + enter” to run that line of code



~~~
---
title: "preliminary analysis"
output: html_document
date: "`r Sys.Date()`"
---

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

~~~

How to write comments

- Put a # before your comment → will make that text green
- Anything outside the chunks will be treated as text too



~~~
```{r}
#load libraries
library(tidyverse)
library(janitor)
```
~~~

3. Install/load packages you need

…

Why use a project file?

- All of your work associated with the project (e.g. data, scripts) are stored in the same portable folder
- You won’t need to worry about configuring your working directory no matter which computer you are on (helps with reproducibility too!)

Packages

- Packages are essentially extensions (think of Chrome extensions) you can add to your RStudio
    - Make your life easier: you don't need to write code yourself for tasks like predictive modeling (carat) or cleaning variable names (janitor)
- To install a package:
    - Type and run install.packages("packagename") either into your console or R markdown file
    - You just need to do this once, which is when you first download the package onto your computer
- To load the package:
    - Type library(packagename)
        - No quotes!
    - Do this every time you want to use a specific package for your project

Some packages I (almost) always use

- tidyverse
  - “A coherent system of packages for data manipulation, exploration, and visualization that share a common design philosophy”
  - Includes the following packages: dplyr (data manipulation), ggplot2, (data visualization), ...
- janitor
  - Cleans variable names
- Stargazer
  - Pretty(ish) regression tables

4. Read in your dataset(s)

•••

There are many different types of files you can read in. I like to read in csv files because they are universal.

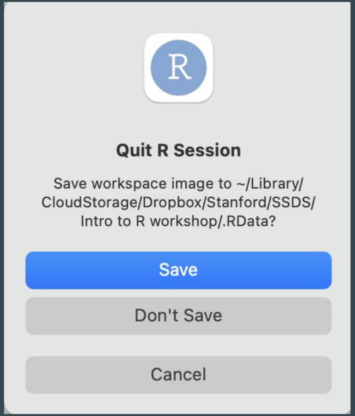
Let’s switch to the R markdown file now to dive into the code!

…

Download the .zip file here: http://tinyurl.com/ssds-intro-to-r

Helpful Tips and Tricks

- Enable rainbow parentheses to improve readability
- To export your file to share with someone else
- Click “don’t save” when R asks you if you want to save your workspace image when you exit out of R.



Quit R Session

Save workspace image to ~/Library/CloudStorage/Dropbox/Stanford/SSDS/Intro to R workshop/.RData?

Resources

- The ultimate guide
- Help files! (? or ??)
- Book an appointment with SSDS