# Topic 1: Getting Started with R & RStudio

1. What is R?
2. What is RStudio?
3. Basics to Using RStudio
4. What is RMarkdown?
5. Getting Started with R scripts and RMarkdown

# 1. What is R

- R is a command-line programming language for statistical computing
- R is a new implementation of S (a similar stat programming precursor to R)
- R has a wide range of packages for statistical analysis and graphing
- R is increasingly popular for data management and analysis
- **Download and Install R for Free!**

# Command-Line Programming

- Can be used as point and click
  - R Commander (Rcmdr)
  - Not covered here (I don’t use it)
  - See Rcmdr https://www.rcommander.com
  - Package https://cran.r-project.org/web/packages/Rcmdr/index.html
- R needs coding instructions
  - Then code needs to be ran before anything happens
  - Some code defines variables etc... which provides no output
  - R Syntax (coding language) requires patience

# Command-line programming

- Helps achieve “reproducible research”
  - Saved code and data files implies analysis & results are preserved
  - Any one with R (with appropriate version and packages installed) can run same analysis
  - Again, R is free (and so are the packages)
- Save your work
  - Forethought on organizing saved scripts and data
  - Organization is important!
    - Save so you can find later for coursework material and your own research
    - Consider cloud options to access and share work in R

# R Background

- Created by Ross Ihaka and Robert Gentleman at University of Auckland
  - Beta release in 2000
  - Named after first letter of first names (play on S language)
- Comprehensive R Archive Network (CRAN)
  - https://cran.r-project.org
  - “R is ‘GNU S’, a freely available language and environment for statistical computing and graphics which provides a wide variety of statistical and graphical techniques: linear and nonlinear modelling, statistical tests, time series analysis, classification, clustering, etc. Please consult the R project homage for further information”
  - spend time at CRAN website for more background/documentation/versions

# Packages and Popularity of R

- Packages are bundles of R code that perform certain tasks
  - Authors contribute for free
  - Over 15,000 packages available
  - Packages make R awesome
  - Packages make R more challenging
  - More on packages in this R Bootcamp
- R has been in and out of top 20 most popular programming languages
  - https://www.tiobe.com/tiobe-index/ currently at 22nd
  - SAS vs R
  - R flexibility, free, AND packages

# Downloading R (forever free)

- https://cran.r-project.org
- Linux, Mac, Windows options for downloading
- Follow Instructions for downloading latest version of R
  - Ideally, all students will have same version ("replace" earlier versions)
  - Frequent updates (multiple per year)
  - Some packages/functions may not be compatible with older/newer versions
- Just Install R per instructions
- No need to open this application of R...at all. Not this 
- We will use RStudio to open R files. USE THIS 

# 2. What is RStudio

- Integrated Development Environment (IDE) for R
- What is IDE?
  - You have R, the programming language, but...
  - You interface with the program thru RStudio, i.e. open RStudio, not R
- Chief developer Hadley Wickham, available in 2009
- RStudio is open source and free!
- Spend time on Website https://www.rstudio.com
  - Explore functionality of RStudio
  - Useful Resources e.g. cheat sheets: https://www.rstudio.com/resources/cheatsheets/
- "R" will refer to RStudio (or R code) from now on
  - e.g. opening R implies opening R code in RStudio

# Downloading RStudio (also forever free)

- https://www.rstudio.com/products/rstudio/download/
- Choose most recent version for operating system
- Again just install RStudio per Instructions
- Do open this RStudio application 
- Make sure it is running ok
  - i.e. note the version of R is shown in Console window
- This will be our interface with R from now on

# 3. Basics to Using RStudio

- The 4 Windows in RStudio (clockwise)
  - Editor (will have tabs for each open application)
  - Environment/History/Connections
  - Plots/Help/Packages
  - Console
- Explore the Menu options
- Commonly Used “Buttons”
  - Create new application
  - Open File
  - Save
  - Run (maybe)

# Basics to Using RStudio (Applications)

- Click New Application Button 
  - Lists many different options
- R Script
  - Basic R code editor
  - Most commonly used to start coding in R
  - May use this or R Markdown for coursework
- R Markdown
  - Is the application we will be using most often for this Bootcamp
  - Saved as .Rmd files
- Note the many other options
  - Will not be using other applications in Bootcamp or initial coursework
  - Consider exploring later (a reminder at end of Bootcamp)

# 4. What is R Markdown

- A markup language
- Performs 4 main tasks
  - Editor for R code
  - Runs R code (and displays output)
  - Able to write narrative along with R code (which can be 'marked up')
  - Produce static (and dynamic) output formats in a file (a report)
- Output to MS Word or pdf documents covered in later topic of this Bootcamp (Topic 7)
- there is much more info at https://rmarkdown.rstudio.com

# 5. Basics to Using R Script and RMarkdown

- Open Saved R Markdown File from Canvas (Topic1: Getting Started)
- Read the narrative
  - this content is similar to what is given when creating a new R Markdown file
- Note the shaded area (these are called code chunks, or just chunks)
  - This is where R code lives for R Markdown file
  - Per instructions (which there will be throughout Bootcamp)
  - In the upper right of the shaded area click play button
- Create a new R Script by clicking 
- Note there are now two tabs
  - Click on GettingStarted.Rmd tab
  - Copy content of first shaded area for summary function
  - Paste in new Untitled R Script and click run 
  - Same output, but now in Console window

# Basics of Using R Markdown

- Back to the R Markdown Code
- Note three ticks and brace with ‘r’, then close with 3 ticks.
- Run the next chunk of code for a scatter plot
  - Note that there is also a Run button in tools above Editor for R Markdown
  - Explore the different run options and note the short-cut keys
- The ‘cars’ in the brace is a built-in dataset
  - Most often you will load data into R from a file
  - There are many built-in datasets in R (more on that later)
- For now, there is text, R code, and R output all within an R Markdown application file in the editor
- Will come back to description on **Knitting** later in Bootcamp

# More Basics to R Scripts/R Markdown

- When closing RStudio...
  - Tabbed scripts that are red indicate unsaved scripts
  - **Save scripts with organization** (and your future self) in mind
  - Quit an R Session from File Menu or by closing RStudio window (or similar)
  - Note that if R Script or R Markdown is not saved, RStudio will ask about saving
  - If items appear in the Global Environment window, RStudio will also ask about saving Workspace Image
    - Generally, no need to save Workspace and click this checkbox off
    - Many Objects will get stored which may cause problems for rerunning/reusing code in future sessions
    - Though some find saving everything useful, RStudio and you will have to deal with clutter
- When opening R
  - You can 'typically' pick up where you left off in code
  - Need to reload objects/packages  (more on this later)