Introduction to R

Jack Cordes

SummR Camp 2020

Agenda

- What is R?

- Installing R and RStudio

- Getting started in R

- Basics in R

- Tips and tricks for when you get stuck

What is R?

- "R is a language and environment for statistical computing and graphics."[1]
- Open source
  - Careful: anyone can create a package
- Free
- Flexible/Adaptable
  - Packages
  - Integration with other software
- Designed specifically for data management, manipulation, calculation, and graphics

1. https://www.r-project.org/about.html

Installing R and RStudio

- Install R

  - https://cran.r-project.org/mirrors.html

  - Choose any CRAN mirror and download according to your operating software

  - Install after download

Installing R and RStudio

- Install RStudio
  - https://rstudio.com/products/rstudio/download/#download
  - It should recognize your operating software automatically
  - Install after download

RStudio Desktop 1.3.1073 - Release Notes

1. Install R.   RStudio requires R 3.0.1+.

2. Download RStudio Desktop.   Recommended for your system:

DOWNLOAD RSTUDIO FOR MAC
1.3.1073 | 148.66MB

Requires macOS 10.13+ (64-bit)

# Getting started in R

- Open RStudio and open a new R script under File>New File>R Script

- Explore the 4 windows

Script window: this is where R code and comments are written

Console window: this is where R code is executed. Print outputs, warnings, and errors appear here

Environment window: keeps track of R objects

Viewer window: displays file directories, plots and figures, animations, and help
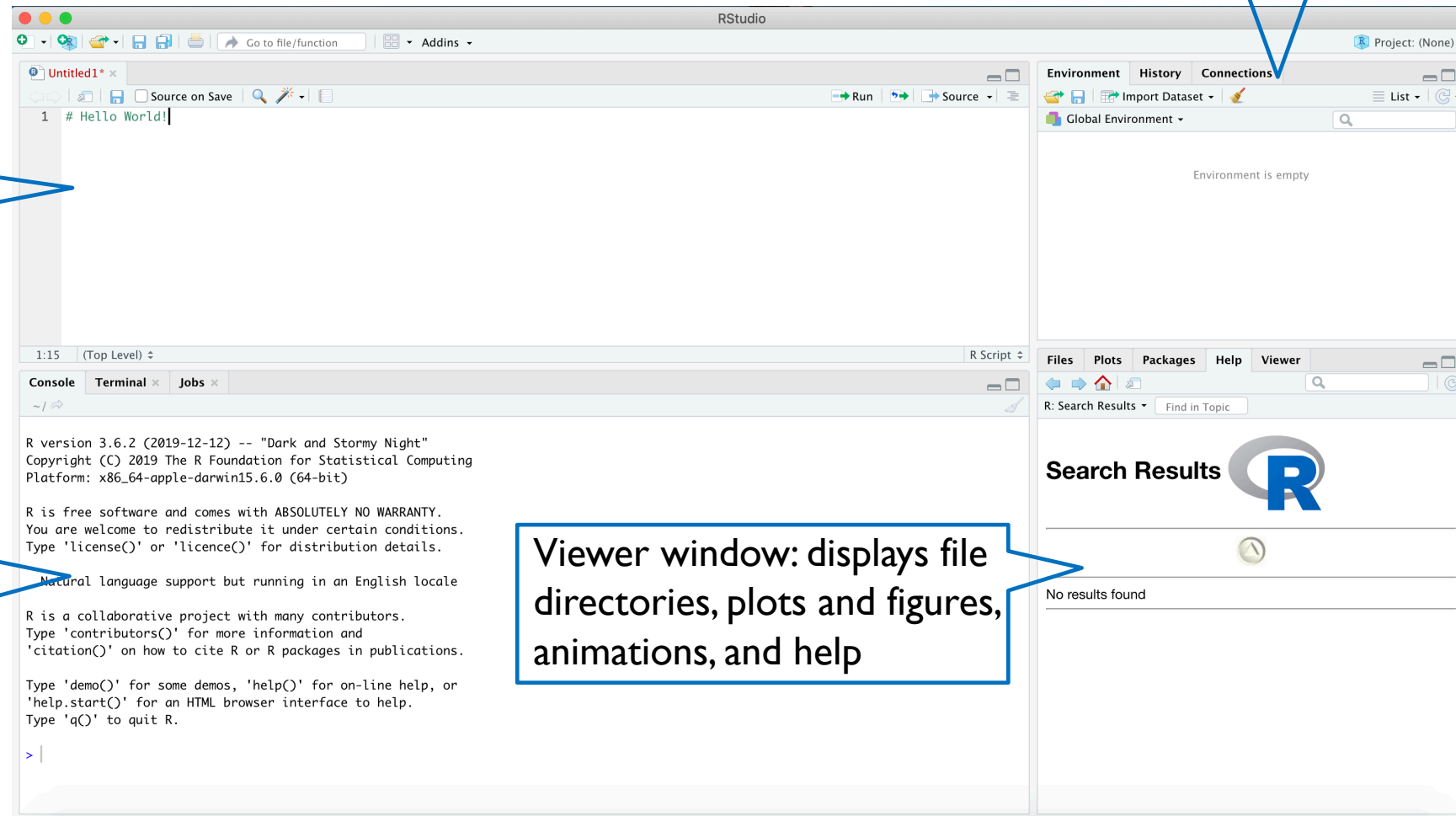
Basics in R

- Vectors, matrices, dataframes, and objects
  - Vector_example <- c(1, 2, 3)
  - Matrix_example <- matrix(data = c(1,2,3,4,5,6), nrow = 3, ncol = 2)
  - Dataframe_example <- data.frame(Matrix_example)
- Importing/Exporting data – read.csv(), write.csv()…
  - Specialized functions for Excel, SAS data, Stata data, etc.
- File paths/RProjects/Here package
  - Example: "Users/jrcordes/Desktop/PHS 2000A TF 2020"
  - read.csv(file = here::here("Desktop", "PHS 2000A TF 2020", "example_csv.csv")

Basics in R

- Installing packages and using libraries
    - install.packages("tidyverse")
    - library(tidyverse)
- Commenting/Annotating (#SummRCampSoMuchFun)
    - A conversation with future you

Basics in R

- Common operations
    - mean(), rep(), table(), head(), summarise(), var(), sd(), cor(), cbind(), png()
- Subsetting dataframes
    - Pull out a specific value: dataframe[row_number, colum_number]
    - Pull out a specific column: dataframe$column_name
    - Create a new variable using mutate()
    - Subset a dataframe without contiguous chunks using select()
    - Subset a dataframe based on conditions using filter()
- Other useful packages
    - ggplot2, forcats, readr, dplyr – all of which are in the tidyverse

Basics in R

- Visualization using ggplot
  - Create a base plot and then add features with "+"
  - The aesthetic ("aes") takes data values and features used for the entire plot
  - Customizable themes, legends, colors, facets, etc.
    - https://www.r-graph-gallery.com/38-rcolorbrewers-palettes.html
  - Highly flexible
    - https://www.r-graph-gallery.com
  - Can save plots using ggsave()

Tips and Tricks

- Unsure about what a function does or arguments it takes
  - ?mean and ??mean will bring up help and documentation
- Copy/paste error messages into Google
- Case matters
- Avoid using existing R function names for objects
- Restart R periodically and run the whole script from the beginning
- R Cheat Sheets
  - https://rstudio.com/resources/cheatsheets/

Questions and Practice