

Getting Started With R and RStudio Guide

What is R?

R is a language and environment for statistical computing and graphics.

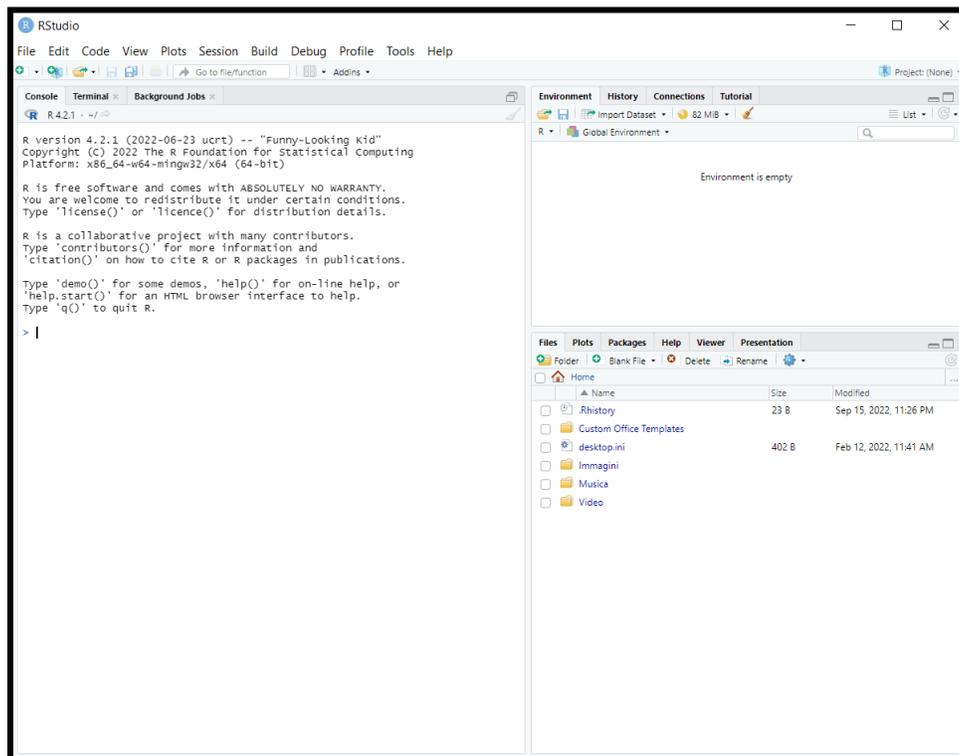
What is RStudio?

RStudio is an integrated development environment for R. RStudio is a must-know tool for everyone who works with the R programming language. It's used in data analysis to import, access, transform, explore, plot, and model data, and for machine learning to make predictions on data.

Getting Started

1. Download R: <https://cran.r-project.org/>
 - o Make sure to select the download based on what type of computer you're using
 - o The newest version should work fine
 - Note: RStudio will need version R 3.3.0 or greater to work
2. Open the download to install
3. Download RStudio Desktop:
<https://support--rstudio-com.netlify.app/products/rstudio/download/>
4. Open the download to install
 - o RStudio should now be installed, enjoy programming!

RStudio Setup



Console:

- You can virtually run anything here, but would not recommend writing your code here because it doesn't save permanently
- Use to:
 - Install and load R packages
 - Perform simple (or complex) mathematical operations
 - Import data
 - Create data frames
- Main purposes:
 - Install R packages
 - Test code

Environment:

- Variables are stored here
 - Can find more details, like data types
- Can easily view data sets by clicking variable names
- Click tick box + broom to delete variable

Terminal: to run commands (you will most likely not use)

Files: to see the structure of the working folder, reset the working folder, navigate between folders, etc.

Plots: to preview and export created data visualizations

Packages:

- To view what packages were loaded
- Load or unload packages by selecting or deselecting the tick boxes

Useful Tips

- Comment a line of code (using): `#`
- Find documentation of function: `?funct_name`
- Install package: `install.packages("package_name")`
 - Install packages in console rather than in the R file, so it will save to the computer and you will only need to install once
 - You can also install packages in the "Packages" tab by clicking the "Install" button and selecting the package(s) you would like to install
- Load package: `library("package_name")`
 - Load necessary packages in the R file, not in the console
- Remove package: `remove.packages("package_name")`

R Markdown Guide: <https://rmarkdown.rstudio.com/lesson-1.html>

R for Data Science Textbook: <https://r4ds.had.co.nz/>

Common R Packages to Install:

- [tidyverse](#)
 - Includes multiple packages: ggplot2, dplyr, tidyr, tibble, etc.
 - [dplyr](#)
 - [tidyr](#)
 - [ggplot2](#)
 - [rmarkdown](#)
 - [shiny](#)
-

Scatter Plots

- Create scatter plot: `plot()`
 - `plot(dataset_name$column_xAxis, dataset_name$column_yAxis, main="title of plot", xlab="x-axis title", ylab="y-axis title")`
 - Ex: `plot(airquality$Wind, airquality$Ozone, main="Ozone and Wind in NYC, 1973", xlab="Wind (mph)", ylab="Ozone (ppb)")`
- Linear regression line: `lm()`
 - `var_name = lm(dataset_name$column_xAxis ~ dataset_name$column_yAxis)`
 - Ex: `air_trend <- lm(airquality$Ozone ~ airquality$Wind)`
- Add straight line to a plot: `abline()`
 - Ex: `abline(air_trend, col = "red")`

Histograms

- Create histogram: `hist()`
 - Ex: `hist(airquality$Ozone, breaks = 20, main = "Histogram of Ozone", xlab = "Ozone")`

Box Plots

- Create box plot: `boxplot()`
 - Ex: `boxplot(Ozone ~ Month, airquality, xlab = "Month", ylab = "Ozone (ppb)")`
-

Common Terminal Commands

- Clear console: `Ctrl + L`
 - or click broom icon in console
- Save file: `Command + S`
- Escape/cancel running code: `Ctrl + C`
- Run current line of code: `Command + Enter`
- Run all lines of code: `Command + A + Enter`

*Note: all keyboard shortcuts are for Mac computers, therefore they may differ depending on your machine