

The

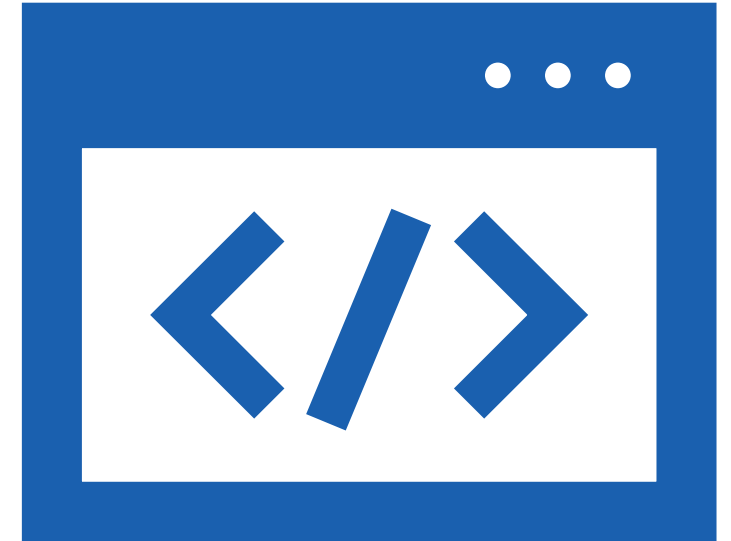


Essentials



# By the end of this workshop, you'll know how to ...

1. Create R markdown (.Rmd) / Script (.R)
2. Load packages
3. Use ?help to find functions in base R & packages
4. Create an R object
5. Learn different variable types
6. Utilize data viewing functions



# What R they?

- R is a programming language used for data visualization and statistical programming
- R Studio is an integrated development environment (IDE) used with R to provide a user-friendly interface for code development



# How to access R & RStudio?



<https://posit.co/download/rstudio-desktop/>

## **Using OnDemand local**

- If your home directory is set up on the HPC and you plan to use OnDemand, use this link:
- <https://ondemandlocal.hpc.vai.org/>

# R Studio start up

CONSOLE

The screenshot shows the R Studio interface with the following components and annotations:

- Console:** Displays the R version (3.2.0), copyright information, and workspace status. The prompt is `>`.
- Environment:** Shows the Global Environment. An annotation "Saved objects found here" points to this panel.
- Files & folders here:** Points to the Files tab in the Workspace & History panel.
- View figures here:** Points to the Plots tab in the Workspace & History panel.
- Package & functions information:** Points to the Packages tab in the Workspace & History panel.
- Package information:** Points to the Help tab in the Workspace & History panel.
- WORKSPACE & HISTORY:** A red label at the bottom of the right-hand panels.

The Workspace & History panel shows a file list with the following details:

Name	Size	Modified
..		
.RData	2.5 KB	Apr 24, 2015, 11:39 AM
.Rhistory	4.2 KB	Apr 24, 2015, 12:38 PM
MainFroc.R	805 B	Apr 24, 2015, 10:51 AM
MainReadDataFiles.R	1.6 KB	Apr 24, 2015, 10:51 AM
MainRoc.R	4.3 KB	Apr 24, 2015, 12:36 PM
MainRoi.R	296 B	Apr 24, 2015, 10:51 AM
TestRJaProc.Rproj	204 B	Apr 24, 2015, 10:51 AM

Saved  
objects  
found here

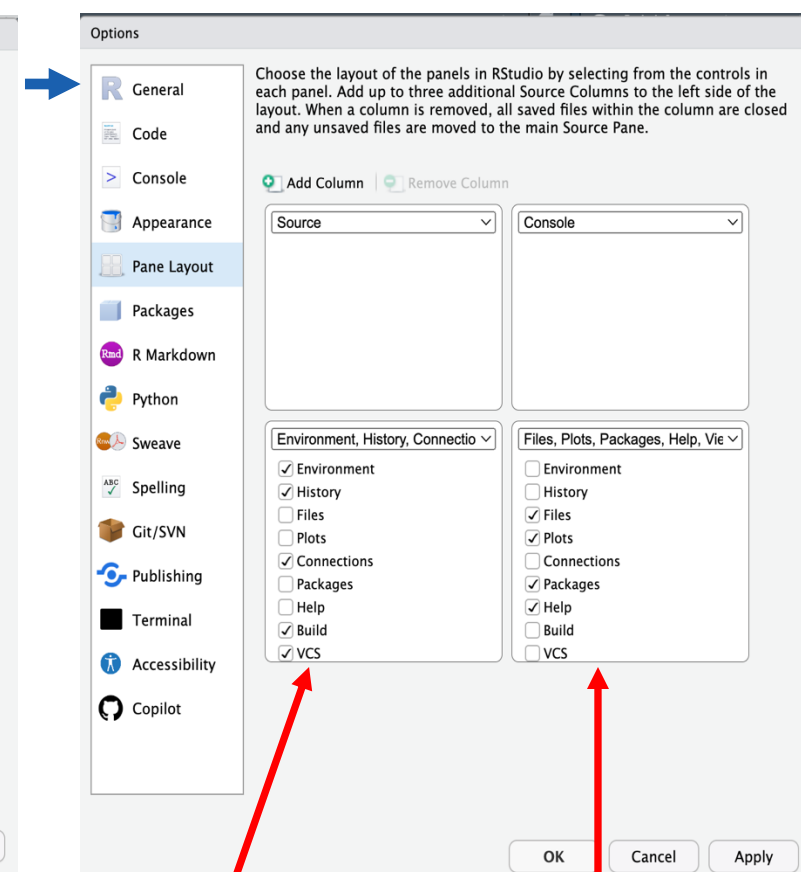
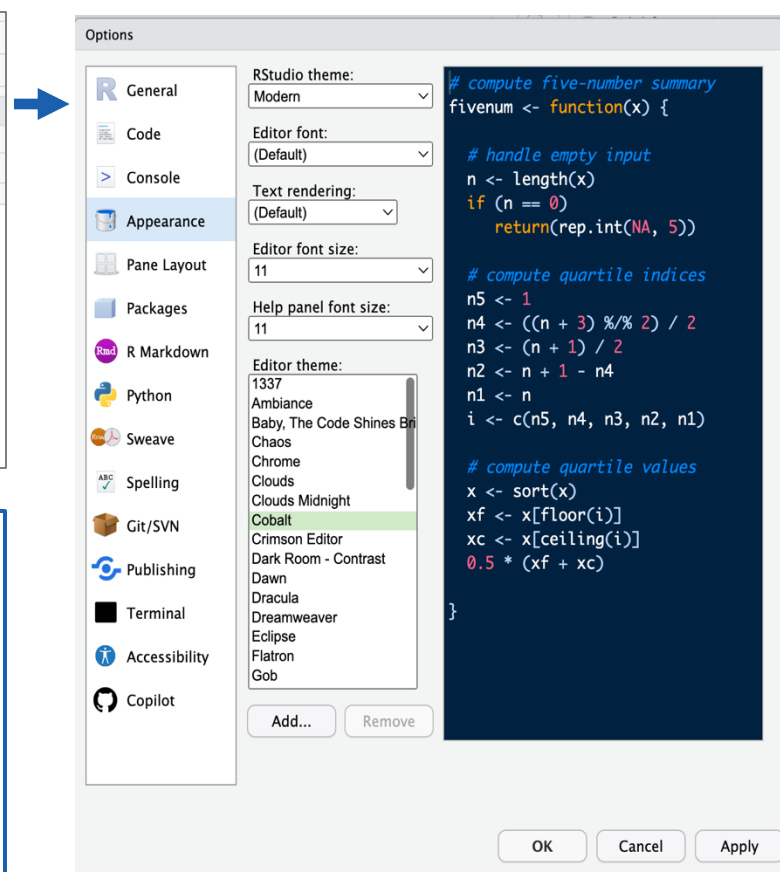
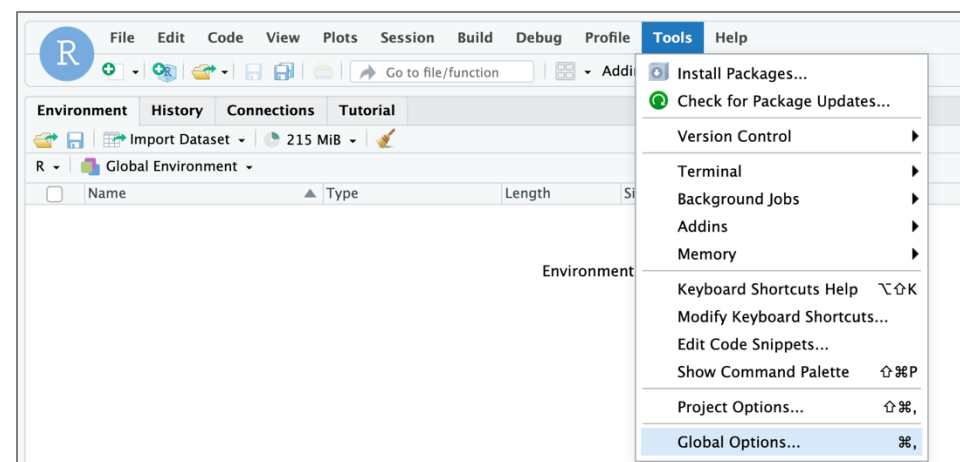
ENVIRONMENT

Package &  
functions  
information

Package  
information

WORKSPACE & HISTORY

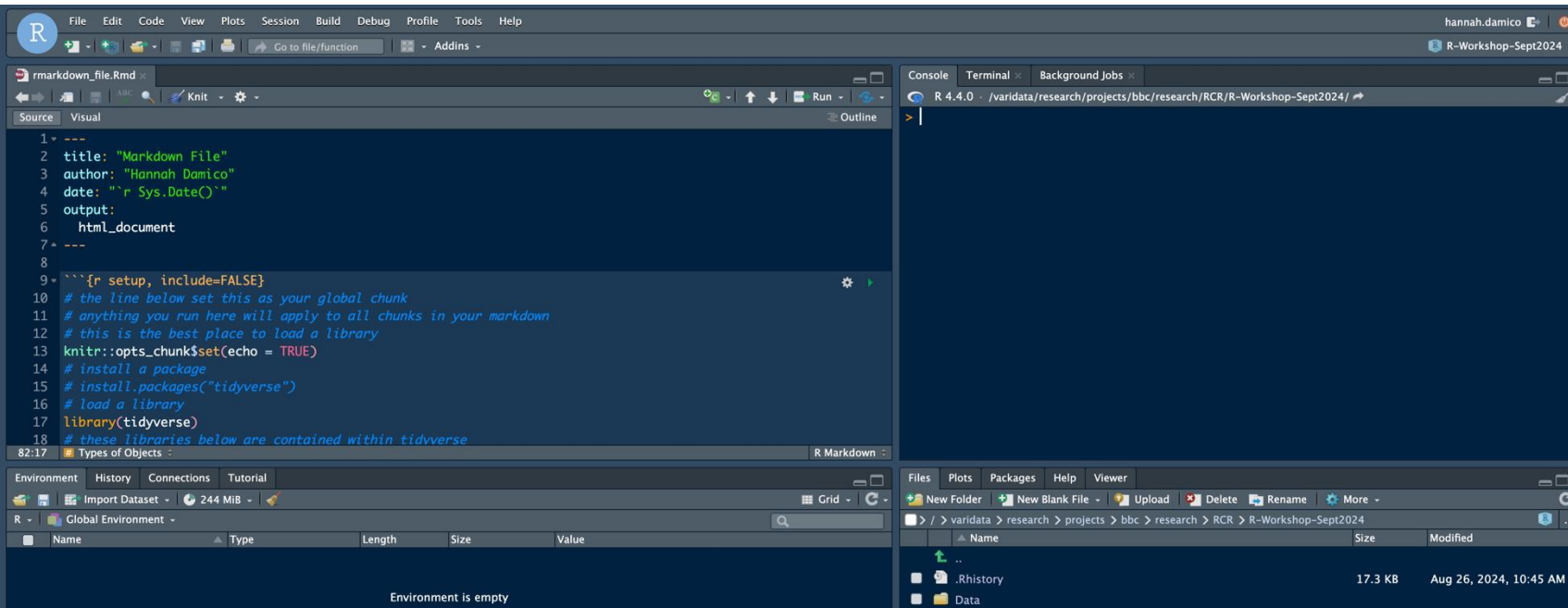
# Now that we have that figured out, let's change it...



1. Tools
2. Global Options
3. Appearance (change as you'd like, but we'll be working in dark mode today)
4. Hit apply
5. Panel Layout
  1. Source (top left)
  2. Console (top right)
  3. Environment (bottom left)
  4. Workspace + history (bottom right)

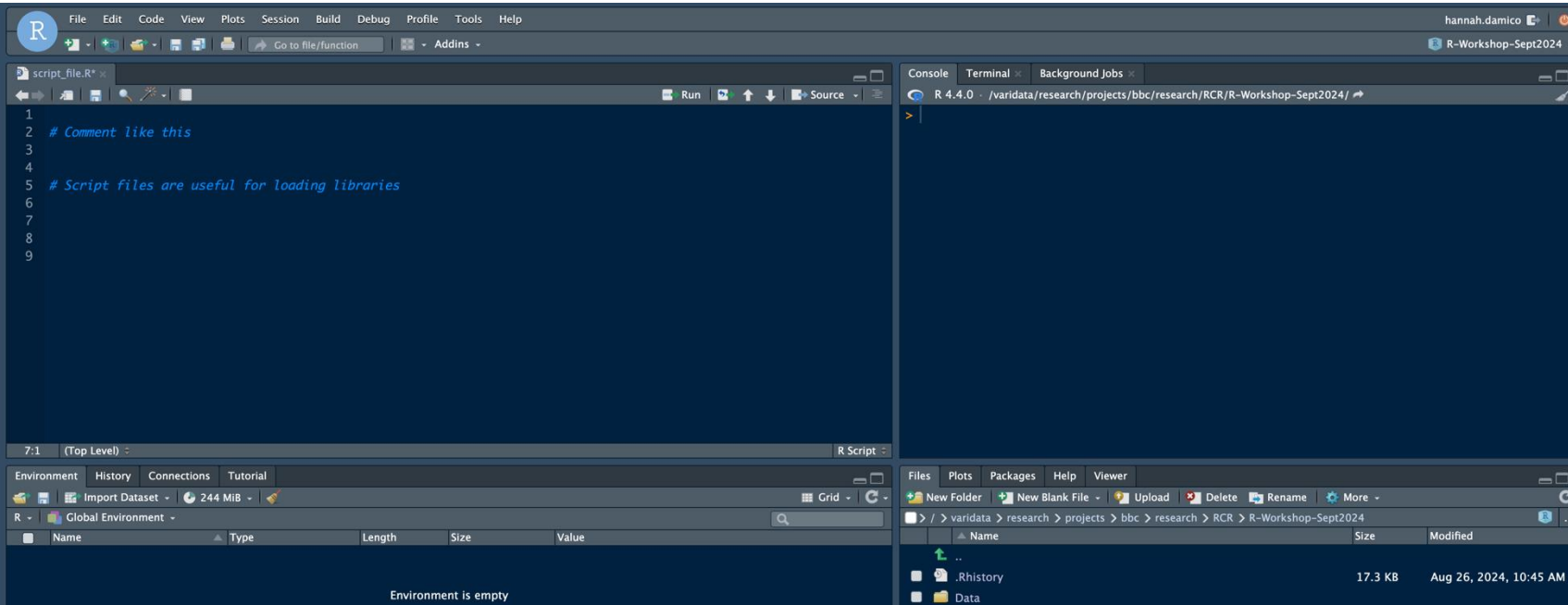
✓ Environment, History, connections, build, VCS, tutorial

✓ Files, plots, packages, help, Viewer



## R Markdown

- Main focus today
- Create code “chunks”
- Write text paragraphs in between code



## R script

- Not covered in this workshop
- More suitable for long lines of code



# Base R functions

---

- Do not require anything other than R itself
- Some covered today:
  - `data.frame()` – data formatted as a table
  - `c()` – combines elements
  - `matrix()` – arrange data in rows and columns
  - `factor()` – turn a variable into a factor
  - `$` – access objects stored someplace

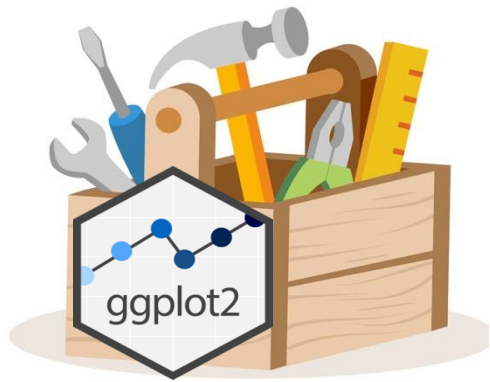
- Other useful functions:
  - `min()` / `max()` – find the minimum or maximum value of vector or dataframe
  - `sum()` – take the sum of elements in a vector
  - `plot()` – general plot function
  - `hist()` – general histogram function
  - `paste()` – combine multiple elements into one
  - `grep()` – search for patterns
  - `sub()` / `gsub()` – replace patterns in strings

# Packages

- Toolboxes filled with functions for specific types of tasks and analyses
- Extend the basic R functionality
- Some packages contain other packages, for example: tidyverse

Remove the # before  
this command to  
install the package  
tidyverse

```
Console | Render x | Background Jobs x
R 4.4.0 · /varidata/research/projects/bbc/research/RCR/R-Workshop-Sept2024/
> # install.packages("tidyverse")
> library(tidyverse)
> tidyverse_packages()
[1] "broom"      "conflicted" "cli"      "dbplyr"
[5] "dplyr"      "dtplyr"     "forcats"  "ggplot2"
[9] "googledrive" "googlesheets4" "haven"    "hms"
[13] "httr"       "jsonlite"   "lubridate" "magrittr"
[17] "modelr"     "pillar"     "purrr"     "ragg"
[21] "readr"      "readxl"     "reprex"    "rlang"
[25] "rstudioapi" "rvest"      "stringr"   "tibble"
[29] "tidyr"      "xml2"       "tidyverse"
>
```



### **ggplot2**

- Visualize your data
- Create histograms, barcharts, scatterplots
- Edit the aesthetics of plots



### **dplyr**

- Manipulate your data
- Select variables
- Filter data frames
- Create new variables

### **readr**

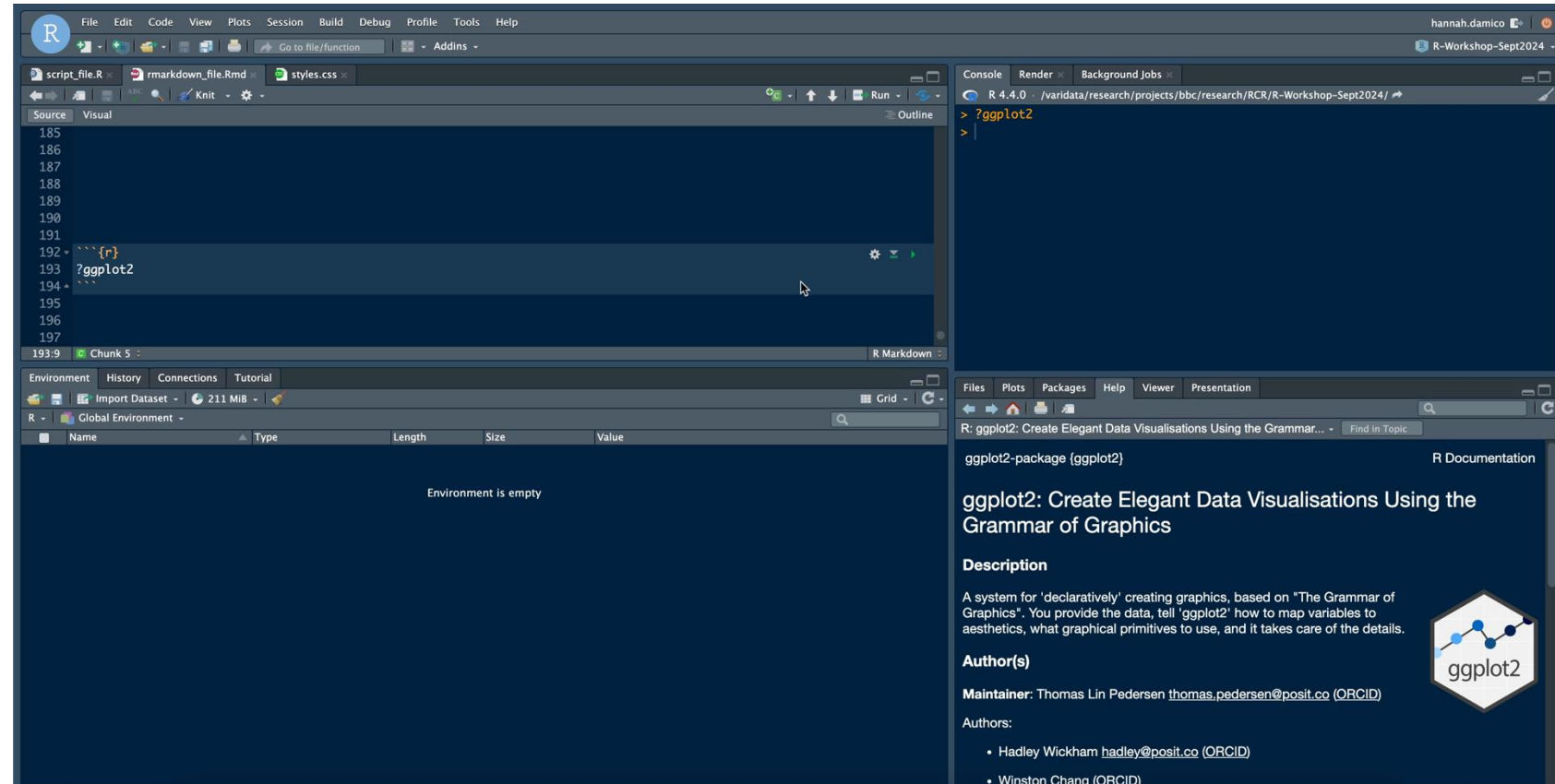
- Read in data from .csv or .tsv files
- Can specify variable type in function call

### **tidyr**

- Clean up data
- Variables are columns
- Observations are rows
- Row1: single mouse's weight and tumor size are measured for one timepoint
- Transpose data
- Wide -> long
- Long -> wide

# Need ?help with a package or function?

Running ?  
before a function  
or package will  
take you to its  
documentation!



# R objects can look like...



**Characters & Strings:** something in quotes, “A”, “1”, “bio”, “statistics”



**Factor:** A variable with levels, Treatment: [“A”, “B”, “C”]



**Numeric:** fractions, decimals, positive, negative numbers



**Logical:** TRUE or FALSE, sometimes 1 or 0

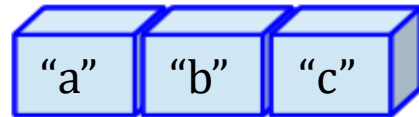
# Check in!



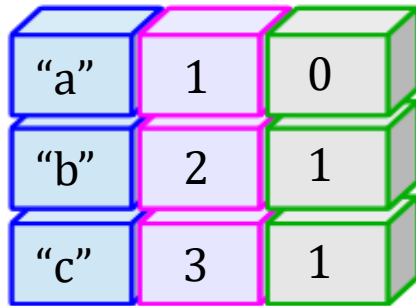
<https://www.menti.com/al9axfxzmymo>

# Larger data types in R

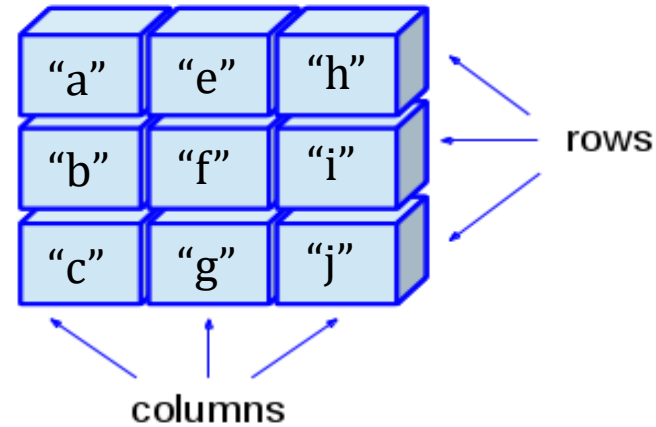
Vector



Data Frame  
(Table)



Matrix



Simply put, these hold information in different ways...

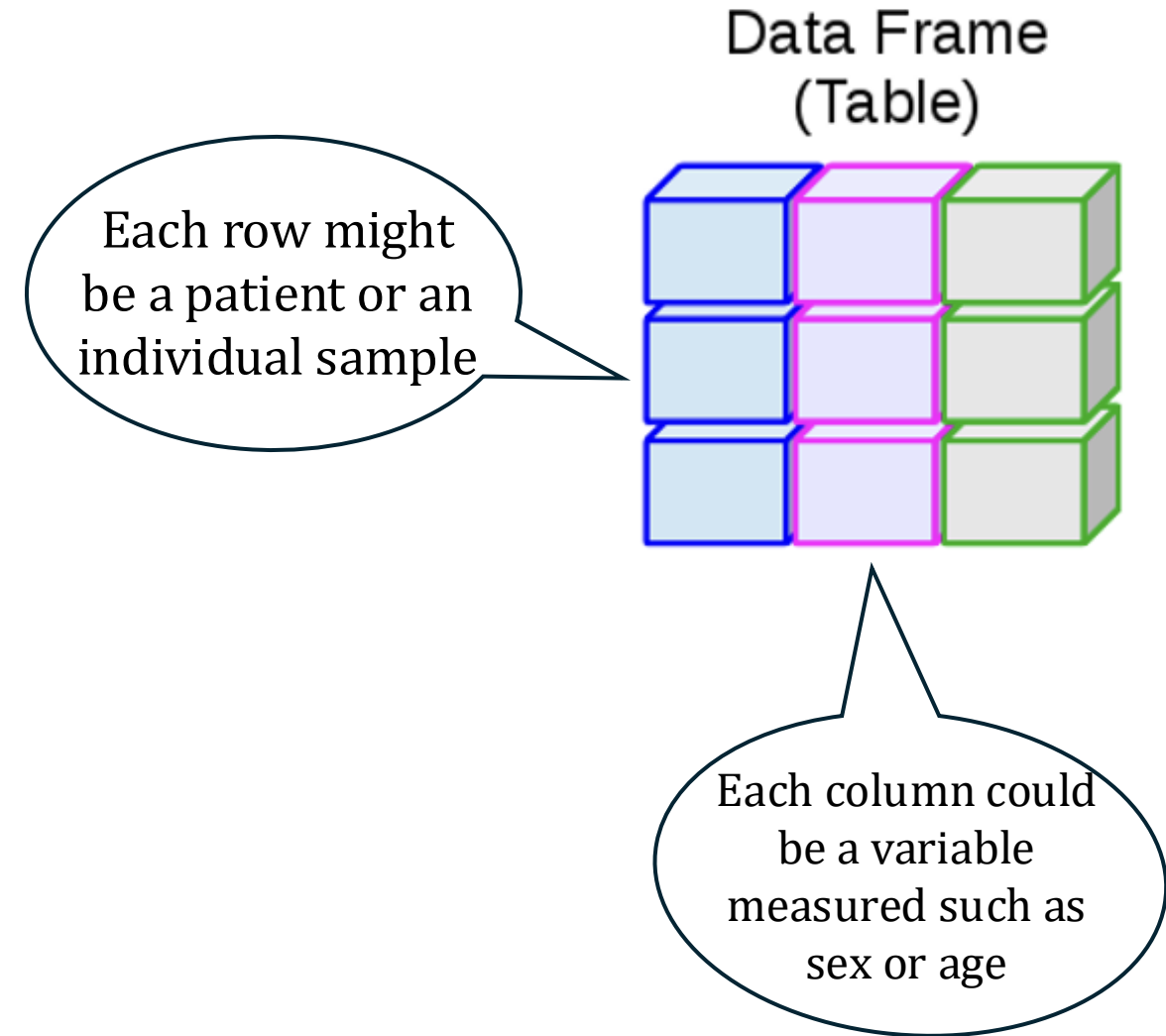
# Data frames

## Creation vs. Coercion

- `data.frame()`: create a data frame from scratch
- `as.data.frame()`: convert an existing object into a data frame

## Input Types:

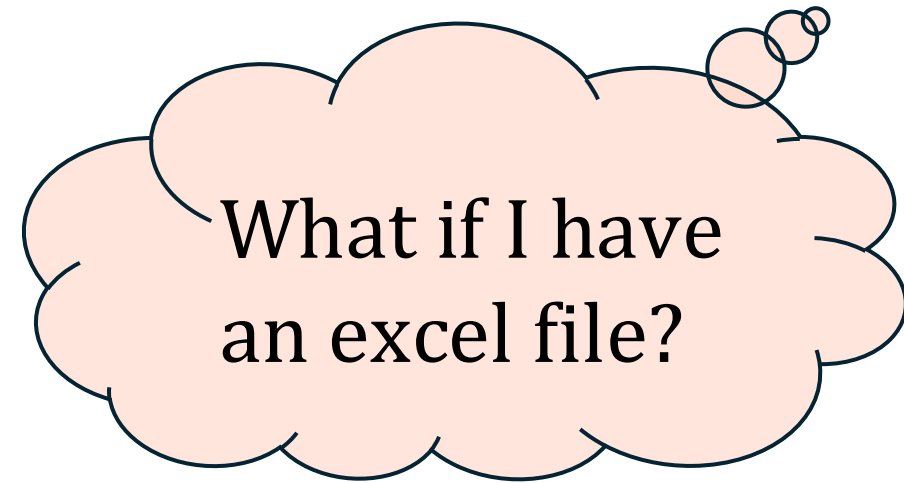
- `data.frame()` takes vectors or variables and combines them into a data frame
- `as.data.frame()` takes an object (like a matrix, or even another data frame) and attempts to convert it into a data frame

























# Reading in data

- `read.csv` (base) vs. `read_csv` (readr/tidyverse)
  - Sometimes these read in differently, so it's important to check your data after you've read it in!
- `read_tsv`



`readxl::read_excel`

# Viewing your data

Abilities	dim()	glimpse()	str()	nrow()	ncol()	head()	tail()	View()
Provide # of rows								
Provide # of columns								
Display data type								
Display data as data.frame()								
Select # of rows to view								
Filter data by column								

# Final Questions



<https://www.menti.com/alsbv87coqxi>