# Intro to R

Sawyer Newman

Data Librarian for the Health Sciences

*Cushing/Whitney Medical Library*

# What is R?

R is a language and environment for data manipulation, calculation and graphical display

Includes:

* Data handling and storage
* Suite of operators
* Tools for data analysis
* Conditionals, loops, user-defined functions, input and output facilities

# R vs Python

| | R | Python |
|---|---|---|
| **Purpose** | Focuses on user friendly data analysis, statistics and graphical models | Emphasizes productivity and code readability |
| **User base** | R has been used primarily in academics in research.<br>* Data scientists * Statisticians | Python is used by programmers that turn to data science.<br>* Developers *Programmers |
| **Specialty** | * Statistical tests and models are easy in R | * Flexible for creating new tools |

Citation for this slide and more info

# R vs RStudio

**R: The programming language**

* Open source programming language
* You do not need to use Rstudio to run R

www.r-project.org

**Rstudio: The graphical user interface**

* Integrated development environment (IDE)
* Privately run company
* You do need R to use Rstudio

www.rstudio.com

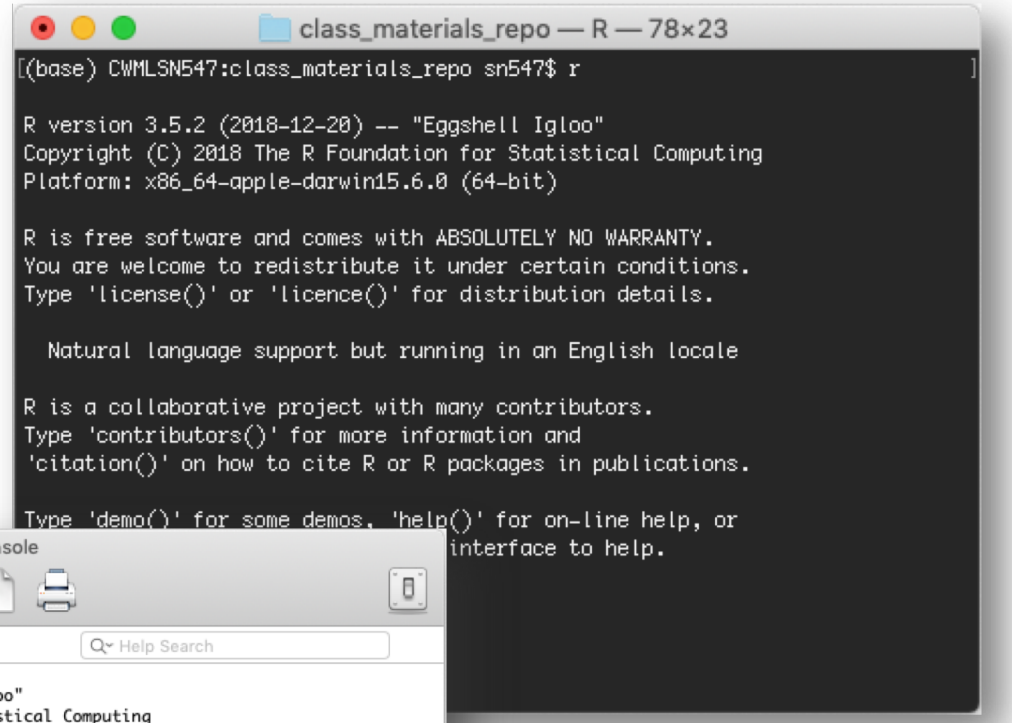
# Where can you run R?

**GUI**

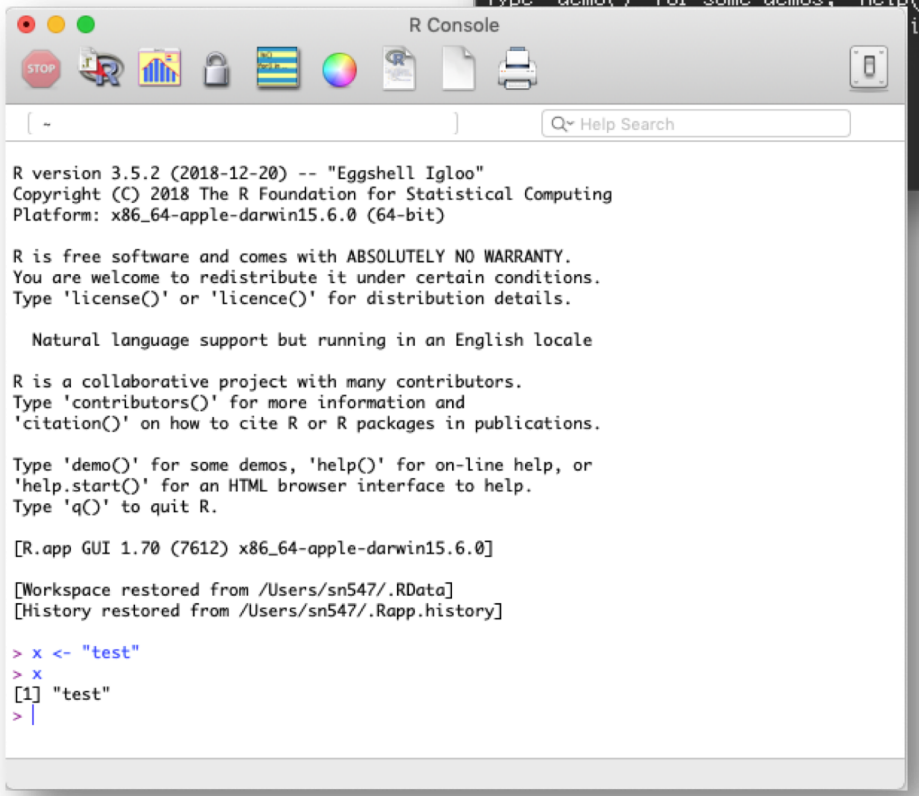
* Rstudio

**CUI**

* Command prompt (Windows)
* Terminal (Mac & Linux)
* R Console




^Terminal window




^R console

# R Packages

**Search through all R packages**

* https://www.r-pkg.org/

**Packages of note**

* bioconductor (and biomaRt)
* Seurat for scRNA analysis (https://satijalab.org/seurat/v3.1/pbmc3k_tutorial.html)
* tidyverse (dplyr, tidyr, stringr, lubridate, ggplot2)
* readxl
* haven

# R Syntax

```
> myString <- “Hello, World!”
> print(myString)
[1] “Hello, World!”


> install.packages(”package_name”)
> library(package_name)
```

Note: R is case sensitive

# Continue learning after this class

Recommended tutorials

- Analyzing Patient Data
- Genomics Data Analysis and Visualization
- Introduction to R and Bioconductor

Texts

- A Little Book for Biomedical Statistics (2016)
- PH525x series – Biomedical Data Science

# Finding help with R

**Cushing/Whitney Medical Library**

* StatLab consultations
* Data Librarian consultations

**Within R**

* ?function_name
* help(function_name)

**On the web**

* R Project Manuals and FAQs
* Stack Overflow | Use the R tag
* Rseek | R search engine
* R-bloggers | R news and tutorials
* Rstudio cheat sheet printouts

# Demo topics

1. Mathematic operators (larger list here)
2. Comparison operators
3. Vectors
4. Reading in files
5. Dataframes