Day 1 Cheatsheet

Intro

Major concepts

- **Package** - a package in R is a bundle or "package" of code (and or possibly data) that can be loaded together for easy repeated use or for sharing with others.
- **Function** - a function is a particular piece of code that allows you to do something in R. You can write your own, use functions that come directly from installing R, or use functions from additional packages.
- **Argument** - an option that you specify to a function.
- **Object** - an object is something that can be worked with in R - can be lots of different things!
- **Variable** something measured or counted that is a characteristic about a sample.
- **Sample** individuals that you have data about - e.g. people, houses, viruses etc.

- **Data frames** A collection of variables and samples in a row/column format. Sample = Row, Variable = Column
- **Tidyverse** - This is a newer set of packages designed for data science that can make your code more intuitive as compared to the original older Base R.

Functions

Library/Package	Piece of code	Example of usage	What it does
Base `R`	`install.packages()`	`install.packages("remotes")`	Installs packages
Base `R`	`library()`	`library(remotes)`	Loads and attaches additional packages to the R environment.

RStudio

Major concepts

- **RStudio** - an Integrated Development Environment (IDE) for R - makes it easier to use R.
- **Source/Editor** - "Analysis" Script + Interactive Exploration - In a .R file (we call a script), code is saved on your disk
- **R Console** - Where code is executed (where things happen) - Code is not saved on your disk
- **Workspace/Environment** - Tells you what objects are in R. What exists in memory/what is loaded?/what did I read in?
- ██
- **R Project** - Helps you organize your work. Helps with working directories (discussed later). Allows you to easily know which project you're on.
- RStudio Keyboard shortcuts

Functions

Library/Package	Piece of code	Example of usage	What it does
Base `R`	`View()`	`View(mtcars)`	allows you to view data in a spreadsheet format.

Library/Package	Piece of code	Example of usage	What it does
Base R	`head()`	`head(mtcars)`	Returns the top 6 rows of an object in the environment by default. You can specify how many rows you want by including the `n =`argument.
Base R	`tail()`	`tail(mtcars)`	Returns the bottom 6 rows of an object in the environment by default. You can specify how many rows you want by including the `n =` argument.

Reproducibility

Major concepts

- **Reproducibility** - A different analyst re-performs the analysis with the same code and the same data and obtains the same result.
- **Repeatable** - keeping everything the same but repeating the analysis - do we get the same results
- **Reproducible** - using the same data and analysis but in the hands of another researcher - do we get the same results?
- **Replicable** - with new data do we obtain the same inferences?

Functions

Library/Package	Piece of code	Example of usage	What it does
Base R	`sessionInfo()`	`sessionInfo()`	Returns the R version information, the OS, and the attached packages in the current R session.

More resources

- The RMarkdown book
- Jenny Bryan's organizational strategies.
- Write efficient R code for science.
- Reproducibility in Cancer Informatics course

* This format was adapted from the cheatsheet format from AlexsLemonade.