Lecture 1: Intro to R and RStudio

STAT 450, Eric Fox

Installing R

To download the latest version of R go here:

https://cloud.r-project.org/

Click on the "Download R" link for your operating system (Windows, Mac) and follow the instructions for installation.

Installing RStudio

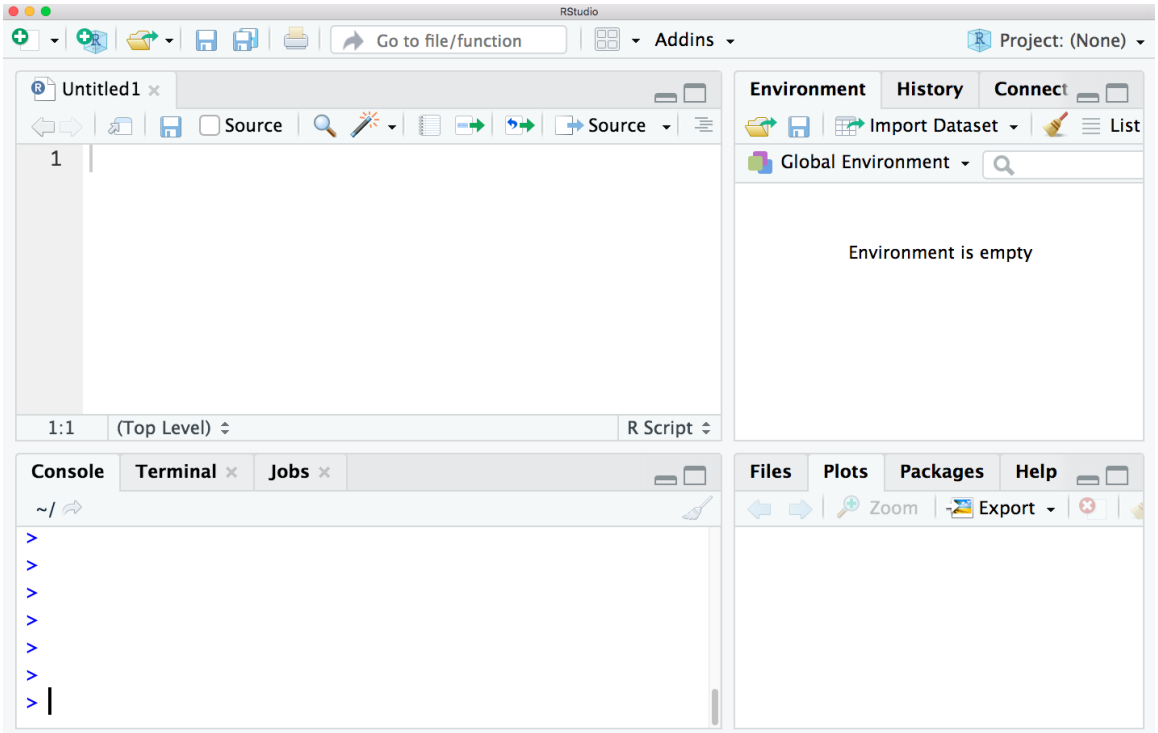
To download the latest version of RStudio go here:

https://rstudio.com/products/rstudio/download/

Click on the button to download "RStudio Desktop Open Source License" (free version). Next click on the link to download RStudio for your operating system (Windows, Mac) and follow the instructions for installation.

Getting started with RStudio

Go ahead and open RStudio, what you see should look like the screenshot below. For clarification: R is the name of the programming language and RStudio is a convenient interface.



RStudio is divided into four panels:

- Top left panel: This is the code editor, this is where you write your code. To create a new R script go to File → New File → R Script. Make sure to save your R scripts so that you can access them later.
- Bottom left panel: This is the console, this is where you run your code. You can either copy and paste code from an R script to the console, or use the shortcut Ctrl+Enter (Windows) or Command+Enter (Mac) to run a line of code.
- Bottom right panel: Any plots you create will show up in this pane. This pane also contains the help menu, where you can read documentation about R functions.
- Top right panel: This contains the Environment and History panes. The Environment pane shows everything that is currently in your workspace (e.g., objects you create or data sets you have loaded for analysis). The History pane provides a history of all previous commands you have run in the console.

Using R as a calculator

You can use R for basic calculations: addition, subtraction, multiplication, and exponentiation

```
2 + 2
```

```
## [1] 4
```

```
5 - 2
```

```
## [1] 3
```

```
2 * 3
```

```
## [1] 6
```

```
2^4
```

```
## [1] 16
```

```
2^(1/2)
```

```
## [1] 1.414214
```

```
sqrt(2)
```

```
## [1] 1.414214
```

```
(1/51 + 1/49)^(1/2)
```

```
## [1] 0.20004
```

```
7.2 + 2 * 2.1 / sqrt(101)
```

```
## [1] 7.617916
```

If you want to edit a previous command, press the up arrow key. For instance, I may want alter the last computation:

```
7.2 - 2 * 2.1 / sqrt(101)
```

```
## [1] 6.782084
```

Variable Assignment

The <- symbol is called the assignment operator. It assigns values to variables.

```
x <- 7 # assign 7 to the variable x
x # print value of x
```

```
## [1] 7
```

```
y <- 3
y
```

```
## [1] 3
```

```
x + y
```

```
## [1] 10
```

The = symbol can also be used for assignment.

```
z = -1
z
```

```
## [1] -1
```

The # symbol is used to write comments. Anything to the right of # is not evaluated in the console. Comments are useful when sharing code with others (or to help me understand my code in the future!).

Exercise:

a) Assign the value 47.5 to the variable `price`. Then create a variable called `tip`, which is `price` multiplied by 0.2. Print out the value of the variable `tip`.

```
# solution
price <- 47.5
tip <- price * 0.2
tip
```

```
## [1] 9.5
```

b) Next create a variable called `total`, which is the sum of `price` and `tip`. Print out the value of the variable `total`.

```
# solution
total <- price + tip
total
```

```
## [1] 57
```

Introducing Vectors

A `numeric vector` in R is just a collection of numbers. Vectors are created using `c()`, which is short for combine. For example, shown below is a vector containing the ages of five individuals.

```
age <- c(38, 23, 21, 35, 63)
```

R has some easy-to-use functions for computing numerical summaries of vectors such as the length, mean, median, and standard deviation.

```
length(age)
```

```
## [1] 5
```

```
mean(age)
```

```
## [1] 36
```

```
median(age)
```

```
## [1] 35
```

```
sd(age)
```

```
## [1] 16.79286
```

Documentation on these functions is provided in the help menu. Enter the following command to read about the `sd` function in the help menu:

```
help(sd)
```

Exercise: Calculate the length, mean, median, and standard deviation of the vector `weight`, which contains the weights of the five individuals.

```
weight <- c(140, 139, 187, 181, 131)
```

```
# solution
length(weight)
```

```
## [1] 5
```

```
mean(weight)
```

```
## [1] 155.6
```

```
median(weight)
```

```
## [1] 140
```

```
sd(weight)
```

```
## [1] 26.245
```