Week 1 – January 27

DAT 204 – R for Analytics

What is R?

- R is a programming language
- R specializes in statistical computing and is excellent for mathematics, statistics, and analytics
- R is growing in popularity for research and business analytics
- R is a useful tool for any analyst and can help you understand the world around you!

Where can I download R?

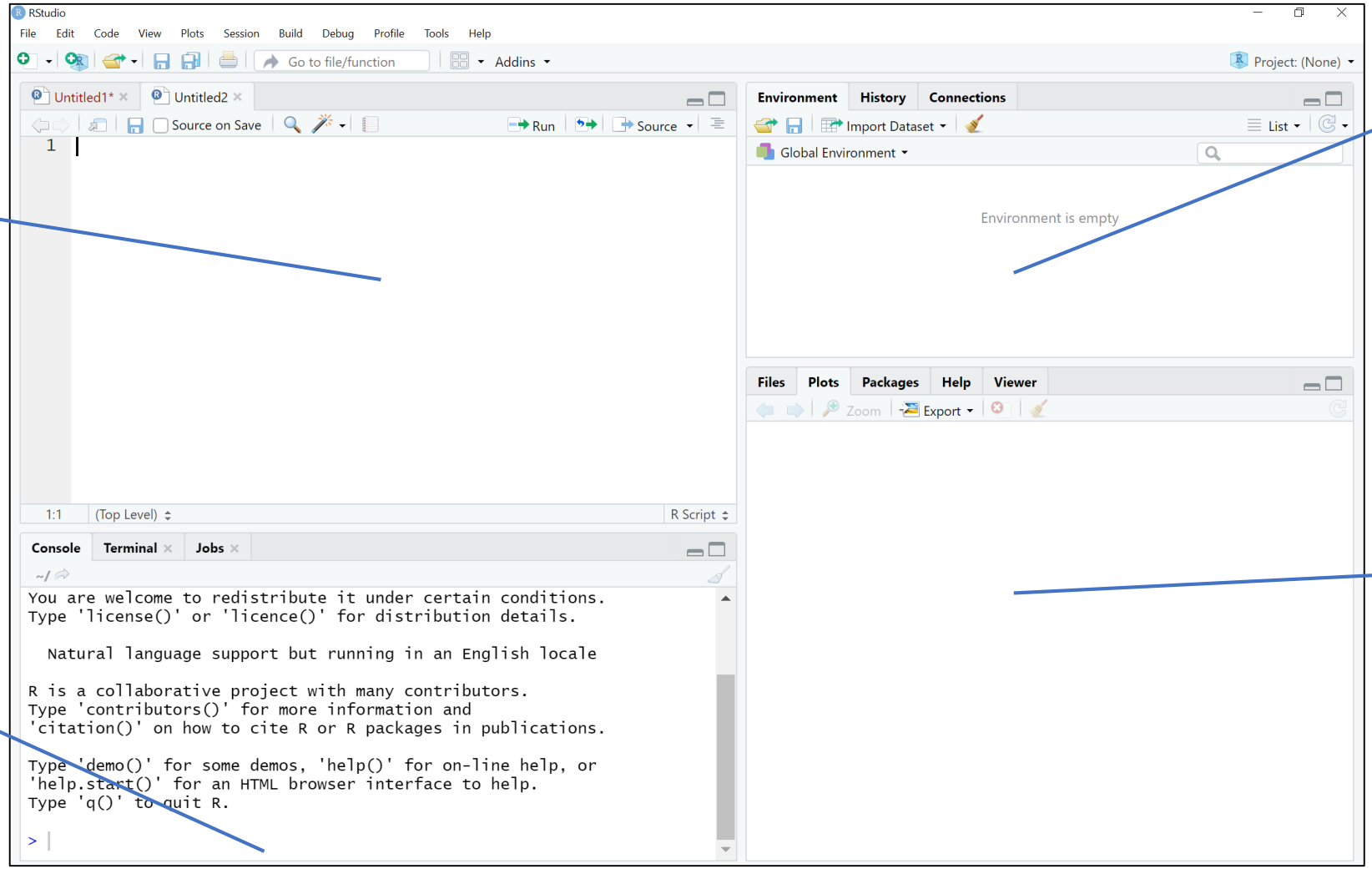
- https://rstudio.com/products/rstudio/download/

Or

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

RStudio

Script pane:
Type code here



Environment pane:
Useful information appears here – you can also import data here

Console pane:
Code runs here

Plots pane:
Graphs and visuals appear here

Keyboard Shortcuts

- Ctrl + s: Save
- Ctrl + z: Undo
- Ctrl + c: Copy
- Ctrl + v: Paste
- Home: Move cursor to beginning of line
- Ctrl + Home: Move cursor to beginning of text pane
- End: Move cursor to end of line
- Ctrl + End: Move cursor to end of text pane
- Shift + Up Arrow or Down Arrow: Highlight one line at a time
- Ctrl + Enter: Run highlighted code
- Ctrl + Shift + p: Run previous code

Basic Math Operators

Operator	Function
+	Adds
-	Subtracts
*	Multiplies
/	Divides
** or ^	Exponent (remember that fractions in the exponent behave as roots)
()	Group expressions for order of operations
%%	Modulo
%/%	Integer division

R follows traditional order of operations but it is best to allows use parenthesis to group your calculations properly

Creating Objects in R

- Use <- to create assignments (= also works but stick with <- when using R)
- Examples:
  - num1 <- 7
  - num2 <- 20
  - str1 <- “Hello”
  - str2 <- “World”
  - numbers <- 2:20
- You can perform operations on the values stored in objects
- Rules for object names:
  - Can not start with a number
  - Can not start with a special symbol (!@#$%^&*)
  - Starting with a capital letter is allowed but could be confusing
- To overwrite the value of an object, simply use the <- assignment operator again

Comments

- You can create a comment using #
- R will ignore everything after the #
- Comments are important to ensure your code is readable
- Comments can also be used to “comment out” code to temporarily remove it when troubleshooting

Common and Useful Built-in R Functions

A function takes 0 or more inputs and returns something

Function	Think….	Arguments	Returns
c(x,…)	Concatenate	x: The things you want to concatenate	A vector with the items you provided
mean(x,…)	Mean	x: The numbers you want to take the mean of	The mean of the numbers you provided
median(x,…)	Median	x: The numbers you want to take the median of	The median of the numbers you provided
seq(from=1,to=1,by=1,…)	Sequence	from: where to start to: where to finish by: how many to count by	A sequence starting with the number you provided
round(x, digits = 0)	Round	x: The value you want to round digits: How many digits to round to	A rounded number
typeof(x)	What is the type of…	x: The object that contains the value you want to check the type of	The data type of the value stored in object x
sample(x, size=1, replace = FALSE)	Sample	x: The object you want to sample from size – How many to sample replace: Replace after each draw?	A sample from the object you provided
replicate(m, exp)	Replicate	n: The number of times you want to replicate exp: What you want to replicate	A vector with the result of each replication

If you want to learn more about a function, type ?<functionName>() in R

Creating your own function

```
myFunction <- function(args) {
…function body…
}
```

Creating your own function with default args

```
myFunction <- function(arg1 = value, arg2 = value) {
…function body…
}
```

Default args make arguments optional!

More RStudio Functionality

- Ctrl+2: Jump to console
- Up and Down arrows when in console: Repeat previous code
- Remove all objects from environment
- ‘Extract function’ tool

For Loops in R

```
for (i in 1:20) {
…body of loop
}
```

*Note that “i” could be any variable and “1:20” could be any vector

Dice Challenge

- Determine the probability of each possible result when rolling two dice by:
  - Calculation
  - Simulation

Pythagorean

- Write a formula that takes the length of two sides of a right triangle and returns all possible lengths of the third side