

# R and RStudio review

## Set up

1. Make a STAT 311 folder (suggested)
2. Making a new document
  - File -> R Markdown... or Quarto Document... (You don't have to name it yet, but do put your name!)
3. Saving work via save or knitting/rendering
  - Save to proper location!
  - HTML or PDF

## Basic coding

R code in a .Rmd or .qmd goes in a code chunk. Make comments using the pound symbol

```
# simple arithmetic  
1 * 4
```

```
[1] 4
```

Ways to run code:

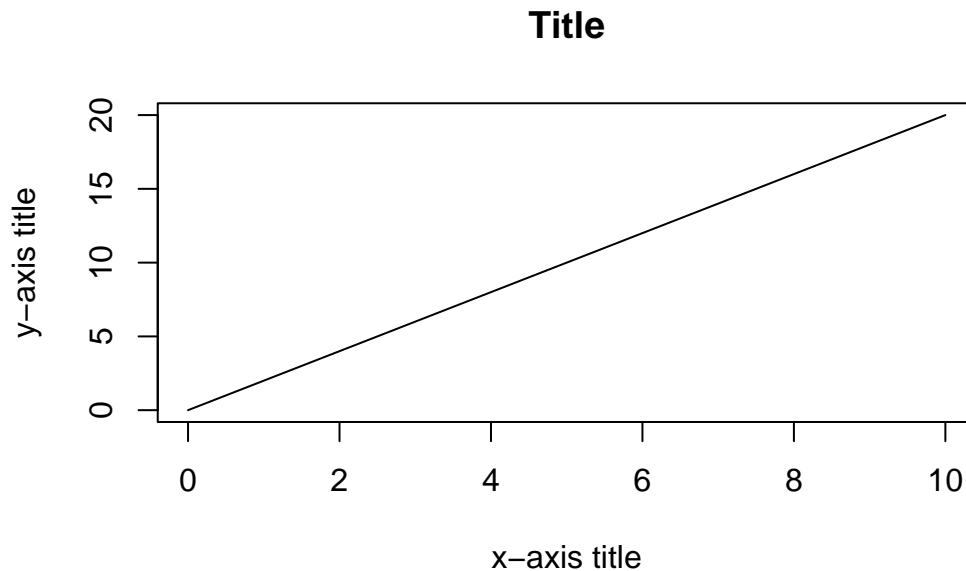
- Run a specific line of code by clicking anywhere on that line and hitting Cmd + Enter
- Run all code in a specific code chunk by clicking on green triangle
- Run all code in previous code chunks (sequentially) but NOT including this code chunk by clicking on gray triangle
- Render/knit document

```
# create a vector using c() and store as variable  
x <- c(1, 3, 5)  
  
# R is vectorized language  
2 * x
```

```
[1] 2 6 10
```

## Basic plotting

```
xx <- seq(0, 10, 0.5)
plot(xx, 2*xx,
      type = "l", xlab = "x-axis title", ylab = "y-axis title", main = "Title")
```



## Functions related to named distributions

Suppose  $X \sim \text{Binom}(8, 0.75)$ .

```
# P(X = 6)
dbinom(6, size = 8, prob = 0.75)
```

```
[1] 0.3114624
```

```
# P(X <= 6)
pbinom(6, size = 8, prob = 0.75)
```

```
[1] 0.6329193
```

```
# What is the median of this distribution? i.e. what is the 50th percentile?
qbinom(0.5, size = 8, prob = 0.75)
```

```
[1] 6
```

```
# Let's generate 10 realizations from this binomial reproducibly
set.seed(311)
rbinom(10, size = 8, prob = 0.75)
```

```
[1] 5 5 5 4 7 5 8 4 6 5
```

```
# let's approximate (empirically) the mean of the this binomial using Law of Large Numbers
samps <- rbinom(1000, size = 8, prob = 0.75)
mean(samps)
```

```
[1] 6.041
```

R has these functions for all named distributions (e.g. Poisson `dpois()`, Hypergeometric `dhyper()`, Normal `dnorm()`, etc.) But be careful with parameterizations!

### Optional: ggplot

```
library(tidyverse) # may need to install
data.frame(x = xx, y = 2*xx) |>
  ggplot(aes(x = x, y = y)) +
  geom_point()
```

