

# ENTMLGY 6707 Entomological Techniques and Data Analysis

## Intro to R

Not submitted for grading

1. Create a vector called `my_vector` with integer values from 1 to 10.
2. Use `log()` to log-transform `my_vector` and store (using `<-`) the transformed values as `my_vector_ln`. You should complete this step entirely in R - I am not asking you to store or save a new file. Note that in R, `log()` is the natural logarithm ( $\log_e$ ) and not  $\log_{10}$ .
3. Create a vector called `my_vector_new` by adding 2 to all values of `my_vector_ln`.
4. Create a data frame called `my_df` using the following code.

```
my_df <- data.frame(variable1 = my_vector, variable2 = my_vector_new)
```

5. Run the `summary()` command on your data frame. Your output should look exactly the same as the below summary.

```
##      variable1      variable2
## Min.   : 1.00   Min.   :2.000
## 1st Qu.: 3.25   1st Qu.:3.171
## Median : 5.50   Median :3.701
## Mean   : 5.50   Mean    :3.510
## 3rd Qu.: 7.75   3rd Qu.:4.046
## Max.   :10.00   Max.    :4.303
```

6. Use `plot()` to create a scatterplot of `variable2` as a function of `variable1` from the `my_df` data frame. Hint: `plot(...~..., data=...)` or `?plot()`

I will use this phrasing - “Y as a function of X” - a lot this semester. So, in this case, `variable2` would be your response variable and displayed on the y-axis and `variable1` would be your predictor and displayed on the x-axis.

7. Reproduce the graph from the previous step, but change the x-axis and y-axis titles to, respectively, “My X axis” and “My Y axis”.
8. When working with data frames in R, we can refer to specific columns using a `$`. Below is an example of how to calculate a mean in R. Calculate the standard deviation for `variable2` in `my_df`.

```
my_df$variable1
```

```
## [1] 1 2 3 4 5 6 7 8 9 10
```

```
mean(my_df$variable1)
```

```
## [1] 5.5
```