Prepare for the Analysis

In this section of pre-lab, you will be shown the R script needed to run the analysis. You will then be asked a series of questions to help you understand what that code is doing. Note that you will **always** be given the R code that you need to run an analysis. When you get to Lab, you will be able to **copy and modify** this code for your new lab question. Follow the directions below.

Primary Research Question

How many of the cyclists were students, how often did they ride, and what was the average distance they rode?

Let's break this analysis into its required steps:

1. Find the number of students in the dataset.
2. Pull out the student data into a separate dataframe for analysis.
3. Make a table to find how often the students ride.
4. Find the average distance ridden.

Here is the R script you will use: *(Remember that # indicates a comment)*

```
# Find the number of students in the dataset
table(bike$student)

# Pull out student data into a new dataframe
student <-bike[bike$student==1,]
```

```
# Find how often the students ride, using the new dataframe
table(student$cyc_freq)

# Create a vector for the distance variable
distance <-student$distance

# Find average distance ridden
mean(distance)
```

(2/2 points)
Look at the first line of code:
table(bike$student)

1a. What is the name of the dataframe?

bike ▾ ✔

1b. What is the name of the variable?

student ▾ ✔

You have used 1 of 1 submissions

(1/1 point)
This line creates a new dataframe called "student":
student <- bike[bike$student==1,]

2. What is the role of the **comma**?

- ◉ It tells R to include all the variables (columns) for the riders that are students. > ✔
- ○ It tells R to include all the students (rows), beginning with the first one.

You have used 1 of 1 submissions

(1/1 point)
3. What does "**student**" refer to in this line of code?:
table(student$cyc_freq)

- ○ the original variable called "student"

- ◉ the new dataframe called "student" ✔

You have used 1 of 1 submissions

(1 point possible)
This line of code creates a vector called "distance:"
distance <-student$distance

4. What does this vector consist of?

- ○ It is a string of numbers, each of which is the distance ridden by a student in the dataset.
- ◉ It is a matrix of numbers, showing all the distances ridden by each student. ✖

You have used 1 of 1 submissions

(1/1 point)
Suppose we have run the following code to try to create a list of the times.

```
#Read in the dataset and name it bike
bike<-BikeData
#Create a vector of the times
rider_times<-bike$times
#Check the contents of our rider_times vector
rider_times
```

NULL

5. What error has caused the vector of rider_times to be empty? (Examine the data set in R for help.)

- ○ We should have named the vector "time" instead.
- ◉ The variable "times" is not spelled the same way in our datafile. ✔
- ○ We cannot check the contents of vectors in R.

- We cannot use an underscore when naming vectors.

You have used 1 of 1 submissions