Name_____ Date_____

LAB 1A: Data, Code & RStudio
Response Sheet

Directions: Record your responses to the lab questions in the spaces provided.

Welcome to the labs!

So let's get started!

- **Describe the data that appeared after running `View(cdc)`:**

- *Who* **is the information about?**

- **What sorts of information about them was collected?**

Data: Variables & Observations

- **Based on the data, describe a few characteristics about the first observation.**

- **What does the first column tell us about our observations?**

Uncovering our Data's Structure

- **How many students are in our `cdc` data set?**

- **How many variables were measured for each student?**

Name_____ Date_____

LAB 1A: Data, Code & RStudio
Response Sheet

Type the following commands into the console

- **Which of these functions tell us the number of observations in our data?**

- **Which of these functions tell us the number of variables?**

First Steps

Syntax matters

- **Run the following commands and write down what happens after each. Which does R understand?**

R's most important syntax

Syntax in action

- **Which one of these plots would be useful for answering the question:** *Is it unusual for students in the CDC dataset to be taller than 1.8 meters?*

- **Do you think it's unusual for students in the data to be taller than 1.8 meters? Why or why not?**

On your own:

- **What is** *public health* **and do we collect data about it?**

Name_____ Date_____

LAB 1A: Data, Code & RStudio
Response Sheet

- **How do you think our data was collected? Does it include every high school aged student in the US?**

- **How might the CDC use this data? Who else could benefit from using this data?**

- **Write the code to visualize the distribution of weights of the students in the CDC data with a histogram. What is the *typical* weight?**

- **Write the code to create a bargraph to visualize the distribution of how often students ate fruit. About how many students did not eat fruit over the previous 7 days?**