Name:_____ Date:_____

LAB 3D: Are you sure about that?
Response Sheet

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

Confidence and intervals

In this lab

(1) The United States has an estimated population of 336,302,171 (as of April 15, 2024 9:10 a.m. PDT). How many people were surveyed for this particular dataset?

(2) Why is it important that the ATUS is a random sample?

(3) Use our `atus` data to calculate an estimate for the average age of people older than 15 living in the U.S.

One bootstrap

Our first bootstrap

(4) Fill in the blanks to `sample` the row numbers we'll use in our *bootstrapped* sample.

```
bs_rows <- ___________(1:___________, size = ___________, replace = ___________)
```

(5) Write and run code using the `slice` function to create a new dataset that includes each row from our `sample`.

```
bs_atus <- slice(atus, bs_rows)
```

Take a look

(6) Write a paragraph that explains to someone who's not familiar with `R` how you created `bs_rows` and `bs_atus`. Be sure to include an explanation of what the *values* of `bs_rows` mean and how those values are used to create `bs_atus`. Also, be sure to explain what each argument of each function does.

Name:_____ Date:_____

LAB 3D: Are you sure about that?
Response Sheet

One strap, two strap

(7) Write and run code calculating the `mean` of the `age` variable in your *bootstrapped* data, then use a different value of `set.seed()` to create your own, personal *bootstrapped* sample. Then calculate its `mean`.

(8) Compare this second *bootstrapped* sample with three other classmates and write a sentence about how similar or different the *bootstrapped* sample means were.

Many bootstraps

Bootstrap function

(9) Fill in the blank space below with the 3-steps needed to create a *bootstrapped* sample `mean` for our `atus` data.

```
bs_func <- function() {




}
```

Visualizing our bootstraps

(10) Once your function is created, fill in the blanks to create 500 *bootstrapped* sample means:

```
bs_means <- do(_____) * bs_func()
```

(11) Create a `histogram` for your bootstrapped samples and describe the *center*, *shape* and *spread* of its distribution.

Name:_____ Date:_____

LAB 3D: Are you sure about that?
Response Sheet

Bootstrapped confidence intervals

(12) Using your `histogram`, fill in the statement below:

The lowest 5% of our estimates are below _____ years and the highest 5% of our estimates are above _____ years.

(13) Write and run code using the `quantile()` function to check your estimates.

(14) Based on your *bootstrapped* estimates, between which two ages are we 90% confident the actual `mean` age of people living in the U.S. is contained?

On your own

(15) Using your *bootstrapped* sample means, write and run code creating a 95% confidence interval for the `mean` age of people living in the U.S.

(16) Why is the 95% confidence interval wider than the 90% interval?

(17) Write down how you would explain what a 95% confidence interval means to someone not taking *Introduction to Data Science*.