

Name: _____

Date: _____

LAB 2F: The Titanic Shuffle Response Sheet

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

Previously ...

The Titanic

- (1) Write and run code using the `data` function to load the `titanic` passenger and survival data.

- (2) Write and run code creating a boxplot of the fares paid by passengers and facet the plot based on whether the passenger survived or not.

- (3) Based on the plot, do you believe that passengers who paid a higher fare on the Titanic were more likely to survive? Explain why and describe how certain you are of being correct.

The search begins!

- (4) Write and run code to visualize the distribution of fares paid.

- (5) Which numerical summary might be preferred to describe the *typical* value?

- (6) What was the *typical* fare paid by survivors? Non-survivors? How much more did the typical survivor pay?

Do the shuffle!

- (7) Write and run code using the `do` and the `shuffle` functions to shuffle the passenger's survival status 500 times.

Name: _____

Date: _____

LAB 2F: The Titanic Shuffle Response Sheet

(8) After shuffling your data, write and run code using the `mutate` function to create a variable called `diff` which is the median fare of survivors minus the median fare of non-survivors.

Put your simulations to use

(9) Using your shuffled data, answer the research question we posed at the beginning of the lab.
Is there evidence to suggest that those who survived paid a higher fare than those who died?

(10) Write up your answer as a statistical analysis. Create a plot and explain how the plot supports your conclusion. Be sure to also explain why shuffling your data is important.

Comparing Mean Fares

(11) If we did this 500 times, what do you predict the distribution of differences will look like?

(12) Write and run code using the `do` and the `shuffle` functions to shuffle the passenger survival status 500 times.

(13) What does the shuffled data reveal? Does the answer to the research question below change when using the mean fares instead of the median fares?

Is there evidence to suggest that those who survived paid a higher fare than those who died?