Name:_____ Date:_____

LAB 4B: What's the score?
Response Sheet

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

Previously

Predictions using a line

(1) Write and run code creating an `xyplot` with `height` on the y-axis and `armspan` on the x-axis.

(2) Fill in the blanks below to create a function that will make predictions of people's `heights` based on their `armspan`:

```
predict_height <- function(armspan) {

  _____ * armspan + _____

}
```

Make your predictions

(3) Fill in the blanks to include your predictions in the `arm_span` data.

```
_____ <- mutate(_____, predicted_height = _____(_____))
```

Sums of differences

(4) Fill in the blanks to add a column of residuals to `arm_span`:

```
_____ <- mutate(_____, residual = _____ - _____)
```

(5) What do the residuals measure?

(6) Fill in the blanks below to calculate our accuracy summary.

```
summarize(_____, sum(_____))
```

(7) Describe and interpret, in words, what the output of your accuracy summary means.

Name:_____ Date:_____

LAB 4B: What's the score?
Response Sheet

(8) Write down why adding positive and negative errors together is problematic for assessing prediction accuracy.

Mean squared error

(9) Fill in the blanks below to calculate the MSE of your line.

```
summarize(_____, mean((_____)^2)
```

(10) Compare your MSE with a neighbor. Whose line was more accurate and why?

Regression lines

(11) Fill in the blanks below to create a *regression line* using `lm`, which stands for *linear model*:

```
best_fit <- lm(_____ ~ _____, data = arm_span)
```

Plotting regression lines

(12) Run the code to create the scatterplot of `armspan` vs. `height` again. Then fill in the blanks below to add the line of best fit.

```
add_line(intercept = _____, slope = _____)
```

Predicting with regression lines

(13) Fill in the blanks to make predictions using `best_fit`:

```
_____ <- mutate(_____, predicted_height = predict(_____))
```

The magic of lm()

(14) Calculate the MSE for the values predicted using the regression line.

(15) Compare the MSE of the linear model you fitted to the MSE of the linear model obtained with `lm()`. Which linear model performed better?

(16) Ask your neighbors if any of their lines beat the `lm()` line in terms of the MSE. Were any of them successful?