Lesson 19: ACFT Problem

Is there an association between the Standing Power Throw and Maximum Deadlift raw scores?

We are going to determine if an association exists between `spr_raw` and `mdl_raw` scores.

The relationship we are modeling looks like this:

$$\widehat{mdl} = \beta_0 + \beta_1 \cdot spr$$

1. Provide the hypotheses in symbols and words:

2. Plot the association, including the regression line.

```
# df %>%
#   ggplot(XXXXX
```

3. Is there anything unusual about the data?

4. Remove the outlier using the `filter(spr_raw < 100)` command. How many observations were dropped?

```
# df <- XXXXX
```

5. Create a new scatter plot with updated dataframe.

```
# df %>%
#   ggplot(XXXXX
```

6. Create a linear model to test your hypotheses.

```
# lrmodel <- df %>%
#   XXXXX
# summary(lrmodel)
```

7. Write out the linear model with the coefficients determined by the linear regression.

8. Interpret the intercept term.

9. Interpret the slope term.

10. Conclusion of your hypothesis test at a 5% significance level.

11. Interpret the R-squared value.

12. Check the 4 Validity Conditions (L.I.N.E.). Are any not met?

```
# lrmodel %>%
#   fortify(lrmodel$model) %>%
#   XXXXX
#   XXXXX
#   labs( x = "Predicted Values",
#         y = "Residuals",
#         title = "Residuals vs. Predicted Values")
#
# lrmodel %>%
#   fortify(lrmodel$model) %>%
#   XXXXX
#   XXXXX
#   labs(x = "Order of Occurence",
#        y = "Residuals",
#        title = "Residuals in Order of Occurence")
#
# lrmodel %>%
#   fortify(lrmodel$model) %>%
#   XXXXX
#   XXXXX
#   labs(x = "Residuals",
#        title = "Histogram of Residuals")
```

13. Can we determine causation? Why or why not.

Does adding the variable `sex` change the association between the Standing Power Throw (`spr_raw`) and Leg Tuck (`ltk_raw`) raw scores?

Now, we add `sex` as another variable to the model but also include it as an interaction with `spr_raw`.

`Our updated model looks like this:`

$\widehat{mdl} = \beta_0 + \beta_1 \cdot spr + \beta_2 \cdot sexmale + \beta_3(spr \cdot sexmale)$

14. Plot the data **(including the interaction term)**, including the regression line.

```
# df %>%
#   ggplot(XXXXX)
```

15. Does there appear to be a different relationship between `mdl_raw` and `spr_raw` based on `sex`?

16. Conduct a linear regression with the updated model.

```
# lrmodel <- df %>%
#   lm(XXXXX)
# summary(lrmodel)
```

17. Interpret the coefficient on the interaction term.

18. Is this significant at the 5% level?

19. Interpret the intercept term.

20. Interpret the base slope term for `spr_raw`.

21. Using this model (even with no significance), calculate the following:

- The `mdl_raw` for a male cadet with an `spr_raw` of 6m:

```
#
```

- The `mdl_raw` for a female cadet with an `spr_raw` of 6m:

```
#
```

22. Interpret the R-squared value.

23. Check the 4 Validity Conditions (L.I.N.E.). Are any not met?

```
# lrmodel %>%
#   fortify(lrmodel$model) %>%
#   XXXXXXXX
#   XXXXXXXX
#   labs( x = "Predicted Values",
#         y = "Residuals",
#         title = "Residuals vs. Predicted Values")
#
# lrmodel %>%
#   fortify(lrmodel$model) %>%
#   XXXXXXX
#   XXXXXXX
#   labs(x = "Order of Occurence",
#        y = "Residuals",
#        title = "Residuals in Order of Occurence")
#
# lrmodel %>%
#   fortify(lrmodel$model) %>%
#   XXXXXXXXXX
#   XXXXXXXXXX
#   labs(x = "Residuals",
#        title = "Histogram of Residuals")
```