Exercises

1. Create a scatterplot of faceoff\_win\_pct against win\_loss\_pctg with a least squares regression line. What is the trend of this relationship?

A graph with a line and dots

Description automatically generated

Moderate positive correlation. (r = 0.63)

1. Fit a linear model with faceoff\_win\_pct predicting win\_loss\_pctg. Is faceoff\_win\_pct an effective predictor? Write down the equation of the model. What is the R-squared value?

A white paper with black text

Description automatically generated

1. Create another scatterplot of faceoff\_win\_pct against win\_loss\_pctg but with separate least squares regression lines for the Mens and Womens Divisions. How do the two lines compare to each other?

A graph with red and blue dots

Description automatically generated

The slope for Women is greater than the slope for men. The intercept is higher for Men than women. The lines meet at a faceoff win percent of roughly 50%.

1. Write the appropriate code to make Division an indicator variable with 1 representing “Mens” and 0 representing “Womens”.

lax <- lax%>%

mutate(genderIND = ifelse(Division == "Mens", 1, 0))

1. Fit a model that would allow us to have different intercepts for the Mens and Womens divisions. Is there evidence that the two divisions have significantly different intercepts?

A white sheet with black numbers and symbols

Description automatically generated

genderIND p-value = 1, so the intercept is not different for men and women when they share the slope. This makes sense because faceoff percentages are not higher for women or men. R-squared value did not change because this adds nothing to our model. The two divisions do not have significantly different intercepts.

1. Now, fit a model that would allow us to have different intercepts **and** slopes for the Mens and Womens divisions.

A screenshot of a computer code

Description automatically generated

1. Write down the equations of the two lines and interpret the estimated slopes and intercepts respectively.

For Women’s Lacrosse, on average, a one percent increase in faceoff win percentage is associated with a 2.45 increase in win loss percentage.

For Women’s Lacrosse, a faceoff win percentage of zero is associated with a win loss percentage of -0.729.

For Men’s Lacrosse, on average, a one percent increase in faceoff win percentage is associated with a 1.50 increase in win losspercentage.

For Men’s Lacrosse, a faceoff win percentage of zero is associated with a win loss percentage of 0.257.

1. Is there evidence that the lines for the two groups have significantly different intercepts?

There is moderate evidence that Men’s and Women’s Lacrosse have different intercept in the linear model predicting win loss percentage from faceoff win percentage (p-value = 0.009)

1. Is there evidence that the lines for the two groups have significantly different slopes?

There is moderate evidence that Men’s and Women’s Lacrosse have significantly different slopes in the linear model predicting win loss percentage from faceoff win percentage (p-value = 0.008).

1. Is the model using two completely separate lines (i.e., different intercepts and slopes) better than the single model ignoring Division entirely?

Yes, there is evidence that the model using two completely separate lines is better than the model ignoring Division based on the Adjusted R-squared value. A nested F-test provides marginal evidence in support of the model using two separate lines.

1. Why do you think the lines meet at a faceoff win percentage of roughly 50%?

Because in their respective leagues, faceoff percentages will always average out to roughly 50% because they are competing against each other in a binary faceoff competition.

1. Filter the data to make separate sets for the Mens and Womens divisions. Now, calculate the correlation between shots\_per\_game and faceoff\_win\_pct for the men and women respectively. Which is higher? What conclusions can you draw?

Mens: r = 0.39

Womens: r = 0.66

For women, there is a higher correlation between winning a faceoff and shooting the ball. This implies that winning the faceoff has a greater and more direct impact scoring opportunities for women than it does for men.

1. Using the filtered data sets again, calculate the mean shot\_pctg (total goals divided by total shots) of Men and Women respectively. Which is higher? How can this add to our concusions?

Mens: mean = 0.30

Womens: mean = 0.42

For women, there is a higher proportion of shots made by women than there is for men. Seeing that women have a much higher correlation between winning faceoff wins and shots, it seems that women are not only being afforded more shot opportunities by winning faceoffs, but also, they are capitalizing on these opportunities by scoring goals more than the men.