Week 1 & 2 Lab

Quiz, 10 questions

1 / 1 point

1.
What type of plot would you use to display the relationship between runs and one of the other numerical variables?

- histogram
- box plot
- scatterplot

Correct

- bar plot

---

1 / 1 point

2.
Plot the relationship between runs and at_bats, using at_bats as the explanatory variable. The relationship appears to be _____.

- negative
- linear

Correct

- u-shaped (∪)
- horseshoe-shaped (∩)

Week 1 & 2 Lab

Quiz, 10 questions

1 / 1 point

3.

Looking at your plot from the previous exercise, which of the following best describe the relationship between these two variables?

- The relationship is negative, linear, and moderately strong. One of the potential outliers is a team with approximately 5520 at bats.
- The relationship is positive, linear, and moderately strong. One of the potential outliers is a team with approximately 5520 at bats.

Correct

- The relationship is positive, linear, and very weak. There are no outliers.
- The relationship is positive, linear, and very weak. One of the potential outliers is a team with approximately 5520 at bats.

---

1 / 1 point

4.

Fit a new model that uses homeruns to predict runs. Using the estimates from the R output, write the equation of the regression line. What does the slope tell us in the context of the relationship between success of a team and its home runs?

- For each additional home run, the model predicts 1.83 more runs, on average.

Correct

- Each additional home run increases runs by 1.83.
- For each additional home run, the model predicts 1.83 fewer runs, on average.
- For each additional home run, the model predicts 415.24 more runs, on average.
- For each additional home run, the model predicts 415.24 fewer runs, on average.

---

1 / 1 point

Week 1 & 2 Lab
Quiz, 10 questions

5.
What is the residual for the prediction of runs for a team with 5,579 at-bats? Choose the closest answer.

- [x] -15.32

Correct

- [ ] 15.32
- [ ] 713
- [ ] 5,579

1 / 1 point

6.
Which of the following statements about the residual plot is **false**?

- [ ] The residuals appear to be randomly distributed around 0.
- [x] The residuals show a curved pattern.

Correct

- [ ] The plot is indicative of a linear relationship between runs and at-bats.
- [ ] The team with a very high residual compared to the others appears to be an outlier.

1 / 1 point

7.
Which of the following is true?

- [x] The residuals are fairly symmetric, with only a slightly longer tail on the right, hence it would be appropriate to deem the the normal distribution of residuals condition met.

Correct

- [ ]

Week 1 & 2 Lab
Quiz, 10 questions

The residuals are extremely left skewed, hence the normal distribution of residuals condition is not met.

The residuals are perfectly symmetric, hence the normal distribution of residuals condition is met.

The residuals are extremely right skewed, hence the normal distribution of residuals condition is not met.

---

1 / 1 point

8.

[TRUE / FALSE] Based on the residual plot from earlier, the constant variability condition appears to be met.

True

Correct

False

---

1 / 1 point

9.

Now that you can summarize the linear relationship between two variables, investigate the relationships between runs and each of the other five traditional variables. Which variable best predicts runs? Support your conclusion using the graphical and numerical methods we've discussed.

at bats

hits

wins

batting average

Correct

---

1 / 1 point

Week 1 & 2 Lab

Quiz, 10 questions

10.

Now examine the three newer variables. These are the statistics used by the author of Moneyball to predict a teams success. In general, are they more or less effective at predicting runs that the old variables? Explain using appropriate graphical and numerical evidence. Of all ten variables we've analyzed, which seems to be the best predictor of runs?

- [x] on-base plus slugging (new obs)

Correct

- [ ] slugging percentage (new slug)
- [ ] on-base percentage (new onbase)