

[Course](#) > [Section 2: Linear M...](#) > [2.4: Regression and...](#) > [Assessment: Regre...](#)

Assessment: Regression and baseball, part 2

Question 9 has two parts. Use the information below to answer both parts.

Use the **Teams** data frame from the **Lahman** package. Fit a multivariate linear regression model to obtain the effects of BB and HR on Runs (**R**) in 1971. Use the **tidy** function in the **broom** package to obtain the results in a data frame.

Question 9a

2/2 points (graded)

What is the estimate for the effect of BB on runs?

✓ Answer: 0.414

Explanation

The estimate can be calculated using the following code:

```
library(Lahman)
library(broom)
Teams %>%
  filter(yearID == 1971) %>%
  lm(R ~ BB + HR, data = .) %>%
  tidy() %>%
  filter(term == "BB") %>%
  pull(estimate)
```

What is the estimate for the effect of HR on runs?

✓ Answer: 1.30

Explanation

The estimate can be calculated using the following code:

```
Teams %>%
  filter(yearID == 1971) %>%
  lm(R ~ BB + HR, data = .) %>%
  tidy() %>%
  filter(term == "HR") %>%
  pull(estimate)
```

Submit

You have used 1 of 10 attempts

i Answers are displayed within the problem

Question 9b

1/1 point (graded)

Interpret the p-values for the estimates using a cutoff of 0.05.

Which of the following is the correct interpretation?

- ☐ Both BB and HR have a nonzero effect on runs.
- ☒ HR has a significant effect on runs, but the evidence is not strong enough to suggest BB also does.
- ☐ BB has a significant effect on runs, but the evidence is not strong enough to suggest HR also does.
- ☐ Neither BB nor HR have a statistically significant effect on runs.



Explanation

The p-value for HR is less than 0.05, but the p-value of BB is greater than 0.05 (0.06), so the evidence is not strong enough to suggest that BB has a significant effect on runs at a p-value cutoff of 0.05.

Submit

You have used 1 of 2 attempts

i Answers are displayed within the problem

Question 10

1/1 point (graded)

Repeat the above exercise to find the effects of BB and HR on runs (`R`) for every year from 1961 to 2018 using `do` and the `broom` package.

Make a scatterplot of the estimate for the effect of BB on runs over time and add a trend line with confidence intervals.

Fill in the blank to complete the statement:

The effect of BB on runs has  **✓ Answer:** increased over time.

Explanation

The scatterplot with trendline can be made using the following code:

```
res <- Teams %>%
  filter(yearID %in% 1961:2018) %>%
  group_by(yearID) %>%
  do(tidy(lm(R ~ BB + HR, data = .))) %>%
  ungroup()
res %>%
  filter(term == "BB") %>%
  ggplot(aes(yearID, estimate)) +
  geom_point() +
  geom_smooth(method = "lm")
```

Submit

You have used 1 of 1 attempt

 Answers are displayed within the problem

Question 11

2/2 points (graded)

Fit a linear model on the results from Question 10 to determine the effect of year on the impact of BB.

For each additional year, by what value does the impact of BB on runs change?

✓ Answer: 0.00355

Explanation

The value can be calculated using the following code:

```
res %>%
  filter(term == "BB") %>%
  lm(estimate ~ yearID, data = .) %>%
  tidy() %>%
  filter(term == "yearID") %>%
  pull(estimate)
```

What is the p-value for this effect?

0.00807

✓ Answer: 0.00807

0.00807

Explanation

The p-value can be calculated using the following code:

```
res %>%  
  filter(term == "BB") %>%  
  lm(estimate ~ yearID, data = .) %>%  
  tidy() %>%  
  filter(term == "yearID") %>%  
  pull(p.value)
```

Submit

You have used 1 of 10 attempts

i Answers are displayed within the problem