Assessment: Baseball as a Motivating Example

Question 1

1/1 point (graded)

What is the application of statistics and data science to baseball called?

- ○ Moneyball
- ◉ Sabermetrics ✔
- ○ The “Oakland A’s Approach”
- ○ There is no specific name for this; it’s just data science.

Answer

Correct:

Correct. The term “sabermetrics” was coined by Bill James, and is derived from the acronym SABR: the society for American baseball research.

Submit You have used 1 of 2 attempts

Question 2

1/1 point (graded)

Which of the following outcomes is not included in the batting average?

- ○ A home run
- ◉ A base on balls ✔
- ○ An out

A single

Answer
Correct: Correct. A base on balls is not considered a hit and is excluded from the at-bat total.

Submit You have used 1 of 2 attempts

Question 3

1/1 point (graded)
Why do we consider team statistics as well as individual player statistics?

- The success of any individual player also depends on the strength of their team. ✔
- Team statistics can be easier to calculate.
- The ultimate goal of sabermetrics is to rank teams, not players.

Submit You have used 1 of 1 attempt

Question 4

1.0/1.0 point (graded)
You want to know whether teams with more at-bats per game have more runs per game.

What R code below correctly makes a scatter plot for this relationship?

```
Teams %>% filter(yearID %in% 1961:2001 ) %>%
    ggplot(aes(AB, R)) +
    geom_point(alpha = 0.5)
```

```
Teams %>% filter(yearID %in% 1961:2001 ) %>%
    mutate(AB_per_game = AB/G, R_per_game = R/G) %>%
    ggplot(aes(AB_per_game, R_per_game)) +
    geom_point(alpha = 0.5)
```

✔

```
Teams %>% filter(yearID %in% 1961:2001 ) %>%
    mutate(AB_per_game = AB/G, R_per_game = R/G) %>%
    ggplot(aes(AB_per_game, R_per_game)) +
    geom_line()
```

```
Teams %>% filter(yearID %in% 1961:2001 ) %>%
    mutate(AB_per_game = AB/G, R_per_game = R/G) %>%
    ggplot(aes(R_per_game, AB_per_game)) +
    geom_point()
```

Answer
Correct: Correct. This makes a scatter plot of runs per game (y-axis) vs. at-bats per game (x-axis).

Submit You have used 1 of 2 attempts

Question 5

1.0/1.0 point (graded)
What does the variable “SOA” stand for in the Teams table?
Hint: make sure to use the help file (`?Teams`).

- sacrifice out
- slides or attempts
- strikeouts by pitchers ✔
- accumulated singles

Answer
Correct: Correct.

Submit You have used 1 of 2 attempts

Question 6

1/1 point (graded)
Load the **Lahman** library. Filter the `Teams` data frame to include years from 1961 to 2001. Make a scatterplot of runs per game versus at bats (`AB`) per game.

Which of the following is true?

- There is no clear relationship between runs and at bats per game.
- As the number of at bats per game increases, the number of runs per game tends to increase. ✔
- As the number of at bats per game increases, the number of runs per game tends to decrease.

Submit You have used 1 of 1 attempt

Question 7

1/1 point (graded)
Use the filtered `Teams` data frame from Question 6. Make a scatterplot of win rate (number of wins per game) versus number of fielding errors (`E`) per game.

Which of the following is true?

- There is no clear relationship between win rate and errors per game.
- As the number of errors per game increases, the win rate tends to increase.
- As the number of errors per game increases, the win rate tends to decrease. ✔

Submit You have used 1 of 1 attempt

Question 8

1/1 point (graded)

Use the filtered `Teams` data frame from Question 6. Make a scatterplot of triples (`X3B`) per game versus doubles (`X2B`) per game.

Which of the following is true?

- (●) There is no clear relationship between doubles per game and triples per game. ✔
- ( ) As the number of doubles per game increases, the number of triples per game tends to increase.
- ( ) As the number of doubles per game increases, the number of triples per game tends to decrease.

Submit You have used 1 of 1 attempt

© All Rights Reserved