Lake Placid Triathlon Correlations

The dataset **ironman1819.csv** contains data on female finishers of the Lake Placid Ironman Triathlon from 2002 to 2021. The motivation for this data analysis is to explore the relationships between swim times, bike times, and run times (in minutes) in order to gain insights into the performance patterns of the athletes. By analyzing these relationships, we can understand the interplay between different segments of the race and potentially identify areas of improvement for athletes. For this activity, we will specifically focus on times from finishers in the years 2018 and 2019.

1. Watch the Introductory Video
    a. http://myslu.stlawu.edu/~msch/SCORE/SarahTrueTriathlonModule.mp4

2. Graph and describe the distribution for each leg of the Triathlon
    a. Swim Times
    b. Bike Times
    c. Run Times

3. Graph the following relationships between the legs of the race.
    a. Swim Times vs. Bike Times
    b. Run Times vs. Bike Times
    c. Swim Times vs. Run Times

4. From Question 3, is a correlation an appropriate calculation for these data? If so, guess the correlation between the two legs of the race.

Relationship	Correlation Appropriate?	Guess for correlation
Swim Times vs. Bike Times	Yes No	r = _____
Run Times vs. Bike Times	Yes No	r = _____
Swim Times vs. Run Times	Yes No	r = _____

5. Now, using technology, calculate the correlations and compare them to your guesses

Relationship	Correlation
Swim Times vs. Run Times	r = _____
Run Times vs. Bike Times	r = _____
Swim Times vs. Run Times	r = _____

6. Which relationship has the largest correlation? What is a possible reason why we may see this?

7. Which relationship has the smallest correlation? What is a possible reason why we may see this?

8. These data are for finishers of the triathlon. How might these relationships be different for athletes that did not finish or were disqualified?

9. What do these correlations suggest for athletes about the aspects of their race they want to concentrate on?