The dataset **ironman\_lake\_placid\_female.csv** contains data on female finishers of the Lake Placid Ironman Triathlon from 2002 to 2022. 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 focus on finishers from 2022.

1. Watch the Introductory Video
   1. <https://www.youtube.com/watch?v=zYG2qdJRolM>
2. Graph and describe the distribution for each leg of the Triathlon
   1. Swim Times
   2. Bike Times See Page 3
   3. Run Times
3. Graph the following relationships between the legs of the race.
   1. Swim Times vs. Bike Times
   2. Run Times vs. Bike Times See Page 4
   3. Swim Times vs. Run Times

|  |  |  |  |
| --- | --- | --- | --- |
| **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 = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |

1. 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**  0.587 |
| Swim Times vs. Bike Times | r = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  0.698 |
| Run Times vs. Bike Times | r = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  0.430 |
| Swim Times vs. Run Times | r = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |

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

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

r = 0.698. A possible explanation for this would be that running and biking require similar cardio and leg strength. They also have similar motions.

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

r = 0.430. A possible explanation for this would be that swimming and biking have very different motions. Swimming requires much more arm strength.

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

Answers may vary

Sample Response: The times for non-finishers way be very high and that may alter the linear relationships between the legs of the triathlon.

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

Answers may vary

Sample Response: Focus on swimming and one of either biking or running.

A picture containing text, screenshot, diagram, plot

Description automatically generated2a)

Unimodal and Symmetric

Center: around 80 minutes

Spread: Most values between 55 and 105 minutes

Possible outliers above 105 minutes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Min | Q1 | Median | Q3 | Max |
| 50.95 | 72.63 | 81.45 | 89.00 | 112.5 |

Mean: 81.116  
SD: 11.31

A picture containing text, screenshot, diagram, plot

Description automatically generated

Unimodal and Symmetric

Center: around 420 minutes

Spread: Most values between 320 and 520 minutes

Possible outliers below 320 minutes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Min | Q1 | Median | Q3 | Max |
| 292.35 | 392.67 | 425.05 | 459.6 | 531.42 |

Mean: 424.74  
SD: 48.29

2b)

A picture containing text, diagram, screenshot, plot

Description automatically generated

Unimodal and Symmetric

Center: around 320 minutes

Spread: Most values between 225 and 420 minutes

Possible outliers below 220 and above 420 minutes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Min | Q1 | Median | Q3 | Max |
| 184.12 | 287.79 | 324.12 | 366.84 | 459.45 |

Mean: 324.62  
SD: 53.52

2c)

A picture containing text, screenshot, diagram, plot

Description automatically generated3a)

A picture containing text, screenshot, diagram, plot

Description automatically generated

3b)

A picture containing text, screenshot, diagram, plot

Description automatically generated

3c)