

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. <https://www.youtube.com/watch?v=zYG2qdJRoIM>
2. Graph and describe the distribution for each leg of the Triathlon
 - a. Swim Times
 - b. Bike Times See Page 3
 - 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 See Page 4
 - 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	<input checked="" type="radio"/> Yes <input type="radio"/> No	$r =$ _____
Run Times vs. Bike Times	<input checked="" type="radio"/> Yes <input type="radio"/> No	$r =$ _____
Swim Times vs. Run Times	<input checked="" type="radio"/> Yes <input type="radio"/> No	$r =$ _____

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

Relationship	Correlation
Swim Times vs. Bike Times	$r =$ <u>0.5478</u>
Run Times vs. Bike Times	$r =$ <u>0.6556</u>
Swim Times vs. Run Times	$r =$ <u>0.3943</u>

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

$r = 0.6556$. A possible explanation for this would be that running and biking require similar cardio and leg strength. They also have similar motions.

Lake Placid Triathlon Correlations

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

$r = 0.3843$. A possible explanation for this would be that swimming and biking have very different motions. Swimming requires much more arm strength.

8. 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:

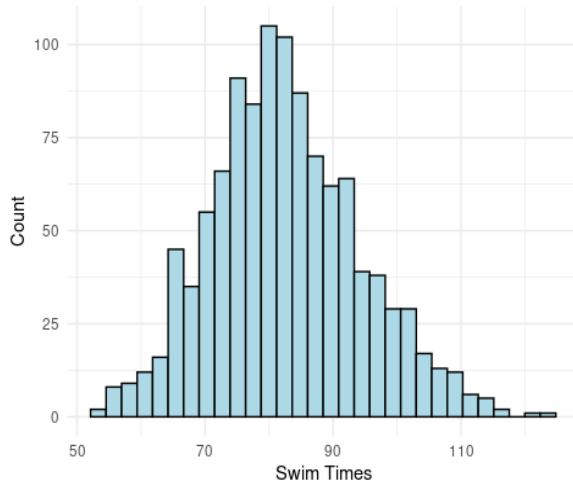
9. What do these correlations suggest for athletes about the aspects of their race they want to concentrate on?

Answers may vary

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

Lake Placid Triathlon Correlations

2a)



Unimodal and Symmetric

Center: around 80 minutes

Spread: Most values between 55 and 115 minutes

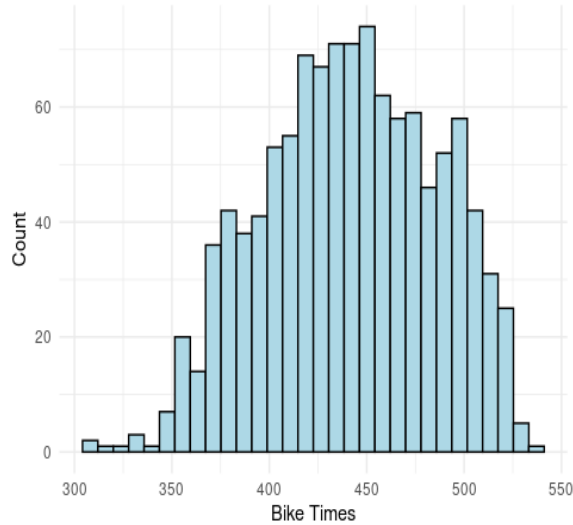
Possible outliers above 120 minutes

Min	Q1	Median	Q3	Max
53.36	74.68	81.7	90.02	123.65

Mean: 82.58

SD: 11.625

2b)



Unimodal and Symmetric

Center: around 440 minute

Spread: Most values between 350 and 525 minutes

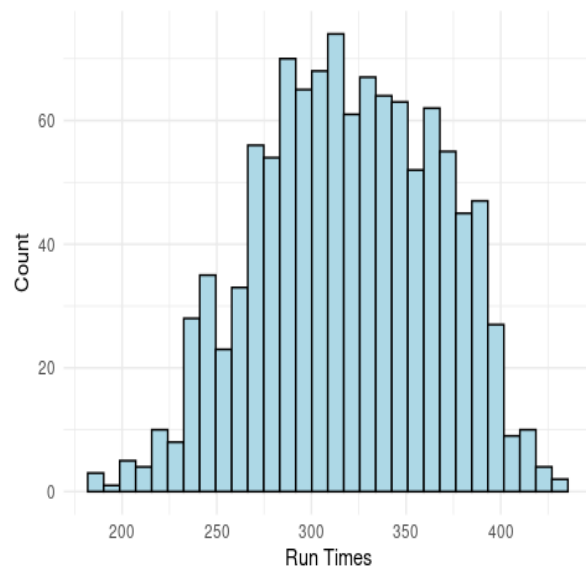
Possible outliers below 350 minutes

Min	Q1	Median	Q3	Max
304.73	410.12	411.83	475.27	533.82

Mean: 441.30

SD: 44.339

2c)



Unimodal and Symmetric

Center: around 300 minutes

Spread: Most values between 225 and 400 minutes

Possible outliers below 200 and above 400 minutes

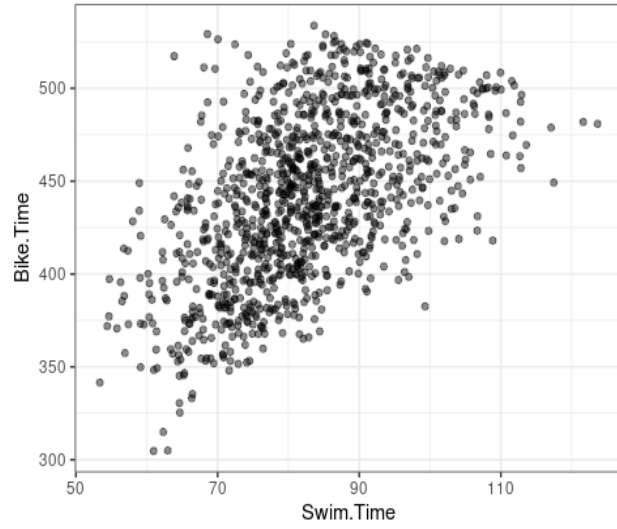
Min	Q1	Median	Q3	Max
187.52	285.38	318.8	356.15	432.75

Mean: 319.30

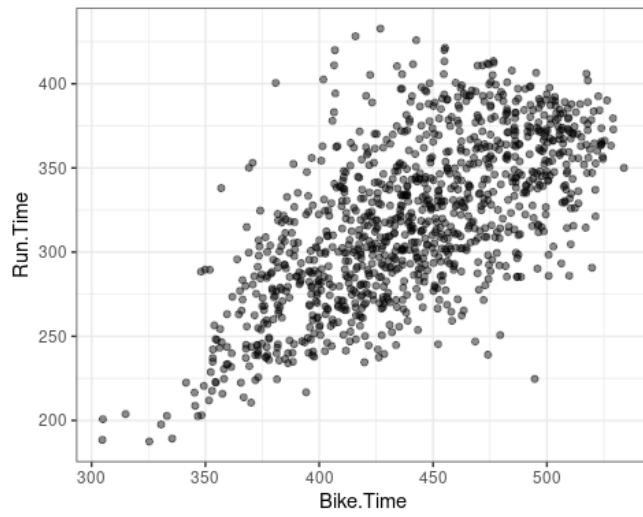
SD: 47.373

Lake Placid Triathlon Correlations

3a)



3b)



3c)

