

⚠ Lagunita is retiring and will shut down at 12 noon Pacific Time on March 31, 2020. A few courses may be open for self-enrollment for a limited time. We will continue to offer courses on other online learning platforms; visit <http://online.stanford.edu>.

Course > EDA: Examining Distributions > One Quantitative Variable: Measures of Spread - Boxplots >
Learn By Doing Activity

🔖 Bookmark this page

Learn By Doing Activity

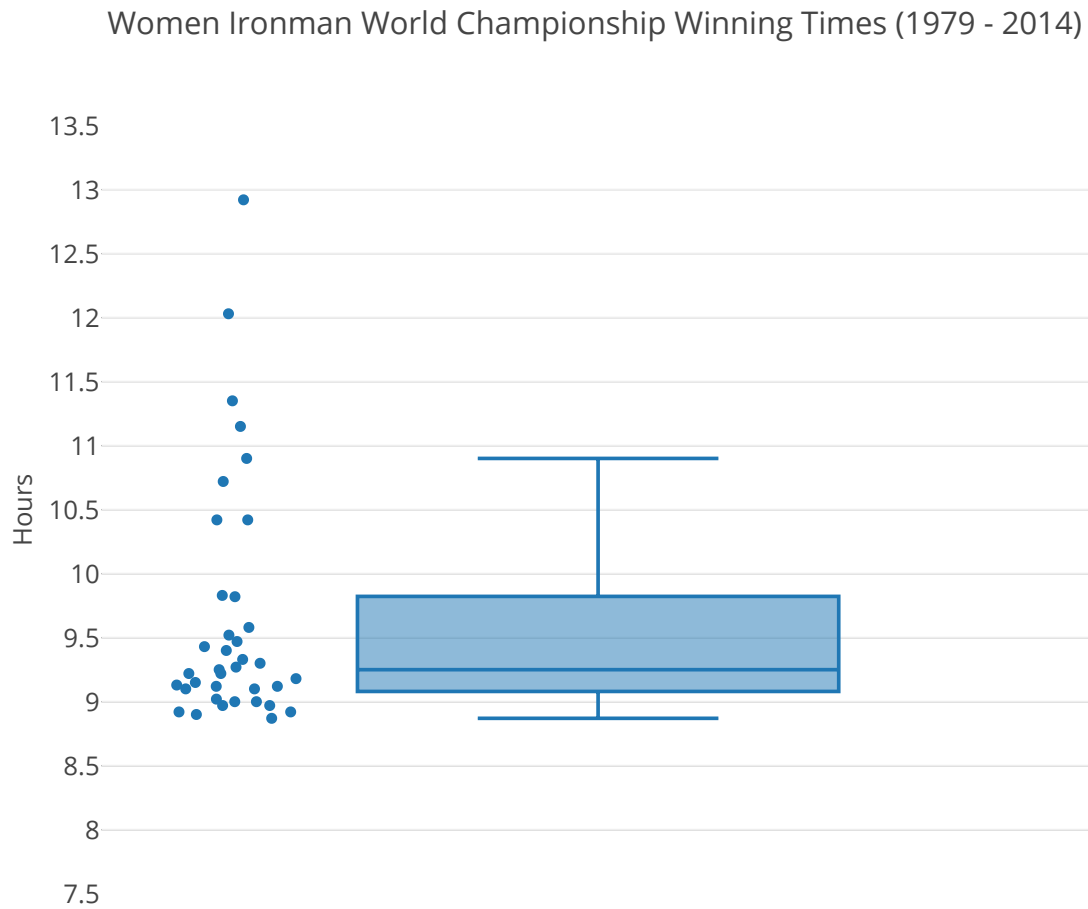
SCENARIO: Ironman Triathlon winners

Background

An Ironman Triathlon is a long-distance triathlon race consisting of a 2.4-mile (3.86 km) swim, a 112-mile (180.25 km) bicycle ride and a marathon 26.2-mile (42.2 km) run, raced in that order and without a break. It is widely considered one of the most difficult one-day sporting events in the world.

The name "Ironman Triathlon" is also associated with the original Ironman triathlon which is now the Ironman World Championship. Held in Kailua-Kona, Hawaii, the world championship has been held annually since 1978. Ironman World Championship has become known for its grueling length and harsh race conditions.

The following boxplot displays the winning times for women Ironman champions from 1979 to 2014 (note the race was conducted twice in 1982). You can mouse over the boxplot to see the values of Max, Q3, Median, Q1, and min. Note that in this graph, we have displayed all of the data points (not just the outliers) next to the boxplot.



Learn By Doing

1/1 point (graded)

What is the value of Q1 in this boxplot (hover over the image to see the five-number summary)?

☐ 8.87

☒ 9.08 ✓

☐ 9.25

☐ 9.8225

☐ 10.9

☐ 12.92

Answer

Correct: 9.08 is the value for Q1.

Submit

Learn By Doing (1/1 point)

Look at the data points on the left side of the graph, and compare them to the boxplot on the right. What do you notice about how the data are grouped, and the corresponding parts of the boxplot?

Your Answer:

The clump of data points is very near the IQR as illustrated by the boxplot.

Our Answer:

We can observe a couple of interesting features by comparing the data points to the boxplot: First, we notice that where the data points are tightly grouped, the distance between the lines in the boxplot (for example, between Q1 and M) is much smaller: Second, it is interesting to note that there is not much of a gap between the datapoint at the top of our boxplot and the four outliers (circled in red below). This is because...

Resubmit

Reset

Learn By Doing

1/1 point (graded)

What percentage of the winners had a time of less than 9.5 hours (hover over the image to see the five-number summary)?

☐ 25%

☐ 50%

☒ most of them ✓

☐ all of them

Answer

Correct:

The median is less than 9.5 hours, so we know that more than 50% of the winners had a time of less than 9.5 hours.

Submit

Learn By Doing

1/1 point (graded)

What is the IQR of this dataset?

☐ 0.5 hours

☒ just under 1 hour ✓

☐ 1 hour

☐ 1.5 hours

Answer

Correct:

We can see that Q1 is just over 9, and Q3 is just under 10, so we know that Q3 - Q1 must be less than 1 hour.

Submit

Learn By Doing

1/1 point (graded)

Within which interval of times would you expect to find the largest number of winners?

☒ 9 - 9.5 hours ✓

☐ 9.5 - 10 hours

☐ 10 - 11 hours

☐ 11 - 13 hours

Answer

Correct:

9 is below the first quartile and 9.5 is above the median, so more than 25% of the winning times are in this interval.

Submit

Open Learning Initiative [🔗](#)



[🔗](#) Unless otherwise noted this work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License [🔗](#).

© All Rights Reserved