

Warm-Up

Find the 5 number summary of the age of death of these U.S. Presidents and construct a box and whisker plot. Are there any outliers?

President	Age of Death
Washington	67
Adams	90
Jefferson	83
Madison	85
Monroe	73
Adams	80
Jackson	78
Van Buren	79
Harrison	68
Tyler	71
Polk	53
Taylor	65

$min = 53$
 $Q_1 = 67.5$
 $med = 75.5$
 $Q_3 = 81.5$
 $max = 90$
 $IQR = Q_3 - Q_1 = 81.5 - 67.5 = 14$
 $1.5 \cdot IQR = 1.5(14) = 21$
 $67.5 - 21 = 46.5$
 $81.5 + 21 = 102$

Apr 20-8:03 AM

Feb 7-11:19 AM

The Correlation Coefficient (r)

- a statistic that measures the strength and direction of a linear (straight line) relationship
- The symbol is "r" and it can take any value from -1 to 1

Direction

The direction of the relationship, positive or negative, is given by the sign of the r value

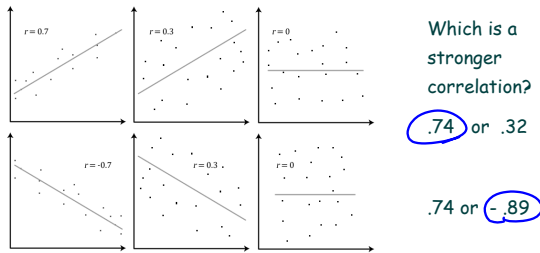
- a positive value for r (+) indicates that the relationship is positive (up and right)
- a negative value for r (-) indicates that the relationship is negative (down and right)

Feb 7-11:22 AM

Feb 7-11:47 AM

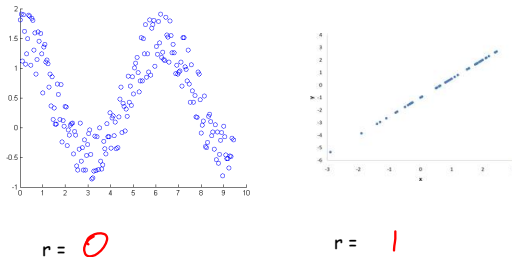
Strength (how close the points are to forming a perfect line)

- An r value of exactly 1 or -1 has a perfect correlation and forms a perfect line.
- The closer the r value is to either +1 or -1, the stronger the linear relationship is. Similarly, as the r value gets closer to 0, the weaker the linear relationship.



Feb 7-11:55 AM

What is the approximate r value for each of these scatterplots?



Feb 7-12:13 PM

Important Facts About Correlation:

- It only describes the strength of straight line relationships.
- It is represented by the letter r
- Positive r means positive association between the variables and negative r means negative association between the variables.
- Is always between -1 and +1.
- Does not change when we change the units of measurement (if we measure something in inches rather than centimeters, the correlation does not change)
- If we reverse x and y , we get the same correlation
- Strongly affected by outliers

Feb 15-10:33 AM

Find the correlation of the following data:

height	weight
60	120
62	135
65	135
68	140
68	150
68	135
67	130
61	150

What does the correlation tell you?
 $r = .7$
 borderline of weak as height increases, the weight increases

How does the correlation change if you add someone who is 61 inches tall and weighs 150 pounds? $r = .28$

Feb 23-9:51 AM

Causation vs. Correlation

Just because two variables have a strong correlation, does not mean that one caused the other.

Sometimes, it can just be a coincidence, but in most cases where one event does not cause the other, there is a lurking variable.

A lurking variable is neither your x or your y value but has an influence on the relationship between them.

Correlation:

Ex: There is a strong positive relationship between the amount of time a person spends watching TV and their body fat percentage. Therefore, watching TV causes you to get fatter.

Causation:

Ex: There is a strong positive relationship between the number of beers a person drinks and their blood alcohol content. Therefore, drinking more beer causes an increase in your blood alcohol content.



Feb 7-7:43 PM

Feb 16-2:46 PM

OVER 2 MILLION AMERICANS EXPOSED TO DRINKING WATER WILL DIE THIS YEAR



Homework:

Section 6.2 Worksheet

Talk About EXTRA CREDIT

Feb 16-2:47 PM

Oct 17-8:56 PM