



Causation =
9 month old eating Spaghetti
and red sauce on her face

Correlation =
age of Piper and the
location of the dogs
during dinner

Unit 11: Comparing Linear, Quadratic, & Exponential Functions

Day 2 - Correlation vs Causation

Correlation: implies a mutual relationship between two or more things. It is very IMPORTANT to understand that just because two variables are strongly correlated does NOT imply a cause and effect relationship. A strong relationship between two variables could be a coincidence or caused by additional factors. Typically, correlations use the words noticed and showed.

Correlations only show relationships... they cannot be used to make conclusions!!

Causation: implies a relationship in which one action or event is the direct consequence of another (cause and effect).

Correlation	Causation
<ul style="list-style-type: none"> Smoking is correlated with alcoholism (but it doesn't cause it). The more ice cream consumed on a beach, the increased number of people who go in the water (eating ice cream doesn't cause you to go in the water more). 	<ul style="list-style-type: none"> The more you smoke, the chances of developing lung cancer increase. (Does smoking cause lung cancer?) The less calories you eat, the more weight you lose (Does eating less cause you to lose weight?)

Example: Determine if the following relationships show a correlation or causation:

A. A recent study showed that college students were more likely to vote than their peers who were not in school.

Correlation

B. Dr. Shaw noticed that there was more trash in the hallways after 2nd period than 1st period.

Correlation

C. You hit your little sister and she cries.

Causation

D. The number of miles driven and the amount of gas used on your trip to Disneyworld.

causation

E. The age of a child and his/her shoe size.

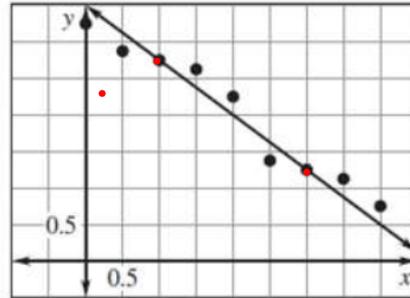
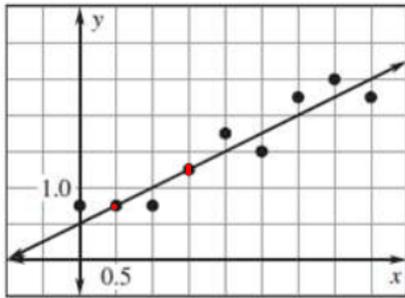
caus.

corr

F. The amount of cars a salesman sells and the amount of commission he makes during the month of July.

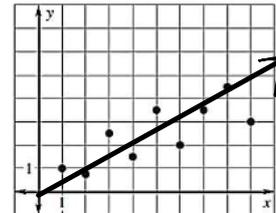
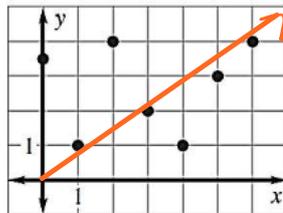
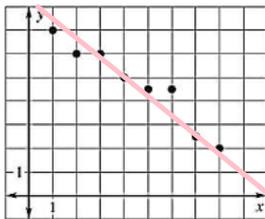
causation

Yesterday, we drew trend lines to help us see if a scatter plot had any types of correlation. A **trend line** is a line that closely models the data. A **line of best fit** is the line that comes closest to all of the points in the data set. The line of best fit provides the predicted values for a set of data.



If a line is a good line of best fit, it will have data points above and below the line.

Example: Draw a line of best fit for each graph:



strong -

weak +

strong +

Example: The table shows test averages of eight students. The equation that best models the data is $y = 0.77x + 18.12$ and the correlation coefficient is 0.87. Discuss correlation and causation for the data set.

U.S. History Test Average	90	70	75	100	90	85	80	90
Science Test Average	80	75	72	95	92	82	80	92



history grades
don't CAUSE

Correlation strong + science grades..

Example: Eight adults were surveyed about their education and earnings. The table shows the survey results. The equation that models the data is $y = 0.59x + 30.28$ and the correlation coefficient is 0.86. Discuss correlation and causation for the data set.

Years of Education	12	16	20	14	18	16	16	18
Earnings Last Year (thousand \$)	40	65	75	44	70	50	54	85



Correlation strong +

education might help get a job, but does not necessarily CAUSE ↑ earnings