

Algebra RH

Essential Questions: How can we model a relationship between two quantities graphically? How can we determine the strength of the relationship?

Do Now: Do you think the number of hours you study for a test affects the grade you get? Explain.

Scatter Plots and Correlations



Recall that.... *Statistics is the study of the collection, organization, analysis, interpretation and presentation of data.*

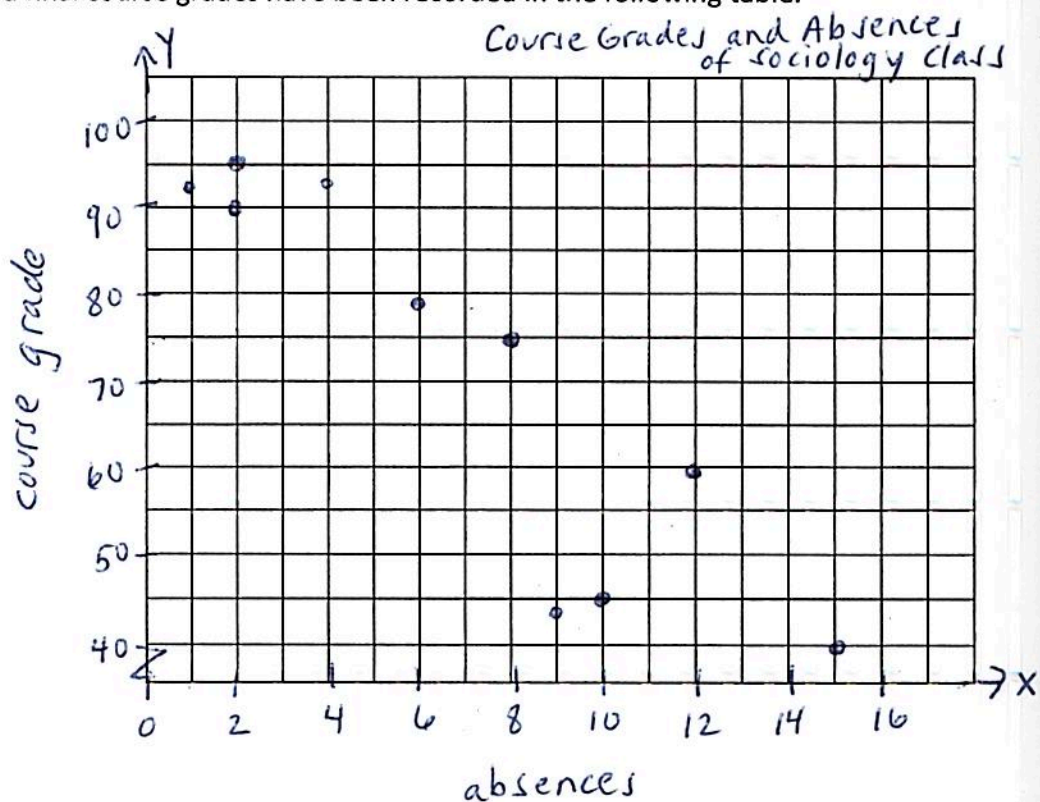
Statisticians gather data to determine if **correlations** (relationships) exist between events.

Scatter Plots will often show at a glance whether a relationship exists between **two** sets of data (**bivariate data**).

Do you think absences from school have an effect on grades?

The following data represents 10 students in a college level sociology class at a community college. Their absences and final course grades have been recorded in the following table.

X	Y
Absences	Grade
1	92
4	93
2	90
10	45
15	40
6	79
8	75
2	95
9	44
12	60



Does a relationship (correlation) exist?

yes, indirect (negative correlation)

What conclusion can be made?

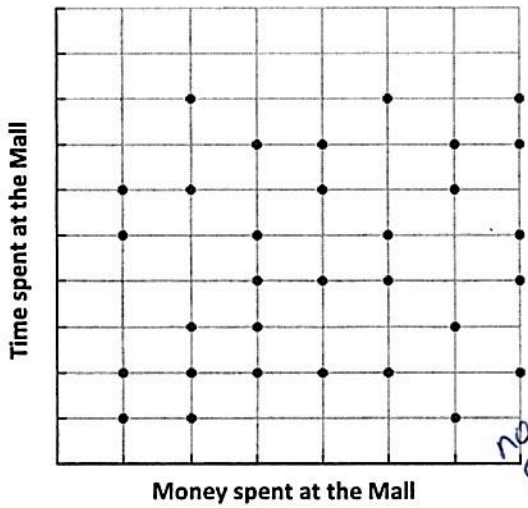
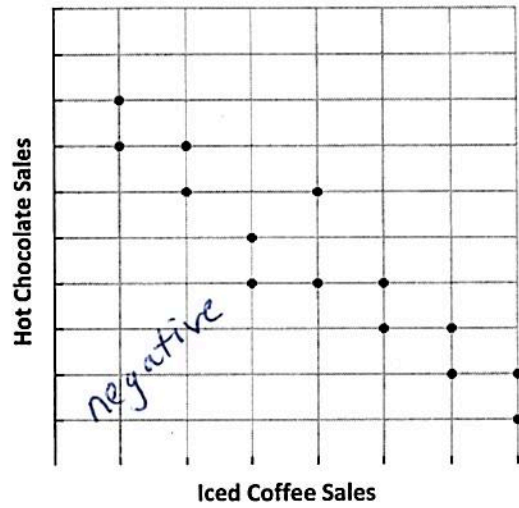
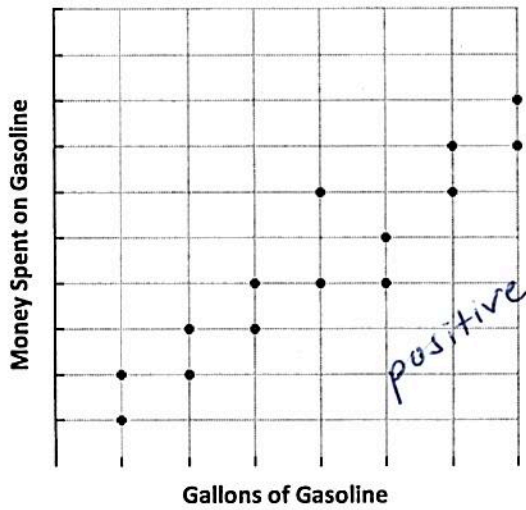
As the number of absences increases, the grades decrease.

A **Positive Correlation** shows that as x increases, y also increases.

A **Negative Correlation** shows that as x increases, y decreases.



Examples: State whether each scatter plot has a positive correlation, negative correlation or no correlation.



Keep in mind, two sets of data may correlate but not show causation.

Causation: Bivariate data shows causation when one variable causes the other. The final grade attained in a course is impacted (caused) by the number of absences.

Examples: Read the following statements and determine if a correlation exists. Be prepared to identify the type of correlation you believe would exist and discuss whether or not the correlation is causal.

1. Level of education attained compared to salary earned.

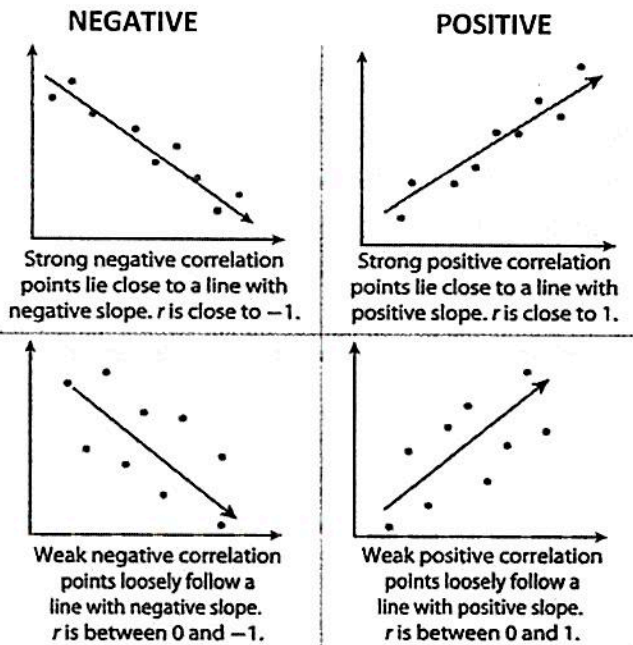
positive, causal

2. Ice cream sales compared to swimsuit sales.

positive, not causal

Correlation Coefficients

One measure of strength and direction of a correlation is the correlation coefficient, denoted by r . The value of r ranges from -1 to 1 .



STRONG

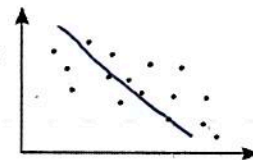
The stronger the correlation, the closer the correlation coefficient will be to -1 or 1 .

WEAK

If there is no correlation between the two variables in a data set, r is close to zero.

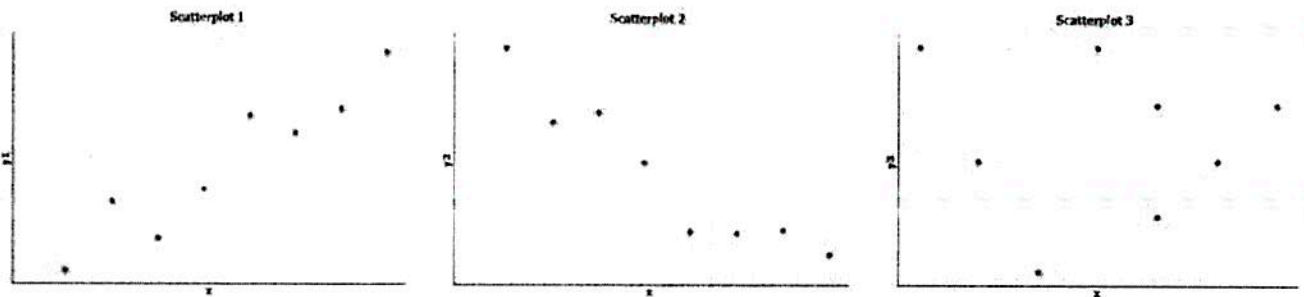
IT'S YOUR TURN NOW...

- Describe the correlation represented by the scatter plot?



negative weak

Use the scatter plots below to answer #'s 2 and 3.



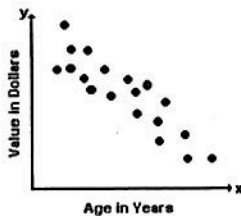
- If one of these scatter plots represents the height and weight for eight adults, which plot do you think it is and why?

#1 as height increases, weight increases

- If one of these scatter plots represents the relationship between height and SAT math score for eight high school seniors, which scatter plot do you think it is and why?

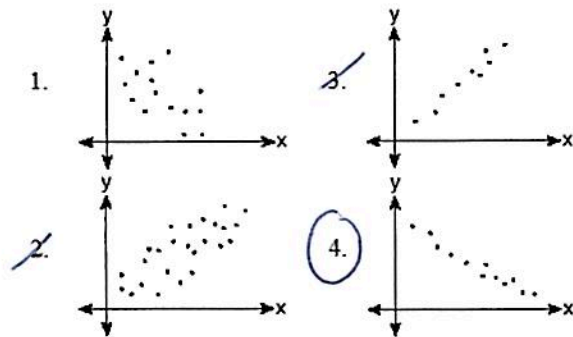
#3 no correlation

4. Based on this scatter plot, it would be reasonable to conclude:



- (1) Age and value have a coefficient of correlation less than zero.
- (2) Age and value have a coefficient of correlation that is equal to zero.
- (3) Age and value have a coefficient of correlation that is between zero and 0.5.
- (4) Age and value have a coefficient of correlation that is greater than 0.5.

5. Which graph represents data used in a linear regression that produces a correlation coefficient closest to -1.



6. Consider the correlation coefficients below that represent various sets of data. List them in order from strongest correlation to weakest correlation.

-0.72

0.92

0.05

-0.5

.92 -.72 -.5 .05

The TAKEAWAY

A scatter plot represents a set of data between two variables. The relationship between the two variables is known as the correlation. The strength of the relationship is represented by a number known as the correlation coefficient. This number ranges from -1 to 1.