GSE Algebra 1
Unit 6 D14
Describing Correlation & Correlation Coefficient
A **correlation** describes a relationship between two data sets. A graph may show the correlation between data. The correlation can help you analyze trends and make predictions.

There are three types of correlations between data.
Correlations

Positive Correlation
Both sets of data values increase.

Negative Correlation
One set of data values increases as the other set decreases.

No Correlation
There is no relationship between the data sets.
Identify the correlation you would expect to see between the pair of data sets. Explain.

the number of people in an audience and ticket sales

You would expect to see a positive correlation. As ticket sales increase, the number of people in the audience increases.
Identify the correlation you would expect to see between the pair of data sets. Explain.

the average temperature in a city and the number of speeding tickets given in the city

You would expect to see no correlation. The number of speeding tickets has nothing to do with the temperature.
Identify the correlation you would expect to see between the pair of data sets. Explain.

A runner’s time and the distance to the finish line

You would expect to see a negative correlation. As time increases, the distance to the finish line decreases.
Identify the type of correlation you would expect to see between the pair of data sets. Explain.

the temperature in Houston and the number of cars sold in Boston

You would expect to see no correlation. The temperature in Houston has nothing to do with the number of cars sold in Boston.
Identify the type of correlation you would expect to see between the pair of data sets. Explain.

the number of members in a family and the size of the family’s grocery bill

You would expect to see a positive correlation. As the number of members in a family increases, the size of the grocery bill increases.
Identify the type of correlation you would expect to see between the pair of data sets. Explain.

the number of times you sharpen your pencil and the length of your pencil

You would expect to see a negative correlation. As the number of times you sharpen your pencil increases, the length of your pencil decreases.
Choose the scatter plot that best represents the relationship between the age of a car and the amount of money spent each year on repairs. Explain.

**Graph A**

![Graph A](image)

**Graph B**

![Graph B](image)

**Graph C**

![Graph C](image)
Choose the scatter plot that best represents the relationship between the number of minutes since a pie has been taken out of the oven and the temperature of the pie. Explain.

Graph A

Graph B

Graph C
Correlation DOES NOT imply causation. If two variables have a strong correlation to one another, it does not mean one variable *causes* the other to increase or decrease. This is a common misunderstanding and is seen often in the media.

Example of causation: Smoking and Lung cancer
Example of correlation: Smoking and Alcoholism

Discuss: Plastic surgery and depression
Social media use and grades
Ice cream sales and murder rates
Correlation Coefficient \( r \)

Measures the **strength** and **direction** of the linear association between the variables.

The value of a correlation coefficient ranges between -1 and 1.

The greater the absolute value of a correlation coefficient, the stronger the **linear** relationship.

The **strongest** linear relationship is indicated by a correlation coefficient of -1 or 1.

The **weakest** linear relationship is indicated by a correlation coefficient equal to 0.
Scatter Plots and Trend Lines

Correlation Coefficient examples

- $r = 1$: Maximum positive correlation
- $r = 0.7$: Strong positive correlation
- $r = 0$: No correlation
- $r = -1$: Maximum negative correlation
- $r = -0.5$: Weak negative correlation
- $r = 0.8$: Strong correlation & outlier
Correlation Coefficient $r$

When the slope of the line in the plot is negative, the correlation is negative; and vice versa.

The strongest correlations ($r = 1.0$ and $r = -1.0$) occur when data points fall exactly on a straight line.

The correlation becomes weaker as the data points become more scattered.

If the data points fall in a random pattern, the correlation is equal to zero.

Correlation is affected by outliers. Compare the first scatterplot with the last scatterplot. The single outlier in the last plot greatly reduces the correlation (from 1.00 to 0.8).
Estimate the correlation coefficient.

\[ r = 1 \]  
\[ r = 0.5 \]  
\[ r = 0 \]  
\[ r = -1 \]  
\[ r = -0.7 \]  
\[ r = 0 \]
Estimate the correlation coefficient.

\[ r = -0.5 \]

\[ r = 0 \]