

3.5/3.6 Guided Notes – Scatterplots and Trend Lines

Vocabulary

Scatter Plot:	A type of graph to display how 2 variables relate for a data set
Trend Line:	A line used to represent the behavior of a set of data ; used to make predictions
Correlation (Coefficient):	A value measuring the strength of relationship between 2 variables , denoted by r . r is between -1 and 1
Residual:	The difference between our predicted output and actual output
Interpolate:	Using trend line to predict an output within the domain of collected data
Extrapolate:	Using trend line to predict an output outside of the domain of collected data

Ex. 1 Create a Trend Line

- * visualize line passing through 'middle' of data points
- * try to select 2 pts to pass through



- a) Calculate the slope of the trend line using 2 points on (or closest to) the trend line.

$$(62, 4) \text{ \& } (67, 9) \quad m = \frac{9-4}{67-62} = \frac{5}{5} = 1$$

- b) Calculate the y-intercept of the trendline.

$$y = mx + b \quad \begin{matrix} 9 = 67 + b \\ 9 = 1(67) + b \end{matrix} \quad \begin{matrix} b = -58 \end{matrix}$$

- c) Write a slope-intercept equation for the trendline.

$$y = 1x - 58$$

- d) **Interpolate** What shoe size would we predict a person who is 63 inches tall would have?

$$y = 1(63) - 58 = 5 \quad \dots \quad \begin{matrix} \text{predict someone } 5'3'' \\ \text{has shoe size } 5 \end{matrix}$$

- e) **Extrapolate** What shoe size would we predict a person who is 76 inches tall would have?

$$y = 1(76) - 58 = 12 \quad \dots \quad \text{predict someone } 6'4'' \text{ has shoe size } 12$$

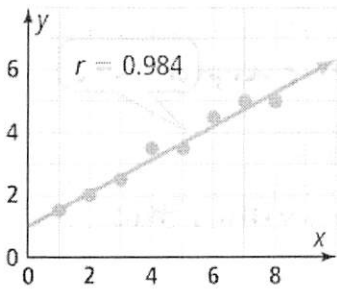
Ex. 2 Interpreting Correlation Coefficients

Strong Correlation: Data that is closely grouped and varies little.
 r values will be close to -1 or $+1$

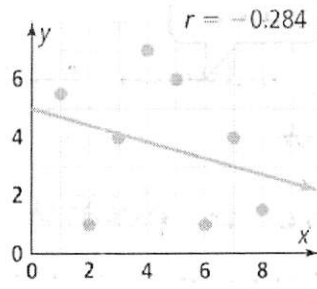
Weak Correlation: Data that is varied, but still shows a trend
 r values will be between -0.6 to -0.2 or 0.2 to 0.6

No Correlation: Data shows no predictive relationship
 r values between -0.2 to 0 and 0 to 0.2

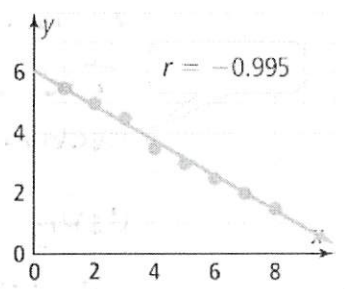
Positive or negative correlation:



Strong, positive

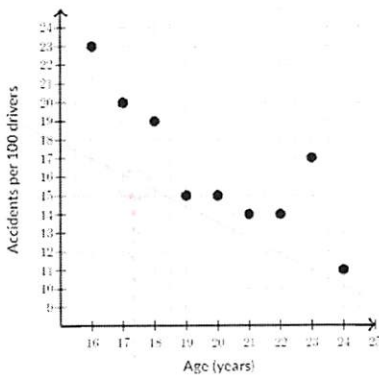


Weak, negative



Strong, negative

You Try!



What type of correlation does the data in this scatterplot represent?

Strong, negative

Circle the correlation coefficient from below that best matches the data

$r = 0.85$

$r = -1$

$r = 0.03$

$r = -0.72$

Ex. 3 Interpreting Residual Plots

Data that can be represented w/ a trend line should have random residuals. They should be clustered close to zero and be randomly dispersed above and below zero.

