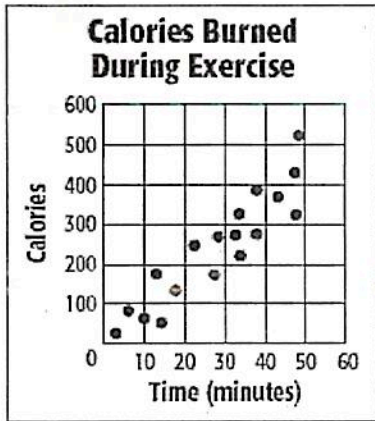Algebra RH

Essential Question: How can we represent a data set with a linear model?

Do Now: Does the graph show a positive correlation, negative correlation, or no correlation? If there is a positive or negative correlation, describe its meaning in the situation.

Calories Burned During Exercise

(scatter plot: Calories vs Time (minutes), x-axis 0–60, y-axis 0–600)

positive correlation

as minutes exercising increases, the number of calories increases

Is there a linear association? Do the points appear to form a line?

Line of Fit

When a ==bivariate data== set displays a *strong positive* or *negative correlation,* you can use a <u>linear</u> equation to represent the data. The process is called finding a ==line of fit== for the data. This line is referred to as a ==trend line== or ==linear regression model==.

Drawing a Line of Fit for Data

Draw a line that passes as close as possible to the plotted points.
- Your line does not necessarily have to pass through any of the plotted points.
- You should try to have about the same number of points above and below the line.

Finding the Equation of the Line of Fit

Choose two points on the line of best fit. $(2, 70)$ $(4, 85)$

Find the slope

graders typically go up 7.5 points per hour worked

$$\frac{\Delta y}{\Delta x} =$$

$$ROC = \frac{7.5 \text{ grades}}{1 \text{ hours}}$$

Find the y-intercept

$$y = mx + b$$
$$70 = 7.5(2) + b$$
$$70 = 15 + b$$
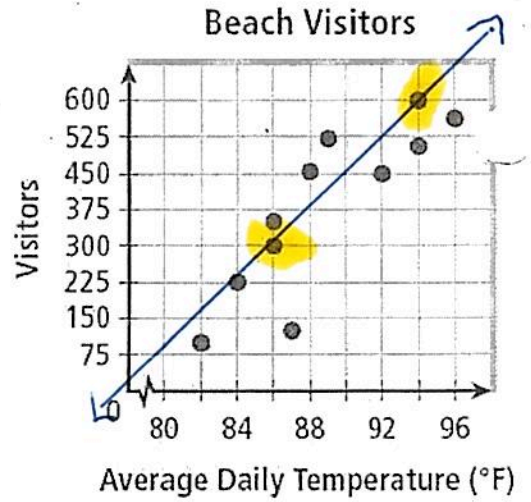$$55 = b$$

initial grade (without studying) is typically 55

Time Spent and Report Grades

(scatter plot: Grade vs Hours Worked, x-axis 0–6, y-axis 60–100)

Equation of Best Fit Line $\underline{\quad y = 7.5x + 55 \quad}$
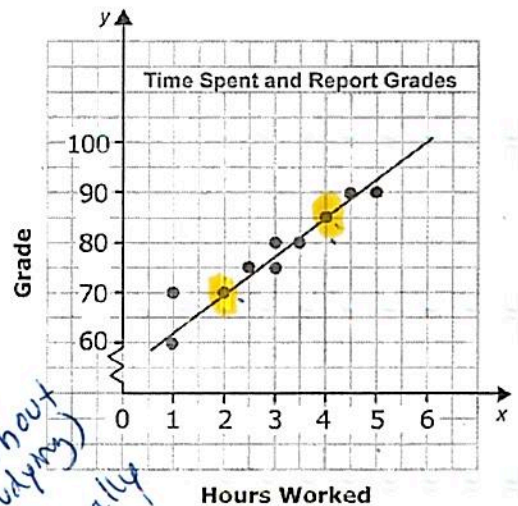
Let's Try Another!

➢ Draw a line of best fit.
➢ Choose two points on the line.
➢ Find slope and the y-intercept.

$$x_1, y_1 \quad x_2, y_2$$
$$(86, 300) \quad (94, 600)$$

Beach Visitors

Visitors: 600, 525, 450, 375, 300, 225, 150, 75

Average Daily Temperature (°F): 80, 84, 88, 92, 96

slope	y intercept
$\dfrac{\Delta y}{\Delta x} = \dfrac{300 - 600}{86 - 94}$	$y = mx + b$
$= \dfrac{-300}{-8}$	$600 = 94(37.5) + b$
$= 37.5$	$600 = 3525 + b$
	$-2925 = b$

typically, there will be 37.5 more beach visitors per increase of 1°F

➢ Write the equation of the line of fit. $\quad y = 37.5x - 2925$

Using the Line of Fit to Make Predictions

When you use a trend line or its equation to predict a value within the data points, you (interpolate) the predicted value.

between

When you make a prediction that is *outside* the data, you (extrapolate) the predicted value. *go beyond*

Using the equation of the line of fit for the beach visitors,

(a) Predict the number of beach visitors if the temperature is 90° (interpolation).

$$y = 37.5x - 2925$$
$$y = 37.5(90) - 2925$$
$$y = 450$$

it's predicted that we'll have 450 visitors when the temp is 90°F

(b) Predict the number of beach visitors if the temperature is 102° (extrapolation).

$$y = 900$$

it's expected the beach will have 900 visitors when the temp is 102°

There are two types of data: _one_ univariate and _two_ bivariate data.

Univariate Data	Bivariate Data
involving a single variable	involving two variables
does not deal with causes or relationships	deals with causes or relationships
the major purpose of univariate analysis is to describe	the major purpose of bivariate analysis is to explain
central tendency - mean, mode, median dispersion - range, variance, max, min, quartiles, standard deviation. frequency distributions bar graph, histogram, pie chart, line graph, box-and-whisker plot	analysis of two variables simultaneously correlations comparisons, relationships, causes, explanations tables where one variable is contingent on the values of the other variable. independent and dependent variables
Sample question: How many of the students in the freshman class are female?	Sample question: Is there a relationship between the number of females in Computer Programming and their scores in Mathematics?

4. Determine if the problem deals with univariate or bivariate data.
    A. Determine if the number of hours a student studies will improve his/her final _bivariate_ examination scores.
    B. Determine the mean score on the last math test. _univariate_
    C. Determine whether lowering cholesterol lowers the risk of heart disease. _bivariate_
    D. Prepare a frequency table for the number of students scoring A, B, C and D on the Chemistry test. _univariate_

5. Which situation should be analyzed using bivariate data?
    A. Ms. Fonseca keeps a list of the amount of time her daughter spends on her social studies homework.
    B. Mrs. Huntley tries to see if her students' shoe sizes are directly related to their heights.
    C. Mr. Chang records his customers' best video game scores during the summer.
    D. Mrs. Gizzi keeps track of her daughter's algebra grades for the quarter.

6. Which is an example of univariate data?
    A. measuring the heart rate of different people at different ages
    B. measuring heart rates with different numbers of hours of sleep
    C. measuring heart rates with different numbers of hours of exercise per week
    D. measuring heart rates of ten 14 year old girls

A ___line of fit___ can help us summarize a set of data. The line and its equation can help us make predictions about the relationship between the two variables. When you use the trend line or its equation to predict a value within the set of data, you are ___interpolating___.

When you make a prediction outside the set of data, you are ___extrapolating___.

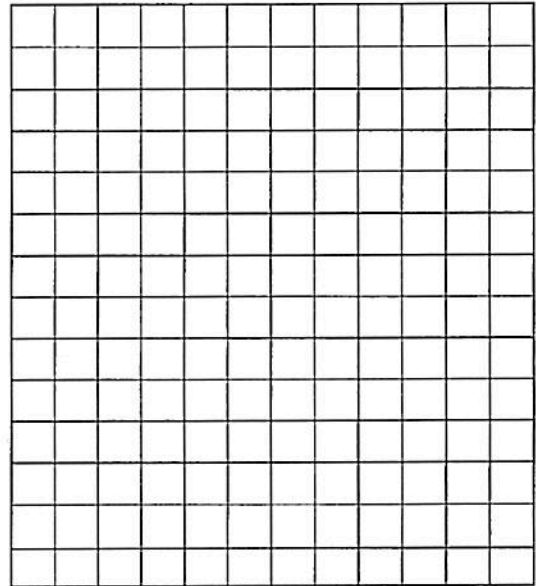Algebra RH   HW # _____

Joyce is training for a 10K race. For some of her training runs, she records the distance she ran and how many minutes it took her.

(a) Make a scatter plot of Joyce's running data.

Distance (mi)	Time (min)
4	38
2	25
1	7
2	16
3	26
5	55
2	20
4	45
3	31

(b) Describe the correlation.

(c) Draw a trend line and write the equation of the line.