

**LEARNING OBJECTIVE:** We will fit a straight line to data in a scatter plot and determine its equation (Lesson 62)

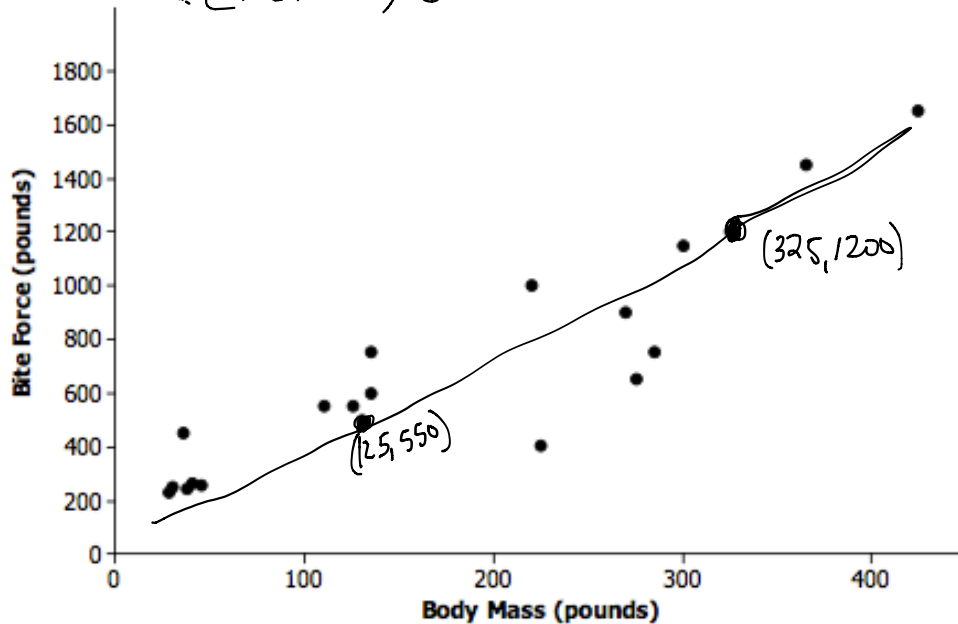
**CONCEPT DEVELOPMENT:**

A **Scatter Plot** is a graph of bivariate numerical data.

*two pieces of*

Why do we view data in a scatterplot?

- Visual representation of data.
- Easier (than a table of data) to see trends.
- *Patterns & trends*
- *To see how one variable affects the other*
- *See negative/positive, linear/non-linear & clusters, outliers*



**Line of Best Fit:** When scatter plots reveal a linear relationship, we can draw a line that represents the trend in the data. Our line should be drawn as close to as many points on the graph as possible.

We can also write an **equation** for this line—by identifying two points on the line, finding a slope and then a y-intercept.



**Dependent Variable** (also called **response variable**) is the variable for which you are trying make predictions. Its value is dependent on the value of the independent variable. This is the y-value, on the vertical axis.

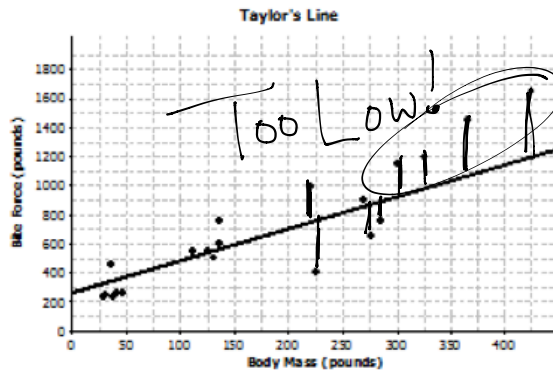
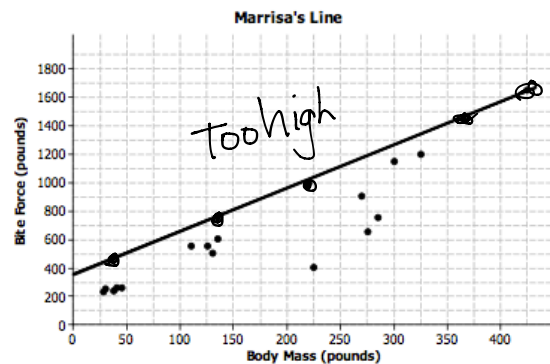
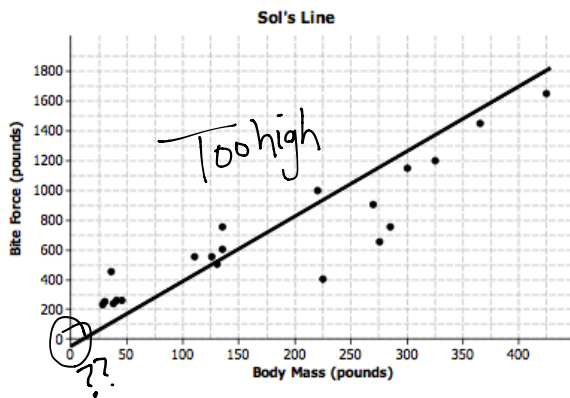
Example #1: The price you pay for a home depends on how big it is.



Example #2: The force with which a crocodile bites down is dependent on how much it weighs.



**Independent Variable** (also called **explanatory variable**) is the variable that is not changed by the other variables. This is the x-value on the horizontal axis.



We can measure the distance between a point on the graph and the line of best fit to see if data points behave as we would predict they would.

**GUIDED PRACTICE:****Steps for Determining Equations of Lines of Best Fit**

1. Evaluate your scatter plot and evaluate the linearity of the data. (Does the data reveal a linear relationship?)
2. Draw a line over the data that goes as close to as many points as possible.
3. Identify two points on the line to find out the slope, and then use that figure out the y-intercept.
4. Answer any other questions in the context of the question.

Examine the table below and the scatter plot on the next page which measures the body mass and bite force of different kinds of crocodiles.

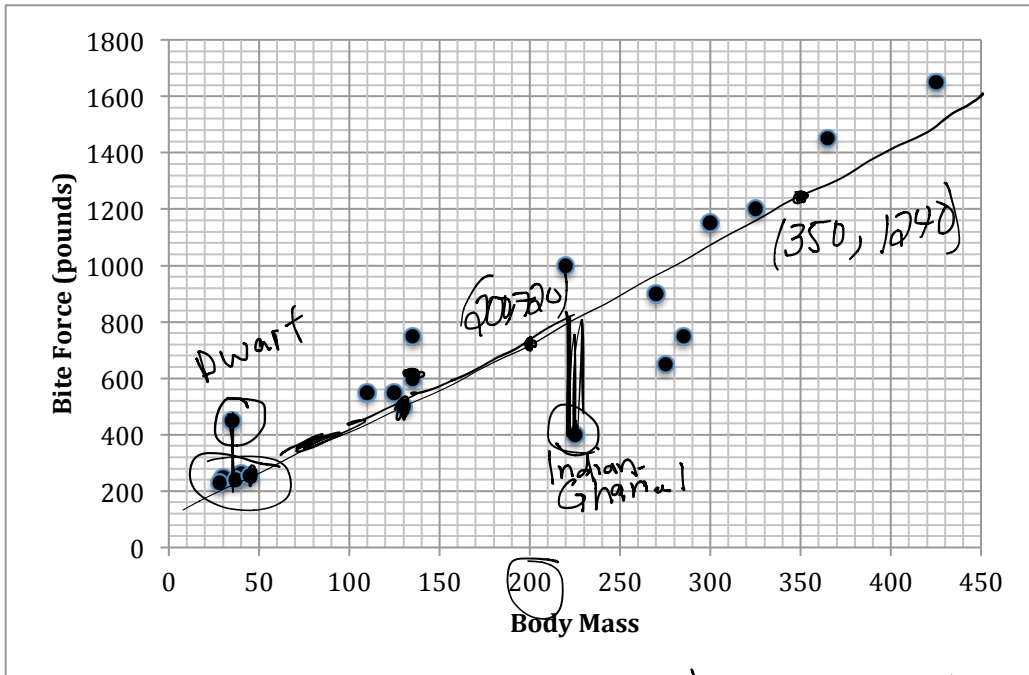
Species	Body Mass (pounds)	Bite Force (pounds)
Dwarf Crocodile	35	450
Crocodile F	40	260
Alligator A	30	250
Caiman A	28	230
Caiman B	37	240
Caiman C	45	255
Nile Crocodile	275	650
Croc A	110	550
Croc B	130	500
Croc C	135	600
Croc D	135	750
Caiman D	125	550
Indian Gharial Croc	225	400
Crocodile G	220	1,000
America Croc	270	900
Croc D	285	750
Croc E	425	1,650
American Alligator	300	1,150
Alligator B	325	1,200
Alligator C	365	1,450

The heavier the  
Croc, the stronger  
the bite.

Strongest bite  
force

What do you notice about the table?

- Lots of different species
- Dwarf croc has strong jaws.
- Alligators have stronger bites. Alligators weigh more



$$\text{slope} = \frac{720 - 1240}{200 - 350}$$

$$= \frac{-520}{-150} = 3.46$$

What do you notice about the scatter plot?

Positive, linear.

Can you draw a line of best fit and find the equation for the line?

$$y = 3.46x + 28$$

$$y = 3.46x + b$$

$$720 = 3.46(200) + b$$

$$720 = 692 + b$$

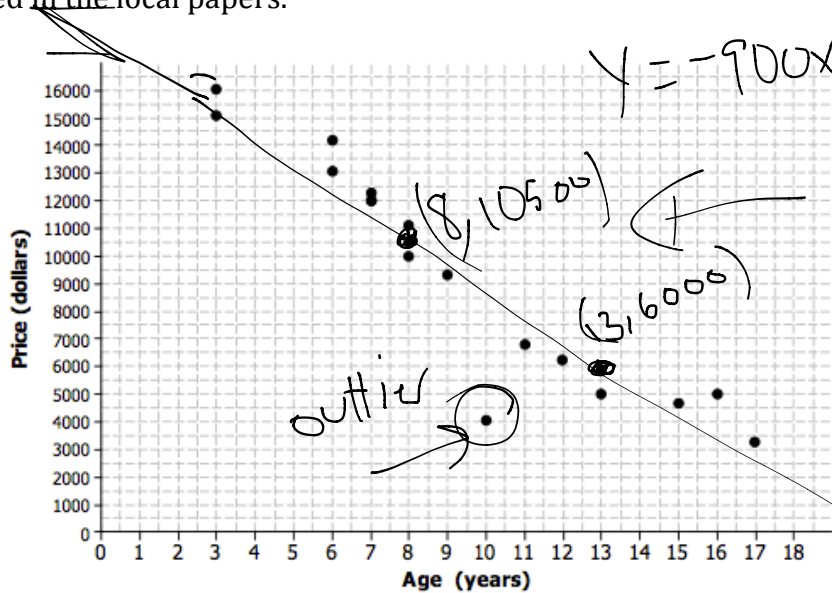
$$28 = b$$

Find the point that represents the Dwarf Crocodile and the point that represents the Indian Gharial Croc. Which observation was closer to the predicted line you drew?

How do you know?

Dwarf Croc. closer to predicted line

Below is a scatter plot that the age in years and price in dollars for used Honda Civic cars advertised in the local papers.



What do you notice about the data?

negative linear, the older the car the less money it costs.

Draw a line of best fit. What is the equation for the line?

$-1142, -1085, -1000, -917, -900$  882  
857  
800

Which car in the data set has a predicted value farthest from the actual line?

18000, 18300, 19000, 18500, 19414

What does your equation predict the cost for a 10-year old car? How close is the prediction to the actual observation in the scatter plot?

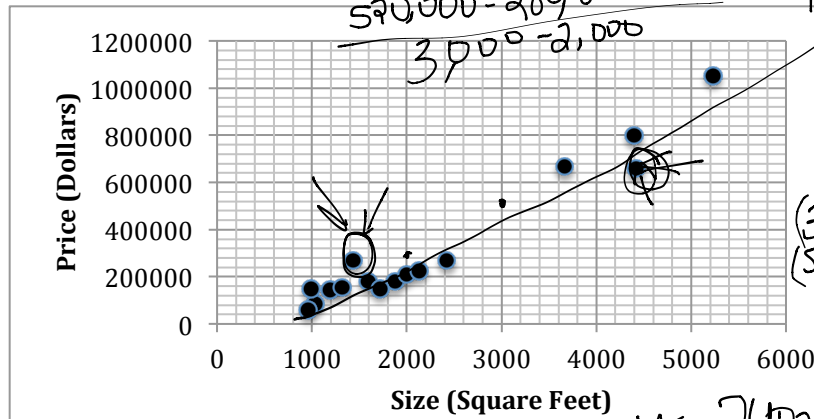
\$8,700

How much would you expect a 5-year old car to cost based on the line you drew?

$y = -900x + 17700$  (15,450) -10500 slope  
(3, 15000) 12

$300000 = 2400(1250)$

This is the data relating to house size and price in Milwaukee.



Draw a line of best fit. What is the equation for the line?

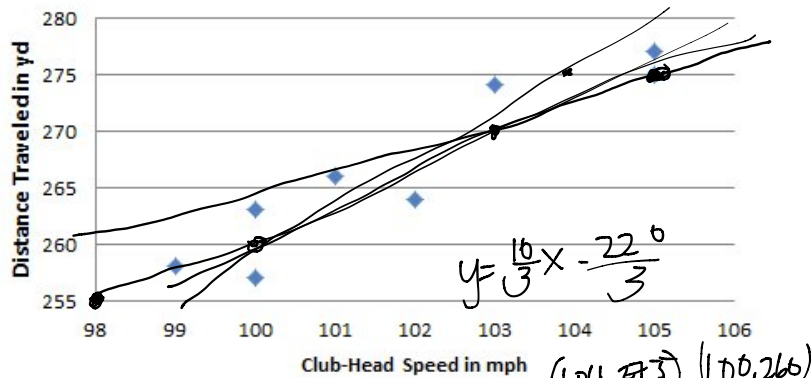
$y = 240x - 200000$   
 $y = 200x - 100,000$   
 $y = 1733x - 66,667$

Which house is the farthest away from the predictions based on the line of best fit? What do you think would account for these differences?

Below is a scatter plot that shows how far a golf ball travels based on the speed of the club head.

$105, 275$   
 $103, 270$

**Golf Ball Travel and Club-Head Speed**



What do you notice?

What is the equation for the line of best fit?

$y = \frac{15}{4}x + b$   
 $260 = \frac{15}{4} \cdot 100 + b$   
 $y = \frac{15}{4}x - 115$

NAME: \_\_\_\_\_

Math 7.2, Period \_\_\_\_\_

Mr. Rogove

Date: \_\_\_\_\_

**INDEPENDENT PRACTICE:**

Have students do research on home prices in Mountain View.

**ACTIVATING PRIOR KNOWLEDGE:**

**CLOSURE:**

Have students look

**NOTES:**

slope of line  
that passes through  
(125, 550) and (325, 1200)

$$\frac{4}{13} \quad \frac{13}{4} \quad \frac{1200 - 550}{325 - 125} = \frac{650}{200} = \frac{13}{4} \checkmark$$

$$y = mx + b$$

$$y = \frac{13}{4}x + b$$

$$550 = \frac{13}{4} \cdot \frac{125}{1} + b$$