Notes Section 4.6

Fit a Line to Data

Vocabulary

A **scatter plot** Correlation is a graph used to determine whether there is a relationship between paired data.

If y tends to increase as x increases, the paired data are said to have a **positive correlation**

If y tends to decrease as x the paired data are said to have a **negative correlation.**

If x and y have no apparent relationship, the paired data are said to have **relatively no correlation**.

When data show a positive or negative correlation, you can model the trend in the data using a **line of fit.**

Common Student Errors

• **Believing that a line of fit must pass through the leftmost and rightmost data points**

Tip Stress that a line of fit does not necessarily pass through all data points and that the same number of data points should lie above the line as below the line.

• **Choosing two data points to write the equation of the line of fit before drawing the line of fit**

Tip Stress that a line of fit must be drawn first before writing the equation of the line

Big Ideas

1. How to determine if a scatter plot shows positive, negative or no correlation in the data.

2. How to find the line of best fit.

3. How to use the relationship between the dependent and independent variables to determine the correlation of the data.

Example 1
Describe the correlation of data

Describe the correlation of data graphed in the scatter plot. The data on the graph shown the result of a forest ranger's annual survey on rabbit and red-tail hawk population.

a.

Rabbit Population and Snow Depth
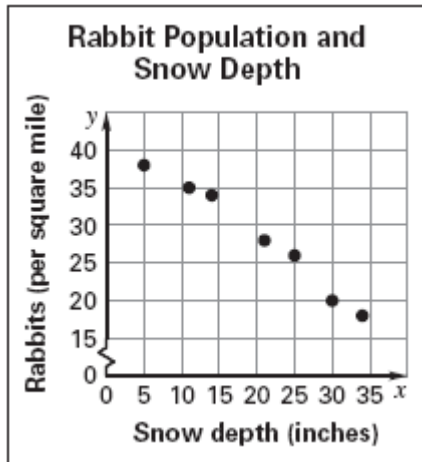
a. The scatter plot shows a negative correlation between the depth of snow cover during the winter and the number of rabbits counted in the spring. This means that as the depth of the snow increased, the number of rabbits counted decreased.

EXAMPLE 2
Make a scatter plot

Softball The table shows the number of girls signed up for a summer softball league each year for 5 years.

Year	2001	2002	2003	2004	2005
Players	105	113	120	132	148

a. Make a scatter plot of the data.
b. Describe the correlation of the data.
c. Write an equation that models the number of girls signed up for a summer softball league as a function of the number of years since 2000.

Solution

Softball Sign-ups
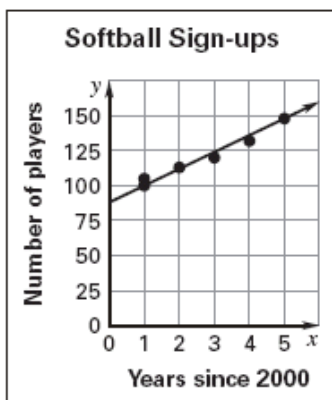
a. Treat the data as ordered pairs. Let x represent the number of years since 2000. Let y represent the number of softball players. Plot the ordered
pairs as points in a coordinate plane.

b. The scatter plot shows a positive correlation, which means that more players have signed up each year since 2000.

c. Draw a line that appears to fit the points in the scatter plot closely. Write an equation using two points on the line. Use (1, 100) and (5, 148). Find the slope of the line.

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{148 - 100}{5 - 1} = \frac{48}{4} = 12$$

Find the intercept of the line. Use the point (5, 148).

$y = mx + b$	Write slope-intercept form.
$148 = 12(5) + b$	Substitute 12 for m, 5 for x, and 148 for y
$88 = b$	Solve for b.

An equation of the line of fit is $y = 12x + 88$.