5.6 Fit a Line to Data

Goal • Make scatter plots and write equations to model data.

Your Notes

VOCABULARY

Scatter plot A graph used to determine whether there is a relationship between paired data

Correlation The relationship between two data sets

Line of fit A model used to represent the trend in data showing a positive or negative correlation

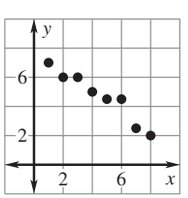
CORRELATION

- 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 increases, 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.

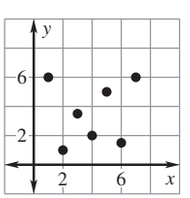
Example 1 ***Describe the correlation of data***

Describe the correlation of data graphed in the scatter plot.

a.



b.



Solution

a. negative correlation

b. relatively no correlation

Your Notes

Example 2 ***Make a scatter plot***

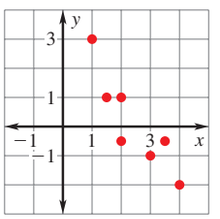
a. Make a scatter plot of the data in the table.

x	1	1.5	2	2	3	3.5	4
y	3	1	1	−0.5	−1	−0.5	−2

b. Describe the correlation of the data.

Solution

a. Treat the data as ordered pairs. Plot the ordered pairs as points in a coordinate plane.



b. The scatter plot shows a negative correlation.

USING A LINE OF FIT TO MODEL DATA

Step 1 **Make** a scatter plot of the data.

Step 2 **Decide** whether the data can be modeled by a line.

Step 3 **Draw** a line that appears to fit the data closely. There should be approximately as many points above the line as below it.

Step 4 **Write** an equation using two points on the line. The points do not have to represent actual data pairs, but they must lie on the line of fit.

Your Notes

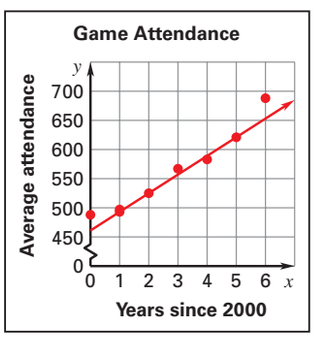
Example 3 ***Write an equation to model data***

Game Attendance The table shows the average attendance at a school's varsity basketball games for various years. Write an equation that models the average attendance at varsity basketball games as a function of the number of years since 2000.

Year	2000	2001	2002	2003	2004	2005	2006
Avg. Game Attendance	488	497	525	567	583	621	688

Solution

Step 1 **Make** a scatter plot of the data. Let x represent the number of years since 2000. Let y represent average game attendance.



Step 2 **Decide** whether the data can be modeled by a line. Because the scatter plot shows a positive correlation, you can fit a line to the data.

Step 3 **Draw** a line that appears to fit the points in the scatter plot closely.

Step 4 **Write** an equation using two points on the line. Use (1, 493) and (5, 621).

Find the slope of the line.

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{621 - 493}{5 - 1}$$

$$= \frac{128}{4}$$

$$= 32$$

Your Notes

Find the *y*-intercept of the line. Use the point (5, 621).

$y = mx + b$ **Write slope-intercept form.**

$621 = 32(5) + b$ **Substitute** 32 **for** *m*, 5 **for** *x*, **and** 621 **for** *y*.

$461 = b$ **Solve for** *b*.

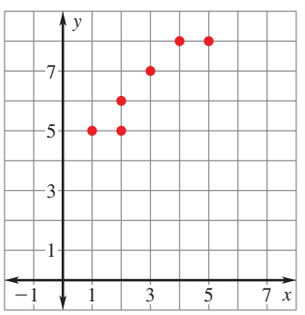
An equation of the line of fit is $y = 32x + 461$.

The average attendance *y* of varsity basketball games can be modeled by the function $y = 32x + 461$ where *x* is the number of years since 2000.

Checkpoint **Complete the following exercises.**

1. Make a scatter plot of the data in the table. Describe the correlation of the data.

x	1	2	2	3	4	5
y	5	5	6	7	8	8



positive correlation

2. Use the data in the table to write an equation that models *y* as a function of *x*.

x	1	2	3	4	5	6
y	65	76	82	86	92	97

$y = 6x + 62$

Homework