

4-6 Regression and Median-Fit Lines

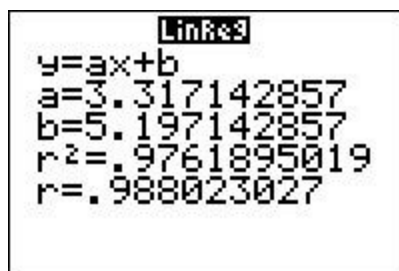
6. **RETAIL** The table gives the sales of jeans at a clothing chain since 2008. Let x be the number of years since 2008.

Year	2008	2009	2010	2011	2012	2013	2014
Sales (Millions of Dollars)	684	76	109	154	176	212	265

- Write an equation of the regression line.
- Graph and analyze the residual plot.

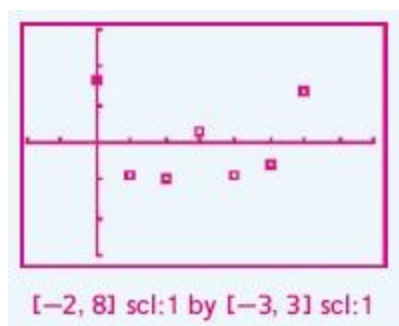
SOLUTION:

- Use a calculator to find the equation of the regression line.



The equation is about $y = 3.32x + 5.20$.

- Use a calculator to graph the residual plot for the data.



The residuals appear to be randomly scattered. Therefore, the regression line fits the data well.

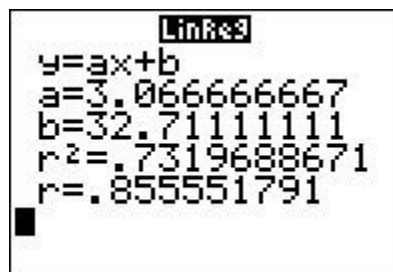
8. **CAMPING** A campground kept a record of the number of campsites rented over the week of July 4 for several years. Let x be the number of years since 2005.

Year	2007	2008	2009	2010	2011	2012	2013	2014	2015
Sites Rented	34	45	42	53	58	47	57	65	59

- Find an equation for the regression line.
- Predict the number of campsites that will be rented in 2017.
- Predict the number of campsites that will be rented in 2025.

SOLUTION:

- Use a calculator to find an equation for the regression line.



$$y = 3.07x + 32.71$$

- To predict the number of campsites that will be rented in 2017, evaluate the regression equation for $(2017 - 2005)$ or $x = 12$

$$y = 3.07x + 32.71$$

$$y = 3.07(12) + 32.71$$

$$y = 36.84 + 32.71$$

$$y = 69.55$$

$$y \approx 70$$

There will be about 70 campsites rented in 2017.

- To predict the number of campsites that will be rented in 2025, evaluate the regression equation for $(2025 - 2005)$ or $x = 20$.

$$y = 3.07x + 32.71$$

$$y = 3.07(20) + 32.71$$

$$y = 61.4 + 32.71$$

$$y = 94.11$$

$$y \approx 94$$

There will be about 94 campsites rented in 2025.

10. **ORGANIZE IDEAS** The prices of the eight top-selling brands of jeans at Jeanie's Jeans are given in the table below.

Sales Rank	1	2	3	4	5	6	7	8
Price (\$)	43	44	50	61	64	135	108	78

- Find the equation for the regression line.

4-6 Regression and Median-Fit Lines

- b. According to the equation, what would be the price of a pair of the 12th best-selling brand?
 c. Is this a reasonable prediction? Explain.

SOLUTION:

- a. Use a calculator to find an equation for the regression line.

Step 1: Enter the data by pressing **STAT** and selecting the **EDIT** option. Enter the sales rank into List 1 (**L1**). These will represent the x -values. Enter the price into List 2 (**L2**). These will represent the y -values.

Step 2: Perform the regression by pressing **STAT** and selecting the **CALC** option. Scroll down to **LinReg(ax+ b)** and press **ENTER**.

```
LinReg
y=ax+b
a=9.797619048
b=28.78571429
r²=.5312671823
r=.7288807737
```

$$y = 9.8x + 28.79$$

- b. To predict the price of a pair of the 12th bestselling brands of jeans, evaluate the regression equation for $x = 12$.

$$\begin{aligned} y &= 9.8x + 28.79 \\ y &= 9.8(12) + 28.79 \\ y &= 117.6 + 28.79 \\ y &= 146.39 \end{aligned}$$

The price of a pair of the 12th bestselling brands of jeans would be about \$146.39.

- c. Sample answer: No; the correlation between sales rank and price is too weak to make any reasonable prediction. The 12th bestselling brand of jeans could just as easily be a poorly made pair of jeans selling for \$15.

sold at auction each year. Let x be the number of years since 1990.

Years	1990	1995	2000	2005	2010	2015
Paintings Sold	8	5	25	21	9	22

- a. Find the equation for the linear regression line.
 b. How many paintings were sold in 2001?
 c. Is the linear regression equation an accurate model of the data? Explain why or why not.

SOLUTION:

- a. Use a calculator to find an equation for the regression line.

Step 1: Enter the data by pressing **STAT** and selecting the **EDIT** option. Let $x = 0$ for the year 1990.. Enter the years since 1990 into List 1 (**L1**). These will represent the x -values. Enter the number of painting sold into List 2 (**L2**). These will represent the y -values.

Step 2: Perform the regression by pressing **STAT** and selecting the **CALC** option. Scroll down to **LinReg(ax+ b)** and press **ENTER**.

```
LinReg
y=ax+b
a=.4457142857
b=9.428571429
r²=.2349034749
r=.4846684175
```

$$y = 0.446x + 9.43$$

- b. To estimate the number of paintings sold in 2001, evaluate the regression equation for 2001 when $x = 11$.

$$\begin{aligned} y &= 0.446x + 9.43 \\ y &= 0.446(11) + 9.43 \\ y &= 4.906 + 9.43 \\ y &= 14.336 \\ y &\approx 14 \end{aligned}$$

There were about 14 paintings sold in 2001.

14. **ART** A count was kept on the number of paintings

4-6 Regression and Median-Fit Lines

c. No, the correlation coefficient is 0.48, so the linear model is not a good fit for the data. It does not appear from the data that there is a relationship between the number of paintings sold and the year sold.

16. **ANALYZE RELATIONSHIPS** For a class project, the scores that 10 randomly selected students earned on the first 8 tests of the school year are given. Explain how to find a line of best fit. Could it be used to predict the scores of other students? Explain your reasoning.

SOLUTION:

Apply a linear regression model to the data. Use the number of each test as the independent variable and the score on each test as the dependent variable. If there is no correlation, the r value will not be close to 1 or -1 . If this is the case, the line of fit could not be used to predict the scores of the other students.

19. Atma used his graphing calculator to find a best-fit equation for some data. In the equation $w = 0.32r + 4.8$, w represents the weight of a cantaloupe in pounds and r represents its radius in inches. Predict the weight of a cantaloupe with a radius of 5 inches.
- A** 0.625 pound
B 1.6 pounds
C 4.8 pounds
D 6.4 pounds

SOLUTION:

Substitute 5 for r and find w .

$$\begin{aligned}w &= 0.32r + 4.8 \\&= 0.32(5) + 4.8 \\&= 1.6 + 4.8 \\&= 6.4\end{aligned}$$

So, a cantaloupe with a 5 in radius would be 6.4 pounds, the correct choice is D.

20. The table shows the canceled check fee for a bank from 2013 to 2016.

Canceled check fee, years since 2013				
Years	0	1	2	3
Fee (\$)	5	10	10	15

Which of the following best predicts the fee in the year 2020?

- F** \$15
G \$20
H \$26.50
J \$30

SOLUTION:

Enter Years in **L1** and Fee in **L2**.

Then fit a linear regression line to the data.

LinReg
$y = ax + b$
$a = 3$
$b = 5.5$
$r^2 = .9$
$r = .9486832981$

The regression equation is $y = 3x + 5.5$.

Find the fee when $x = 7$.

$y = 3(7) + 5.5 = 21 + 5.5 = 26.5$. The fees will be about \$26.50 in 2020.

So, the correct choice is H.