



Checkpoint 9

Problem 9-50

Scatterplots and Association

Answers to problem 9-50:

a: box plot

b: scatterplot

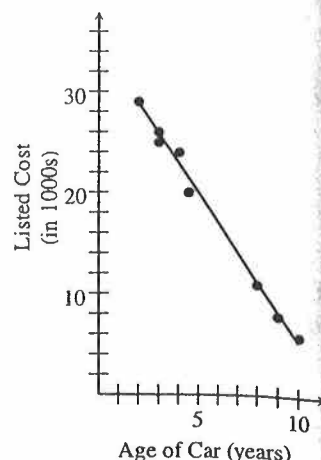
c: See graph at right.

d: strong linear negative association

e: $y = -3x + 35$

f: \$17,000

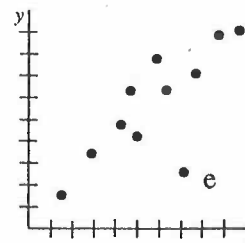
g: A slope of -3 means the car is losing \$3000 in value each year, y-intercept of 35 means the cost when new was \$35,000.



An association (relationship) between two numerical variables on a graph can be described by its form, direction, strength, and outliers. When the association has a linear form, a line a best fit can be drawn and its equation can be used to make predictions about other data.

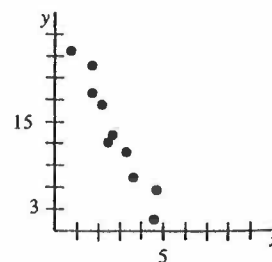
Example 1: Describe the association in the scatterplot at right.

Solution: Looking from left to right, except for point (e), the points are fairly linear and increasing. This is a moderate, positive linear association. Point (e) is an outlier.

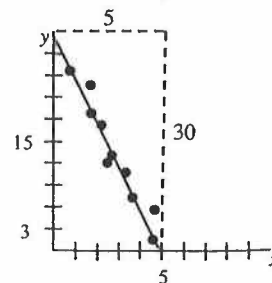


Example 2: For the scatterplot, draw a line of best fit and determine the equation of the line.

Solution: Use a ruler or straightedge to draw the line that approximates the trend of the points. If it is not a perfect line, approximately the same number of points will be above and below the line of best fit.



To determine the equation of the line, draw in a slope triangle and determine the ratio of the vertical side to the horizontal side. In this case it is $\frac{-30}{5} = -6$. Estimate the y-intercept by looking at where the line intersects the y-axis. In this case, it is approximately 30. The equation of any non-vertical line may be written in the form $y = mx + b$ where m is the slope and b is the y-intercept.



$$y = -6x + 30$$

Core Connections, Course 3

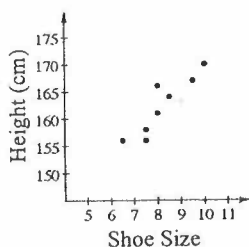
Now we can go back and solve the original problem.

- Since the costs are a single set of data, a box plot is a convenient way to show the distribution.
- Age and cost are two sets of related data so a scatterplot is appropriate.
- See the graph given in the answers.
- Reading from left to right, the scatterplot is decreasing, linear and the points are close to the line of best fit. This is a strong, linear, negative association.
- Looking at the line of best fit, the slope triangle has a ratio of $-\frac{3}{1}$ and the y-intercept is approximately 35. Placing that information into the equation of a line, $y = mx + b$, yields $y = -3x + 35$.
- Substituting $x = 6$ into the equation of part (e) yields $y = -3(6) + 35 = 17$, \$17,000.
- Slope represents the rate of change. A rate of change of -3 means that the value is decreasing by 3 units (in this case each unit is \$1000) per year. The y-intercept represents the value at year zero or a new car.

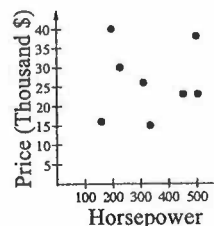
Here are some more to try. *Do the following problems justifying!*

In problems 1 through 4, describe the association.

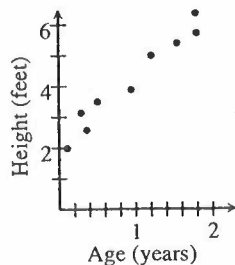
1.



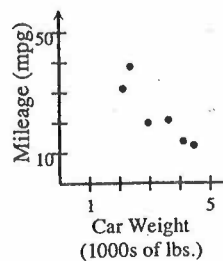
2.



3.



4.



In problems 5 through 8 plot the data, draw a line of best fit, and approximate the equation of the line.

5.	Distance to Airport (mi)	5	10	15	20	25	30
	Cost of Shuttle (\$)	14	17	21	31	33	40

6.	Exercise/Month (hours)	3	6	9	12	15	18
	Rate of Heart Attack/1000	24	21	18	12	6	0

7.	Time Spent Studying (hours)	0	2	2.5	2.8	3	4.5	5
	Score on Test	65	70	70	85	80	95	100

8.	Time Since Purchase (hours)	0	2	4	6	8	10
	Number of Cookies	24	20	14	11	5	0

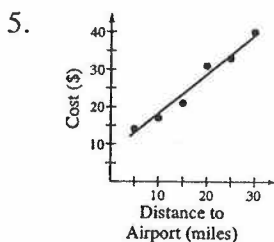
Answers

1. strong positive association

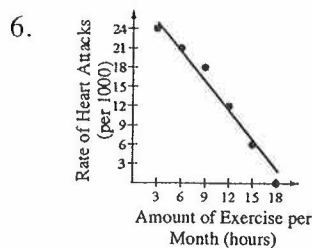
2. no association

3. strong positive association

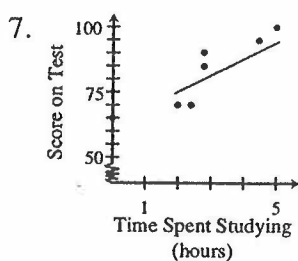
4. strong negative association



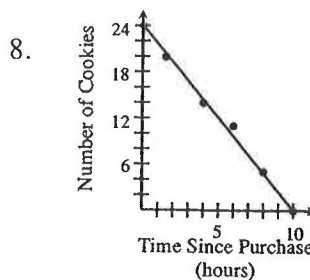
$$\approx y = 1x + 8$$



$$\approx y = -1.5x + 30$$



$$\approx y = 6x + 63$$



$$\approx y = 2.4x + 10$$

