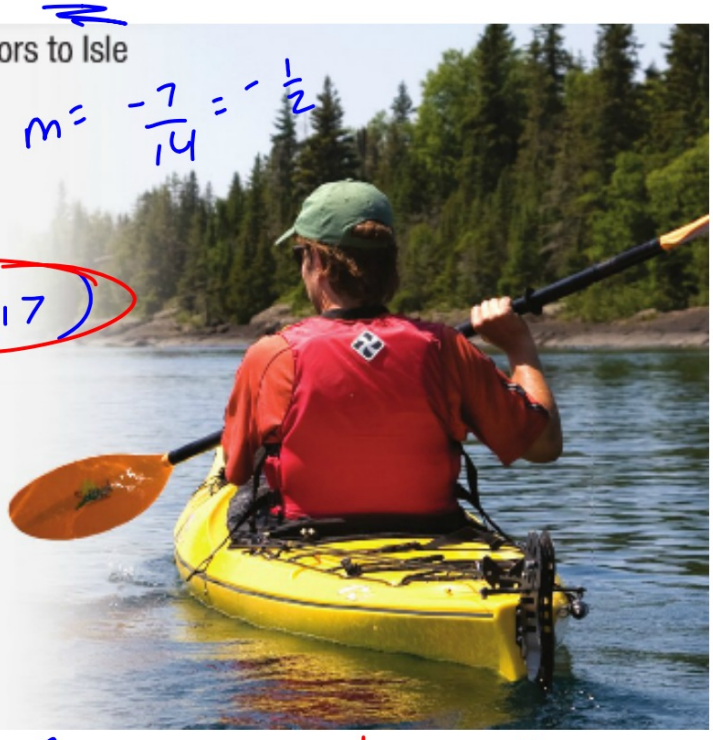
- bivariate data (x, y)
- scatter plot
- dot plot
- correlation (positive, negative, none)
- → line of fit best line
- → prediction equation 2 pt. $y = mx + B$
- regression equation
- correlation coefficient
- graphing calculators (on computer) & scatter plots

The scatter plot shows the number of visitors to Isle Royale National Park in Michigan per year.



Source: National Park Service

$m = \frac{-7}{14} = -\frac{1}{2}$

(2, 24) (16, 17)

$y = mx + B$

$17 = -\frac{1}{2}(16) + B$

$17 = -8 + B$

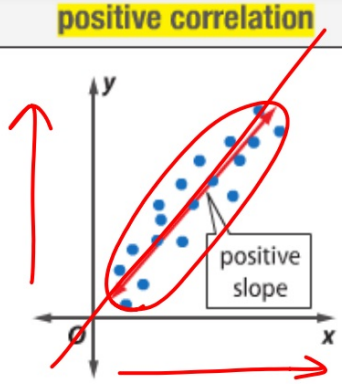
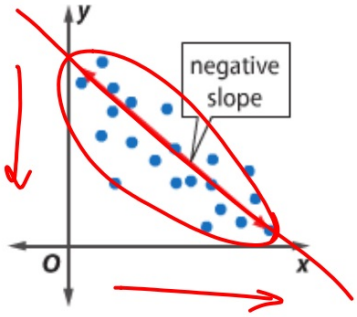
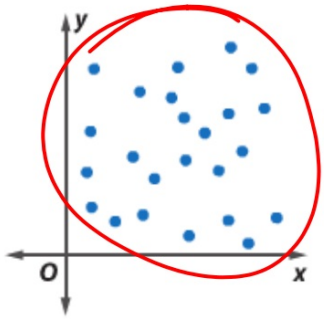
$+8 \quad +8$

$B = 25$

$y = -\frac{1}{2}x + 25$

vis. (1000s) $= -\frac{1}{2}$ (yr since 1990) $+25$

KeyConcept Scatter Plots

positive correlation	negative correlation	no correlation
		
Strong Positive Correlation	**Weak Negative Correlation**	**No Relative Correlation**
The slope of the line is positive and the points are close to the line.	The slope of the line is negative and the points are not close to the line.	There is no obvious pattern of increase or decrease for the given data.

Real-World Example 1 Use a Scatter Plot and Prediction Equation

TECHNOLOGY The table shows the percent of U.S. households with Internet access.

Year	1997	2000	2001	2003	2007
Percent	18.0	41.5	50.4	54.7	61.7

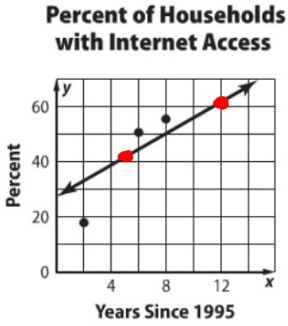
Source: U.S. Census Bureau

a. Make a scatter plot and a line of fit, and describe the correlation. Let x be the number of years since 1995.

b. Use two ordered pairs to write a prediction equation.

c. Predict the percent of households with Internet access in 2020.

d. How accurate does your prediction appear to be?



Consider:

- Scatter in the data...
- Could it continue indefinitely?

1. **HOUSING** The table shows the mean selling price of new, privately-owned, single-family homes for six consecutive years.

Year	0	1	2	3	4	5
Price($1000)	154.5	166.4	181.9	207.0	228.7	273.5

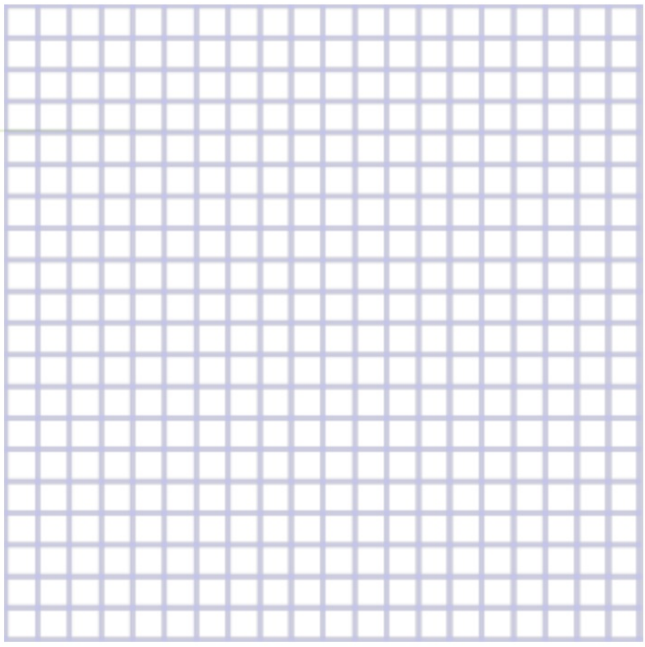
A. Make a scatter plot and a line of fit, and describe the correlation.

B. Write a prediction equation.

C. Predict the selling price of a new home for year 8.

D. How accurate does your prediction appear to be?

p 96 3-6



Graphing calculator (on computer...might need a consult from Ms. Kruse)

2 Lines of Regression Another method for writing a line of fit is to use a line of regression. A **regression line** is determined through complex calculations to ensure that the distance of all data points to the line of fit are at a minimum. Most graphing calculators and spreadsheets can perform these calculations easily.

The **correlation coefficient** r, $-1 \leq r \leq 1$, is a measure that shows how well data are modeled by a linear equation.

- When r is close to -1, the data have a negative correlation.
- When $r = 0$, the data have no correlation.
- When r is close to 1, the data have a positive correlation.

The only way to get a number for correlation!

Basic calculator operation:
Turn on & check settings
Contrast adjustments
Mode key
y=
Statplots
Clear screen
Lists

graphing calculators
note: keystrokes are given on p. 94

Real-World Example 2 Regression Line

The table shows the life expectancy for people born in the United States.

Year of Birth	1980	1983	1990	1995	2000	2006
Life Expectancy (yr)	73.7	74.6	75.4	75.8	76.8	77.7

Source: U.S. CDC

Use a graphing calculator to make a scatter plot of the data. Find an equation for and graph a line of regression. Then use the equation to predict the life expectancy of a person born in 2025.

GuidedPractice

2. MUSIC The table at the right shows the percent of sales that were made in music stores in the United States for the period 1999–2008. Use a graphing calculator to make a scatter plot of the data. Find and graph a line of regression. Then use the function to predict the percent of sales made in a music store in 2018.

Music Store Sales	
Year	Sales (percent)
1999	44.5
2000	42.4
2001	42.5
2002	36.8
2003	33.2
2004	32.5
2005	39.4
2006	35.4
2007	31.1
2008	30.0

Source: Recording Industry Association of America