6.3 Lines of Best Fit

line of best fit: a straight line that best approximates the trend in a scatter plot

regression function: a line or curve of best fit, developed through statistical analysis of data, that best models the data.

interpolation: the process used to estimate a value within the domain of a set of data, based on a trend

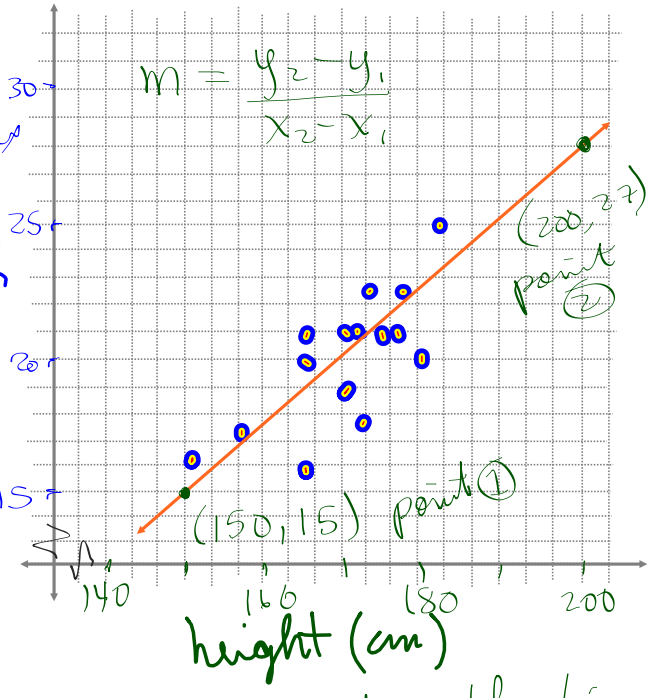
extrapolation: the process used to estimate a value outside the domain of a set of data, based on a trend

Investigate pg. 401

Height (cm)	Hand Span (cm)	Height (cm)	Hand Span (cm)
165.0	20.0	182.5	25.0
172.5	21.1	172.5	23.0
172.5	17.6	180.0	20.2
153.8	16.5	177.5	21.1
157.5	17.5	165.0	20.7
170.0	19.0	165.0	16.0
168.8	20.8	175.0	21.2
177.5	22.5		

A. → hand spans varies directly with height

C. D: $\{x \in \mathbb{R} \mid 153.8 \le x \le 182.5\}$
R: $\{x \in \mathbb{R} \mid 16.5 \le y \le 25\}$



$m = \frac{y_2 - y_1}{x_2 - x_1}$

$m = \frac{27-15}{200-150} = \frac{12}{50} = \frac{6}{25} = 0.24$

$\therefore y = 0.24x + b$ ← use any point on the line to solve for "b"... (150, 15)

$15 = 0.24(150) + b$

$b = -21$

E. $y = 0.24x - 21$

D. → $y = 0.2257x - 18.305$

G.

H.

EXAMPLE 1 Using technology to determine a linear model

The one-hour record is the farthest distance travelled by bicycle in 1 h. The table below shows the world-record distances and the dates they were accomplished.

Continuous data

Year	1996	1998	1999	2002	2003	2004	2007	2008	2009
Distance (km)	78.04	79.14	81.16	82.60	83.72	84.22	86.77	87.12	90.60

International Human Powered Vehicle Association

a) Use technology to create a scatter plot and to determine the equation of the line of best fit.
b) **Interpolate** a possible world-record distance for the year 2006, to the nearest hundredth of a kilometre.
c) Compare your estimate with the actual world-record distance of 85.99 km in 2006.

Your Turn

If there had been a world-record distance record in the year 2000, what would you expect this distance to have been?

EXAMPLE 2 Using linear regression to solve a problem that involves discrete data

Matt buys T-shirts for a company that prints art on T-shirts and then resells them. When buying the T-shirts, the price Matt must pay is related to the size of the order. Five of Matt's past orders are listed in the table below.

Number of Shirts	Cost per Shirt ($)
500	3.25
700	1.95
200	5.20
460	3.51
740	1.69

Matt has misplaced the information from his supplier about price discounts on bulk orders. He would like to get the price per shirt below $1.50 on his next order.

a) Use technology to create a scatter plot and determine an equation for the linear regression function that models the data.
b) What do the slope and y-intercept of the equation of the linear regression function represent in this context?
c) Use the linear regression function to **extrapolate** the size of order necessary to achieve the price of $1.50 per shirt.

Homefun: Pg. 407 # 1-7, 11