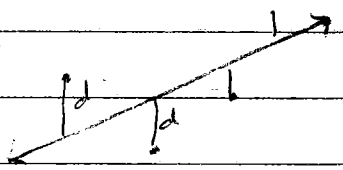
4.2 Linear Regression

pg. 142 3-6, 8

Least squares regression line - line of best fit
Line that fits the data such that the sum of the squares of the vertical distances to the line is as small as possible.



distance below the line are negative
distances above are positive

STAT | CALC

4: LinReg(ax+b) → $y = ax + b$
a = slope
b = y-intercept

8: LinReg(a+bx) → $y = a + bx$
a = y-intercept
b = slope

Text: uses $a+bx$

x = explanatory variable
y = response variable

$\bar{x}$ = average of x values
$\bar{y}$ = average of y values
} $(\bar{x}, \bar{y})$ will be on the line of best fit

STAT | CALC | 2-VAR Stats → give $\bar{x}$ and $\bar{y}$

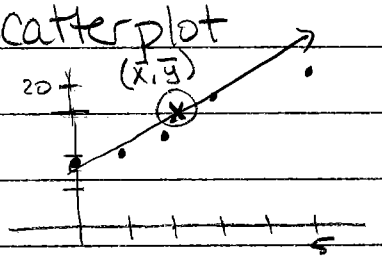
Slope - how many units y changes for each unit of x

Graph - Graph line using Y=

Example 1 #1 pg. 142 Assembly line Defects

x	3	5	0	2	1
y	16	20	9	12	10

(a) scatterplot



(b) $\bar{x} = 2.2$

$\bar{y} = 13.4$

b (slope) $= 2.338$

least squares line:

$y = 2.338x + 8.257$

(c) $(\bar{x}, \bar{y}) = (2.2, 13.4)$ is on the least squares line

8.257 = y-intercept

(d) $x = 4 \Rightarrow y = 2.338(4) + 8.257$

$= 17.609$

When 4 workers are absent the regression line predicts 17.609 defects.