Correlation Regression

NO CLASS on Monday!
Hint: Problem #6 on HW (FPP Sec 9.4)

CS. #17: Correlation

| \( Y \) | \( X \)
|---|---|
| \( Y_1 \) | \( X_1 \)
| \( Y_2 \) | \( X_2 \)
| \( Y_3 \) | \( X_3 \)

\[ X = \text{Ht. of father} \]
\[ Y = \text{Ht. of son} \]

Outcome (Response) Variable (Dependent)

\[ \bar{X} = 68 \text{ in.} \quad \bar{X}_x = 2.7 \text{ in} \quad n = 10 \]
\[ \bar{Y} = 69 \text{ in.} \quad \bar{Y}_y = 2.7 \text{ in} \]

Mean SD \( \Delta y \approx \Delta x \)

\[ \text{pt. of averages } (x, y) \]

Positive association

\[ Y \sim (r < 0) \]

Negative association

\[ Y \sim (r = 0) \]

No assoc.

\[ r = \text{Strength of straight-line (linear) assoc. between } X \& Y \]

\[ \sum \frac{(x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2(n-1)} = r \]

Ave. of products of variables in standard units

\[ \Delta x = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2} \]
\[ \Delta y = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \bar{y})^2} \]
2/6 AM 5 Friday

Basic Facts about $r$

1) $r$ always comes out between $-1$ & $+1$

-1 - 0 - +1

Perfect * Perfect pos. r, linear assc,
reg. linear 
nonlinear

Linear assc,

outlier

healthy, no assc.

2) $r$ can be Footed by 2 things: non linearity & outliers

- outliner can confuse $r$, especially when $n$ is small.

3) units of $r$? ex. $x = \text{income ($)}$, $Y = \text{ht (in.)}$
R is a pure # w/out units.

4) switching $x < Y$ leaves $r$ unchanged

Y stays the same

5) adding a const. to $x$
also leaves $r$ unchanged

6) multiplying $x$ or $Y$ by a pos. const. 
seems to change $r$ visually 
but in fact this also leaves $r$ unchanged.