

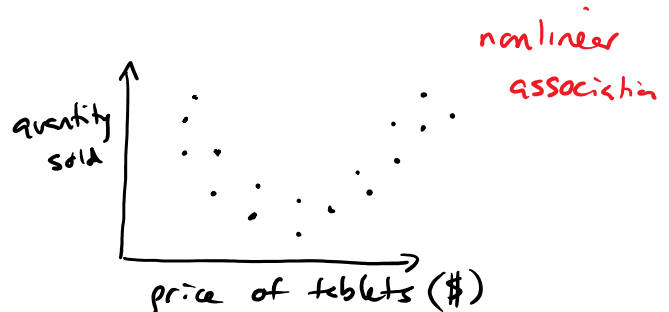
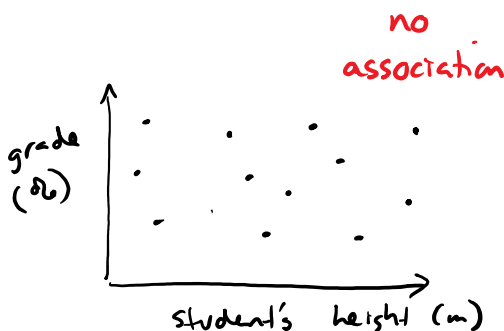
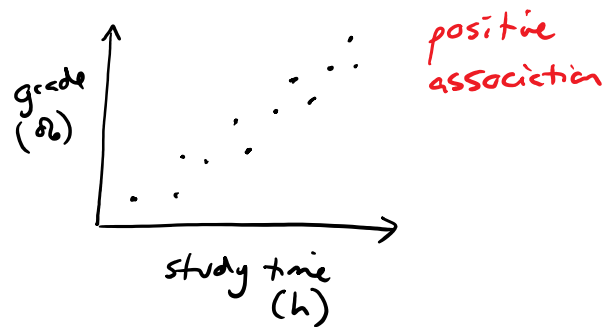
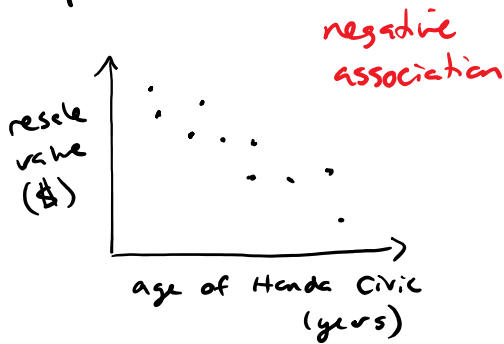
Section 9: Linear Regression

Monday, April 10, 2017 1:19 PM

we will consider bivariate data (x, y)
(two variables)

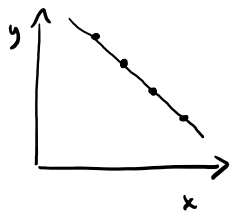
question: Is there a relationship between x and y ?

scatterplots:

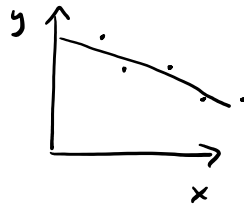


The correlation coefficient r measures how linear a scatterplot is

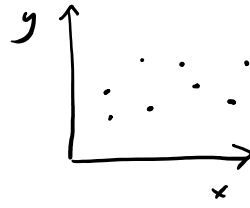
$$-1 \leq r \leq 1$$



$$r = -1$$



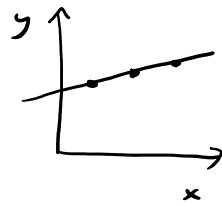
$$r = -0.8$$



$$r \approx 0$$



$$r \approx 0.8$$



$$r = +1$$

Rule of Thumb: $|r| \geq 0.7$ means that the data has a linear association

example: bivariate data set A has $r = 0.73$
 B has $r = -0.85$

which has a stronger linear association?

B since $|r_B| > |r_A|$

note: CORRELATION DOES NOT IMPLY CAUSATION

- just because there's a relationship doesn't mean one causes the other