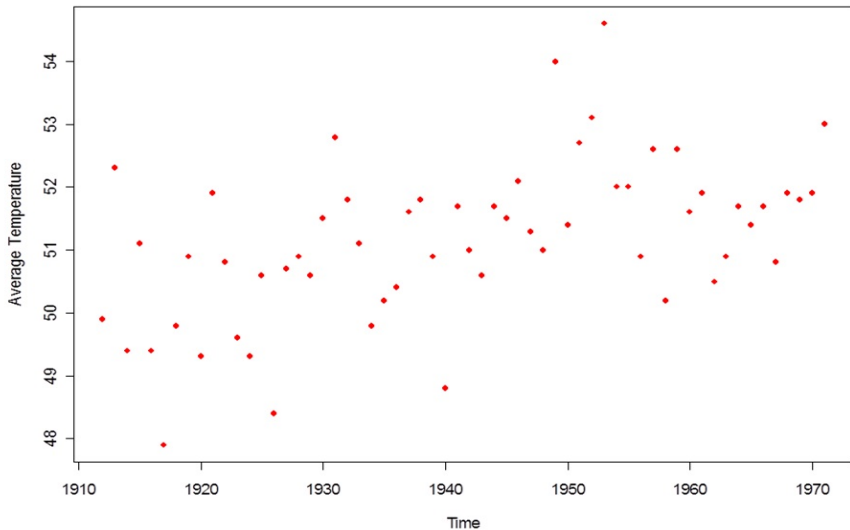


Relationship between two Random Variables

Lecture 14

02/13/2012

Average Temperature through the years



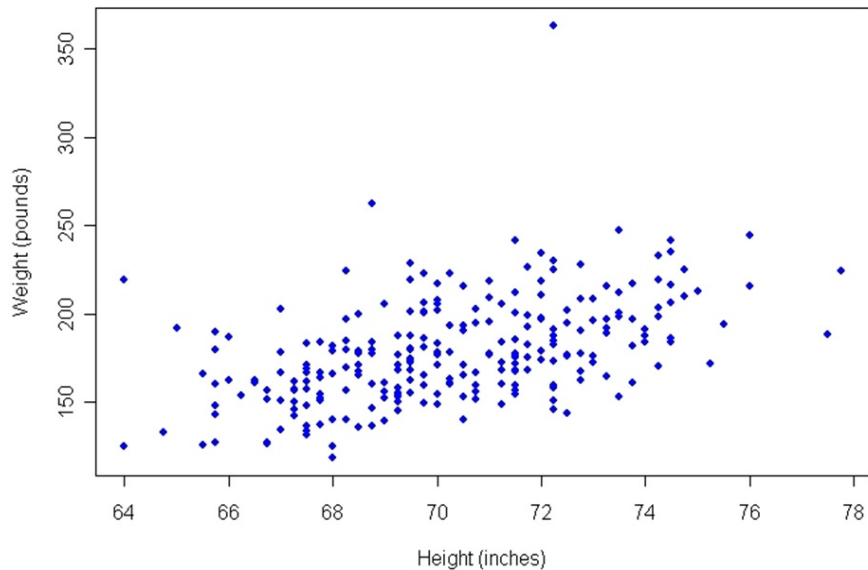
What to look for

- Linear relationship
 - ▶ Positive Association
 - ▶ Negative Association
- Non-linear relationship
- Outliers

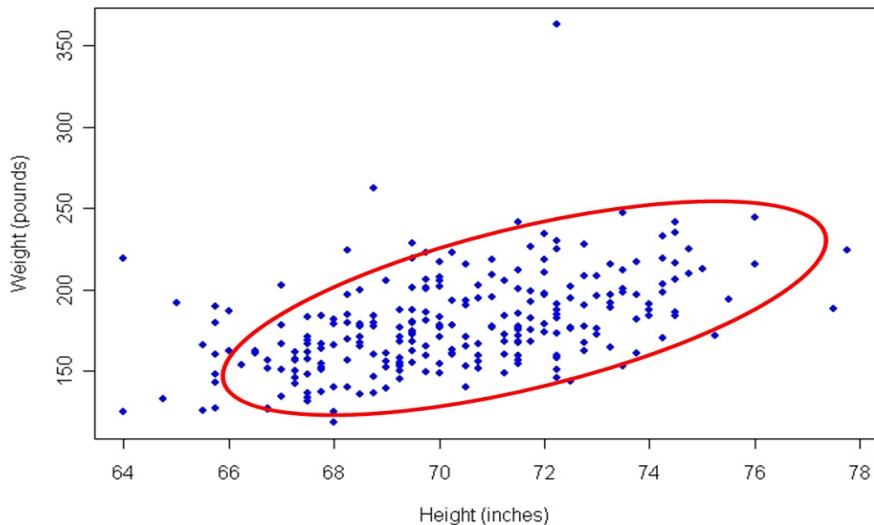
Linear Relationship

- Explanatory variable increases or decreases
- Response increases or decreases proportionally

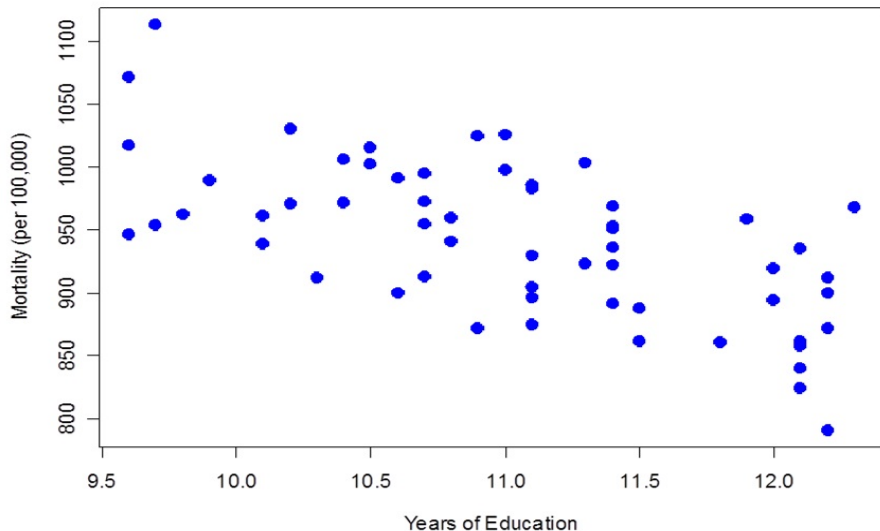
Positive Association



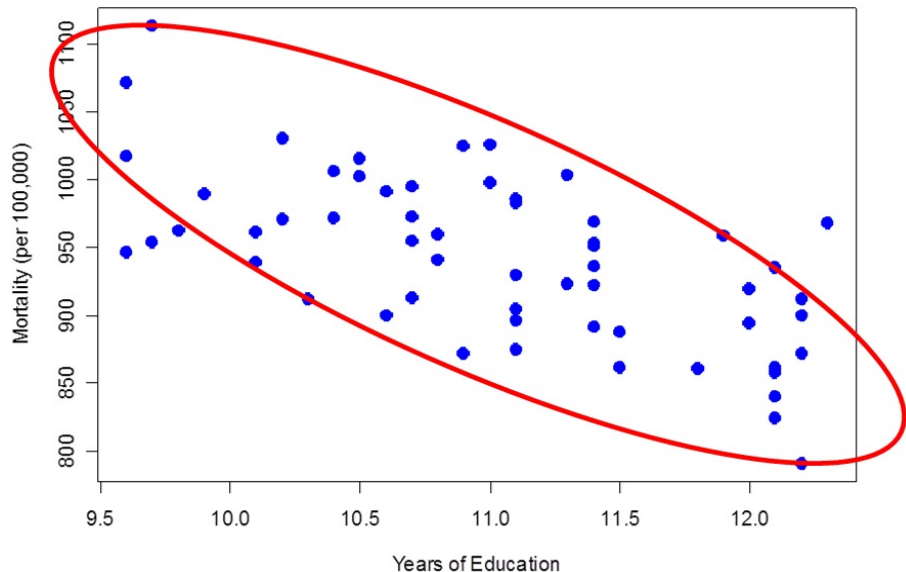
Positive Association



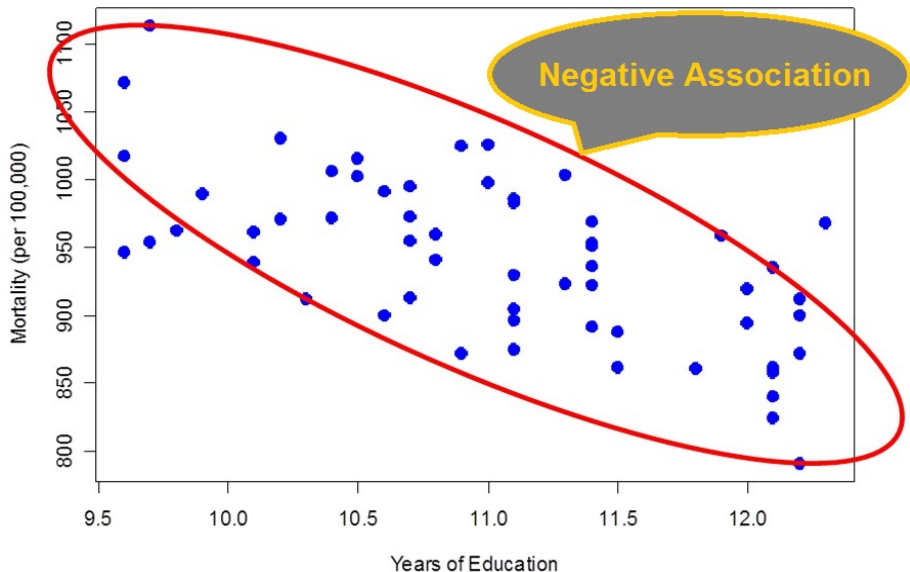
How does education affect health?



How does education affect health?



How does education affect health?



Linear Association

	Explanatory Variable	Response Variable
Positive Association	Increases	Increases
	Decreases	Decreases
Negative Association	Increases	Decreases
	Decreases	Increases

Correlation

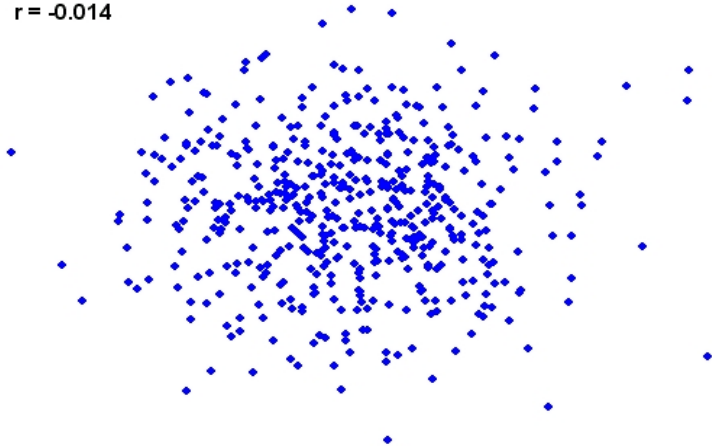
- Do X and Y depend on each other?
- How strong is the relationship?

$$r = \frac{1}{(n-1) s_x s_y} \left[\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y} \right]$$
$$r = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{y_i - \bar{y}}{s_y} \right)$$

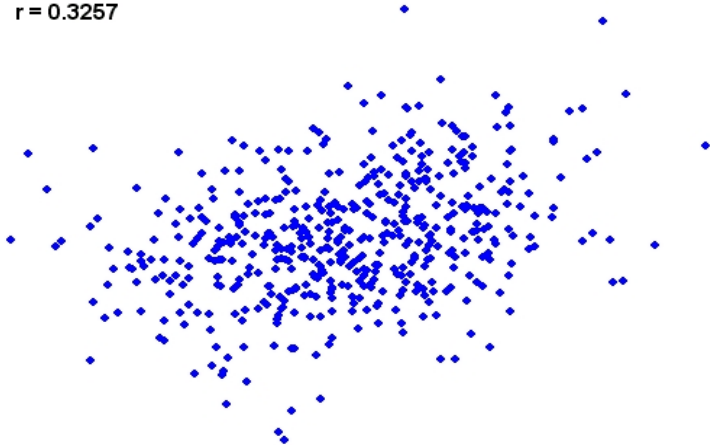
Correlation Coefficient (r)

- Measures strength of association
- r is a number between 1 and -1
- $r = 0$ means no linear association
 - ▶ Independence implies $r = 0$
- $r = 1$ or -1 means exact line
 - ▶ $Y = aX + b$

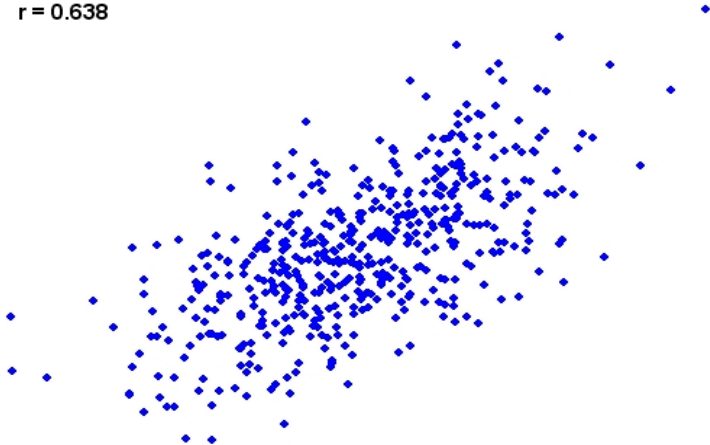
$r = -0.014$



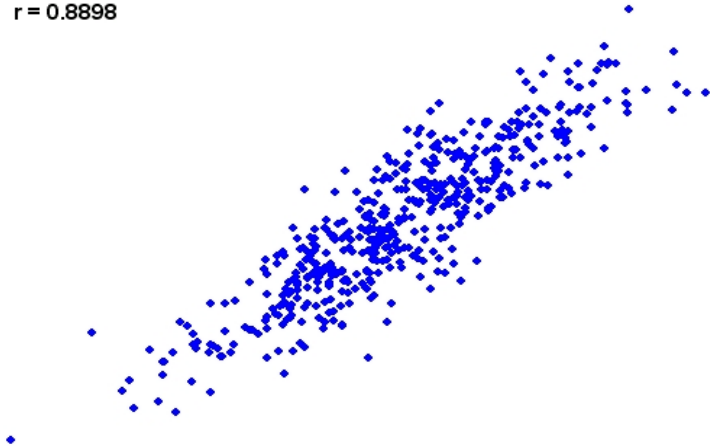
$r = 0.3257$



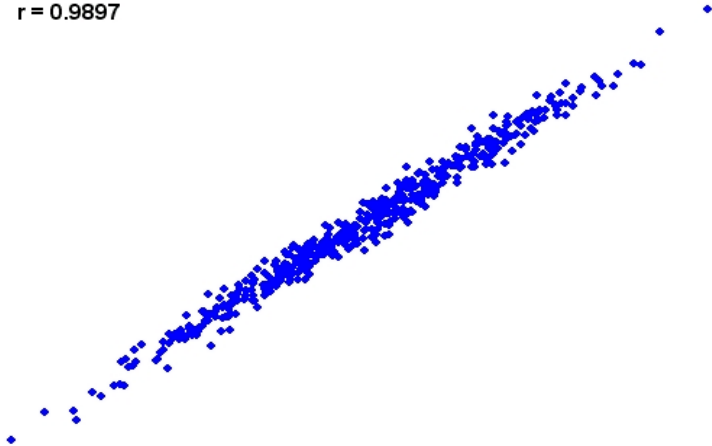
$r = 0.638$



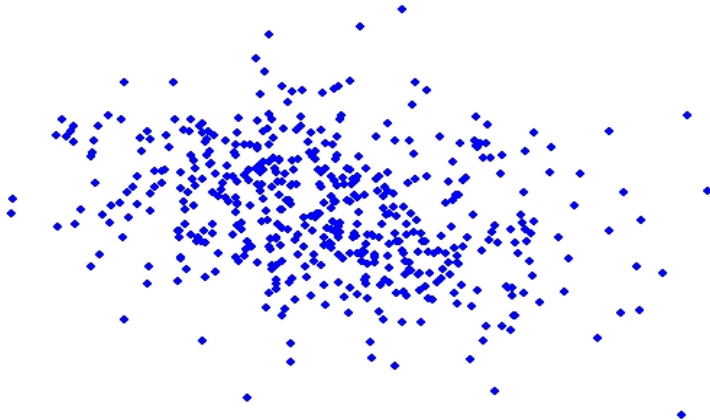
$r = 0.8898$



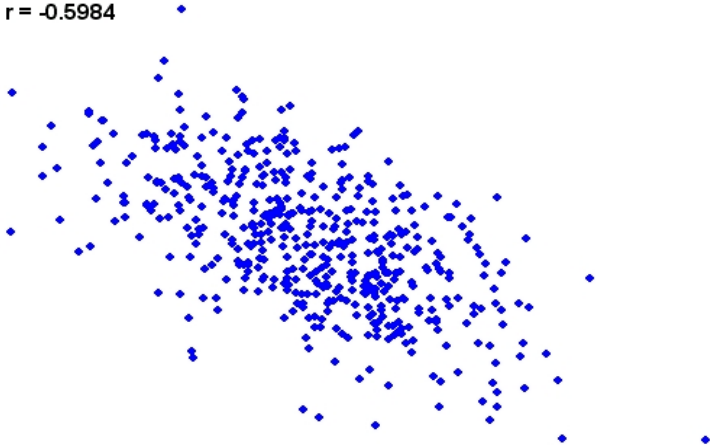
$r = 0.9897$

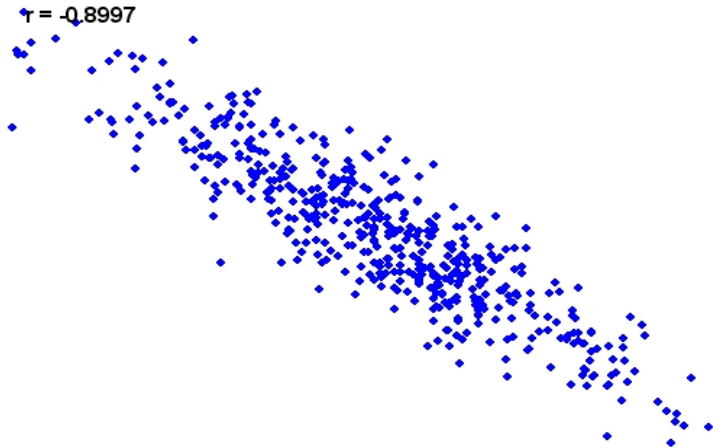


$r = -0.3011$ ♦

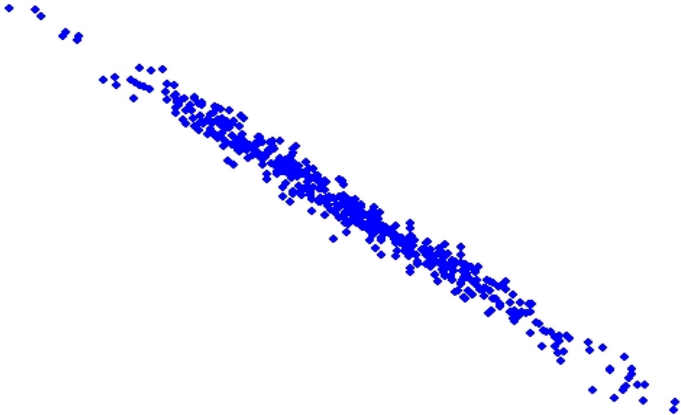


$r = -0.5984$





• $r = -0.9892$



Turn on your clickers!

Are the following two variables positively associated, negatively associated or not associated?

X : the amount of gas you have in your tank (in gallons)

Y : the cost to fill your gas tank.

- (a) The two variables are positively associated.
- (b) The two variables are negatively associated.
- (c) The two variables are not associated.

Turn on your clickers!

Are the following two variables positively associated, negatively associated or not associated?

X : the amount of gas you have in your tank (in gallons)

Y : the cost to fill your gas tank.

- (a) The two variables are positively associated.
- (b) The two variables are negatively associated.
- (c) The two variables are not associated.

Turn on your clickers!

- A recent article in an educational research journal reports a correlation of $+0.8$ between math achievement and overall math aptitude. It also report a correlation of -0.8 between math achievement and a math anxiety test. Which of the following interpretations is the most correct?
 - (a) The correlation of $+0.8$ indicates a stronger relationship that the correlation of -0.8 .
 - (b) The correlation of $+0.8$ is just as strong as the correlation of -0.8 .
 - (c) It is impossible to tell which correlation is stronger.

Turn on your clickers!

- A recent article in an educational research journal reports a correlation of $+0.8$ between math achievement and overall math aptitude. It also report a correlation of -0.8 between math achievement and a math anxiety test. Which of the following interpretations is the most correct?
 - (a) The correlation of $+0.8$ indicates a stronger relationship that the correlation of -0.8 .
 - (b) The correlation of $+0.8$ is just as strong as the correlation of -0.8 .
 - (c) It is impossible to tell which correlation is stronger.

How to calculate r

- 1 Calculate sample statistics: \bar{x} , \bar{y} , s_x , s_y .
- 2 Multiply means: $\bar{x} \times \bar{y}$
- 3 Find the average product: $\sum_{i=1}^n y_i x_i$
- 4 Take the difference: $[\sum_{i=1}^n y_i x_i] - n(\bar{x}\bar{y})$
- 5 Multiply

$$r = \frac{1}{(n-1) s_x s_y} \left[\sum_{i=1}^n y_i x_i - n(\bar{x}\bar{y}) \right]$$