Homework: # 8, 15, 19, 22, 23

8. **Association.**

a) Altitude is the explanatory variable, and temperature is the response variable. As you climb higher, the temperature drops. The association is negative, straight, and strong.

b) At first, it appears that there should be no association between ice cream sales and air conditioner sales. When the lurking variable of temperature is considered, the association becomes more apparent. When the temperature is high, ice cream sales tend to increase. Also, when the temperature is high, air conditioner sales tend to increase. Therefore, there is likely to be an increase in the sales of air conditioners whenever there is an increase in the sales of ice cream. The association is positive, straight, and moderate. Either one of the variables could be used as the explanatory variable.

c) Age is the explanatory variable, and grip strength is the response variable. The association is neither negative nor positive, but is curved, and moderate in strength, due to the variability in grip strength among people in general. The very young would have low grip strength, and grip strength would increase as age increased. After reaching a maximum (at whatever age physical prowess peaks), grip strength would decline again, with the elderly having low grip strengths.

d) Blood alcohol content is the explanatory variable, and reaction time is the response variable. As blood alcohol level increase, so does the time it takes to react to a stimulus. The association is positive, probably curved, and strong. The scatterplot would probably be almost linear for low concentrations of alcohol in the blood, and then begin to rise dramatically, with longer and longer reaction times for each incremental increase in blood alcohol content.

15. Matching.

a) - 0.977 **b)** 0.736 **c)** 0.951 **d)** - 0.021

19. Performance IQ scores vs. brain size.

a) It is appropriate to calculate correlation. The scatterplot of IQ scores *vs.* Brain Sizes is positive, but scattered, with no outliers, and both variables are quantitative.

b) The correlation between IQ scores and Brain Size would be low, perhaps $r = 0.1$.

22. Correlation errors.

a) If the association between GDP and infant mortality is linear, a correlation of –0.772 shows a moderate, negative association. Generally, as GDP increases, infant mortality rate decreases.

b) Continent is a categorical variable. Correlation measures the strength of linear associations between quantitative variables.

23. More correlation errors.

a) Correlation must be between –1 and 1, inclusive. Correlation can never be 1.22.

b) A correlation, no matter how strong, cannot prove a cause-and-effect relationship.

Insane Things That Appear to Be Related

http://www.tylervigen.com/spurious-correlations

Correlation ≠ Causation

- Whenever we have a strong correlation, it is tempting to explain it by imagining that the predictor variable has caused the response to help.
- Scatterplots and correlation coefficients never prove causation.
- A hidden variable that stands behind a relationship and determines it by simultaneously affecting the other two variables is called a lurking variable.

A/c Sales vs. ice cream sales

What Can Go Wrong?

linear & data only! & quantitative

- Don't say "correlation" when you mean "association."
  - More often than not, people say correlation when they mean association.
  - The word "correlation" should be reserved for measuring the strength and direction of the linear relationship between two quantitative variables.

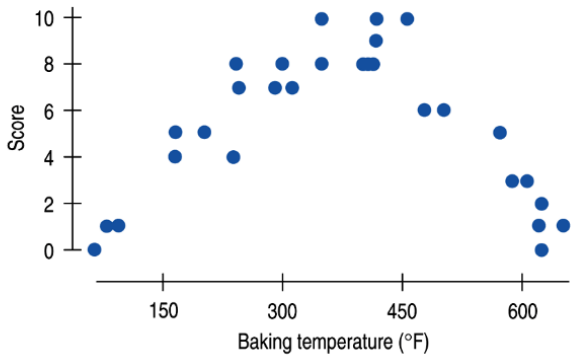
What Can Go Wrong?

- Don't correlate categorical variables.
  - Be sure to check the Quantitative Variables Condition.
- Don't confuse "correlation" with "causation."
  - Scatterplots and correlations never demonstrate causation.
  - These statistical tools can only demonstrate an association between variables.

correlation ≠ causation

What Can Go Wrong? (cont.)

- Be sure the association is linear.
  - There may be a strong association between two variables that have a nonlinear association.

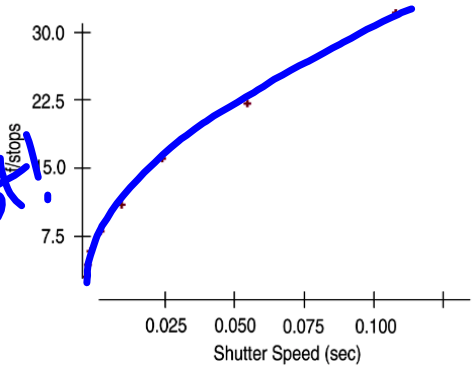


Strong assoc. but weak (no) (linear) correl.

What Can Go Wrong? (cont.)

- Don't assume the relationship is linear just because the correlation coefficient is high.

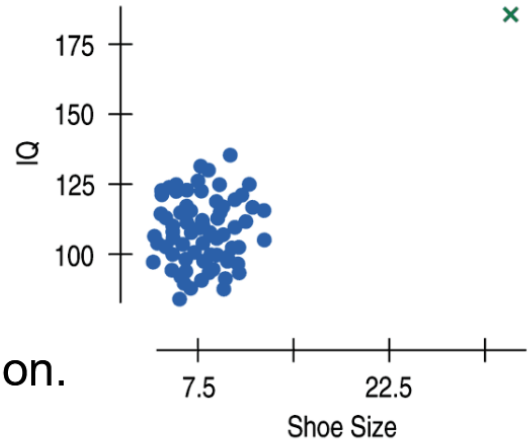
- Here the correlation is 0.979, but the relationship is actually bent.



look at Scatterplot!

What Can Go Wrong? (cont.)

- Beware of outliers.
  - Even a single outlier can dominate the correlation value.
  - Make sure to check the No Outliers Condition.



What have we learned?

+ Explain/describe

- We examine scatterplots for *direction*, *form*, *strength*, and *unusual features*.
- Although not every relationship is linear, when the scatterplot is straight enough, the *correlation coefficient* is a useful numerical summary.
  - The sign of the correlation tells us the direction of the association.
  - The magnitude of the correlation tells us the *strength* of a linear association.
  - Correlation has no units, so shifting or scaling the data, standardizing, or swapping the variables has no effect on the numerical value.

What have we learned? (cont.)

- Doing Statistics right means that we have to *Think* about whether our choice of methods is appropriate.
  - Before finding or talking about a correlation, check the Straight Enough Condition.
    - Watch out for outliers!
- Don't assume that a high correlation or strong association is evidence of a cause-and-effect relationship—beware of lurking variables!

+ Quantitative
+ Outliers

Homework:

pg. 157 #20, 25, 29

~~and~~

~~Multiple Choice Practice Worksheet~~