Scatter Diagram

• Statistics is about … variation.
• Recognize, quantify and try to explain variation.
• Variation in two quantitative variables is displayed in a scatter diagram.

Scatter Diagram

• Numerical variable on the vertical axis, \( y \), is the response variable.
• Numerical variable on the horizontal axis, \( x \), is the explanatory variable.

Scatter Diagram

• Example: Body mass (kg) and Bite force (N) for Canidae.
  – \( y \), Response: Bite force (N)
  – \( x \), Explanatory: Body mass (kg)
  – Cases: 28 species of Canidae.
Positive Association

• Positive Association
  – Above average values of Bite force are associated with above average values of Body mass.
  – Below average values of Bite force are associated with below average values of Body mass.

Scatter Diagram

• Example: Outside temperature and amount of natural gas used.
  – Response: Natural gas used (1000 ft³).
  – Explanatory: Outside temperature (° C).
  – Cases: 26 days.
Negative Association
- Above average values of gas are associated with below average temperatures.
- Below average values of gas are associated with above average temperatures.

Correlation
- Linear Association
  - How closely do the points on the scatter diagram represent a straight line?
  - The correlation coefficient gives the direction of and quantifies the strength of the linear association between two quantitative variables.
Stat 104 – Lecture 8

Correlation

• Standardize y
  \[ z_y = \frac{y - \bar{y}}{s_y} \]

• Standardize x
  \[ z_x = \frac{x - \bar{x}}{s_x} \]

Bite Force vs Body Mass of Canidae

Correlation Coefficient

\[ r = \frac{\sum z_x z_y}{n - 1} \]
\[ r = \frac{\sum (x - \bar{x})(y - \bar{y})}{(n-1)s_x s_y} \]
Correlation Coefficient

• Body mass and Bite force

\[ r = \frac{\sum z_x z_y}{n - 1} = \frac{26.4796}{27}, \]

• \( r = 0.9807 \)

Correlation Coefficient

• There is a strong correlation, linear association, between the body mass and bite force for the various species of *Canidae*.

JMP

• Analyze – Multivariate methods – Multivariate

• Y, Columns
  – Body mass
  – BF ca (Bite force at the canine)
Correlation Properties

- The sign of $r$ indicates the direction of the association.
- The value of $r$ is always between $-1$ and $+1$.
- Correlation has no units.
- Correlation is not affected by changes of center or scale.

Correlation Cautions

- Don’t confuse correlation with causation.
  - There is a strong positive correlation between the number of crimes committed in communities and the number of 2nd graders in those communities.
- Beware of lurking variables.