# Math 207: Statistics

## Chapter 9: More about Correlation
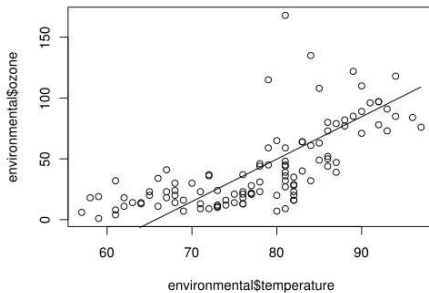


SHENANDOAH® UNIVERSITY

## Features of the Correlation Coefficient

- Given lists $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$, the correlation coefficient $r$:
    - is a measure of linear association between the lists,
    - is a measure of the clustering of the $(x_i, y_i)$ points around a line,
    - is a number between $-1$ and $1$ and
    - is defined by:

    $$r = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{x_i - \text{mean}_x}{\text{SD}_x} \right) \left( \frac{y_i - \text{mean}_y}{\text{SD}_y} \right)$$

    $=$ average of the $x$ and $y$ values measured in standard units

- $r$ is a number with no units (if $x$ is in feet, for example, then $\text{mean}_x$ and $\text{SD}_x$ are also in feet so the units cancel).

- Transformations that don't change the correlation:
    - Reversing the variables: $\text{cor}(x, y) = \text{cor}(y, x)$
    - Shifting one (or both) variables: $\text{cor}(x + A, y) = \text{cor}(y, x)$
    - Scaling one (or both) variables: $\text{cor}(B\,x, y) = \text{cor}(y, x)$

## Transforming Variables
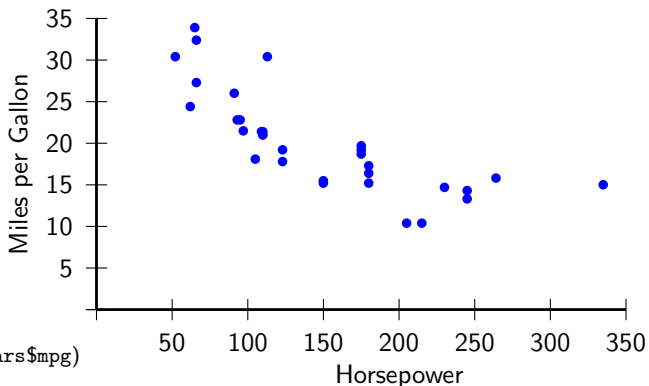


- Compare:
  ```
  > cor(mtcars$hp, mtcars$mpg)
  [1] -0.7761684
  ```
  ```
  > cor(mtcars$hp, mtcars$mpg − 30)
  [1] -0.7761684
  ```
  ```
  > cor(mtcars$hp, 0.425 ∗ mtcars$mpg)    # Convert to km / liter.
  [1] -0.7761684
  ```
  ```
  > cor(mtcars$mpg, mtcars$hp)
  [1] -0.7761684
  ```

# Nonlinearity

- Correlation measures linear association between variables.

- It does not caputure nonlinear association between variables.

- Example: The variables shown below have a strong quadratic association but their correlation is $-0.017$.
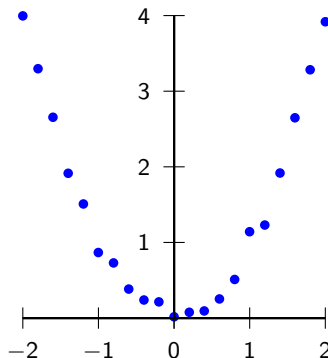
# Association is Not Causation!

- Correlation measures association. But association is not the same as causation.
- Two variables may have a strong positive or negative correlation without having one causing the other to change. The change may be due to some other, confounding factor.
- Example: gold and government bond prices tend to move in opposite directions. There is a strong negative correlation. Other economic factors can lead investors to move funds out of bonds into something that will hold a stable value.



Daily gold prices and 10-year real yield rates