STAT 101  PROF. GOSET  10/11

- Measures of Central Tendency:
  - Arithmetic mean: "average" $\bar{x} = \frac{\sum x_i}{n}$
  - Median = middle value
  - Mode = most common value
  - Skewness to right $\rightarrow$ and to left $\leftarrow$

- Measures of Dispersion (or "spread" of the data):
  - Variance: $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2$ (unbiased with n-1 in denominator)
  - Standard Deviation: $s = \sqrt{s^2}$
  - Coefficient of Variation (C.V.) = $100 \left( \frac{s}{\overline{x}} \right)$

- Sampling Distribution of the Mean:
  - $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$. We estimate $\sigma$ with $s$, taking large $n$.
  - For $n \rightarrow \infty$, $\bar{x} \sim N(\mu, \sigma_{\bar{x}})$
  - (68% within 1 std. error, 95% within 2)

- Least-Squares Regression:
  - Data $X$, estimates $b$, dependent values $Y$
  - $Y = X \cdot b$ $\Rightarrow$ $b = (X'X)^{-1} X'Y$
  - Minimizes $\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$