# Data Science I: Notes

Neal Kuperman

January 8, 2026

**Standard Error**

$$\text{Var}(\hat{\mu}) = \text{SE}(\hat{\mu})^2 = \frac{\sigma^2}{n} \tag{1}$$

where $\sigma$ is the standard deviation of each of the realizations of $y_i$ of Y. Roughly speaking, the standard error tells us the average amount that this estimate $\hat{\mu}$ deviates from the true mean $\mu$.

**Standard Error $\hat{\beta}_0$ and $\hat{\beta}_1$**

$$\text{SE}(\hat{\beta}_0)^2 = \sigma^2 \left( \frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \right) \tag{2}$$

$$\text{SE}(\hat{\beta}_1)^2 = \frac{\sigma^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \tag{3}$$

where $\sigma^2$ is the variance of the error term $\text{Var}(\varepsilon)$. It is assumed that the errors $\varepsilon_i$ are independent and identically distributed (i.i.d.) with mean 0 and variance $\sigma^2$.

**Residual Standard Error**

$$\text{RSE} = \sqrt{\frac{RSS}{n-2}} = \sqrt{\frac{1}{n-2} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{4}$$

where $\hat{y}_i$ is the predicted value of $y_i$ given the independent variables $x_i$. n is the number of observations and p is the number of predictors.