APPLIED STATISTICAL METHODS
COMPARISON OF SIMPLE LINEAR REGRESSION AND ONE-WAY ANOVA

	Simple Linear Regression	One-way Analysis of Variance
Response Variable type	Binary ~~Categorical~~ Numerical	Binary ~~Categorical~~ Numerical
Explanatory Variable type	Binary Categorical Numerical	Binary Categorical Numerical
Model Equation	$Y = \beta_0 + \beta_1 X + \varepsilon$	$Y = \mu + \alpha_k + \varepsilon$
Description of model terms	β_0 = intercept β_1 = slope ε = error term	μ = grand mean α_k = group k effect ε = error term $k = 1, \dots, K$ ← total # of groups
Model Assumptions	• Linearity b/t X + Y • independence of error terms • $\varepsilon \sim N(0, \sigma_\varepsilon)$ — normal, mean=0, constant variance across X • data is representative of population	• independence of error terms • $\varepsilon \sim N(0, \sigma_\varepsilon)$ — normal, mean=0, constant var. for each group ($SD_{max}/SD_{min} \le 2$) • data is representative of population
Estimated model equation	$\hat{y} = \hat{B}_0 + \hat{\beta}_1 x$	$\hat{y} = \hat{\mu} + \hat{\alpha}_k = \bar{y} + (\bar{y}_k - \bar{y})$
Description of estimated model terms	$\hat{\beta}_0$ = least-squares estimate of intercept $\hat{\beta}_1$ = least-squares est. of slope	$\bar{y}$ = mean of all obs. $\bar{y}_k$ = mean of obs. in group k
Equation of residuals	resid = observed − predicted $\hat{\varepsilon}_i = y_i - \hat{y}_i$ $= y_i - (\hat{\beta}_0 - \hat{\beta}_1 x_i)$	resid = observed − predicted $= y_i - \hat{y}_i = y_i - \bar{y}_k$, for points (i) in the k^{th} group

	Simple Linear Regression	One-way Analysis of Variance
Deviation about the overall mean (in symbols)	$y-\bar{y}$	$y-\bar{y}$
Deviation about the model (in symbols)	$y-\hat{y}$	$y-\hat{y}=y-\bar{y}_k$
Deviation of the model about the overall mean (in symbols)	$\hat{y}-\bar{y}$	$\hat{y}-\bar{y}=\bar{y}_k-\bar{y}=\hat{\alpha}_k$
SS(Model)	$\Sigma(\hat{y}-\bar{y})^2$	$\Sigma(\hat{y}-\bar{y})^2=\Sigma(\hat{\alpha}_k)^2$
SS(Error)	$\Sigma(y-\hat{y})^2$	$\Sigma(y-\hat{y})^2=\Sigma(y-\bar{y}_k)^2=\Sigma\hat{\varepsilon}^2$
SS(Total)	$\Sigma(y-\bar{y})^2$	$\Sigma(y-\bar{y})^2=\Sigma y^2-\Sigma\bar{y}^2$
MS(Model)	$\frac{\text{ssmodel}}{1}$	$\frac{\text{ssmodel}}{K-1}$
MS(Error)	$\frac{SSE}{n-2}$	$\frac{SSE}{n-K}$
F-statistic	$\frac{\text{msmodel}}{\text{msE}}\sim F_{1,n-2}$	$\frac{\text{msmodel}}{\text{msE}}\sim F_{K-1,n-K}$
Hypotheses for ANOVA-based test (in symbols)	$H_0: \beta_1=0$ $H_1: \beta_1\neq 0$	$H_0: \alpha_1=\dots=\alpha_k=0$ H_1: at least one $\alpha_k \neq 0$
Hypotheses for ANOVA-based test (in words)	H_0: not a linear relationship between X and Y / model is not useful H_1: model is useful	H_0: no group effects / all group means equal H_1: at least one group is different