0.1 Checking Assumptions in ANOVA and Linear Regression Models

- The assumptions of normality and homogeneity of variance for linear models are not about Y, the dependent variable.
- The distributional assumptions for linear regression and ANOVA are for the distribution of Y—X (Y given X).
- You have to take out the effects of all the Xs before you look at the distribution of Y. As it turns out, the distribution of Y—X is, by definition, the same as the distribution of the residuals. So the easiest way to check the distribution of Y—X is to save your residuals and check their distribution.

What are those distributional assumptions of Y—X?

1. Independence
2. Normality
3. Constant Variance

These assumptions can be checked with a few residual plots a Q-Q plot of the residuals for normality and a scatterplot of Residuals on X or Predicted values of Y to check 1 and 3.