Demo: Evaluating Model Assumptions

In this video, we use the Cleaning data to fit a regression model for Removal and ID using Fit Y by X, and then conduct a residual analysis to evaluate model assumptions.

We select Fit Y by X from the Analyze menu. We'll use Removal as Y, Response, ID as X, Factor, and click OK. Then, we'll select Fit Line from the red triangle to fit the regression model. As we saw earlier, the fitted regression equation is displayed under Linear Fit.

To understand whether a linear model makes sense, we'll analyze the residual plots. To do this, we select Plot Residuals from the red triangle next to Linear Fit. This produces a variety of residual plots.

Let's look at these one at a time. The residuals in the Residual by Predicted plot appear randomly scattered around the center line of zero, with no obvious pattern. The distribution of the residuals in the histogram doesn't look overly normal, but we only have 50 observations. Instead, we look at the Residual Normal Quantile Plot. We don't see any obvious departures from normality in this plot. The points in the Actual by Predicted Plot appear randomly scattered around the line of fit, so this is another good sign.

What about the Residual by Row Plot? Let's assume that our data are time-ordered. There aren't any obvious systematic patterns in the Residual by Row Plot, so it doesn't look like sequential observations are correlated with one another.

We also don't see outliers in any of the plots.

Since our regression assumptions have not been violated, we can proceed to interpret our regression model. We'll talk about the statistical output, and how to use our models to make predictions, in another video.

Statistical Thinking for Industrial Problem Solving

Close