

Perform Regression with Outliers Set by User-defined Criteria in PE

Outliers can be established either via standard deviation criteria or by user definition. This topic describes how to create user-defined outliers.

To create user-defined outliers, follow the steps below:

1. Open the Performance Models window.
2. Right-click the node of the data set on which linear regression will be performed and select **Analyze Model**. The system displays the Analyze Model window with a condition versus age plot of all data points in the data set.
3. In the Graph tab, right-click and select **Edit Outliers**. The system displays the Edit Outliers dialog box. This dialog box allows you to specify the criteria that determine whether a data point is an outlier. You can insert as many records in this dialog box as desired.
4. In the Edit Outliers dialog box, define the conditions that determine what data points are outliers. Functions are selected by clicking the down arrow. Values are typed directly in the provided field. To insert additional records (criteria), right-click and select **Insert**. (You can also delete records by using the **Delete** command.) When the dialog box shows the desired criteria, click **OK**. The system closes the dialog box and colors the data points identified as outliers black.
5. To perform regression analysis without the identified outliers, click the Models tab. The system displays a message saying that the definition of the outliers has changed and gives you the option of continuing or aborting.
6. Click **Yes** to continue.
7. In the right pane of the Models tab, locate the record showing 0-Selected Model.

Note: The system only performs regression on the model identified with a Model Run type of 0-Selected Model. If the model so identified is not the one on which you wish to perform regression analysis, right-click the record that shows the desired model and select **Move for Analysis**. This changes the Model Run type of this model to 0-Selected Model, which allows it to be analyzed.

8. Right-click the row identified in step 7 and select **Define Outliers**. The system updates the table to show the current amount of outliers in the NUM OUTL column. (Note: The value in the NUM SAMPLE column is also decremented by the amount of outliers found.)
9. Right-click the row identified in step 7 and select **Calculate Excluding Outliers**. The system performs the regression analysis on the data set (minus the outliers) utilizing the current model. Click the Graph tab to view the results of this analysis.