

# Inspect and Modify the Prediction Model

## Before you Begin

This lab shows you how to review related datasets produced by the predictive model to determine its accuracy and implement changes to improve results.

## Background

Related datasets are generated when you train and create a predictive model. These related datasets contain details about the model such as prediction rules, accuracy metrics, confusion matrix, and key drivers for prediction depending on the algorithm type. You can examine the rules used by the predictive model to help tune the model to get better results. Use this information to iteratively adjust the model settings to improve accuracy and predict better results.

This is the second lab in *Train and Apply Predictive Models* in Oracle Analytics. Complete the labs in the order listed.

- Create a Random Sample Dataset and Train a Prediction Model
- Inspect and Modify the Prediction Model
- Apply a Predictive Model


## What Do You Need?

- Access to Oracle Analytics

When using Oracle Analytics Desktop, you must install machine learning (DVML) to use Diagnostics Analytics (Explain), Machine Learning Studio, or advanced analytics.

- Access to the `elastic_train_df` data flow
- Access to the `sample_donation_data` dataset

## Review the Model Input List

1. On the Home page, search for your `donations_random_sample` workbook.
2. In the `donations_random_sample` workbook, click the **Actions Menu** , and then select **Open**.
3. In the `donations_random_sample` workbook, right-click **elastic\_model\_1** in the Data panel, and then select **Edit Scenario**.
4. In Edit Scenario - Map Your Data, scroll to **PROJECTID** in the Model Input list.

## Edit Scenario - Map Your Data



Select which Data Set you want to use with the Model

Data Set

For each model input listed on the left, select a corresponding data element from your workbook

| Model Input             | Map To                    |
|-------------------------|---------------------------|
| SECONDARY_FOCUS_AREA    | * SECONDARY_FOCUS_AREA    |
| SECONDARY_FOCUS_SUBJECT | * SECONDARY_FOCUS_SUBJECT |
| TCHR_PREFIX             | * TCHR_PREFIX             |
| TCHR_TEACH_FOR_AMERICA  | * TCHR_TEACH_FOR_AMERICA  |
| NUM_DONORS              | * NUM_DONORS              |
| SCH_LATITUDE            | * SCH_LATITUDE            |
| SCH_LONGITUDE           | * SCH_LONGITUDE           |
| STUDENTS_REACHED        | * STUDENTS_REACHED        |
| TEACHER_ACCTID          | TEACHER_ACCTID            |

\* Required Fields

Done

Cancel

PROJECTID doesn't show up in the list, along with TEACHER\_ACCTID, and TOTAL\_DONATIONS because they weren't used as variables in the model.


5. Click **Cancel**. Click **Go back**

## Edit the Training Model Data Flow

1. On the Home page, enter elastic\_train\_df, and then click **Search**.
2. In the elastic\_train\_df data flow, click the **Actions menu** , and then select **Open**.
3. Click **Add a step** on the line before the Train Numeric Prediction step, and then click **Select Columns**. In the Selected list, click the following:
  - **PROJECTID**
  - **TEACHER\_ACCTID**
  - **SCH\_LATITUDE**
  - **SCH\_LONGITUDE**
  - **STUDENTS\_REACHED**
  - **NUM\_DONORS**

- **FUNDING\_STATUS**
  - **DATE\_COMPLETED**
4. Click **Remove Selected**.
  5. Click the **Save Model** node in the data flow. In **Name**, enter elastic\_model\_2.

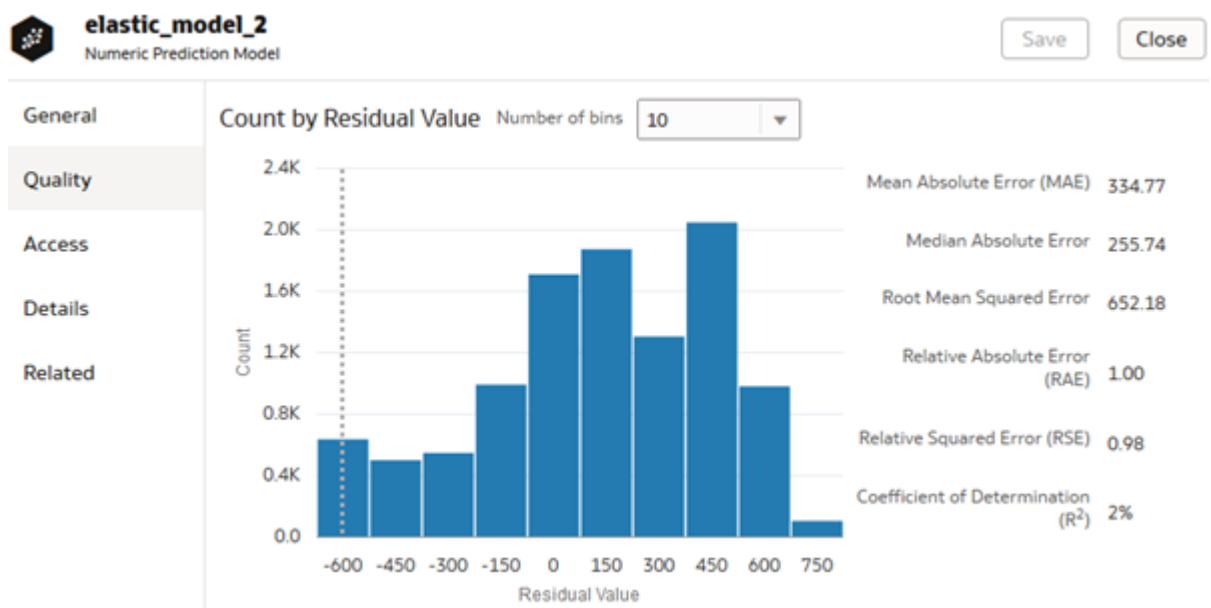


6. Click **Save**. Click **Run Data Flow** .

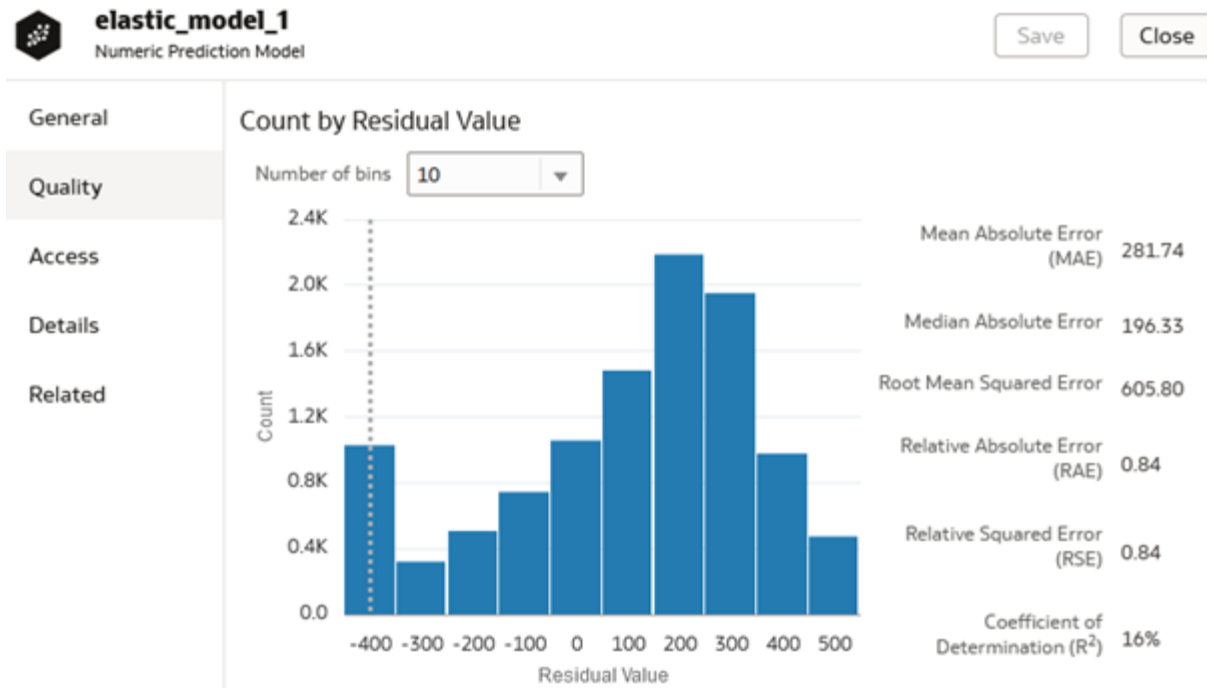
## Review the Revised Machine Learning Model

1. On the Home page, click **Machine Learning**. In the elastic\_model\_2, click the **Actions** menu, and then select **Inspect**.
2. In elastic\_model\_2, click **Quality**.

The residual value distribution in elastic\_model\_2 is slightly different those in elastic\_model\_1.




The Coefficient of Determination value is 2% in elastic\_model\_2. In elastic\_model\_1, the Coefficient of Determination value is 16%. Your values are different due to the random selection of donation data when the data flow is executed.

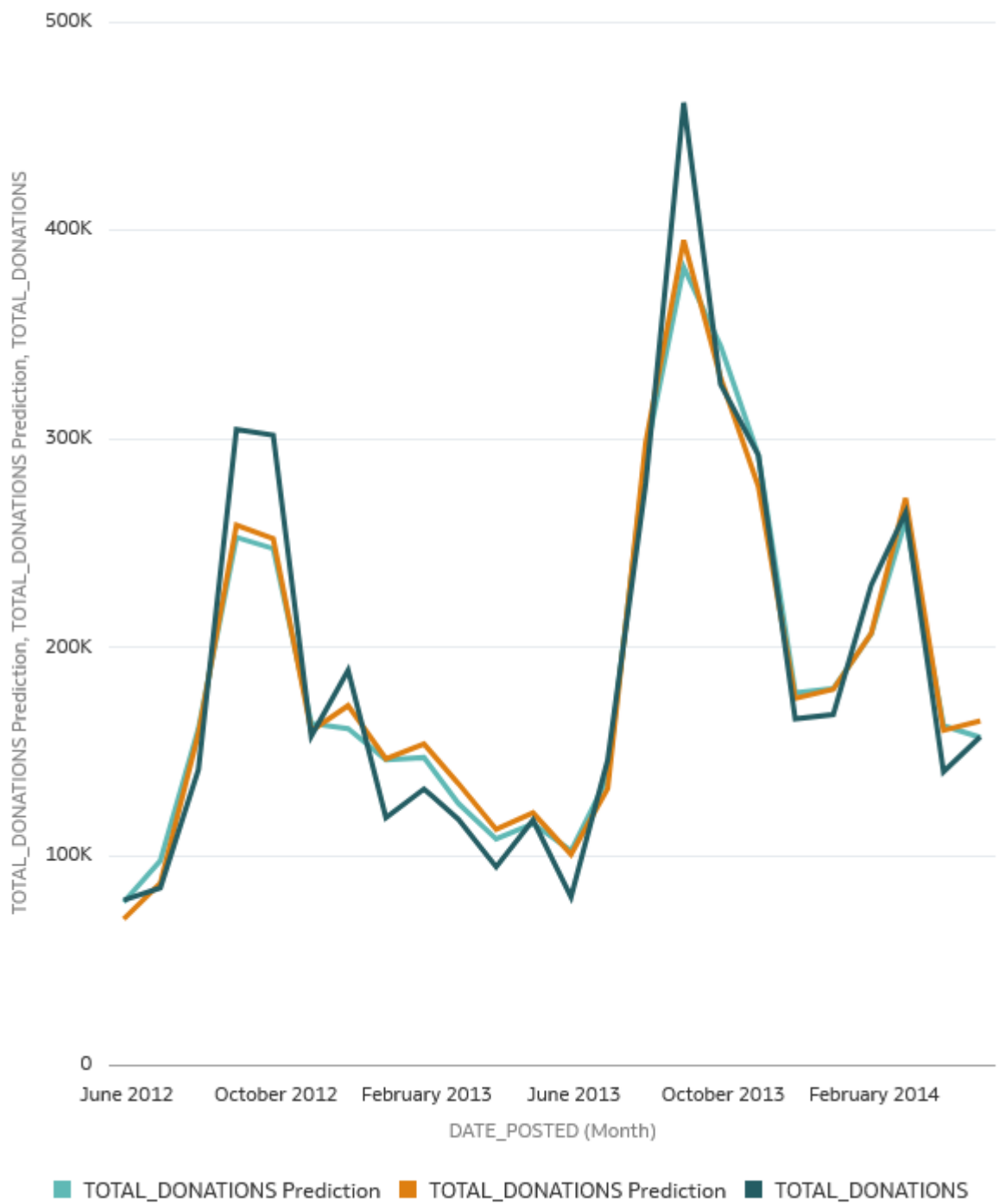


3. Click **Close**.

## Compare Predictive Model Scenarios

1. Click **Home**, click **Workbooks and Reports**, click **donations\_random\_sample**, click the **Actions menu** , and then click **Open**.
2. In the Data panel, click **Create Scenario**. In Create Scenario - Select Model, click **elastic\_model\_2**, and then click **OK**.
3. In the Data panel, expand **elastic\_model\_2**, select **TOTAL\_DONATIONS Prediction**, and drag it in **Values (Y-Axis)** in the Grammar panel.  
The two models produce similar results even though they use different model variables.

TOTAL\_DONATIONS Prediction, TOTAL\_DONATIONS Prediction, TOTAL\_DONATIONS b...



4. Click **Save**.