## Contents

### Chapter 1 / Setting Up Your Project
- Preparing an Input Data Set .............................................. 1
- Creating Your First Project ........................................... 2
- Using the Data Tab ..................................................... 5
- Working with Attributes .............................................. 12
- Adding Events to Your Project ..................................... 15
- Segmenting Project Data ............................................ 18
- Working with Projects ............................................... 22

### Chapter 2 / Working with Pipelines
- Overview of Pipelines .................................................. 31
- Using the Pipelines Tab .............................................. 32
- Modeling Strategy Nodes ............................................ 37
- Generating Forecasts .................................................. 57
- Working with Segmented Pipelines ....................... 58
- Selecting a Champion Pipeline ................................ 64
- Working with Pipeline Templates ............................ 67

### Chapter 3 / Viewing Project Data
- Working with Filters ..................................................... 71
- Viewers for Time Series in Your Project .................... 77

### Chapter 4 / Working with Overrides
- Locking the Data Definition ...................................... 83
- Overrides Overview .................................................. 84
- Working with Overrides ............................................. 87
- Exporting Output Data from a Project .................... 93

### Chapter 5 / Forecasting Concepts
- Understanding Time Series Data .............................. 95
- Understanding Time Intervals ................................ 97
- Demand Classification ............................................... 100
- Understanding Hierarchy Reconciliation .................. 102

### Chapter 6 / Reference
- Assign Variable Roles ............................................... 105
- Time Interval Accumulation Settings ....................... 106
- Project Settings ...................................................... 112
- Settings for Modeling Strategy Nodes .................... 114
- Derived Attributes .................................................. 116
- Model Selection Criteria ......................................... 149
- Weighted Model Comparison Selection Criteria .... 158
- Reserved Names for Variables and Models ............. 162
- Promoting Projects and Templates from Previous Releases to Model Studio 8.5 163
- Troubleshooting ...................................................... 168
Setting Up Your Project

Preparing an Input Data Set ................................................................. 2
  Overview of Data Sources for Forecasting Projects ......................... 2
  Preparing Time Series Data ............................................................... 3
  Preparing Attributes for Your Project .............................................. 4

Creating Your First Project ................................................................. 5
  Creating a Forecasting Project ............................................................ 5
  Assigning Variable Roles ................................................................. 7
  Defining the Hierarchy ...................................................................... 11
  Checking the Project Settings ............................................................ 11

Using the Data Tab .............................................................................. 12
  Left Data Pane .................................................................................. 12
  Middle Data Pane ............................................................................. 14
  Right Data Pane ................................................................................ 14

Working with Attributes ..................................................................... 15
  Importing Attributes to Your Project ............................................... 15
  Selecting Attributes for Filters ........................................................ 16

Adding Events to Your Project ............................................................ 18
  Overview ......................................................................................... 18
  Adding Predefined Events ................................................................ 19
  Importing Custom Events ................................................................ 20
  Changing Event Settings .................................................................. 21
  Creating a Customized List of Events .............................................. 21

Segmenting Project Data ..................................................................... 22
  Defining Project Segments ............................................................... 23

Working with Projects ....................................................................... 24
  Exporting Projects ........................................................................... 24
  Importing Projects .......................................................................... 24
  Sharing Projects .............................................................................. 26
  Working with External Forecast Projects ........................................ 27
  Refreshing Project Data .................................................................... 28
  Running Batch Code ........................................................................ 28
Preparing an Input Data Set

Overview of Data Sources for Forecasting Projects

In SAS Visual Forecasting, you can create a project using one of two types of data sets:

Time series
This is the source time-stamped data that you need to form time series vectors and to determine the best modeling strategies to use and to create forecasts. The input data set requires at least one time variable using a SAS date or datetime format. For more information, see "Preparing Time Series Data" on page 3.

The input data set must also include the dependent variable, which is the numeric variable that you want to forecast. Independent variables that can explain variations in data patterns must also be included in the input data set. However, inclusion of independent variables is optional.

You also want to have several variables that define the individual time series in your data. For example, in an inventory scenario, you might want a variable that represents different regional warehouses and another variable that represents the store locations in those regions. These variables will be assigned as BY variables. They enable you to group observations into time series that can be used for hierarchical forecasting. Even more importantly, these variables will also be used as the default attributes for the project. Any other attributes that you want to include must be imported using a separate attribute data set. For more information, see "Preparing Attributes for Your Project" on page 4.

External forecast
This is an output data set that already contains forecasts along with the historical data. Typically, this data has been exported from another forecasting project as an OUTFOR data set. External forecasts are used primarily to work with any overrides that need to be applied to the forecasts. For more information, see "Working with External Forecast Projects" on page 27.

The data can be imported into a new SAS Visual Forecasting project by loading it into Cloud Analytic Services (CAS). See SAS Cloud Analytic Services: User’s Guide for more information about using CAS.

Sample data sets are provided, which you can use to experiment with SAS Visual Forecasting. You can access this data at the following location:

Preparing Time Series Data

This topic describes the general requirements for getting source data into SAS Visual Forecasting. For further information about time series data, see “Understanding Time Series Data” on page 95.

Requirements for Time Series Data

SAS Visual Forecasting projects require input data that is in time series format. Each observation or row in the table should be using a SAS date or datetime format.

Some time series data is equally spaced. That is, successive observations are a fixed time interval apart, and the data can be described by a single interval, such as daily, every three weeks, or twice a month. You might already have this time series data, or you might have time-stamped data at irregular intervals. You can use the accumulation options in SAS Visual Forecasting to convert the time-stamped data into a time series.

Your time series data must meet the following requirements:

- The data set contains the dependent variable that needs to be modeled and forecast. If multiple variables need to be forecast, you can set that up in separate projects. Multiple projects can be run using the same source data set.
- The data set contains a variable that is in a SAS date or datetime format. When SAS Visual Forecasting creates a project, the data is sorted by the time variable so that the observations are in order according to time.
- The data is loaded into a table in Cloud Analytic Services (CAS). You can use SAS Data Explorer to import data into CAS, or to load a SAS data set into CAS. See SAS Data Explorer: User’s Guide for more information.
- The data must have a time series that includes at least ten observations.

Note: The names of the variables cannot match any of the reserved variable names that are used in the output data set. For more information, see “Reserved Names for Variables and Models” on page 163.

- For general information about working with time series data, see Working with Time Series Data in the SAS/ETS User’s Guide.
- For more information about creating time series data from transactional data, see "The TSMODEL Procedure" in the SAS Visual Forecasting: Programming Guide.

Converting Dates from Character to SAS Date Format

Often a data source might have the date or date and time in character format. For example, consider a variable with the name Sale_Date with this construction:

\[ ddd, yyyy \ mm \ \ dt \ hh:mm \]

- \( ddd \) = 3-letter day of the week
- \( yyyy \) = 4-digit year
Prefered Attributes for Your Project

An important consideration for many forecasting projects involves selecting variables from the data source to be used as attributes. Attributes are used to create filters. Use these filters to query the time series based on selected values for one or more attributes. For more information about filters, see “Working with Filters” on page 71.

Variables in the input data set for a project can be assigned as BY variables. BY variables are the default attributes for the project. Data from the dependent variable is aggregated based on the unique values of the BY variables assigned to the project.
Note: SAS Visual Forecasting requires that the values for the BY variables do not contain any leading or trailing spaces.

If you want to add more variables as attributes, you can import them from an external data source. As you are preparing your data for forecasting projects, determine which variables should be included in the input data set as default attributes, and which variables you want to import as additional attributes.

The following guidelines are required for the external data source for the attributes.

- The data source with the additional attributes must include the same variable names as the existing BY variables in the original data source for the project. For example, if the original BY variables in the time series project are Location and Category, then the attribute table must include columns for Location and Category.

- The imported attributes must contain a row for each unique combination of the default attributes (BY variables) as defined in the input time series or external forecast table. For example, assume the data source for the project has observations with this combination of default attributes:
  - Location = Miami
  - Category = Shoes

  The imported attributes table must have exactly one row with this same combination. When the table is imported, it is checked to verify that all of the unique values for the default attributes match.

If you intend to use segmentation in your project, see “Defining Project Segments” on page 23 for additional considerations for the attributes data set.

See Also

“Working with Attributes” on page 15

Creating Your First Project

Most projects are created using time series input data. You can also create a project using output data from another forecasting project. See “Overview of Data Sources for Forecasting Projects” on page 2 to learn more about the different data sources for forecasting projects.

To open a forecasting project, click on the link in the project from the list on the Model Studio Project page. Follow the steps below to set up a new project.

Creating a Forecasting Project

1. From the Model Studio Project page, click New Project. The New Project dialog box is displayed.
2 For **Name**, give the project a unique name.

3 For **Type**, select **Forecasting**.

4 For **Template**, the template that you choose determines the initial pipeline for the project. For projects using time series data source, the Auto-forecasting template is good for a quick start. See “Pipeline Templates Provided by SAS Visual Forecasting” on page 67 before you select any other template for your project. Some of the templates are not suitable for use when creating a new project.

   If your data source is output data from another forecasting project, the only valid template is *External Forecasts*. If any other template is chosen, the project is automatically created using the *External Forecasts* template. For more information about external forecasts and projects based on output data, see “Working with External Forecast Projects” on page 27.

5 For **Data Source**, click **Browse**. The Browse Data dialog box is displayed.

   A list of data sets is displayed in the left-side Available tab. These are data sets that are available in CAS and ready for use in a Model Studio project.

6 If your source data set is not listed, select the Data Sources tab and search through the libraries for your data set. When you select a data set, the variables are displayed to the right.
Note: If you have any data sets in SAS7BDAT or SASHDAT format, you can right-click the data set and select Load. The data is loaded into CAS memory and is ready to be used in a project.

Note: If you have a local file with a valid time variable, select the Import tab and click Local File to upload it. Then click Import to load the data set into CAS.

For more information about how to use this dialog box to access data, see “Understanding the Available, Data Sources, and Import Tabs” in SAS Data Explorer: User’s Guide.

7 Click OK to load the selected data set and then click Save to create the new project.

The new project that you created is opened with the Data tab selected. Your next steps are to assign the time and dependent variables for your project.

Note: If SAS Visual Forecasting cannot successfully create the project, the project is still added to the Model Studio project list. In this case, you should manually delete the project from the Projects view. Typically, project creation is not successful when a valid time variable cannot be detected in the data source.

See Also

- “Working with Pipeline Templates” on page 67
- “Working with External Forecast Projects” on page 27

Assigning Variable Roles

When you create a project with time series data, the project is opened to the Data tab. In the left pane, the time series data set is selected.

Overview

Use the middle pane to select variables and the right pane to assign roles to a selected variable.

- The time and dependent variables are required and should be assigned before performing any other activities in the project.
- The BY variables are optional but are strongly recommended. These variables are used to set the project up so that you can work with different cross sections of your data based on specific attribute values that you specify.
- You might want to assign some of your variables as independent. Independent variables can be used to model and forecast the dependent variable.
- After making your variable assignments, review the project settings on page 114.

Note: After you run one or more pipelines and open the Overrides tab, the variable assignments and settings are committed and cannot be subsequently updated for
the project. If you need to change any of these settings after they are committed, you can create a new project using the same data source. For more information, see “Locking the Data Definition” on page 83.

Assigning the Time Variable

For new projects, if the system detects only a single time variable candidate, then that variable is already assigned. If the system detects multiple time variable candidates, you must choose one of them as the time variable. The variable must be using either a date or datetime format.

1 Select the variable in the middle pane and, in the right pane, select Time for Role.

![Time Variable Selection](image)

The settings for the time variable are displayed on the right.

2 Review the time variable settings and make sure they are correct for this project. For a complete description of each field, see “Time Variable” on page 106.

If you need to reassign the time role to another variable, you must first remove the assignment for the current time variable.

Assigning the Dependent Variable

Dependent variables are the variables that you want to model and forecast. You must assign one numeric variable to this role. For example, you want to forecast the sales for each product, you assign the Sales variable to the dependent variable role.

1 Select the variable in the middle pane that you want to use to obtain forecasts and, in the right pane, select Dependent for Role.

2 Review the dependent variable settings and make sure they are correct for this project. For a complete description of each field, see “Dependent Variable” on page 109.

If you need to reassign the dependent role to another variable, you must first remove the assignment for the current dependent variable.

Assigning the Default Attributes

BY variables are used to create attributes to uniquely identify each time series. For example, your data source might have variables for Region, Product line, and
Product. By assigning these as BY variables, SAS Visual Forecasting collects the dependent variable into individual time series using each unique combination specified by these attributes.

See “BY Variables” on page 112 for more information about assigning default attributes.

You can also order your BY variables to create a forecast hierarchy. For more information about hierarchical forecast projects, see “Defining the Hierarchy” on page 11.

---

**Note:** For performance reasons, you should not assign variables that have a continuous value as BY variables. See “Display as range” on page 17 for a description of continuous values.

---

When you are certain about which variables you want to use as default attributes, follow these steps.

1. Select one or more variables in the middle pane that you want to use as an attribute.

2. In the right pane, select **BY Variable** for Role.

3. In the right pane, examine order of the BY variables under **Selected hierarchy**. If you use hierarchical forecasting in the project, use ↓ and ↑ to arrange selected variables in the best order for the hierarchy. For more information about hierarchical forecasts, see “Defining the Hierarchy” on page 11.

4. Examine the **Reconciliation level** setting. Select the default reconciliation method for hierarchical modeling nodes to use. This setting can be overridden by settings in the modeling node, if that option is provided.
   
   To use top-down reconciliation, select **Top**. Or you can select a specific BY variable to generate the forecasts and then let the forecasting model reconcile the forecasts for the other BY variables.

   If you make any change to the structure of the hierarchy, either adding or removing BY variables, or changing the order, the **Reconciliation level** is reset to **Top**.

   For more information about reconciliation methods, see “Understanding Hierarchy Reconciliation” on page 102.

---

When you have completed the default attributes assignments, select **Default attributes** in the left pane. The default attributes are listed in the middle pane.

If you import additional attributes to your project, the default attributes are merged with the imported attributes table and the **Default attributes** selection is removed from the left pane. See “Working with Attributes” on page 15 for information about importing attributes.

---

**Assigning Independent Variables**

After you have assigned your time and dependent variables, you might want to assign some independent variables that should be considered for model generation. Independent variables are the explanatory, input, predictor, or causal variables that can be used to model and forecast the dependent variable. You can assign only numeric variables to this role. When creating the system-generated models, SAS Visual Forecasting tries to use the independent variables in the model generation.
Follow these steps to assign independent variables.

1. Select the variable in the middle pane that you want to use to obtain forecasts and, in the right pane, select Independent for Role.

2. Complete the following fields:

   **Hierarchy aggregation**
   Select from the following values:

   **Standard deviation of values**
   Aggregation uses the standard deviation of values for all of the time series in each hierarchy level.

   **Minimum of values**
   Aggregation uses the minimum of the values for all of the time series in each hierarchy level.

   **Maximum of values**
   Aggregation uses the maximum of the values for all of the time series in each hierarchy level.

   **Average of values**
   Aggregation uses the average of the values for all of the time series in each hierarchy level. For example, select this option if your data set contains the price of each product, and you want to know the average price for a product line.

   **Sum of values**
   Aggregation uses the sum of the values for all of the time series in each hierarchy level.

   **Number of non-missing values**
   Aggregation uses the number of nonmissing values for all of the time series in each hierarchy level.

   **Number of missing values**
   Aggregation uses the number of missing values for all of the time series in each hierarchy level.

   **Uncorrected sum of squares**
   Aggregation uses the uncorrected sum of squares for all of the time series in each hierarchy level.

   **Corrected sum of squares**
   Aggregation uses the corrected sum of squares for all of the time series in each hierarchy level.

   **Time interval accumulation**
   Accumulation combines data within the same time interval into a summary value for that time interval. For a complete description of the accumulation options, see “Time Interval Accumulation Settings” on page 112.

   **Missing value interpretation**
   Once the data has been accumulated to form a time series, missing value interpretation is performed. If any time series contain missing values for variables other than the time variable (such as the dependent or independent variables), you can specify how SAS Visual Forecasting should to interpret these missing values. For a complete description of these options, see "Missing interpretation" on page 111.
Usage in system-generated models
Check this option if SAS Visual Forecasting should use this variable to generate forecasts.

See Also
- “Time Variable” on page 106
- “Dependent Variable” on page 109
- “BY Variables” on page 112
- “Defining the Hierarchy” on page 11

Defining the Hierarchy
To forecast hierarchically, you must assign BY variables and use one of the hierarchical models in the pipeline for the project. The ordering of the BY variables constructs the forecasting hierarchy.

An example of a hierarchy is Region > Product Line > Product Name.

Figure 1.1 Example of Hierarchy

![Diagram showing hierarchy: Project_Name, regionName, productLine, productName]

After you set up a project hierarchy, you can specify how SAS Visual Forecasting reconciles forecasts between levels in the hierarchy. Make this specification in the hierarchical forecast model in the pipeline. See "Understanding Hierarchy Reconciliation" on page 102 for more information.

Checking the Project Settings
Before working with any pipelines, verify that the global settings for the project are correct.

1. From the Data tab, click ☰ and select **Project Settings**.
   The Project Settings dialog box is displayed.
Changes to these settings can affect your pipelines and overrides. Make sure these settings are correct so that you do not have to rerun pipelines or resubmit overrides. For more information about these fields, see “Project Settings” on page 114.

See Also

“Project Settings” on page 114

### Using the Data Tab

The Data tab is used to work with the data source for your project. After you create your project, the project is opened to the Data tab so that you can assign roles to the variables that are key to the project. The Data tab displays the data that was loaded into the project CAS library from the original data source.

The left and right panes are collapsible.

### Left Data Pane

The left pane enables you to select different data sources to work with. The toolbar along the top of the pane provides these actions:

- Add or change the data source. If the data source already exists, this action replaces the current data set. For replacement, the data source must match the schema of the original data set that is to be replaced.

  For attributes, this action imports an external table with additional attributes for the project. You must assign BY variables before you can add an attributes table. For more information, see “Working with Attributes” on page 15.

  For events, this action supplements the predefined events. For more information, see “Importing Custom Events” on page 20.

  After any data is changed, any pipelines that have been run become out-of-date and need to be run again.

- Check the data source for updates. You can also perform this action by right-clicking the data source and selecting Refresh all data.

- Open the selected data source in SAS Visual Analytics for further investigation. This action is not available for predefined events.

  You can also perform this action by right-clicking the data source and selecting Explore and visualize.

You can work with these types of data in the left pane:

- **Source data**
  
  The source data for any project must be one of these types.
**Time Series**
Time series are the most common input data source for forecasting projects. Each record has a date or date-time stamp that is used to assess the historical record of the data and to create forecasts. The source table name is listed under Time series in the left Data pane. Selecting the table time series displays the variables from the table in the middle pane. For more information, see "Understanding Time Series Data" on page 95.

**External Forecasts**
Sometimes you need to work only with the output data from another forecasting project. For example, you can generate forecasts for a time series data set and then export the data. Using the output data, you can create multiple projects, experiment with different pipelines and modeling strategies, and work with overrides in the forecast horizon.

The source table name is listed under External forecasts in the left Data pane. Selecting the table time series displays the variables and data from the table in the middle pane.

For more information, see “Working with External Forecast Projects” on page 27.

**Attributes**
Attributes provide a way for you to visualize and work with subsets of project data based on specific attribute values, or filters, that you define. Each project can include these types of attributes.

**Project attributes**
Project attributes are based on variables from your project data. There are two types of project attributes:

- **Default attributes**
  BY variables are the initial, default attributes for the project. After assigning the BY variables, they are listed when you select the Default attributes. For more information, see “Assigning the Default Attributes” on page 8.

  If you import additional attributes to your project, the default attributes are merged with the imported attributes table and the Default attributes selection is removed from the left pane.

- **Imported attributes**
  If you import attributes from another data source, the source table name is listed under Attributes in the left Data pane. Selecting the source table displays the variables from the table in the middle pane along with the default attributes.

  To add more attributes, see "Working with Attributes" on page 15.

**Descriptive statistics**
SAS Visual Forecasting generates summary statistics for the dependent variable as determined by the value of the forecast horizon. You can work with filters based on time series that meet specific values of these statistics.

For more information, see "Descriptive Statistics" on page 153.

**Demand classification attributes**
These attributes are defined by patterns in the time series that can help improve forecast accuracy. Demand classification uses time series information, hierarchical information, and configuration information as input and generates the output with additional segmentation and series statistics information. For more information, see “Demand Classification Attributes” on page 149.
Forecast attributes
You can select filters of time series in the Forecast viewer based on forecast attributes. These attributes are generated by each modeling strategy in the pipeline and are not available in the Time series viewer.

For more information, see “Forecast Attributes” on page 154.

Events
An event is an incident that disrupts the normal flow of a process that generates the time series. Examples of events are holidays, retail promotions, and natural disasters. You can create events from a list of predefined events and you can import events from an external data source with custom event definitions. Predefined events are available by selecting Predefined events under Events.

For more information, see “Adding Events to Your Project” on page 18.

Middle Data Pane
The middle pane shows the content of the data source that is selected in the left pane. The toolbar along the top of the pane provides the following actions.

Note: These actions are not available if an External Forecast data set is selected in the left pane.

Filter
For large lists of data, you can type into Filter to limit the table to a subset of the content. The filter matches the string that you enter to any string in the table. Filter is not case-sensitive.

Filter is disabled if (View table) is selected.

Show the variables in the selected data source, their role assignments, and other properties. This is the default view for time series projects.

Some of the columns in the table can be sorted when you click on them.

Show the data source table, with the variables as column headings and the observations in each row.

This is the only view available for external forecasts unless you import attributes to the project. For more information, see “Working with External Forecast Projects” on page 27.

The columns in the table can be sorted when you click on them.

You can arrange and remove columns in the table, as described in “Working with Tables in Model Studio” in SAS Visual Forecasting: Overview.

Right Data Pane
The right pane shows the properties of the current selection from the other panes in the Data tab. Use this pane to assign properties to variables in the project data set, imported attributes, or events. After you assign the time and dependent variables,
these roles cannot be assigned to another variable unless you first remove the original assignment.

- If a data set is selected in the left pane, the right pane shows the properties of that data set.
- If a variable is selected in the middle pane, the right pane shows the properties of that variable. If you select multiple variables, the right pane shows blank values if the variables do not share the same property values.

The right pane is not available for external forecasts unless you import attributes to the project. For more information, see “Working with External Forecast Projects” on page 27.

Working with Attributes

Attributes provide a way for you to visualize and work with subsets of project data based on specific values of those attributes, or filters.

Importing Attributes to Your Project

BY variables are the initial, default attributes for the project. You can add additional attributes to a project from another data source. See “Preparing Attributes for Your Project” on page 4 for more information.

Before importing attributes to the project, make sure you have assigned the BY variables as described in “Assigning Variable Roles” on page 7. These are the default attributes for the project. The default attributes should have discrete values, such as character data or numeric data with a limited set of values.

Follow these steps to add attributes from another data source to your project.

1. From the Data tab, click [ ], and select Attributes.
   The Browse Data dialog box is displayed.

2. Locate the data source with the additional variables to use as attributes with this project. Select the data source and click OK to load it.
   The imported attributes must contain a row for each unique combination of the default attributes (BY variables) as defined in the input time series or external forecast table. When the table is imported, each row is checked to verify that the combination of default attribute values is unique. For large data sets, this process can take a long time.

Whether an attributes table is imported determines how the project attributes are selected in the project. The following figure shows how the Data tab is changed after importing attributes. On the left, the assigned BY variables are listed as Default attributes when attributes are not imported. The image on the right shows an attributes table with the name ATTRIBUTES_SEG that has been imported. Selecting this table shows all of the default and imported attributes.
Selecting Attributes for Filters

Use discretion in the number of attributes to use as filters. For performance reasons, it is better not to have more than 20 attributes as filters in a project. When selecting which attributes you want to use as filters, it is helpful to see how attribute values can be specified as filters.

For imported attributes, follow these steps for selecting the best attributes to use as filters by exploring a time series in a pipeline.

1. In the Data tab, select the table listed under Attributes in the left pane.
2. Select an attribute in the middle pane and examine these fields in the right pane.

**Activate attribute**

This setting enables you to use the attributes to create filters. Filters are used to subset the data and create overrides for forecasted values based on the filter settings. Filters can also be used to display filtered views of the historical data in the project. Activate any attributes that you want to use for creating filters.

**Activate attribute** is available only for imported attributes. All derived attributes are activated.

**Display attribute by default**

You can work with time series based specific values that you select from attributes in your project. Attributes are listed in a filter pane that is used to select attribute values. By default, all active attributes are displayed in the filter pane. You can choose to hide some attributes by default if you need to reduce the visual clutter of the filter pane. The filter pane enables you to hide or show attributes as you are working with them.
Display as range

Turn on this setting for any numeric filters that have a continuous value. Continuous variables can have a limitless number of values within a range. For example, if you are working with a Weight attribute that could be any decimal range between 20 and 200 pounds, you would turn this setting on.

With this setting turned on, you have these tools to select the lowest and highest values of the range for the filter.

- For numeric attributes with formatted date or datetime values, you can choose the earliest and latest date in the valid range using a selector. In this image, it shows that dates outside of the range cannot be selected:

  ![Calendar](image)

- For all other numeric attributes with a continuous range, you can click and enter the low and high values of the range using a slider:

  ![Slider](image)

This is explained in more detailed in “Creating Filters” on page 72.

**Note:** Display as range is enabled only for imported attributes. This setting is fixed for derived attributes and cannot be changed. Derived attributes include Descriptive statistics, Demand classification attributes, and Forecast attributes.

**Note:** If any filters are created using imported attributes, subsequent changes to Activate attribute and Display as Range can invalidate the filters. For more information, see “Disabled Filters” on page 75.

3 Select some variables that are good candidates for filtering the project data. In the right pane, select Yes for Activate attribute.

4 For any numeric variables with the Activate attribute setting turned on, examine the Display as Range setting and make any necessary changes.

5 Select the Pipelines tab.

For new projects, the default pipeline is loaded.
6 Right-click the Data node at the top of the pipeline and select Run.

7 When the Data node has finished running, right-click the Data node again and select Time series viewer.
   This opens the plot of the project’s historical data along with a list of attributes that you can use to filter the data in the plot. See “Filters” on page 77 for more information.

8 Explore different combinations of filter values and observe how the aggregation plot changes to reflect the data that you have selected.
   After you run pipelines on the project and a champion pipeline is selected, you can use any combination of these filter values to specify overrides on forecast values. However, after you start working with overrides, your selection of which attributes to use as filters is limited. You can still add attributes to the list of filters, but you cannot remove them.

9 Return to the Data tab and select the attributes table again.
   Make any necessary changes to your filter selections. Here are some considerations for setting Display as Range for different data types.
   - You cannot use Display as Range for attributes with a character data type.
   - Attributes with integer values with a very limited range are not good candidates for Display as Range. The following example shows a numeric attribute, type, with only three values. Using a slider to specify a range of values for this attribute is not useful.

   ![Type Attribute Example](image)

   - Attributes with decimal values or a very large integer range are good candidates for Display as Range.
   - Use Display as Range for attributes with a SAS date or datetime type.

10 Review your settings carefully. After your assignments are locked down, the Activate attribute setting cannot be turned off for an attribute. For more information, see “Locking the Data Definition” on page 83.

Adding Events to Your Project

Overview

The most likely problem with your initial forecast model is that the model ignores a known event. You can improve the accuracy of your model by adding events. When
you add events in a project, SAS Visual Forecasting must reforecast the project in order for the models to reflect these changes.

An event is an incident that disrupts the normal flow of a process that generates the time series. Examples of events are holidays, retail promotions, and natural disasters. Defining an event enables you to model the effect that special events have on the dependent time series. When you apply an event to a time series, SAS Visual Forecasting creates an indicator variable. This variable indicates the occurrence of the event at the time period for which the event is defined. This indicator variable is used as a regressor variable for time series modeling and forecasting.

For example, retail sales data follows a fairly steady pattern depending only on the day of the week. However, when you include events such as a one-time New Year's Day promotion, the forecasting model can predict a temporary increase in sales and then return to normal sales levels after the event. Some yearly events that occur on a fixed date, such as New Year's Day, can be modeled as part of the regular seasonal model. Seasonal events that are most effectively modeled as regressors are those that occur on a different date each year, such as Thanksgiving.

You can use SAS Visual Forecasting to predict the effect of a future event. Before the forecasting model can predict the effects of a future event, your data must include the effects of past occurrences of the event. For example, after reviewing your December sales data in the United States, you decide to create an event for next Christmas. SAS Visual Forecasting must have data that explains how retail sales were impacted by previous Christmas events. Without this data, SAS Visual Forecasting cannot predict the effect of the next Christmas event on retail sales.

SAS Visual Forecasting includes a list of defined events. However, you can provide a customized list of events. This is described in “Creating a Customized List of Events” on page 21.

Note:

- You cannot add events to an external forecast project.
- The Naive Model Forecasting modeling strategy does not support events.

---

**Adding Predefined Events**

SAS Visual Forecasting is shipped with a list of predefined events that you can add to any project. You assign events at the project level.

Follow these steps to select from this list which events to add to your project.

1. On the Data tab for a project, select **Predefined events** under Events in the left pane.
   
   An empty table is shown in the middle pane.

2. Click **Add Predefined Events** at the top of the middle pane. If this button is not visible, click ‥ and select **Add Predefined Events**.
   
   The Add Predefined Events dialog box is displayed.

3. Select events from the list and click **Add**.
The events are added to the table in the middle pane. The Event Name is a system name, for example, MOTHERS is the name for Mothers Day.

After a predefined event is added to the project, it cannot be removed. If you do not want to use it in the project, you can change the event usage to Do not use. For any predefined events that you add to the project, the default is Use if significant.

To change settings for predefined events, follow the steps in “Changing Event Settings” on page 21. You can change these settings for specific models in a pipeline. For more information, see “Changing Event Usage in Modeling Strategies” on page 43.

Note: If you import custom events that include an event with the same name as a predefined event, you cannot add the predefined event.

---

Importing Custom Events

If you have your own set of custom events to add to the project, you can add an event definition table to your project. Follow these steps to import your custom events.

If a custom event definition table is already added, these steps are used to replace it. The data set that is used to replace the custom events must include the same event definitions. More events can be added in the replacement, but existing custom events cannot be removed.

Note: If the data set with custom events includes an event with the same name as one that you have already added from predefined events, the data set cannot be added.

You can also add custom events from a Forecast Server project. You will need to import the data set that contains the event definitions. For more information, see Save Events to a SAS Data Set in the SAS Forecast Studio: User’s Guide.

1  In the left pane of the Data tab, click and select Events.

   The Browse Data dialog box is displayed.

2  Locate the data source with the custom events to use with this project. Select the data source and click OK to load it.

   When the data source is successfully loaded, the new events are listed in the middle pane.

By default, each custom defined event that you add is set to Do not use. To change settings for custom events, follow the steps in “Changing Event Settings” on page 21. You can change these settings for specific models in a pipeline. For more information, see “Changing Event Usage in Modeling Strategies” on page 43.

You can click to replace the custom events data set. The new data set must contain at least the same event definitions as the original custom events data set. More event definitions can be added. If you update the original event data set, use instead to update the events.
Changing Event Settings

In the left pane of the Data tab, selecting a data source under Events shows the list of events. The data source could be the predefined events provided by SAS Visual Forecasting or a data set with custom event definitions that was imported for the project.

In the middle pane, each event for the data source is listed along with the event settings, under the column Usage in System Models. These settings are the default settings used by the models in each pipeline, unless you change the event setting for any individual models within the pipeline, as described in “Changing Event Usage in Modeling Strategies” on page 43.

Follow these steps to change the default event settings.

1. With the event data source selected in the left pane, select an event in the middle pane. You can select multiple events if you want to assign the same setting to more than one event.

2. In the right pane, click Usage in system-generated models to view the settings that you can apply to the event. If the right pane is not visible, click << to display it.

3. Select one of the following settings:
   - Do not use
     Do not include the event in the model.
   - Try to use
     Include the event in the model as long as the parameters of the events are significant and the increment of the value of the criterion exceeds a specified threshold.
   - Use if significant
     Include the event in the model as long as the parameters of the events are significant.
   - Force use
     Include the event in the model as long as the model does not fail to be diagnosed.

The new settings are displayed in the middle pane under Usage in System Models. If you have run any pipelines, they become out-of-date and must be run again.

Creating a Customized List of Events

SAS Visual Forecasting includes a list of predefined events. However, you might need to add to this list of events or create a customized list of events for your company. To define an event, you must create the event using the HPFEVENTS procedure.

See Also

“Creating a Customized List of Events” on page 21
To create a customized list of events:

1  From the SAS Drive menu bar, click ➞ and select **Develop SAS Code**.

2  Create an events data set using HPFEVENTS procedure. Here is an example.

   Example Code 1.1  HPFEVENTS Procedure

   ```sas
   proc hpfevents data=sashelp.air lead=12; /* 1 */
      id date interval=month;
      eventdef fallpromo= '01oct2018'd; /* 2 */
      eventdef back2school= '07aug2018'd to '30aug2018'd by day;
      eventdef superbowl = '06FEB2011'D '05FEB2012'D '03FEB2013'D
                           '02FEB2014'D '01FEB2015'D '07FEB2016'D
                           '05FEB2017'D '04FEB2018'D '03FEB2019'D
                           '02FEB2020'D;
      eventdata out=eventds (label='Event List'); /* 3 */
   run;
   ```

   **a**  The DATA option requires an input data set. For this example, **sashelp.air** is used because it is already sorted. The goal of this step is to get the output data set with event definitions. Any sorted data set is sufficient.

   **b**  Use the EVENTDEF statement for each custom event definition. Each definition is in the form of **SAS-variable-name=timing-value-list**.

   In this case, the **fallpromo** defines a single date. Most events are likely to be repeated, such as the **superbowl** definition. Since the Superbowl is always on a Sunday, there is a specific date for each year. The **back2school** event definition spans each day from August 7 to August 30. Like **fallpromo**, this event is defined only for a single year. Adding another **fallpromo** event for the following year would require another definition.

   The event definition can also include optional **qualifier-options**. See the SAS Forecast Server Procedures User’s Guide for more information about event definitions.

   **c**  The event data set is created by specifying the data set name using the OUT option in the EVENTDATA statement.

   ```sas
   proc hpfevents data=sashelp.air lead=12; /* 1 */
      id date interval=month;
      eventdef fallpromo= '01oct2018'd; /* 2 */
      eventdef back2school= '07aug2018'd to '30aug2018'd by day;
      eventdef superbowl = '06FEB2011'D '05FEB2012'D '03FEB2013'D
                           '02FEB2014'D '01FEB2015'D '07FEB2016'D
                           '05FEB2017'D '04FEB2018'D '03FEB2019'D
                           '02FEB2020'D;
      eventdata out=eventds (label='Event List'); /* 3 */
   run;
   ```

   **Note:** Do not use the CONDENSE option in the EVENTDATA statement. This can cause errors when trying to import the event definition to a project.

For more information about defining events using the EVENTDEF statement or using the HPFEVENTS procedure, see **The HPFEVENTS Procedure** in the SAS Forecast Server Procedures User’s Guide.

---

**Segmenting Project Data**

Often, you need to partition data into different groups based on the nature of the data (for example, slow moving items, new products, and so on). Segmentation is a process for creating segments that you can run through different modeling strategies in a pipeline.
Defining Project Segments

To segment your data, you first have to define how you want the data segmented. For example, you might want to run different modeling strategies on seasonal products as opposed to products that are sold throughout the year.

After you have your segmentation strategy defined, you need to add the segment definitions to an external attributes data set. As described in "Preparing Attributes for Your Project" on page 4, each row in the external attributes data set must have a unique combination of values of the default attributes. Add a _SEG_ variable to each row to define how you want the data segmented. Each unique value of the _SEG_ variable creates a new segment that is used in segmented pipelines. Each segment name uses the _SEG_ value from the imported attributes. If any rows in the attributes data set has a missing value for _SEG_, the data is placed in a segment named Remaining Series. The _SEG_ variable can take numeric or character values.

SAS Visual Forecasting supports a maximum of 1,000 segments.

The following code provides a simple example for adding segments to the SKINPRODUCT_ATTRIBUTES data set. This data set is available with other sample data sets that you can download, as described in "Preparing an Input Data Set" on page 2.

```plaintext
   cas casauto;                                       /* 1 - Start a CAS session             */
   libname mycas cas;                                 /* 2 - Create a CAS engine libref      */
   options cassessopts=(caslib="Public");           /* 3 - Specify Public as active caslib */
   data mycas.segment_attr (promote=yes);             /* 3 - Specify the new data set        */
      set mycas.skinproduct_attributes;                /* 4 - Read in the attributes table    */
      length _seg_  $20;                               /* 5 - New variable for segments       */
      if Venue = 'Catalog' then _seg_ ='Catalog';      /* 6 - Assign Segments                 */
      if Venue = 'Grocery Store' then _seg_ = 'Grocery Store';
      if Venue = 'Internet' then _seg_ = 'Internet';
      if Venue = 'Outlet Store' then _seg_ = 'Outlet Store';
      if Venue = 'Third Party Vendor' then _seg_ = 'Third Party Vendor'  ;
   run;
```

This code simply shows how to assign _SEG_ values based on the value of a single variable, Venue. How you segment your project data depends on the characteristics and complexity of your data source, and how you might need to process different segments.

Follow the steps in "Importing Attributes to Your Project" on page 15 to add the data set with the segments and attributes to the project.

After the data set with the segment definitions have been added, you can create segmented pipelines to run the project data. For more information, see "Working with Segmented Pipelines" on page 58.
Working with Projects

When you sign in to Model Studio, the page displays a list of projects that have already been created. To create a new project, see “Creating Your First Project” on page 5.

To open a project, click the link on the project name.

Exporting Projects

Users in the SAS Administrators group can download projects and export them to another SAS Visual Forecasting server. The remote server must have access to the same data sources as the original project. If the project has any pending overrides, they will not be included in the downloaded project.

1. From the Model Studio Projects page, select one or more projects to export.
2. Click the menu icon (⋮) over the right side of the project list and select Export.
   
   A message indicates that the project is being prepared for export.

3. When the project is ready for export, select the location on your local drive for the project. If multiple projects are selected, repeat this step until all selected projects are downloaded.

Importing Projects

Model Studio can be used to import projects that have been created from another forecasting application.

- Model Studio users can import projects that have been archived from SAS Forecast Studio using a FAR file extension. The imported project is placed in your private or personal CASUSER directory.
  
  - Scenario analyses are not supported in this release of SAS Visual Forecasting and are not migrated.
  
  - Overrides and custom models are supported, but are not migrated.
  
  - Events must be imported separately. See “Importing Custom Events” on page 20 for more information.
  
  - For the independent variable, these values are not supported for hierarchy aggregation: First value, Last value, Median of values, Total number of values.

The data set for the project must contain a variable that is in a SAS date or datetime format. For more information, see “Requirements for Time Series Data” on page 3.
Users in the SAS Administrators group can use Model Studio to import files from another SAS Visual Forecasting installation. The project must be in ZIP format.

Projects from a previous release of SAS Visual Forecasting must be imported using SAS Environment Manager. See “Promoting Projects and Templates from Previous Releases to Model Studio 8.5” on page 168 for complete instructions.

The source data for the project must already be loaded to the CAS server before importing the project. If the project has additional attributes or event definitions from an external data source, those items must also be loaded to the CAS server.

1. From the Projects page, click the menu icon (⋮) over the right side of the project list and select Import ➤ Forecasting.

   The Import Forecasting Project dialog box is displayed.

2. Click Browse to provide the filename and location of the project to be uploaded.

3. Click Browse to provide the location on the target server where the project data is located. The variables in the selected data set must match the variables in the uploaded project.

4. Click Import to complete the process. The uploaded project is opened to the Data tab. The variable assignments on the Data tab are the same as the original project.

5. Depending on the type of project that you are importing, follow these steps.

   SAS Visual Forecasting
   - After importing a project, the pipelines must be run again. If a champion pipeline was manually selected in the original project, that selection is not preserved in the uploaded project. After uploading a project from another SAS Visual Forecasting server and running the pipelines, any existing overrides must be resubmitted.

   SAS Forecast Studio
   - When importing projects from SAS Forecast Studio, the following hierarchy aggregation and time accumulation values are not supported. The imported project sets these values to Sum of values.

   - Median
   - First Occurrence
   - Last Occurrence
   - Number of Observations
   - None

   In the right pane, review the settings for the dependent and independent variables and make any adjustments as needed.

After a project ZIP file is imported, only the original owner of the project can access it. The owner can share the project with other user groups. If the owner is not available, the SAS administrator can share it.
Sharing Projects

When you create a project, you become the project owner and the project is kept in private mode so that only you can view or edit it. When you need to allow other users to access the project, you can use Model Studio to share the project. You can designate groups of users to share projects, not individuals.

The Model Studio implementation of sharing is separate and distinct from project sharing as performed in SAS Drive. Any projects that you share using SAS Drive do not retain the same settings for user groups in Model Studio. Also, any projects that you share using Model Studio do not retain the same settings for users in SAS Drive.

Note: Project sharing can be changed only by the project owner or an administrator.

When you share a project using Model Studio, you must specify how it can be accessed, either read-only or read-write mode.

- When a project is in Read-Only mode, it can be viewed but nobody can make any changes, including the project owner. If you need to make changes to the project, you can change the access to read-write and make any necessary adjustments to the user groups who can access the project. You can also make the project private again while you work on it.

- When a project is in read-write mode, anyone with access to it can make changes to the project. However, the project can be accessed only by a single user at a time. If another user tries to access the project while it is being edited, the user receives a message indicating that the project cannot be accessed until it is released by the current editor.

To share a project that you own, or to change the project back to private access, follow these steps.

1. Open the Projects page in Model Studio and select the project in the table.

2. Click the menu icon (⋮) over the right side of the project table and select Share project. The Share Project dialog box is displayed.

3. Select the project sharing options.
   - To share the project with other users, click Shared Project and select the user groups that should have access.
   - To remove sharing for a project, click Private project. The project is now available only for you to view and edit.

4. For shared projects, set the access level for all users. Select Read-only if you want to prevent any changes to the project. If you want all users to be able to make changes to the project, make sure Read-only is not selected.
Working with External Forecast Projects

External forecasts are projects created from data that already has forecasts along with the historical data for each time series. The data set is typically exported as an OUTFOR table and must follow this schema.

- _NAME_ - the name of the dependent variable in each row
- ACTUAL - values of the dependent variable in the historical data
- PREDICT - predicted values of the dependent variable in the forecast period (horizon)
- ERROR - prediction errors
- STD - prediction standard errors
- UPPER - upper confidence limits
- LOWER - lower confidence limits
- The time variable from the original time series project
- All BY variables in the hierarchy of the original time series project

Note:

When you create an external forecast project, the input data set must include both historical and forecast time periods. After an external forecast project is created, the forecast period is already defined. Refreshing the input data source with actual values placed in the forecast period changes the start of the horizon, which can invalidate the external forecast project.

You can create an external forecast data set by exporting from a time series project in SAS Visual Forecasting and SAS Forecast Studio. For SAS Visual Forecasting, see “Exporting Output Data from a Project” on page 93 for more information.

To create an external forecast project, follow the instructions in “Creating a Forecasting Project” on page 5, and specify the External Forecast template.

When the project is created, it is opened to the Data tab with the OUTFOR table shown. The _NAME_ column is shown using the label Variable Name. The project pipeline is initialized with the External Forecasts modeling node. This is the only model that can be used in the pipeline for an external forecast project.

By importing additional attributes to an external forecast project, you can add more attributes to the project and create filters using the additional attributes. For more information, see “Working with Attributes” on page 15.

Note: For external forecast projects, demand classification attributes are listed on the Data tab. However, they are not available for the project in the Filters pane for the Time series viewer, Forecast viewer, or Overrides.
Refreshing Project Data

SAS Visual Forecasting automatically checks the source data from which a project is created for updates. If the source data has been changed, you are notified and can update the project data. If you choose not to update the project data when you are prompted, you can submit a request to refresh the data at a later time as follows:

1. Navigate to the project’s Data tab.
2. In the left pane, click 📈.

All data sources listed are refreshed if their content has changed.

**Note:** In the original data source, if any of the variable formats are changed or if variables are added or removed, data refresh fails.

The data source for a project is kept in a CAS table in memory. In general, the source for this CAS table comes from a table stored on disk. If the source for the project's CAS table is not stored on disk, and CAS is restarted, you need to reload the table to memory in CAS. This triggers a refresh notification, even if the data has not actually been changed.

If you choose not to refresh data when prompted, the project still displays updated metadata in the right data pane when the data source is selected.

Running Batch Code

You can download code to run a SAS Visual Forecasting project in batch mode.

1. From the Model Studio Project page, select a project to run in batch.
2. Click the menu icon ( ⚙️) and select **Batch API**.
   
   The Batch API dialog box is displayed with the code for the project inline.

   **Note:** You can also download batch API code from within a project by clicking the settings icon 📜 in the project title bar.

3. Select the **Download Type** at the bottom of the dialog box. The options are as follows:
   - **SAS:** Downloads code to run in a SAS programming environment.
   - **Python:** Downloads code to run in a Python programming environment.
   - **REST:** Downloads a text file with example REST calls that you can use in an application.

4. Click **Download** and select the location to save the file.

To run the batch code, you need to supply the host name of the SAS Visual Forecasting server and the user name and password.
**SAS**

Update the host, username, and password macro variables at the end of the code.

```sas
%let protocol = http;
%let host = test.example.com;
%let port = 80;
%let username = my_username;
%let password = my_password;
%let projectId = a0548b2f-a669-4a10-a4dd-052c671c0c00;
```

**Python**

Set these parameters on the command line. For example, if `download.py` is the file name of the Python batch code, invoke the command:

```
$ python download.py --host test.example.com --username my_username
--password my_password
```

If you have overrides in the project, they might generate conflicts when they are submitted by the batch code. You can update the batch code to automatically resolve these conflicts by adding `autoresolve="true"` to the code.

**SAS**

Find the PROC HTTP procedure that includes the following IN option with `firstTransaction` and `lastTransaction`.

```sas
in="{"firstTransaction":"@first","lastTransaction":"@last"}"
```

Add the autoResolve setting, as follows:

```sas
in="{"firstTransaction":"@first","lastTransaction":"@last",
"autoResolve":true}"
```

**Python**

Find the `resubmit_overrides` function:

```python
def resubmit_overrides(env):
    body = '{"firstTransaction":"@first","lastTransaction":"@last","autoResolve":true}'
```

For more information about running SAS code in batch mode, see *SAS Companion for UNIX Environments*.
## Overview of Pipelines

Model Studio projects are built around one or more pipelines. A **pipeline** is a process flow diagram that can be used to represent a sequence of analytical tasks. These analytical tasks are represented as individual nodes in a pipeline.
By default, the initial pipeline for a project uses the template that was specified when the project was created. You can create new pipelines using different templates, and you can make changes to the initial pipeline.

Using the Pipelines Tab

Initial Pipeline for a New Project

Pipelines consist of a series of nodes that are run in a sequence to produce forecasts based on the models and settings that you choose for the project. After you first create a project, the pipeline is set up in the Pipelines tab. The pipeline is labeled Pipeline 1.

By default, the initial pipeline for a project uses the template specified when the project was created. You can create new pipelines using different templates and you can make changes to the initial pipeline.

The pipeline cannot be run until you assign the required variable roles as described in "Assigning Variable Roles" on page 7.

See Also

- "Nodes Pane" on page 36
- "Options Pane" on page 36

Creating a New Pipeline

1. Click + to create a new pipeline. The New Pipeline dialog box is displayed.

2. For Name, change the value to any meaningful string. You can change the pipeline name at any time. After the pipeline is created, you can change the pipeline name by double-clicking it the name in the tab.
3 Select a template and click Save. See "Pipeline Templates Provided by SAS Visual Forecasting" on page 67 for a description of the templates that are available.

4 Optional. When the new pipeline is loaded to the page, you can add a formatted description to it. Click and select Expand header. See “Expand header” on page 35 for more information.

Node Status

As you look at a pipeline, hold your pointer over each node to see its status.

- When you first create the pipeline, all nodes have a status of Initialized.
- After a node in the pipeline has successfully finished execution, the node shows a status of Successful.
- A node with a status of Pending is waiting for other nodes in the pipeline to complete before starting.
- When a node starts executing its functions, its status is Running. You cannot make any changes to the node until it finishes running.
- If a pipeline run fails, check to see which node has a Failed status.
- If you make any changes to the project settings, variable assignments, or the project training table the nodes change to out-of-date status, and the pipeline must be run again.
- If a node fails for any reason, while a pipeline is running, the status of the subsequent nodes in the pipeline show Canceled.

Actions on Nodes in the Pipeline

Right-click any node in a pipeline to open a menu of actions.

- Select Run to run the node. Any preceding nodes in the pipeline that do not have a status of Successful are run first. After the preceding nodes complete successfully, the selected node runs.
- Select Rename to rename a node. The new name applies only to the node on this pipeline.
- Select Log to view a log of the node’s processing. This option is enabled after running the node.

Data Node Actions

Right-click the Data node to open these actions in the menu.

- Select Add child node to add a forecasting modeling node to the pipeline. You cannot add child nodes in the External Segmentation or Demand Classification pipelines.
After the Data node has completed a run, select **Time series viewer** to open a plot of the historical data. See “Viewers for Time Series in Your Project” on page 77 for more information.

**Note:** If the input data for the project is very large, running the Data node can take a long time. Data node processing is described in “Running the Forecast Pipeline” on page 57.

### Modeling Strategy Actions

Right-click any modeling node to get these actions in the menu.

- Select **Delete** to remove the node from the pipeline. If it is the only modeling node in the pipeline, then the Model Comparison and Output nodes are also removed.

  **Note:** You cannot delete nodes from the External Segmentation or Demand Classification pipelines.

- Select **Open** to open the code editor for the node. This option is available only for *pluggable* modeling strategy nodes.

  **Note:** The **Open** action is different on nodes in a segmented pipeline. See “Customizing Each Segment” on page 60 for more information.

- Select **Modify Event Usage** to change the settings for events for the node in this pipeline. See “Changing Event Usage in Modeling Strategies” on page 43 for more information.

  After the modeling strategy has completed a run, select **Forecast viewer** to open a plot of the time series, including historical and forecast data. See “Viewers for Time Series in Your Project” on page 77 for more information.

  Select **Save as** to save the node, with any settings you have made in the Options pane on the right, to The Exchange. This makes the saved node available for other projects.

  Select **Results** to view the results of the modeling strategy, including an execution summary, MAPE distribution, and other relevant data. For more information, see “Viewing the Results for a Modeling Strategy Node” on page 44.

  After the modeling strategy has run, you can select **View code** to get a read-only view of the code for the modeling strategy.

### Segmentation Nodes

For segmented pipelines, you have these nodes to work with.

#### Profile Nodes

- **External Segmentation Profile**
- **Demand Classification Profile**

#### Modeling Nodes

- **External Segmentation Modeling**
Demand Classification Modeling

Merge Segments
Merges the results produced by the segments.

For more information about actions on these nodes, see “Customizing Each Segment” on page 60.

Model Comparison Node Actions

Right-click the Model Comparison node to get these actions in the menu.

- Select Results to view the MAPE distribution and summary statistics for each modeling node along with the selected champion model in the pipeline.

- Select Select champion model to choose a champion model other than the one selected by the Model Comparison. If you choose a different champion model, you need to rerun the Model Comparison and Output nodes.

Actions on the Pipeline

Click on the pipeline tab to perform these actions.

Run
Runs the entire pipeline.

Stop
Stops a pipeline that is running.

Duplicate
Creates a duplicate pipeline. The name is appended with a number. You can rename the duplicate after it is created.

The pipeline for a segment of a pipeline cannot be duplicated.

Rename
Renames the pipeline.

Save to The Exchange
Saves the pipeline with the nodes and any settings applied to those nodes as a template to The Exchange. The new templates can be used in other projects.

Delete
Deletes the pipeline from the project. At least one pipeline must exist in a project. If only one pipeline exists, this action is disabled.

Show overview map
Places a map of the pipeline in the upper left corner of the canvas.

Expand header
Provides a space at the top of the pipeline to add a description or other text that might be useful. The text can be formatted. You can add text and use the tools provided to format it.
Nodes Pane

On the left side of the pipeline is the Nodes pane. This pane lists the modeling strategy nodes that you can add to any pipeline.

If the Nodes pane is not displayed, click on the left side of the pipeline to open it. You can collapse the pane again by clicking .

You can drag an analysis node over the Data node in the pipeline to add it. To remove a node, right-click the node and select Delete. Each pipeline requires at least one modeling node.

Options Pane

The Options pane on the right side of the pipeline displays options that you can set for a selected node in the pipeline.

If the Options pane is not displayed, click to open it. You can collapse the pane again by clicking .

- All nodes in the pipeline have a Description option that you can update. Any description you provide is saved for the node only in the pipeline. The description for the node template in The Exchange is not affected by any updates that you make in the pipeline.

- Pluggable models in SAS Visual Forecasting that are nodes have an Open button that you can use to view and edit the code inline.

After running a pipeline, any subsequent changes to the options for a modeling node move the pipeline back to out-of-date state. This is also true if you make a change and then change the setting back to the original value.

When a pipeline is running, the options for each modeling node in the pipeline cannot be changed until the pipeline completes. If you need to make further updates to a modeling node after starting the pipeline, click Stop Pipeline and wait for the pipeline to stop operation before you make any changes.
Modeling Strategy Nodes

Every pipeline must have at least one modeling strategy node. You can select from the models provided by SAS Visual Forecasting or any custom models that might have been created for your site.

To add a modeling node to the pipeline, choose a node from the Nodes pane on the left and drag it over the Data node in the pipeline. If the Nodes pane is collapsed, click to reveal the nodes that are available.

Modeling Strategy Nodes Provided by SAS Visual Forecasting

You can use one or more of the following nodes for your pipelines. These modeling strategies are listed in the Nodes pane on the left of the Pipelines tab. They are also available as templates in The Exchange.

Some modeling strategies provided by SAS Visual Forecasting are pluggable models, which are described in “Customizing a Modeling Strategy” on page 51. These nodes provide a code editor that you can use to customize its settings. After running a pipeline, any change to the code or right pane settings for a modeling node requires the pipeline to be run again.

Auto-forecasting

This node models each time series based on the models that you select (ARIMAX, ESM, IDM, UCM). For each time series, the Auto-forecasting node performs the following tasks.

- Diagnoses the statistical characteristics of the time series
- Generates a list of appropriate time series models
- Selects the model
- Generates forecasts

This node is a pluggable model. You can edit the code to create your own modeling strategy. For more information, see “Customizing a Modeling Strategy” on page 51.

For a complete description of the settings for Auto-forecasting, see “Auto-forecasting Settings” on page 116.

External Forecasts

This node works only with external forecast projects. It cannot be added to the pipeline for projects created from any other type of data source. External forecast projects can use only External Forecasts in the pipeline.
Note: The External Forecasts node is not pluggable and the code cannot be edited. See “Customizing a Modeling Strategy” on page 51 for more information about pluggable modeling strategies. External forecasts do not support events.

Use External Forecasts in a pipeline to generate the summary statistics and to work with overrides in the forecast horizon. For more information, see “Working with External Forecast Projects” on page 27.

Hierarchical Forecasting

For each time series, the Hierarchical Forecasting node performs the following tasks.

- Diagnoses the statistical characteristics of the time series for each level of the hierarchy
- Generates a list of appropriate time series models based on the diagnostic settings that you select
- Selects the champion model from candidate list of models
- Generates forecasts
- Reconciles the forecasts from the reconciliation setting

To use this modeling node, you need at least one BY variable assigned.

Note: This node is not a pluggable model. The code used for this node cannot be edited. If you download this node, it does not include the code to create the node. However, you can make changes in the Options pane and save those changes to The Exchange.

For a complete description of the settings for Hierarchical Forecasting, see “Hierarchical Forecasting Settings” on page 117.

See Also

“Pipeline Fails with Hierarchical Forecasting Model” on page 174

Hierarchical Forecasting (Pluggable)

The Hierarchical Forecasting (Pluggable) node provides similar functionality as the Hierarchical Forecasting node. The main difference is you can use the code editor to make changes to the model inline. Also, if you download this node, it includes the code for further customizing. For a complete description of the settings for this modeling node, see “Hierarchical Forecasting (Pluggable) Settings” on page 123.

The results for this node are expected to be similar but not the same as the results that you would get from Hierarchical Forecasting.

Naive Model

The Naive Model node uses one of the naïve models that you select to generate forecasts for each time series.

The following settings are available in the Options pane for the Naive Model strategy.
Naive model type
Select from the following.

- **Moving average**
  For this option, indicate the size of the moving average window.

- **Random walk**
  For this option, specify if you want the drift option.

- **Seasonal random walk**
  For this option, specify if you want the drift option.

**Drift option for random walk models**
If you choose a Random walk or Seasonal random walk model, you can add the drift option.

**Window size for the moving average**
Specify an integer, greater than 1, for the window size for the moving average model.

This node is a pluggable model. You can edit the code to create your own modeling strategy. For more information, see “Customizing a Modeling Strategy” on page 51.

---

**Note:** The Naive Model does not support events.

For any time series with fewer than six non-missing observations, an ESM model (ESMBEST) is used instead.

---

**Modeling Strategy Nodes for Demand Classification**

The following models are intended for time series that have been segmented based on patterns detected in the time series. They can be applied to other time series data, but the results might not be useful.

For example, the Retired Series modeling strategy is intended for time series that have not recorded any results in the final time periods. It does not produce any forecasts or measurements for statistics of fit.

**Non-seasonal Model**
The Non-seasonal Model Forecasting node generates a list of appropriate non-seasonal model candidates for each time series, based on the model selection criteria that you choose. You can select between ESM, ARIMAX, and UCM models to be evaluated. After model selection, forecasts are generated for each time series using the selected model.

For ESM and UCM models, prebuilt options are available in PROC TSMODEL. For ARIMAX models, the search is performed over the models that have no seasonal AR, MA or differencing.

For a complete description of the settings that you can use, see “Non-seasonal Model Settings” on page 145.

**Retired Series**
This node is intended for time series that have been segmented based on patterns of inactivity as determined by the Demand Classification pipeline. It is not a good
model for project data that has current active records. Retired Series does not support events.

This node does not generate forecasts. By default, all forecast values are set to missing. For information about changing this value, click for Retired Series in the Options pane.

Seasonal Model

This node generates a list of seasonal model candidates based on the model selection criteria that you select and generates forecasts. You can select between ESM, ARIMAX, and UCM as candidate models. ESM and ARIMAX are selected by default. The model selection is limited to seasonal models. For ESM and UCM models, prebuilt options are available in PROC TSMODEL. For ARIMAX, model selection is performed between models that have either seasonal AR, MA, or differencing greater than zero.

Time series that do not fit the seasonal model are forecast using the best ESM candidates. Candidate models include simple, linear, damped trend, additive seasonal, Winters multiplicative, and Winters additive models.

For a complete description of the settings that you can use, see “Seasonal Model Settings” on page 144.

Time Series Regression

This node is normally used in project data that has been segmented based on patterns determined by a Demand Classification pipeline. This modeling strategy is intended for time series classified as INSEASON, which means activity occurs only during certain time periods. This model should be a good fit for project data that consists of short time spans.

An integer index is generated for the time stamps. This index, the index squared, and the index cubed are included as independent variables. Seasonal indices are also included in the model as independent variables. An ARMA model with no intercept, no AR component, and no MA component (which is equivalent to a regression model) is fit to the data.

For a complete description of the settings that you can use, see “Regression for Time Series Settings” on page 147.

Temporal Aggregation Model

This node is intended for time series that are intermittent at the high-frequency level but seasonal at the accumulated low-frequency level. This strategy forecasts high-frequency series using IDM and the low-frequency series are forecasted using one of the seasonal models that you select, such as ESM, ARIMA, or UCM. The results are then reconciled from the low-frequency level to the high-frequency level to yield the final forecasts.

Note: The Temporal Aggregation Model does not support events.

For a complete description of the settings that you can use, see “Temporal Aggregation Model Settings” on page 145.
Modeling Strategy Nodes Using Neural Networks

If your site has a license for SAS Visual Data Mining and Machine Learning, these modeling strategies are available for use in the pipelines.

Panel Series Neural Network

The Panel Series Neural Network modeling strategy provides forecasts by training a neural network based on user settings and develops a model to extract salient features across multiple time series. Neural networks consist of predictors (input variables), hidden layers, an output layer, and the connections between each of those items.

The Panel Series Neural Network modeling strategy is best suited for panels of time series with several interval independent variables. Independent variables for the neural network can be assigned in the data tab or generated within the Feature Generation section of the node options. BY variables are automatically included as categorical inputs. This neural network cannot generate a model for the project without an independent variable.

Note: The code used for this node cannot be edited. If you download this node, it does not include the code to create the node.

For a description of the settings for this node, see “Panel Series Neural Network Settings” on page 131.

Multistage Model

The Multistage Model provides a general framework that combines time series models and feature extraction techniques to build a hierarchy-based forecasting system in two stages.
This diagram shows the process flow of the Multistage Model node.

- For the first stage, you choose a model for generating your stage 1 low-level forecasts.

  If you select a regression or neural network model, features are extracted across the time series and used as an adjustment factor to generate the stage 1 low-level forecasts using a time series model. If you select a time series model for the first stage, the initial stage 1 forecasts are generated without the adjustment factor.

  You can specify a high level in the hierarchy to generate stage 1 high-level forecasts using the time series model. The high-level stage 1 forecasts are then reconciled with the initial low-level stage 1 forecasts to generate the final forecasts. If you do not indicate a high level in the hierarchy, the initial low-level forecasts are used as the final forecasts for stage 1.

- In the second stage, feature extraction is applied again to generate forecasts for each individual time series at lower levels, using either a regression or neural network model.

- Finally, top-down reconciliation is performed between the stage 1 and stage 2 forecasts to generate the final forecasts.

This modeling strategy is intended for a project with a defined hierarchy. If no BY variables are defined for the project, the forecast is generated using stage 1 time series model. For more information about hierarchies, see “Defining the Hierarchy” on page 11.

When working with the settings in the Options pane for Multistage Forecasting, consider any signals you might notice at different aggregation levels in the hierarchy. For example, in a retail scenario, time series signals such as trend or seasonality might be best observed at a REGION or CATEGORY level. Setting this level as the stage 1 high level can better capture the time series signal using the stage 1 high-level time series forecast.

You might capture a stronger signal about how the sales respond to the price or promotion changes at the STORE or BRAND level. Setting this to the stage 1 low level allows the use of regression or neural network to model the price or promotion effect. At the lowest level, each STORE or SKU combination might have its own basic behavior that is not captured using the stage 1 models. These can be modeled separately in stage 2.
You can generate more a robust and reliable forecast by combining these elements:

- the time series signal that is modeled at the stage 1 high level
- the price or promotion effect that is modeled at the stage 1 low level
- the characteristics of each individual time series that is modeled at stage 2

For more information about the individual settings for Multistage-forecasting, see "Multistage Model Settings" on page 125.

**Stacked Model (NN + TS) Forecasting**

The stacked modeling strategy generates forecasts using stacked models that include a neural network model (NN) and a time series model (TS). This modeling strategy captures the nonlinear relationship between the dependent and independent variables as well as time series characteristics in the data, such as seasonality and trend. As this diagram shows, it models the time series in two steps:

- In the first step, the neural network model is used to generate the forecasts.
- In the second step, the residuals from the first step are passed to the time series model to generate residual forecasts.

The first and second steps run sequentially.

The final forecasts are the sum of the forecasts from the neural network and the residual forecasts from the time series model.

**Note:** The code used for this node cannot be edited. If you download this node, it does not include the code to create the node.

For more information about the settings for Stacked Model (NN + TS) Forecasting, see "Stacked Model (NN + TS) Forecasting Settings" on page 139.

---

**Changing Event Usage in Modeling Strategies**

Events that have been added to a project can be further customized for each model in each pipeline. For predefined events, you can change only the settings for
predefined events that have been added to the project. For custom events that have been added to the project, you can change the settings for all of the events.

1 Right-click the modeling node and select **Modify event usage**. The Modify event usage page is displayed.

   **Note:** For segmented pipelines, you must open the modeling node to modify the event usage on the nested pipeline for each segment. For more information, see “Customizing Pipelines for Each Segment” on page 63.

2 Select either **Predefined events** or the custom event table from the list in the left pane.

3 Select an event from the table in the middle pane.

4 In the right pane, change the **Usage in system-generated models** setting.
   - **Do not use**
     Do not include the event in the model.
   - **Try to use**
     Include the event in the model as long as the parameters of the events are significant and the increment of the value of the criterion exceeds a specified threshold.
   - **Use if significant**
     Include the event in the model as long as the parameters of the events are significant.
   - **Force use**
     Include the event in the model as long as the model does not fail to be diagnosed.

The setting that you apply is effective for this modeling strategy in the pipeline. If you save the modeling strategy to The Exchange, the event definitions are not included.

Changing the event usage makes the modeling strategy out-of-date. The pipeline must be run again.

---

**Viewing the Results for a Modeling Strategy Node**

When a modeling node has finished running successfully, it has a green check mark indicating a status of Successful. Right-click the node and select **Results** to view the output of the node. Results displays a Summary and Output Data tab for a modeling strategy.

**Summary Tab**

The following information is shown on the Summary tab. For the Hierarchical Forecasting modeling strategy, this information is provided for each level within the defined hierarchy.

**MAPE Distribution**

This graph shows the MAPE distribution of results.
Model Family
This graph shows the percentage of time series model by each model family. This graph shows how each model type (ARIMA, Combined, Exponential Smoothing, Intermittent Demand, and Unobserved Components) fits the series. Neural network models are included in the other family.

Model Type
This graph shows the percentage of time series used for by each model type. This graph shows whether the model included a dependent transformation, a seasonal component, a trend component, inputs (such as independent variables), events, and outliers. This table is not displayed for non-time series based modeling strategies, such as the Panel Series Neural Network, Stacked Model (NN + TS) Forecasting, and Multistage-forecasting modeling.

Execution Summary
This table shows the results of the following measures.

Number of series
the total number of time series accumulated from the data. For a Hierarchical Forecasting modeling strategy, this corresponds to the number of time series in each level of the hierarchy.

Number of failures
the number of forecasts that failed. For a Hierarchical Forecasting modeling strategy, this corresponds to the number of failed forecasts in each level of the hierarchy.

Number of forecasts equal to zero
the number of forecasts that are 0

Number of intermittent series with flat forecasts
the number intermittent time series that have a constant forecast value over the length of the forecast horizon

Number of seasonal series with flat forecasts
the number in seasonal time series that have a constant forecast value over the length of the forecast horizon

Number of short series with flat forecasts
the number of short time series that have a constant forecast value over the length of the forecast horizon

Number of all other series with flat forecasts
the number of all other time series that have a constant forecast value over the length of the forecast horizon. This number does not include short, intermittent, or seasonal time series with flat forecast values.

Number of forecasts below 70%
the number of time series with forecast values that are below the minimum range of the mean value for that time series. The forecasts for these series can be problematic. The minimum range value is set to 70% by default. You can change this setting using the Forecast exception boundaries in the Project Settings on page 115.

Number of forecasts above 300%
the number of time series with forecast values that are above the maximum range of the mean value for that time series. The forecasts for these series can be problematic. The maximum range value is set to 300% by default. You can change this setting using the Forecast exception boundaries in the Project Settings on page 115.
Output Data

The following output tables are generated by modeling strategies in a pipeline or segment. If the table does not initially load, click View Output Data in the content area. Click 📋 to save any selected table to a location in the CAS library. Click 🕵️‍♂️ to port the table to SAS Visual Analytics for further investigation.

These tables can be sorted when you click the columns. You can also arrange and remove columns in the table, as described in “Working with Tables in Model Studio” in SAS Visual Forecasting: Overview.

OUTFOR

The OUTFOR table provides the forecasted values generated by the modeling strategy. This table includes the BY variables for the project and the following information.

- **_NAME_** - the name of the dependent variable
- **time variable** - the date or datetime value for each time series.
- **Actual Values** - the values of the dependent variable for each time series in the historical record
- **Predicted Values** - the predicted values for each time series in the forecast horizon
- **Prediction Standard Errors** - the standard error estimate from the models
- **Lower Confidence Limits** - the lower confidence limit as determined by the model
- **Upper Confidence Limits** - the upper confidence limit as determined by the model
- **Prediction Errors** - residuals between the actual values and the model predictions

OUTMODELINFO

The OUTMODELINFO table contains detailed information about the selected forecast model. For each BY variable combination, this table contains the following information.

Note: This table is not available for non-time series based modeling strategies such as the Panel Series Neural Network, Stacked Model (NN + TS) Forecasting, and Multistage-forecasting.

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Variable Label</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>MODEL</em></td>
<td>Model</td>
<td>the name of the model selected for this time series.</td>
</tr>
<tr>
<td>Variable Name</td>
<td>Variable Label</td>
<td>Description</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------</td>
<td>-------------</td>
</tr>
<tr>
<td><em>MODELTYPE</em></td>
<td>Model Family</td>
<td>the type of model used (ESM, ARIMA, UCM, or IDM)</td>
</tr>
<tr>
<td><em>DEPTRANS</em></td>
<td>Dependent Variable Transform</td>
<td>the type of dependent variable transformation that was used for this time series.</td>
</tr>
<tr>
<td><em>SEASONAL</em></td>
<td>Seasonal model</td>
<td>indication of seasonal model used for this time series. A value of 1 indicates a seasonal model was used.</td>
</tr>
<tr>
<td><em>TREND</em></td>
<td>Trend Model</td>
<td>indication of trend model used for this time series. A value of 1 indicates a trend model was used.</td>
</tr>
<tr>
<td><em>INPUTS</em></td>
<td>Inputs present</td>
<td>the number of input variables used in the model</td>
</tr>
<tr>
<td><em>EVENTS</em></td>
<td>Events Present</td>
<td>indication of whether events were used in the model. A value of 1 indicates events are present.</td>
</tr>
<tr>
<td><em>OUTLIERS</em></td>
<td>Outliers Present</td>
<td>the number of outliers discovered in the model</td>
</tr>
<tr>
<td><em>SOURCE</em></td>
<td>Model Source</td>
<td>the named source of the model</td>
</tr>
<tr>
<td>Variable Name</td>
<td>Variable Label</td>
<td>Description</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------</td>
<td>-------------</td>
</tr>
<tr>
<td><em>STATUS</em></td>
<td>Model Status</td>
<td>the execution status of the model. The following status codes can be present.</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>The forecast completed successfully.</td>
</tr>
<tr>
<td></td>
<td>3000</td>
<td>Model selection could not be completed. Forecast values are set to missing.</td>
</tr>
<tr>
<td></td>
<td>3001</td>
<td>Model selection could not be completed and NOALTLIST prohibits use of default exponential smoothing. Forecast values are set to missing.</td>
</tr>
<tr>
<td></td>
<td>3002</td>
<td>The forecast was completed subject to qualification that one or more input variables were omitted from the selected model. This can occur only in the context of ARIMAX or UCM models.</td>
</tr>
<tr>
<td></td>
<td>3003</td>
<td>The desired model could not be forecast. The forecast reverted to the default exponential smoothing model.</td>
</tr>
<tr>
<td></td>
<td>3004</td>
<td>The attempt to forecast the desired model produced an arithmetic exception. The forecast is generated by CATCH(ESM) processing.</td>
</tr>
<tr>
<td></td>
<td>3005</td>
<td>The attempt to forecast the desired model produced an arithmetic exception. The forecast is generated by CATCH(RW) processing.</td>
</tr>
<tr>
<td></td>
<td>3006</td>
<td>The attempt to forecast the desired model produced an arithmetic exception. The forecast is generated by CATCH(MISSING) processing.</td>
</tr>
<tr>
<td></td>
<td>3007</td>
<td>The mean value forecast is generated as a result of the MINOBS criterion.</td>
</tr>
<tr>
<td></td>
<td>3008</td>
<td>There were insufficient nonmissing observations in the variable to be forecast. A missing value forecast is produced.</td>
</tr>
<tr>
<td></td>
<td>3009</td>
<td>There were insufficient nonzero observations in the variable to be forecast. A zero-valued forecast is produced.</td>
</tr>
</tbody>
</table>

OUTSELECT

The OUTSELECT table contains selection statistics from the selected model. This information is useful for comparing the performance of various models. The CAS table schema that is used for storing the fit statistics is compatible with the schema that is used by the HPFENGINE procedure for its OUTSTATSELECT= data set.

For each BY variable combination, this table contains the following statistics. This table also provides the measures from each of the statistics of fit that can be used
for model selection. For a description of each statistic, see “Descriptions of Model Selection Criteria” on page 159.

**Model**

The model specification name

**Selection List**

The name of the model selection list to which the model belongs

**Model Label**

Model transfer function

**Selected Status**

Indicates whether the model was chosen to forecast the dependent series or used by the chosen forecast when the chosen forecast is a combined model. Status values include:

- **NO** — indicates that the model is neither selected nor used.
- **YES** — indicates that the model is the primary model selected for the forecast.
- **USED** — indicates that the model is used by the primary model in producing the final forecast.
- **USED_SELECT** — indicates that the model is used by the primary model in the model selection region, but not in producing the final forecast.

**Region**

Region in which the statistics are calculated. Values include:

- **FIT** — indicates that fit statistics were calculated over the fit region.
- **FORECAST** — indicates that fit statistics were calculated over the forecast region.

**Degrees of Freedom Error**

The degrees of freedom error calculated by the model

**Number of Observations**

Total number of observations for this time series

**Number of Observations Used**

The total number of observations used for this time series

**Number of Missing Actuals**

The total number of actual values that are missing

**Number of Missing Predicted Values**

The total number of predicted values that are missing

**Number of Model Parameters**

The total number of parameters used by this model

**Total Sum of Squares**

The total sum of squares calculated by the model

**Corrected Total Sum of Squares**

The corrected total sum of squares calculated by the model

**OUTSTAT**

This table contains the statistics for the current model for the most recent data update. For each BY variable combination, this table includes the following information. In addition to the following statistics, this table also provides the measures from each of the statistics of fit that can be used for model selection. For
a description of each statistic, see "Descriptions of Model Selection Criteria" on page 159.

- **Variable Name** - the name of the dependent variable.
- **Region** - shows the region in which the statistics are calculated.
  - **FIT** indicates that fit statistics were calculated over the fit region.
  - **FORECAST** indicates that fit statistics were calculated over the fit region.
- **Selection List** - the name of the model selection list to which the specified **Model** belongs.
- **Model** - the name of the model specification.
- **Degrees of Freedom Error**
- **Number of Observations**
- **Number of Observations Used**
- **Number of Missing Actuals**
- **Number of Missing Predicted Values**
- **Number of Model Parameters**

### OUTSUM

The OUTSUM data set contains the summary statistics for the dependent variable as determined by the value of the forecast horizon. Each time series is listed based on the unique combination of the default attributes. For each time series, the summary statistics are shown as determined by the number of forecast periods (horizon) for the project. These statistics are also provided as attributes listed on the Data tab.

See "Descriptive Statistics" on page 153 for more information about these attributes.

### MERGED_ATTRIBUTES

The MERGED_ATTRIBUTES tables provides a list of all user-defined and derived attributes that are available in the project. For a complete description of the derived attributes, see “Derived Attributes” on page 149.

### Logs

This section contains the OUTLOG table, VALIDATIONLOG table, or both. The tables contain text messages that arise from the processing of the BY group’s time series data. If no exceptions occur during processing, the content for this table is empty. In addition to the BY variables, this table includes the following information.

- **_ERRNO_ Status** - a numeric variable that stores the _ERRNO_ variable for the BY group. The value of _ERRNO_ might be set by the user-defined program directly or might be set implicitly by calling a function or method that sets the _ERRNO_ value.
- **Length of log content** - a numeric variable that stores the length of the _LOG_ variable text (byte count).
- **Log content** - messages that are logged from the execution of the user-defined program on the BY group’s time series data. All messages from the BY group are concatenated into the variable. End-of-line characters separate the individual messages.
Working with Modeling Strategy Templates

You can customize many modeling strategies and save them as templates in The Exchange. If you are familiar with SAS code for SAS Visual Forecasting, you can create your own modeling strategy templates and upload them to The Exchange. Pluggable modeling strategies provide a code editor for you to customize the node to meet the specific requirements of your projects.

Customizing a Modeling Strategy

Use the Options pane on the Pipelines tab to customize the settings for a selected modeling strategy in your pipeline. If the modeling strategy is a pluggable model, you can open the code editor to make changes.

**Code editor**

Click **Open** to view and change the code for this node. The code editor is available only for pluggable models. See “Modeling Strategy Nodes Provided by SAS Visual Forecasting” on page 37 to determine which modeling strategy nodes are pluggable.

Use *SAS Visual Forecasting: Programming Guide* as a reference when you are working with this code.

The changes that you make to the modeling strategy affect only the current pipeline. Click ![save](#) to save the modeling strategy to The Exchange. This makes the modeling strategy available to be selected in other pipelines and projects, or to be downloaded for use in other deployments of SAS Visual Forecasting.

For more information about The Exchange, see “Working with Model Studio” in *SAS Visual Forecasting: Overview*.

**Note:** The code for modeling strategies that are not pluggable, like Hierarchical Forecasting, cannot be edited inline or downloaded. However, the settings that you choose in a pipeline can be saved. External Forecasts has no settings that can be customized and saved and the code is also not available.

Event definitions for modeling strategies are not saved to The Exchange.

You can also download a modeling strategy and make changes to it. When you are finished testing it and it is ready for use, you can upload it to The Exchange for use in project pipelines.

**Downloading Modeling Strategies**

To download a modeling strategy, follow these steps.

1. Go to The Exchange, as described in “Working with Model Studio” in *SAS Visual Forecasting: Overview*.
   
   The types of modeling strategies are listed on the left under Nodes.

2. In the list of nodes, select **Forecasting Modeling**.
3 Select the modeling strategy that you want to download. You can select multiple modeling strategies to download.

4 Click the Download option from the menu icon ( ).

Each modeling strategy is downloaded as a ZIP file. The files in the downloaded package are explained in “Creating Your Own Modeling Strategy” on page 52.

Creating Your Own Modeling Strategy

There are two ways to create your own modeling strategy. The first way is to write your own using SAS code. The second way is to make changes to an existing modeling strategy to get the files in the required format. The files are packaged in a ZIP file. After making your changes, package the files back to the ZIP format and upload it. See “Modeling Strategy Nodes Provided by SAS Visual Forecasting” on page 37 to determine which nodes are pluggable.

Note: If the modeling strategy is not pluggable, you can only make changes to the template.json to change the default settings.

Creating or Modifying the Code

The code.sas file contains the SAS code to run the model for the input data. Many of the pluggable modeling strategies run procedures like PROC TSMODEL, PROC REGSELECT, PROC NNET, and PROC CAS.

Note: PROC REGSELECT and PROC NNET require a license for SAS Visual Data Mining and Machine Learning.

You can modify the code that was downloaded or you can replace it with SAS code that you have developed. For best results, the SAS code that you write should work using SAS Studio from the same server where you intend to upload the modeling strategy to SAS Visual Forecasting.

The SAS code must generate the OUTFOR table from the source data. The variables in the OUTFOR table are described in “Working with External Forecast Projects” on page 27. If you download the Auto-forecasting modeling strategy, you can see from the code that it also generates the OUTSTAT, OUTSELECT, OUTINFORMATION, and OUTMODELINFO tables.

The SAS code that you use to create the modeling strategy is not limited to the procedures and CAS actions for SAS Visual Forecasting. For more information about a modeling strategy that uses gradient boosting (PROC GRADBOOST), see the SAS Global Forum 2019 paper Writing a Gradient Boosting Model Node for SAS® Visual Forecasting.

Creating the Validation Model

The validation.xml contains the validation model for the modeling strategy. This file is included in the ZIP file only for pluggable models.

The validation model is specified by an XML schema that is validated when you upload the modeling strategy to The Exchange. The specifications in the validation
model define the settings that are displayed in the Options pane when the modeling strategy is selected in a pipeline.

The XML starts with the root `<validationModel>` element with three subordinate elements.

```xml
<validationModel description="description" name="name" revision="number">
  <links/>
  <version>number</version>
  <properties>
    ...
  </properties>
</validationModel>
```

For the `name` and `description` attributes, these should match the root name and description properties in the `template.json` file, which is described in “Identifying the Default Settings” on page 55.

```
"name" : "Custom Model Name",
"description" : "Custom Model description",
```

For the `revision` attribute, enter a number to correspond with different iterations of your modeling strategy. Use the `<version/>` element to track updates that are more substantial. All of this information also needs to be included in the `template.json`.

The main content of the validation.xml file is included in the `<properties>` element. The properties define the settings that users can change in the Options pane, based on their knowledge of the data and scope of the project. Each `<property>` element includes these attributes.

- **type**
  - the data type for the property.
  - `boolean` — creates a check box. If checked, the value is "true".
  - `integer` — requires integer values with no decimal
  - `double` — requires numerical values that can have decimals
  - `string` — requires character values. Usually, these values are a limited set of values that you define in the `<choicelist>` element.

- **name**
  - the name of the macro variable for the property. The value that the user chooses for this property is passed to the macro variable in the code. This value must match the macro variable name used in the code.

  For example, in the Naive Model modeling strategy, `name="_naiveModelType"` is the property that passes its user-chosen value to the `&_naiveModelType` macro variable in the code.

- **required**
  - a true or false value that indicates whether a value for this property is required. The default is true.

- **selector**
  - a true or false value. This should be set to “false” if the property is a Boolean data type.

- **id**
  - You do not need to provide an ID. The ID is assigned to the property by SAS Visual Forecasting. When a pluggable model is downloaded, it includes this ID for each property.
description provides more information about the property that is displayed when users hold
the mouse pointer over the label

array a true or false value.

enabledWhen specifies a condition in which the setting is enabled. The following example from
Hierarchical Forecasting (Pluggable) uses a Boolean property that determines
whether the Sensitivity level for intermittency test is enabled. If the user
unchecks Include IDM models, the value for _idmInclude becomes false and
this field is disabled.

<property type="integer" required="true" selector="false"
  name="_intermittencySensitivity" displayName="Sensitivity level
  for intermittency test" array="false"
  enabledWhen="_idmInclude" enabledWhenValue="false">
  ...
  ...
</property>

You can specify other conditions.

enabledWhen="_holdoutSampleSize &gt;=1"
  ■ If _holdoutSampleSize is greater than or equal to 1

enabledWhen="_modelSelection_criteria != 'MAPE'"
  ■ If _modelSelection_criteria is not equal to the string "MAPE"

enabledWhen="_modelSelection_criteria != 'MAPE' &amp; _esmInclude"
  ■ If both conditions are true. If you use the || operand, then the value is true if
  either condition is true

enabledWhen="_reconcileLevel == 0"
  ■ If _reconcileLevel is equal to 0

Examining the structure of the validation.xml in several pluggable modeling
strategies can demonstrate how to define the properties for your modeling strategy.
The following examples explain many of these techniques for each data type.

numeric data
Generally, for properties with a double or integer data type, specify a constraint
on the minimum or maximum value. The following code sample shows a typical
property definition for a double data type. The property defines the setting for
Percentage of data points used in the holdout sample in the Hierarchical
Forecasting (Pluggable) node. When the property takes a numerical value, you
should include the <constraints> element and specify any minimum or
maximum value on the <range> element.

<property type="double" required="false" selector="false"
  name="_holdoutSamplePercent" displayName="Percentage of data points used in the
  holdout sample" description="Specifies the maximum percentage of data points used
  for validation. ... " array="false" enabledWhen="_holdoutSampleSize &gt;=1"
  enabledWhenValue="false">
  <clientProperties>
    ...
    ...
  </clientProperties>
  <constraints>
    <range min="0" max="100" includeMin="true" includeMax="false"
boolean

Properties of a Boolean data type generate a simple check box. If the box is checked, the value is true. The following shows a typical property definition for a Boolean data type. The property defines the setting for **Include IDM models** in the Auto-forecasting modeling strategy.

```xml
<property type="boolean" required="false" selector="false" name="_idmInclude"
  displayName="Include IDM models" array="false" enabledWhenValue="false">
  <clientProperties/>
  <constraints/>
</property>
```

string

Like numeric data types, strings require constraints that limit the user’s input to data that the forecasting model can process. The following property, **IDM method**, uses a `<choicelist>` element that specifies the options that the user can select from. The `displayValue` is what users can select in the drop-down menu and the value is passed to the & _idmMethod macro variable in Auto-forecasting.

```xml
<property type="string" required="true" selector="false" name="_idmMethod"
  displayName="IDM method" array="false" enabledWhenValue="false">
  <clientProperties/>
  <constraints>
    <choicelist enabledWhenValue="false">
      <choice value="AVERAGE" displayValue="Average"/>
      <choice value="BEST" displayValue="Best"/>
      <choice value="CROSTON" displayValue="Croston"/>
    </choicelist>
  </constraints>
</property>
```

Identifying the Default Settings

The template.json file contains the metadata about the strategy in JSON format, including the name, description, version, and the default values for the settings in the Options pane when the modeling strategy is selected.

The structure of the template.json file is very rigid and should not be changed. The following values should be changed to match your custom modeling strategy. The other values should not be changed or removed.

**name**

Provide the name as it should appear in The Exchange and in a pipeline.

**description**

Provide a brief description. This appears at the top of the Options pane when the modeling strategy is selected in a pipeline.

**revision**

Numerical value indicating the latest revision of the modeling strategy.

**version**

Numerical value indicating the latest version of the modeling strategy.

**prototype**

Here are the only name value pairs in "prototype" that can be changed.
Here is an example of some of the default values that you can change in Hierarchical Forecasting, along with the settings in the Options pane that are affected. All of these properties are under the block of diagnostic properties.

Table 2.2 Default Settings

<table>
<thead>
<tr>
<th>Default definitions compared to Options pane</th>
<th>Diagnostic settings in the Options pane</th>
</tr>
</thead>
</table>

```
"diagnostic" : {
    "intermittencyTest" : true,
    "intermittencySensitivity" : 2,
    "seasonalityTest" : false,
    "seasonalitySensitivity" : 0.01,
    "diagnoseIndependentVar" : "NONE",
    "functionalTransformation" : "NONE",
    "functionalTransformationPARAM" : 0,
    "functionalTransformationOPT" : "MEDIAN",
    "minobs" : 2,
    "minobsTrend" : 2,
    "minobsSeason" : 2
},
```

To check for seasonality in this example, change `seasonalityTest` to “true”. If you change `functionalTransformation` to “AUTOMATIC”, you can also change the value of `functionalTransformationOPT` to “MEAN”, or you can leave it to the original value “MEDIAN”.

Providing Version Information

The metadata.json file defines the version of SAS Visual Forecasting that can support the modeling strategy. After downloading the modeling strategy, do not make any changes to this file.

Uploading Modeling Strategies

The modeling strategy must be packaged in a ZIP file. The contents must include the same files that the download package includes (code.sas, validation.xml, metadata.json, and template.json) in the root directory.

If the modeling strategy has the same name as a modeling strategy that is already installed to The Exchange, the name of the uploaded modeling strategy is appended with a number in parentheses. If you want to replace a modeling strategy
on the Exchange, it must be removed first before you upload the new modeling strategy.

To upload a modeling strategy, go to The Exchange and click the **Upload** option. When the upload operation is complete, the time stamp is displayed by the modeling strategy.

For more information about using The Exchange, see "The Exchange" in SAS Visual Forecasting: Overview.

---

**See Also**

"Modeling Strategies Take a Long Time to Complete" on page 175

---

**Generating Forecasts**

**Running the Forecast Pipeline**

Before running a pipeline, be certain that your settings for variable assignments and project settings are correct. After you run one or more pipelines successfully and open the Overrides tab, many of these settings are committed and cannot be changed. Before running the complete pipeline, you can run just the data node and review the results.

1 Right-click the Data node and select **Run**.

   When the Data node runs, it performs these tasks with the project data:
   - accumulates the input time series
   - generates the attributes for descriptive statistics
   - generates the attributes used for filtering and segmenting project data based on demand classification
   - merges the attributes from the descriptive statistics and demand classification with the attributes from the input data for the project

   If the input data for the project is very large, this process can take a long time.

2 When the Data node has finished processing successfully, right-click the Data node and select **Time series viewer**.

   The Time series viewer shows the plot of the historical data. The data is accumulated using the time interval settings that you specified in the variable assignments. The project attributes are listed on the left. If you need to make any final changes to these settings, make those changes before running the full pipeline.

   When you are certain that the variable assignment and project settings are good, click **Run Pipeline**.
While the pipeline is running, you can create another pipeline or exit the project. If you remain on the Pipelines tab long enough, a message is displayed when the pipeline completes.

If there are any problems running the pipeline, check the logs of the modeling nodes. To open the log, right-click a modeling node and select Log. The log shows the code for the node along with any notes, errors, and warnings.

**Note:** If you notice the modeling nodes running slow or taking a long time to complete, read “Modeling Strategies Take a Long Time to Complete” on page 175 for actions that you can take to improve the performance.

---

# Changes That Invalidate Pipelines

Some changes that you make to your project can make your pipelines out of date. When this happens, the pipelines must be run again to pick up the changes. All pipelines must be run again to correctly reflect any of the following changes:

- changing the forecast settings in the Project Settings dialog box.
- changing or refreshing the data source for a time series or external forecast project, or imported attributes. Any change to the data source must follow the same schema as the original.
- adding attributes to the project using the Attributes view of the Data tab.
- adding or removing any independent variables or changing their settings.
- any changes to events in the project.
- changing the modeling nodes in a pipeline. This includes adding or removing nodes, or changing certain options for a node. Only the affected pipeline needs to be run again.

After you run a pipeline and open the Overrides tab, variable assignments and attribute properties are committed and cannot be changed in the project. For more information, see “Locking the Data Definition” on page 83.

If you need to make any changes to the committed settings, for example, setting a different time interval for the time variable, you can start a new project using the same data source.

---

# Working with Segmented Pipelines

SAS Visual Forecasting provides two types of segmented pipeline templates. These pipelines cannot be modified by adding or removing nodes. The nodes in these pipelines cannot be added to other pipelines.

- **Demand Classification**

  Demand Classification can be used with any time series project. The segments are based on demand patterns in each time series. For more information, see “Demand Classification” on page 100.
External Segmentation

External Segmentation can be created only for a project using defined segments from imported attributes. For more information, see "Segmenting Project Data" on page 22.

Creating Segmented Pipelines

Follow these steps to create a pipeline for segmented data. To use the External Segmentation pipeline, the attributes must already be imported with defined segments. See "Defining Project Segments" on page 23 for more information.

1. From the Pipelines tab, click to create a new pipeline.
   This opens the New Pipeline dialog box.

2. For Name, enter a meaningful name or accept the default.

3. Click the drop-down menu for Template and select Browse templates.

4. In the template list, select the type of segmentation template to use. Choose either External Segmentation or Demand Classification and click OK.

5. Click Save on the New Pipeline dialog box.

The segmented pipeline is added to the Pipelines tab.
Examine the structure of the segmented pipeline. Instead of modeling strategies and Model Comparison, segmented pipelines have the following nodes.

**Profile nodes**
- The profile node for the External Segmentation pipeline is named External Segmentation Profile. The profile node for the Demand Classification pipeline is named Demand Classification Profile.

**Modeling nodes**
- The modeling node for the External Segmentation pipeline is named External Segmentation Modeling. The modeling node for the Demand Classification pipeline is named Demand Classification Modeling.

**Merge Segments**
- This node is different from Model Comparison. In a segmented pipeline, you can view model comparisons only for each segment. Merge Segments collects the results from all of the segments. To view this information, right-click Merge Segments and select Results. The view is similar to the results of a modeling strategy or Model Comparison. The only difference is no champion model or OUTMODELINFO table is provided.

  For a description of the results, see “Viewing the Results for a Modeling Strategy Node” on page 44.

Segmented pipelines cannot be modified by adding or removing nodes. You can open these nodes to customize the segments. Each segment has its own nested pipeline. For more information, see “Customizing Each Segment” on page 60.

---

**Customizing Each Segment**

You must run the profile node before you can open the profile or modeling nodes to view and make changes to the segments.

**Working with the Profile Node**

Right-click the profile node for your pipeline and select Open.

The profile view includes the Advanced Filter pane and a table listing each segment.
Advanced Filter
You can filter the displayed segments by segment name or by the number of time series in each segment.

Segment table
The table listing the segments shows the segment name and the number of time series within each segment. For External Segmentation, the segment names come from the _SEG_ value in the data source.

- In the Demand Classification Profile, there are 11 segments. The name is determined by patterns detected in the time series and classified by the _DEMAND_CLASS attribute.
- In the External Segmentation Profile, there can be any number of segments. The segments are listed based on the value of the _SEG_ attribute, which is where they are defined in the data source.

1 To change the name of a segment, right-click the segment and select Rename. The name that you provide is used for this pipeline only.

2 When you are finished making changes, click Close.

Overview of the Modeling Node
Right-click the modeling node for your pipeline and select Open. The segments are listed in the table on the right. Each segment in the pipeline has its own nested pipeline.

Advanced Filter
You can filter the segments shown based on the following columns:

- _DEMAND_CLASS or _SEG_
- Number of Time Series
- Status
- MAPE

Use these tools to customize the attributes displayed in the Advanced Filter.

Opens Filter Settings for selecting which attributes should be displayed in the Advanced Filter, and in what order.

Expands each attribute so that all values are visible.

Collapses each attribute so that no values are visible.

These attributes are explained below.

Segment table
The table listing the segments shows following information for each segment.

Segment Name
By default, the name of the segment comes from the source. For External Segmentation, the source is the value of the _SEG_ variable from the imported attributes. For Demand Classification, the source is the value of the _DEMAND_CLASS variable, based on the characteristics of the time series.

The segment name can be changed in the profile node.
_SEG_ or _DEMAND_CLASS
The original value of the segment name. This name cannot be changed in the pipeline.

Number of Time Series
This column shows the distribution of time series for all of the segments.
It is common to have some empty segments in the Demand Classification pipeline.

Pipeline
This is the nested pipeline that is assigned to each segment.

<table>
<thead>
<tr>
<th>External Segmentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Each pipeline is assigned using the <strong>Default pipeline</strong> setting for the External Segmentation Modeling node. If the value is changed after running the pipeline, it does not change the assigned pipeline for each segment. To change the assigned pipeline to the new default setting, select one or more segments and click</td>
</tr>
<tr>
<td>After successfully running the segmented pipeline, if you reset the default pipeline, the modeling node and the following nodes show a status of <strong>out-of-date</strong>. If you open the modeling node, the individual segments show a status of <strong>Completed</strong>. However, the nodes that are <strong>out-of-date</strong> should be run again.</td>
</tr>
</tbody>
</table>

**Note:** When importing projects with segmented pipelines from SAS Visual Forecasting 8.3, any customized segments in the pipelines are shown with Auto-forecasting listed as the pipeline. This is incorrect. However, if you right-click the segment and select **Open pipeline**, it correctly shows the customizations that you made in the original project.

<table>
<thead>
<tr>
<th>Demand Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Each pipeline is determined by the patterns detected in each time series, using the best template for each classification. This can be changed to any other template, including any custom templates available for your site. You can also open the pipeline to modify the settings to the modeling strategy, or add other modeling strategies.</td>
</tr>
<tr>
<td>To make specific changes to the pipeline for a segment, see “<strong>Customizing Pipelines for Each Segment</strong>” on page 63. If you make any changes to a segment’s pipeline, the pipeline name indicates <strong>custom</strong> in the modeling node.</td>
</tr>
</tbody>
</table>

**Status**
Each segment has one of the following status indications.

| Created — This value is assigned to all segments when a segmented pipeline is first created. The **Status** cell is empty. |
| Pending — The pipeline is waiting for other processing to complete before it can start execution. |
| Running — The pipeline for the segment is running. |
| Modified — A change has been applied to the pipeline for the segment. The pipeline must be run again. |
| Failed — The pipeline has completed with errors. Open the pipeline for the segment, check the location of the error, and open the log. |
| Canceled — Pipeline execution has been canceled. |
Completed — The pipeline run has completed successfully.

Note: After running one or more segments in the modeling node, segments with zero time series show a status of Completed if the other segments completed successfully. This is necessary for the modeling node to attain a status of Successful.

Subsequent data refreshes can populate the segment with time series. You can open the pipeline for an empty segment and modify it. This changes the segment’s status to Modified. To return the segment’s status to Completed, right-click the segment and select Run pipeline.

Always run segments with zero time series from the modeling node. If you open the segment’s pipeline and run it, the pipeline fails.

Customizing Pipelines for Each Segment

Any changes you make to the pipeline for a segment changes the segment’s status to Modified. The pipeline must be run again.

To make changes to the pipeline for each segment, follow these steps.

1 Right-click the External Segmentation Modeling or Demand Classification Modeling node for your pipeline and select Open.

2 To choose another pipeline template for the segment, right-click the segment and select Replace pipeline. The Browse template dialog box is opened and you can select the replacement template for the segment.

Note: You can replace multiple pipelines that you have selected by using ．

Some pipeline templates in The Exchange are not available for use in a segment pipeline. In External Segmentation Modeling, you can replace the pipeline for a segment with Demand Classification. The External Segmentation template is not available as a replacement for any segment. In Demand Classification modeling, segmented pipelines are not available for replacement.

3 To customize the pipeline for the segment, right-click the segment and select Open pipeline. The nested pipeline for the segment is opened. You can add and remove nodes to this pipeline.

a To add other modeling strategies to the pipeline, right-click the Data node, select Add child node \ Forecast Modeling.

b To remove a modeling strategy, right-click the node and select Delete.

c To change how each modeling strategy uses events, right-click the node and select Modify event usage.

d When you are finished making customizations, click Close. You are returned to the modeling view and the list of segments. The segment’s status changes to Modified.

4 To run a pipeline for a segment, right-click the segment and select Run pipeline.
5 After making changes to the pipeline for a segment in External Segmentation Modeling, you can change it back to the default pipeline by right-clicking the segment and selecting **Reset to default**.

---

**Note:** You can reset multiple pipelines that you have selected to use the default by using **|**.

---

For Demand Classification, do not use **Reset to default**. This sets the pipeline to Auto-forecasting. The default assignments for each segment is listed in “Demand Classification” on page 100.

6 When you are finished making changes to the segments, click **Close** to return to the segmentation pipeline.

---

### Selecting a Champion Pipeline

Before you work on any overrides for your forecasting project, you need to make sure that you are working with the best pipeline for your data. SAS Visual Forecasting selects the best fit model in each pipeline and in each segment of a pipeline. After each pipeline is run, the champion pipeline is selected based on the statistics of fit that you choose for the selection criteria. If necessary, you can change the selected champion pipeline.

Changes to the project or other settings force your pipelines and the selected champion pipeline to become out of date. Make sure all pipelines are up-to-date and the best fit model and pipeline is selected as champion before working with overrides.

---

### Model Comparison

Each pipeline with a modeling strategy has a Model Comparison node. When this node has completed successfully, the Model Comparison node has a green check mark indicating a status of Successful. Right-click the node and select **Results** to view the output of the node. The results display a Summary and Output Data tab for the pipeline.

For segmented pipelines, the modeling strategies and Model Comparison are included in the nested pipelines for each segment.

---

### Summary Tab

The following information is available on the Summary tab for Model Comparison Results.

**Model Comparison**

- **Champion** - The champion modeling strategy is marked with **|**. If you determine that another model should be selected as champion, you can manually select the champion model for the pipeline. To select a champion model, right-click the Model Comparison node in the pipeline and select the model from the **Select champion model** list. If you choose a different
To select a champion model, you need to rerun the Model Comparison and Output nodes.

### Model Comparison Results

<table>
<thead>
<tr>
<th>Model Name</th>
<th>Status</th>
<th>WMAE</th>
<th>WMAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seasonal Model</td>
<td>Successful</td>
<td>2,355.2218</td>
<td>6.6232</td>
</tr>
<tr>
<td>Auto-forecasting</td>
<td>Successful</td>
<td>2,733.3461</td>
<td>7.7334</td>
</tr>
<tr>
<td>Time Series Regression</td>
<td>Successful</td>
<td>26,965.7462</td>
<td>82.5465</td>
</tr>
<tr>
<td>Hierarchical Forecasting</td>
<td>Successful</td>
<td>3,446.2891</td>
<td>8.8222</td>
</tr>
</tbody>
</table>

- **Model Name** - Each model in the pipeline or segment is listed.
- **Status** - Indicates the status of each modeling node in the pipeline or segment. If any models are out of date, rerun the pipeline before selecting the final champion pipeline and working with overrides.
- **Weighted metrics** - These columns display the selection criteria that are available for comparing different pipelines. Not all columns are shown. To view all of the selection criteria, click `X`.

Note: For projects using the External Forecast pipeline, if the values are missing from the ACTUAL column, the weighted metrics will have missing values in the Model Comparison results. See “Working with External Forecast Projects” on page 27 for more information about the composition of external forecast projects.

For a complete description, see “Weighted Model Comparison Selection Criteria” on page 162.

### MAPE Distribution

For each model in the pipeline, this is a graph showing the MAPE distribution of results.

### Execution Summary

For each model in the pipeline, the following summary statistics are provided.

- **Number of series**
  
  the total number of time series accumulated from the data. For a Hierarchical Forecasting modeling strategy, this corresponds to the number of time series in each level of the hierarchy.

- **Number of failures**
  
  the number of forecasts that failed. For a Hierarchical Forecasting modeling strategy, this corresponds to the number of failed forecasts in each level of the hierarchy.
Number of forecasts equal to zero
the number of forecasts that are 0

Number of intermittent series with flat forecasts
the number intermittent time series that have a constant forecast value over the length of the forecast horizon

Number of seasonal series with flat forecasts
the number in seasonal time series that have a constant forecast value over the length of the forecast horizon

Number of short series with flat forecasts
the number of short time series that have a constant forecast value over the length of the forecast horizon

Number of all other series with flat forecasts
the number of all other time series that have a constant forecast value over the length of the forecast horizon. This number does not include short, intermittent, or seasonal time series with flat forecast values.

Number of forecasts below 70%
the number of time series with forecast values that are below the minimum range of the mean value for that time series. The forecasts for these series can be problematic. The minimum range value is set to 70% by default. You can change this setting using the Forecast exception boundaries in the Project Settings on page 115.

Number of forecasts above 300%
the number of time series with forecast values that are above the maximum range of the mean value for that time series. The forecasts for these series can be problematic. The maximum range value is set to 300% by default. You can change this setting using the Forecast exception boundaries in the Project Settings on page 115.

Output Data
The tables on this tab reflect the output generated by the current champion modeling strategy in the pipeline or segment. If the table does not initially load, click View Output Data in the content area. For a full description of these tables, see “Output Data” on page 46.

Pipeline Comparison
The Pipeline Comparison tab shows the different pipelines for a project and selects a champion pipeline. The selected champion pipeline, along with the selected model in that pipeline, are used for the final forecasts for the project. When you create overrides for the forecasts, they are based on forecasts generated from the champion pipeline. You do not want to start creating overrides until you are confident that the best pipeline and model has been selected as champion.

If you determine for other reasons that another pipeline should be used as champion, follow these steps to select a different champion.

1 Select the Pipeline Comparison tab.
2 Select a pipeline from the list.
Working with Pipeline Templates

When you create a new pipeline, Model Studio opens a dialog box that you can use to select a template for your pipeline. Pipeline templates are typically distinguished by the models that are included. You can also view the list of all pipeline templates in The Exchange.

Any changes that you make to a pipeline can be saved as a template to The Exchange. This includes any models that you add to the pipeline and any changes to the settings, names, code, or descriptions that you apply to any of the models. To save a pipeline, click \textcolor{red}{\textendash} on the pipeline tab and select \textcolor{red}{Save to The Exchange}.

Pipeline Templates Provided by SAS Visual Forecasting

The following pipeline templates are provided by SAS Visual Forecasting. Some of these templates have specific project requirements that cannot be established when a new project is created. For this reason, make sure you read these descriptions before selecting a pipeline template for a new project.

General Purpose Templates

These templates are good to use for any forecasting project using a time series data source. You will want to compare the different templates and other modeling strategies to find the template that works best for your data.

- Auto-forecasting
- Hierarchical Forecasting
- Naive Forecasting
- Base Forecasting

This template simply creates a pipeline with a single Data node. You will need to add at least one modeling strategy to complete the pipeline. If you create a project selecting an invalid pipeline template, the project is created using the Base Forecasting template.
External Forecasts

This template is valid only for external forecast projects. This template is automatically loaded as the pipeline for any external forecast projects that you create. The schema for external forecasts is explained in “Working with External Forecast Projects” on page 27.

If a new time series project is created with this template, the project uses the Base Forecasting template instead.

Note:

Do not save this pipeline as a new pipeline template. The new template cannot be used in any project.

Segmentation Templates

These templates are intended for forecasting projects that include established segmentation definitions.

- External Segmentation
  
  This template is valid only for projects with defined segments. If you specify this template during project creation, SAS Visual Forecasting cannot determine the segmentation definitions and the default pipeline is set to use the Base Forecasting template instead.

  For more information, see “Defining Project Segments” on page 23.

- Demand Classification
  
  This template segments the time series based on demand classification patterns. The classification is not performed until after project creation is complete. See “Demand Classification” on page 100 for more information about this pipeline template.

Demand Classification Templates

These templates are intended to be used for data that has been segmented based on demand patterns in the time series. When using a Demand Classification pipeline, these templates are selected for use in nested pipelines within each segment, based on the classification for that segment. Some of these templates might be suitable for time series projects. Other templates are not. For example, Retired Forecasting is intended for time series with large periods of inactivity and does not produce any forecasts or measurements for statistics of fit.

- Auto-forecasting (Intermittent)
- Retired Forecasting
- Naive (Moving Average) Forecasting
- Non-seasonal Forecasting
- Regression Forecasting
- Seasonal Forecasting
Setting the Default Pipeline Template

In Model Studio, the default template is always displayed first in the template field when you create a new project or pipeline. You always have the option to select another template instead of using the default.

SAS Visual Forecasting sets the Auto-forecasting pipeline template as the default.

Follow these steps to choose another pipeline template as the default for forecasting projects.

1. Open The Exchange. A list of node and pipeline templates is displayed.
2. Under the Pipelines templates on the left side, select Forecasting.
3. Select the template that you want to use as the default whenever you create a new pipeline.
4. Click and select Set as default template.

Note: The following templates should never be selected as default templates in The Exchange:

- External Forecasts
- External Segmentation
What Is a Filter?

Use filters to subset project data based on specific value combinations of attributes in your project. At the minimum, the default attributes for the project can be used to create filters. You can also import additional attributes to your project. For more information, see "Working with Attributes" on page 15.

For example, consider a project with the following attributes.

- Region
- Product line
- Product

You could create filters using specific values of these attributes:

- Region = District_12
- Productline = Shoes
- Region = District_12 and District_8 and Productline = Swimwear
- Productline = Shoes and Product = Women's Addidas Tennis Shoes

You can use these filters to view the historical and forecast values for data that meets any of these attribute combinations. You can also override forecast values for time series based on these filters.
Creating Filters

One or more default attributes must be assigned to the project before you can create filters. See “Assigning the Default Attributes” on page 8 for more information.

Use Filters on the left pane of the viewer to create and manage filters. The Default attributes are already displayed and ready for you to make your selections. If additional attributes are imported into the project, those attributes and the default attributes are listed under the name of the table that was used for the import.

1. Click next to Attributes to add or remove one of these attribute types from the pane:
   - Descriptive statistics
   - Forecast attributes
   - Demand classification attributes

   Note: For external forecast projects, demand classification attributes are listed on the Data tab. However, they are not available for the project in the Filters pane for the Time series viewer, Forecast viewer, or Overrides.

2. Use next to each attribute type to show, hide, or reorder the attributes that are listed.

3. Select values from the attributes in the list to create a subset of the data to work with. You can select multiple values from any set of attributes, using any combination of values. The plot in the view is updated to reflect the selections that you make. If no time series exists in the set of attributes that you have selected for a filter, then no data is shown.

Making selections from the attributes depends on whether the attribute values are discrete or continuous.

Selecting from attributes with discrete values

You can make selections from attributes with discrete values using check boxes. If the attribute contains more than five values, click the More link to open a dialog box with all of the values.

For example, perhaps you want to filter the project to select only items where the Product Name attribute includes the string Shampoo. Under the Product Name attribute, click More and enter the search string to list all products with Shampoo in the name. You do not need to enter the full word to get matching results.

To select all of the results in the search request, be sure to manually select each item shown. Do not use the Select all field because this selects all values for the attribute.
The plot is updated to reflect the selected products that contain the string *Shampoo*.

**Selecting from attributes with a continuous range of values**

Use the slider to select from numeric data within a range of values.

- The following image shows the slider for a numeric range for one of the statistics of fit attribute, Mean Absolute Error (MAE). This is a derived attribute that is available on the Forecast viewer and Overrides tab. In this figure, the numeric range of MAE measurements for all of the time series ranges between 1.7453 to 10,148.3699. A total number of 1,337 time series are included in this range.

![Mean Absolute Error slider](image)

To set the slide control to a specific number, follow these steps.

a. Click the end of the slider control on one side of the range.

   - For a simple numeric value, an entry field appears over the value and you can enter a number to use for the range. In the following figure, the number 900 is entered at the low end of the range.

![Mean Absolute Error slider with number 900](image)

b. If the attribute has date values, 🗓️ appears instead of an entry field. Click the icon to open the date picker. The date picker enables you to select starting and ending dates that are only within the range of the filter.

![Date picker](image)
Repeat this step for the other side of the range. Using the entry fields, you can enter a precise number for the lower and upper limits of the range that you want to filter.

In the following figure, the range has been set to 900 to 5,000 for the MAE attribute. This changes the number of time series selected from 1,337 to 379.

4 After creating a filter that you can use, click and provide a name for the new filter.

The filter is saved and can be recalled using the drop-down list of saved filters.

As you make selections from the Filters pane, a token is displayed above the plot for your selection.

- For attributes with discrete values, multiple selections can be chosen. If a single value is chosen for an attribute, the token displays the value. In this example, the DIAG1_ARIMAX1 value is selected from the Model attribute of the Forecast attributes type:

- For attributes with a continuous range of values, a single range can be specified for the attribute. The token for the selected range shows the attribute with the range selected. In this example, a range of 10 to 200 is selected from the Mean Absolute Percent Error attribute of the Forecast attributes type.

See Also

“Filters” on page 77
Working with Saved Filters

Use the Filter drop-down menu to select filters that have been saved. The page is updated to show the plot and time series that match the filter selections.

- If you are using the Time series viewer, any filters that were created using forecast attributes are disabled.
- Any changes that you make to the attributes in the filter can invalidate the filter. For example, consider an attribute with the name Size. If Size is used in a filter and then the attribute is deactivated on the Data tab, the filter is no longer valid.

In this situation, you can recover the filter by going to the Data tab, selecting the attribute, and turning on the Activate attribute setting. If you no longer need the filter, click to open Manage Filters and remove the filter.

As shown in this image, sometimes the filters that are saved are disabled and cannot be selected from the drop-down menu. These are described in “Disabled Filters” on page 75.

See Also

“Filters” on page 77

Disabled Filters

Sometimes a filter that has been saved cannot be selected from the drop-down list of the Filters pane. In this case, you also cannot apply the filter or view its properties from the Manage Filters dialog box. This can be caused by any of the following circumstances.

Working with Filters in the Time Series Viewer

The Time Series viewer can access attributes created from the project's input data, the descriptive statistics, and the demand classification. If a filter is created using forecast attributes, you cannot invoke that filter from the Time Series viewer. Go to the Forecast viewer to work with that filter.
Changing Attribute Settings on the Data Tab

There are two settings on the Data tab for attributes that originated from the project's input data.

**Activate attribute**
- When this setting is turned on, the attribute is available in the Filters pane and can be used for creating filters. If you turn this setting off after saving the filter, the filter is no longer valid. You can turn the setting back on if you need to continue to use the filter. If you no longer need the filter, you can delete it using Manage Filters (図).

After running a pipeline and opening the Overrides tab, the *Activate attribute* setting is locked if it is set to Yes. If any attributes have this set to No, they can still be changed after opening the Overrides tab.

**Display as range**
- This setting determines whether attribute values are selected for filters using a continuous value (range) or discrete value (check boxes). Changing *Display as range* invalidates any filters created with that attribute. You can turn the setting back to the previous value if you need to continue to use the filter. If you no longer need the filter, you can delete it using Manage Filters (図). After running a pipeline and opening the Overrides tab, the *Display as range* setting is locked.

Incompatible Settings in a Modeling Strategy

Some attribute values are derived from settings from a modeling strategy.

For example, many modeling strategies enable you to select which models should be considered for the time series in the project. If you run the Auto-forecasting node with ARIMA, ESM, and UCM all selected for model generation, the result generates the number of time series that are selected to use these models for forecasting. In the Forecast viewer for Auto-forecasting, you can create filters using the *Model family* attribute.
The same pipeline could include a modeling strategy that does not support the Model family attribute, such as Panel Series Neural Network. In this case, the filter is still valid for certain modeling strategies, but is disabled in the Forecast viewer for Panel Series Neural Network.

You can create overrides using model-specific filters. However, if you change the champion pipeline to one that does not support that filter, then submitted overrides using that filter are moved to Archived status.

Note: In general, be aware that unexpected results can sometimes happen when you save filters based on model-specific forecast attributes. See “Saved Model Filter Conflicts with Other Model Attributes” on page 176 for more information.

---

**Viewers for Time Series in Your Project**

Use the Time Series viewer to examine the envelope plot of the historical data for your project. You can filter the display using attributes that have been created for the project, along with descriptive statistics and demand classification attributes. The Time Series viewer is available from the Data node in a pipeline.

Use the Forecast viewer to examine the envelope plot of the historical data and the forecasted values for your project. You can filter the display using attributes that have been created for the project, along with descriptive statistics, demand classification, and forecast attributes. The Forecast viewer is available from any modeling strategy node in the pipeline. For a segmented pipeline, you need to open the nested pipeline for each segment to access the Forecast viewer.

**Opening the Time Series and Forecast Viewers**

Use the Data node to open the Time Series viewer. Use the modeling strategy node to open the Forecast viewer.

1. Right-click the node in the pipeline and select **Run**.
2. When the node completes successfully, choose one of the following actions.
   - To access the Time Series viewer, right-click the Data node again and select **Time series viewer**.
   - To access the Forecast viewer, right-click the modeling strategy node again and select **Forecast viewer**.

The viewers are divided into three panes.

**Filters**

The Filters pane provides a list of attributes that you can use to filter the display of time series. The attributes listed include project variables that have been assigned...
as attributes, and attributes derived from characteristics in the time series and forecast values of the project. If no attributes are assigned for the project, this pane is blank.

Filters is available in three locations:

- On the left pane of the Time series viewer. This view is accessed from the Pipelines tab. After the Data node has been run, right-click the node and select **Time series viewer**. Forecast attributes derived from modeling strategies are not available on the Time series viewer.
- On the left pane of the Forecast viewer. This view is accessed from the Pipelines tab. After the modeling strategy has been run, right-click the node and select **Forecast viewer**.
- On the left pane of the Overrides view of the Overrides tab. For more information, see **Overrides on page 84**.

Use the **Attributes** list to browse the attributes of your data set and create filters based on specific values of those attributes. Filters enable you to work with a subset of the data set. The plot and data on the page are updated as you make your attribute selections.

To make it easier to work with a large list of attributes, the attributes are grouped in the Filters pane by the attribute types. See “**Attributes** on page 13” for a description of the attribute types.

When the Filters pane is initially rendered, the attributes setup by your project are displayed and the derived attributes are hidden by default. This can be changed using the **Display attribute by default** setting. For more information, see “**Selecting Attributes for Filters** on page 16.”

See “**Creating Filters** on page 72” for more information about creating filters.

Note:
There can be a discrepancy between the number of distinct values displayed for an attribute and the corresponding number shown in the Attributes view of the Data tab. In this case, the value on the Data tab is based on the formatted value of the variable. The values shown in the Overrides view are based on the raw, unformatted values of the variable.

Use the following to work with filters:

**Filters**
- Use this drop-down list to select the name of a filter that you have already saved
Saves your selections from the Attributes list as a new filter for subsequent use. You are prompted to provide a unique name for the filter.

Opens Manage Filters. Use Manage Filters to delete, rename, or change the properties of filters that have been saved.

See Also

“Working with Filters” on page 71

Plots

When a viewer is opened, the middle panel shows an envelope plot of the dependent variable for the full time range of the project data. For the Time series viewer, the range is the historical data only. For the Forecast viewer, the range is the historical data plus the forecasts. Click to select or deselect which envelope plots to display. This can be helpful when working with plots for individually selected time series.

By default, the plot reflects the aggregate data from all of the time series in the project. The plot changes to reflect any filters that you select from the left pane or time series that you select from the right pane.

Working with Plots

Follow these steps to explore the plots that are available in the Time series viewer and Forecast viewer. For the Forecast viewer, additional actions are described in “Working with Plots in the Forecast Viewer” on page 80.

1 Click in the upper right corner of the plot.
   The overview axis is displayed below the plot.

2 Use the sliders at each end of the plot to zoom the plot to any range within the time series.

3 When you are finished using the overview axis, you can remove it by clicking again. The plot returns to show the full length of all of the time series.

4 Examine the legend below the plot. The legend describes how the mean, first and second standard deviation, and the upper and lower boundaries (Range) are depicted in the plot.

5 Hold your pointer over the mean line in the plot. The plot displays the date, mean, standard deviations, and range for that data point.
6 Select one of the time series from the right pane. A time plot of the selected time series is added to the envelope plot.

Use Ctrl + click or Shift + click keyboard combinations to select up to 16 time series from the right pane. For each time series that you select, its time plot is added to the envelope plot. See "Time Series Selector" on page 82 for more information about selecting individual time series.

7 Click ↓ to download the plot data.

The data is downloaded to your hard drive in CSV format. The downloaded data includes the following information for each time series.

- Minimum and maximum values by date (using SAS numeric value)
- Two standard deviation range by date (using SAS numeric value)
- One standard deviation range by date (using SAS numeric value)
- Mean by date (using SAS numeric value)

8 To view the plot in a larger frame, click ⬠. Click × or use the Esc key to return to the pane view.

9 Take time to explore these features with different filters selected from the left pane. When you are finished, click Close to return to the pipeline.

Working with Plots in the Forecast Viewer

Overview

The Forecast viewer enables you to customize the time plots for individual series that are selected from the right pane. If envelope plots are displayed, the plots for each selected time series are superimposed over them. The following image shows the Forecast viewer with two time series selected and the envelope plots removed.
The Actuals, Predicted, and Confidence limit buttons over the plot are used to customize the display of any time series that you select from the right pane. By default, the plot shows the actual values from the time series, the time plot of the predicted values from the model, and the range of the confidence limit in the forecast horizon. Using these buttons, you can select or deselect which of these measures to display.

Modeling Tab

The Forecast viewer also includes a tab along the top for the following modeling strategies.
- Auto-forecasting
- Hierarchical Forecasting
- Hierarchical Forecasting (Pluggable)

The Modeling tab lists all of the models included in the modeling strategy and the MAPE statistic for a time series that you select in the right pane. The selected champion model is marked with the indicator.

The Modeling tab is empty for projects upgraded from versions of SAS Visual Forecasting before 8.4. This is because the earlier version does not generate the data sets required to populate the table.
Note: When you navigate back to the Forecast tab, the series is still selected. Click Selected Series to deselect it.

See Also

“Time Series Selector” on page 82

Time Series Selector

The right pane displays the list of time series from the selected filter. When the viewer is first opened, all time series for the project are shown. With each selection that you make in the list, the time plot for the selection is added to the middle pane. The time series descriptor, using the BY group values that define the time series, is added to the legend for the plot, along with a color key for each time plot.

For legibility reasons, you can select only up to a maximum of 16 time series.

You can multi-select or remove selected time series by holding down the Ctrl key and clicking each selection.

After a large number of time series is selected, the listing legend might collapse as it runs out of space to display all of your selections. To view the entire list of selected time series in the legend, click the collapsed icon next to Selected Series under the plot.

To find more information about any single time series, select it in the list and click . The descriptive statistics for the selected time series are displayed.
Locking the Data Definition

There are two important events that occur when you open Overrides.

- When you first open Overrides after running a pipeline, the data definition for the project is locked. The data definition includes the role assignments for the time, dependent, and BY variables, and the settings for imported attributes. The assignments for independent variables are not locked. After the role assignments are committed, you can no longer change these settings for this project.

  Make sure you are satisfied with the variable assignments before opening Overrides for the first time. After the data definition is locked, if you determine that you need to change these role assignments, you can create a new project using the same data source with different settings.

- Every time Overrides is opened, it checks for any changes to the champion pipeline output. If updates are found, you are prompted to refresh the overrides data. If you refresh the data, any applied overrides must be resubmitted and any pending overrides are deleted.

  If you have made any changes to BY variables or attributes, you should select Yes to update the data in Overrides.
After Overrides has been opened and the data definition is locked, you can still work with pipelines, adding or removing new models, or changing settings for the models. After completing your changes, make sure all of your pipelines show a status of successful and the best champion pipeline is selected before returning to the Overrides tab and refreshing the data.

Overrides Overview

Use Overrides to override the forecasts from the champion pipeline. The Overrides tab consists of two views.

Overrides

TheOverrides view is divided into three sections.

Filters

The Filters pane enables you to create and manage filters. See “Working with Filters” on page 71 for more information about how to work with filters.

Time Series Aggregation and Overrides Plot

The Time Series Aggregation and Overrides Plot provides a graphical view of your historical and forecast data. You can focus the plot on a subset of the time series by selecting Show overview axis. This displays a slider that you can use to set a beginning and ending date for the plot.

You can also use the scroll wheel to zoom in or out of the plot time line. When you are zoomed into a portion of the time line, you can scroll the view horizontally or vertically as follows.

- Hold your cursor over the bottom of the plot to display a horizontal scroll bar so that you can view time periods before or after those that are in view.
- Hold your cursor over the right side of the plot to display a vertical scroll bar so that you can view values above or below those that are in view.

Forecasts and Overrides Table

This table shows the historical data for the time series and forecasts for the designated forecast horizon. Use this table to create, edit, and submit, override values for a time series, based on external factors that are not included in the forecast models. For time series within the forecast horizon, the table shows the statistical forecast values from the champion pipeline along with the lower and upper confidence limits (Forecast LCL and Forecast UCL).

In the Override row, right-click an empty cell for a time period to open the menu. Depending on the status of any overrides, the following menu options are available. These options are also available from the menu icon over the table for a selected cell.
Override Calculator
Use this dialog box to set override specifications on a filter for one or more time periods. This action is not available for a time period that already has an applied lock on the current filter. For complete information, see “Override Calculator” on page 92.

Impact Analysis
Open this view on a cell to see information about overrides from other filters in the same time period and their impact on any override you might create. This action is available when overrides from other filters are applied on the same time period. Impact Analysis is not available if another filter has a locked statistical forecast for the same time period.

See “Impact Analysis” on page 87 for more information.

Submit All
After creating pending overrides, this action is available on any cell in the Overrides row. All pending overrides that you have created for this filter are submitted.

Delete Pending Override
This action is available on a cell with a pending override. Only the selected override is deleted.

Remove statistical forecast lock
This action is available for any cell with a statistical forecast lock, regardless of whether the lock has been submitted. After clicking this action, the forecast reverts to the original forecast value.

Override Management
Use this view to edit overrides that have already been submitted or to resolve conflicts between overrides. You can also remove overrides that have already been applied or resubmit overrides that have been impacted by other activity.

Filter
Provides a way to show a subset of the overrides based on the values in the table. This is useful when working with more overrides than can be viewed in the height of the window. For columns with character values, you can filter the results based on any string in a cell. This includes the formatted date values in the Time Period column.

For numerical values in the Forecast, Override, Number of Series, and Final Forecast columns, you can filter based on the first digits in the cells. For example, entering 10 can display rows with values like 104,663.00 or 1,000. For cells with numerical values that do not start with 10 (for example, 510.00), these rows are not displayed when the filter value is 10.

Including the comma separator for numerical values is optional.

View menu
Click this menu to select the overrides to display, based on the override status.

View pending
View overrides that have been created but not submitted, or those that need to be resubmitted (Resubmit Pending).

View applied
View overrides that have been successfully submitted.
View conflicts
View overrides that have conflicts with other overrides for one or more time periods. Click the conflict link to view details about the conflict and resolution options. See “Working with Override Conflicts” on page 90 for more information.

View archived
View overrides that are no longer in the current forecast horizon. Archived overrides cannot be edited or removed. There are two types of archives.

- **Archive - Expired** — The data has been refreshed and overrides that were created in the forecast horizon are now included in the historical data.

- **Archive - Blank** — The forecast horizon has been reset so that the last time interval ends on a date that precedes the time period for the override. The override is no longer valid for this filter and time period.

**Note:** If a filter becomes disabled after it is used to create an override, the status of that override becomes Archive. For more information, see “Incompatible Settings in a Modeling Strategy” on page 76.

View active
View all overrides except for those that are archived.

View all
View all overrides, regardless of status.

Use the following actions to manage overrides. All of these actions are also available when you right-click in the Override Management table.

**Submit Pending**
Submits all pending overrides listed in Override Management. This performs the same action as submitting pending overrides in the Forecasts and Overrides table. After reconciliation, the overrides are moved to applied status if no conflicts are found. If conflicts are found in any overrides, they are moved to conflict status and you must resolve the conflict before you can proceed any further with overrides.

**Undo Last Submitted**
The last submitted overrides are in Applied status. Click this button to move the last submitted overrides to Pending status. After performing this action, you can either delete or resubmit specific pending overrides.

- If the last submitted override has been archived, this action is disabled.
- If you have any overrides with Pending status, they must be removed before you can undo the last submission.

If you have selected a single override in the table, this opens the Override Calculator with the settings for the selected override. You can make changes to these settings using the Override Calculator. For more information, see “Override Calculator” on page 92.

**Delete**
Deletes one or more selected overrides.

**Open Menu**
Opens a menu with the following actions:
Impact analysis
Open this view on a selected pending or applied override to see information about overrides in the same time period. For more information, see "Impact Analysis" on page 87.

Automatically resolve all conflicts
If one or more override conflicts are listed, select this option to let SAS Visual Forecasting resolve the conflicts for you.

Resubmit overrides
Under some circumstances, applied overrides must be resubmitted. This can occur when certain applied overrides are removed, conflicts are resolved, or override data has been refreshed.

View overrides logs
Opens the output log for the overrides listed in Override Management. Click Download Logs to download the log to your local drive. You can select any transaction group on the left, or you can view and download all override logs.

Impact Analysis
Use Impact Analysis to view information about overrides from other filters and their impact on your current filter and time slice selection either before or after submitting an override. It can be used to review the impacts of previous overrides in history or inform a decision on an upcoming set of overrides. It shows the number of time series that intersect with the selected override. It also shows the difference in statistical forecast and final forecast as denoted by the Delta value. Click on any filter link to view the override specification in the Overrides tab or click on a transaction group to view the override in context on the Override Management tab.

Working with Overrides

Overview
An override is used to provide adjustments based on the statistical forecasts and assess the impact of overrides on the final forecast results. Use the Overrides view to create, edit, or delete pending overrides. When you have finished working with an override, you can submit. After submission, the override is moved to applied state and the Final Forecast for that time period reflects the reconciled value.

The statistical forecasts in the time series in Overrides are based on the champion pipeline selected from all of your pipelines. All pipelines should be finished running and the champion pipeline should be selected before you start working with overrides. If a new champion pipeline is selected after you start working with overrides, you are prompted to refresh the overrides data and resubmit any overrides that have been applied.
Creating Overrides

Before you can add any overrides, you must first run one or more pipelines and a champion pipeline must be selected. Any changes you make to the project settings can invalidate the selected champion pipeline. In this case, the pipelines must be run again and a new champion must be selected. When you are certain that the final champion pipeline is selected, you are ready to start working on applying overrides to the statistical forecasts from that pipeline. See “Running the Forecast Pipeline” on page 57 for more information.

When you are specifying overrides, it is important to remember the following:

- You can enter overrides only for the time periods in the forecast horizon. The forecast horizon is the period of time for which forecasts are computed. It starts after the end of the historical data and continues for the number of periods specified by the Number of forecast periods (horizon) option in the Project Settings dialog box. The end date for the series is the largest time ID value with a nonmissing value for the dependent variable.

- You cannot add a negative override if negative forecasts are not allowed in the project. To allow negative forecasts, you must select the Allow negative values for forecasts and overrides check box in the Overrides pane of the Project Settings dialog box. Changes to this setting require you to rerun the pipelines.

- If you create one or more overrides for one filter, you must submit those overrides before you can create overrides on another filter.

- Override values cannot be pasted from the clipboard into the override cell. You can paste the value into the Set to a value field in the Override Calculator.

Follow these steps to add overrides:

1. In SAS Visual Forecasting, select the Overrides tab for the project.

2. Make your selections from the Filters pane on the left. You can select from saved filters in the drop-down list or you can make new selections from the attributes listed below. For example, you could select a single product name if you want to review the forecasts and submit any overrides for that product. Or you might select several store locations within a focused geographical location that often have similar results. See “Creating Filters” on page 72 for more information about saving filters.

   If you do not select any filters, you create an override for all time series in the project.

3. To add an override for any forecasted time series, right-click the Override row for that time series and select Override Calculator. For more information about using this tool to create overrides, see “Override Calculator” on page 92. You also have the option of entering a specific value directly in that table cell.

   For each override, you can specify whether the override is locked or unlocked. This locking determines how SAS Visual Forecasting treats the override during the reconciliation process. For more information, see “Lock and Unlock Overrides” on page 89.

4. If you are creating an override for an unnamed filter, select the Filter tab and enter a name for the filter.
5 When you have finished specifying overrides for a filter, click and select **Submit All**.

When you submit a pending override, SAS Visual Forecasting reconciles the override value with other overrides and constraints imposed on that time period. If no conflicts are found, the override's status is applied and the final forecast shows the reconciled value. If conflicts are found, you must resolve them as described in "Working with Override Conflicts" on page 90.

### Override Status

After you have created and submitted new values for overrides, the cells change depending on the status of the override. The following are status indicators shown on table cells to describe the status of an override:

**Pending override:**

After you create an override, the cell in the Override row displays the override value and the indicator. When you are ready for one or more pending overrides to be calculated in the final forecast, right-click the cell and select **Submit All**.

If you set an aggregate final forecast lock on this time period, the cell is disabled. You cannot enter an override value directly in the cell or delete the pending override, but you can still open the Override Calculator and make changes while the status is in pending.

**Applied override:**

After you submit an override, the override is reconciled and, if no conflicts are found, the status becomes applied. The applied value is shown on the Final Forecast row for that time period.

If the **Aggregate final forecast lock** is turned on for an applied override, the cell in the Override row is disabled.

**Locked statistical forecast:**

If the **Statistical forecast lock** is turned on for a time series, the value is shown as locked in the Forecast row. The icon remains in the Overrides cell until the statistical forecast lock is applied.

**Applied locked statistical forecast:**

After a locked statistical forecast is applied, the status is indicated in the Final forecast row. The value shown matches the value in the Forecast row and the Override cell is disabled.

### Lock and Unlock Overrides

You can choose to lock an override value or the statistical forecast for a time period.
A locked override is a user-supplied value for a forecast that is honored when the hierarchy is reconciled. When you lock an override, SAS Visual Forecasting changes the final forecast to the override value and ensures that the reconciled forecasts add up to the locked override value. If the locked override value cannot be honored, this creates an override conflict that you must resolve, as described in “Working with Override Conflicts” on page 90.

Two types of locks exist in SAS Visual Forecasting:

Statistical forecast lock
The statistical forecast lock preserves the statistical forecast for a time period and filter, ignoring any overrides that might impact the values for that time period. For example, if you lock the statistical forecast for one store in the month of June, any overrides that might apply to the products sold in that store are ignored for the June time period. Those overrides might be applied to other stores for the month of June instead.

Aggregate final forecast lock
The aggregate final forecast is the value calculated for an override that is applied to a time period for a specific filter. The lock is applied to the aggregated value.

To lock an individual override, right-click the Overrides cell and select Override Calculator. Select the lock that you want to use on the right side of the calculator.

If SAS Visual Forecasting cannot resolve any override conflicts, then you can also unlock overrides from the Override Conflicts dialog box. For more information, see “Working with Override Conflicts” on page 90.

Remove Overrides

After you remove an override, you must manually re-enter that override. There is no other way to restore any overrides that you previously specified.

To remove overrides in pending state, right-click the override cell in the Forecasts and Overrides table and select Delete pending override. You can also delete pending overrides in the Override Management tab by selecting the overrides and clicking .

To remove overrides in applied state, select the overrides that you want to remove from the Overrides Management tab and click .

When the operation is complete, any remaining applied overrides are moved to Resubmit pending state. You must resubmit all of these overrides before you can perform any further work with overrides.

Working with Override Conflicts

What Is an Override Conflict?

Override conflicts are detected when you submit pending overrides or resubmit applied overrides. When you submit an override, SAS Visual Forecasting reconciles the override with other overrides and forecasts in the time series. If a conflict is found, you are notified and the conflict must be resolved immediately.

Conflicts can be caused by the following conditions:
Two or more overrides cannot be reconciled for the same time period.

An override falls outside of the **Maximum percentage change of final forecast** setting in the Project Settings.

Any combination of constraints caused by locked statistical or aggregate values, maximum allowed changes, or overrides from filters with overlapping filter specifications.

When the conflict is found, you are prompted to make a choice between **Resolve Automatically** (letting SAS Visual Forecasting resolve the conflict) or **Resolve Manually** (resolving it yourself). You cannot create any new overrides until conflicts are resolved.

**Resolving Override Conflicts Automatically**

If you do not choose to resolve conflicts automatically when you are prompted, you can still choose to take this option using the Overrides Management tab.

Click : over the right corner of the table and select **Automatically resolve all conflicts**. SAS Visual Forecasting resolves the conflicts and updates the table.

**Resolving Override Conflicts Manually**

If you choose to resolve override conflicts yourself, a Conflict dialog box is displayed. You are provided with several resolution options for resolving the conflict.

The following figure shows a conflict with two filters that are already applied with an aggregate final forecast lock.

<table>
<thead>
<tr>
<th>Filter</th>
<th>Time Period</th>
<th>Override Status</th>
<th>Statistical Forecast</th>
<th>Override</th>
<th>Solution 1</th>
<th>Solution 2</th>
<th>Aggregated Final Forecast Lock</th>
<th>Delete</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group B</td>
<td>Sun, Mar 2...</td>
<td>Applied</td>
<td>4,297.5,... 2,477.0,...</td>
<td>2,477.0,...</td>
<td>3,517.5,...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Filter 2</td>
<td>Sun, Mar 2...</td>
<td>Pending</td>
<td>5,884.0,... 10,414,...</td>
<td>7,861.5,...</td>
<td>10,414,...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Filter 1</td>
<td>Sun, Mar 2...</td>
<td>Applied</td>
<td>4,343.7,... 6,515.6,...</td>
<td>6,515.6,...</td>
<td>8,178.0,...</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Solution 1**

This option gives priority to locked overrides that have already been applied. Choosing this solution replaces your pending override value with the best alternative that can be applied.

**Solution 2**

This option is available when you try to apply a locked override that cannot be reconciled with a previously locked value. This solution gives priority to preserving the pending locked override and replaces one or more previously applied locked overrides.

**Update lock(s) and/or delete overrides(s)**

Choose this option to either update a lock or delete an override altogether.
In some cases where the **Optimized Final Forecast** value is the same for both solutions, only **Solution 1** is offered.

After you make your selection and resolve the conflict, go to the Override Management tab and resubmit the overrides.

If you cancel from the Conflict dialog box without resolving the conflict, you can resolve the conflict later using the Override Management tab. The override with the conflict is listed with a conflict link in the Status column. Each conflict in the table is labeled as **Conflict** plus a number. Click the conflict link in the Status column to open the dialog box and resolve the conflict.

<table>
<thead>
<tr>
<th>Filter</th>
<th>Time Period</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filter 2</td>
<td>22Mar2015</td>
<td>Conflict 1</td>
</tr>
<tr>
<td>Group B</td>
<td>22Mar2015</td>
<td>Applied</td>
</tr>
<tr>
<td>Filter 1</td>
<td>22Mar2015</td>
<td>Applied</td>
</tr>
</tbody>
</table>

**Override Calculator**

Use the Override Calculator to create and edit specifications for overrides. Each override is used to update the final forecast for a filter on a given time period.

**Properties**

Use the Properties tab to set the specifications of the override. You can choose from the following specifications.

**Adjust based on an existing forecast value**

Select the forecast value for which you want the override to be evaluated during reconciliation.

- **Statistical Forecast** — Reconcile the override with the statistical forecast as determined by the model pipeline.

- **Final Forecast** — Reconcile the override with the final forecast along with any other overrides or constraints that affect this time period and filter selection.

**Adjustment** — Select an up or down (+ or –) change to the forecast based on a percentage or number of units from the statistical or final forecast value.

**Set to a value**

Provide a specific integer value to be considered for the final forecast value.

**Set to a range**

Specify a lower and upper confidence limit to be considered during reconciliation for the final forecast value.

You can add multiple time periods to your override specifications. Click the + icon in the **Selected time periods(s)** field to move time periods from the list of available items to the list of selected items. When you click **OK**, the time periods are added to the selected field.
Filter

Use the Filter tab to save the selected filters or change the name or description for saved filters.

Name
Provide a name for the filter. This name is used in the drop-down selection of saved filters.

Description
Optional. You can provide more descriptive information about the filter in this field. You can retrieve a filter’s description using these steps.

1. From the Filters pane, click ≥.
2. Select the filter.
3. Click ≥.

Exporting Output Data from a Project

You can export project data that can be used in another project. The exported data is saved as an OUTFOR data set. When this is used to create a new project, that project can be used to view the historical time series and to work with overrides.

There are two places that you can export project data.

Exporting from a Selected Project Pipeline

You can export the project along any pipeline that you choose.

1. Select the Pipeline Comparison tab from your project.
2. Select a pipeline from the list.
3. Click ≥ and select Export output data.
   The Export Data dialog box is displayed.
4. Under Data Sources, select a library to export the data.
5. Enter a table name for the output data set and click Export.

You can now create a new project using the output data as a data source.

Exporting Applied Overrides with Project Data

You can export your project data along with applied overrides. The OUTFOR data set will include an FF column for final forecasts.
1. Select the Overrides tab from your project.

2. Click \( \cdot \) and select **Export all data**.
   
The Export All Data dialog box is displayed.

3. Under **Data Sources**, select a library to export the data.

4. Enter a table name for the output data set and click **Export**.
   
You can now create a new project using the output data as a data source.
How SAS Visual Forecasting Creates Time Series Data

SAS Visual Forecasting creates the time series data through the following process:

1. The data is sorted by the BY variables and the time variable.

2. The data is accumulated to the appropriate time interval if the input is one of the following types:
   - time-stamped data that is recorded at no particular frequency (also called transactional data)
   - data recorded at a higher time interval frequency than needed for forecasting (for example, data recorded on a daily frequency but weekly interval forecast is desired)

3. Your input data set might contain BY variables. If you do not use all of the BY variables in your project, the observations are aggregated.

4. Any gaps in the data are filled in. Gaps appear when there is not an observation for each time period or when the data is not equally spaced. The added observations have the required values of the time variable and the value that you specified for missing values. For more information, see “Missing interpretation” on page 111.
When you create a project, you select the input data set to use, and you must assign variables to the time and dependent variable roles. You can also specify the default attributes (BY variables) and independent variable roles. SAS Visual Forecasting uses this information to create the time series data.

Examples of Input Data Sets

Here are two examples of input data sets:

- This input data set contains monthly sales revenue and price information for the past 12 months. In this example, the time variable is Date, the dependent variable is Revenue, and the independent variables are Avg. Price and Holiday. The variable Holiday indicates the number of holidays during the month.

  **Table 5.1 Monthly Sales for 12 Months**

<table>
<thead>
<tr>
<th>Date</th>
<th>Revenue</th>
<th>Avg. Price</th>
<th>Holiday</th>
</tr>
</thead>
<tbody>
<tr>
<td>JAN2011</td>
<td>18817</td>
<td>26.3</td>
<td>0</td>
</tr>
<tr>
<td>FEB2011</td>
<td>52573</td>
<td>25.3</td>
<td>0</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>DEC2011</td>
<td>44205</td>
<td>20.3</td>
<td>1</td>
</tr>
</tbody>
</table>

- This input data set contains monthly retail sales information for different regions and product categories over the past 12 months. You can use the Region and Product variables to create a hierarchy for the sales forecasts.

  **Table 5.2 Monthly Sales for Different Regions and Categories**

<table>
<thead>
<tr>
<th>Date</th>
<th>Sales</th>
<th>Avg. Price</th>
<th>Holiday</th>
<th>Region</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>JAN2011</td>
<td>355</td>
<td>25.3</td>
<td>0</td>
<td>Region 1</td>
<td>Product 1</td>
</tr>
<tr>
<td>FEB2011</td>
<td>398</td>
<td>25.3</td>
<td>0</td>
<td>Region 1</td>
<td>Product 1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>JAN2011</td>
<td>555</td>
<td>19.8</td>
<td>0</td>
<td>Region 1</td>
<td>Product 2</td>
</tr>
<tr>
<td>FEB2011</td>
<td>390</td>
<td>25.3</td>
<td>0</td>
<td>Region 1</td>
<td>Product 2</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>JAN2011</td>
<td>301</td>
<td>27.1</td>
<td>0</td>
<td>Region 2</td>
<td>Product 1</td>
</tr>
<tr>
<td>FEB2011</td>
<td>350</td>
<td>25.3</td>
<td>0</td>
<td>Region 2</td>
<td>Product 1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Including Future Input Values in the Input Data Source

To include future input (independent) values for the time series in your project, include these future values in the input data set. For example, suppose that you are running a book store. For the next three months, the publisher will sell you copies of a book for $20. Because you know this future input value for the series, you want to use the values from your input data set rather than have SAS Visual Forecasting generate a forecast. Future values for dependent variables must be set to missing.

If future independent variable values are not provided, SAS Visual Forecasting extends the future independent variable series using a smoothing model.

Understanding Time Intervals

Overview

All time intervals must meet the following criteria:

- A discrete time interval has a beginning and an ending SAS date or SAS datetime.
- For SAS date intervals, the ending date is defined as 1 day before the beginning of the next interval.
- For SAS datetime intervals, the ending time is 1 second before the beginning of the next interval.
- All observations with an identifying SAS date or SAS datetime that is between the beginning and the end of the interval $t_i$ correspond to the interval $t_i$.

SAS Visual Forecasting analyzes the variable assigned to the time variable role to detect the time interval of the data. SAS assumes that all of the values in the time variable are either date or datetime values and distinguishes between the values by their magnitude.

For many businesses, their time series data is equally spaced, or any two consecutive indices have the same difference between the time intervals. The following table shows an equally spaced time series with a one-year interval.

<table>
<thead>
<tr>
<th>Date</th>
<th>Sales</th>
<th>Avg. Price</th>
<th>Holiday</th>
<th>Region</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>JAN2011</td>
<td>314</td>
<td>27.2</td>
<td>0</td>
<td>Region 2</td>
<td>Product 2</td>
</tr>
<tr>
<td>FEB2011</td>
<td>388</td>
<td>25.3</td>
<td>0</td>
<td>Region 2</td>
<td>Product 2</td>
</tr>
<tr>
<td>DEC2011</td>
<td>518</td>
<td>20.3</td>
<td>1</td>
<td>Region 2</td>
<td>Product 2</td>
</tr>
</tbody>
</table>
Table 5.3  Equally Spaced Time Intervals

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of Sales</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>42,100</td>
</tr>
<tr>
<td>2006</td>
<td>45,000</td>
</tr>
<tr>
<td>2007</td>
<td>47,000</td>
</tr>
<tr>
<td>2008</td>
<td>50,000</td>
</tr>
</tbody>
</table>

SAS Visual Forecasting accumulates the data into observations that correspond to the interval that you specify. For nontransactional data, you might need to specify the interval and seasonal cycle length if there are numerous gaps (missing values) in the data. In this case, SAS Visual Forecasting supplies the missing values. A validation routine checks the values of the time variable to determine whether they are spaced according to the interval that you specified.

Specifying Time Intervals

Statistical forecasting requires that the historical input data be organized as time series, where successive measurements are taken at points equally spaced in time. SAS Visual Forecasting satisfies this requirement in three ways:

- The input data is already in time series form, and is used without any alteration in the time dimension.
- The input data is already in time series form, but is aggregated over time to a different interval (lower frequency), which you specify.
- The input data is transactional (recorded at no particular interval) and is aggregated over time to form time series at an interval which you specify.

Define the time interval specifications when you first create your project using the Assign Variables dialog box. For more information, see “Assigning Variable Roles” on page 7.

SAS Visual Forecasting represents time intervals using a very flexible notation based on a set of basic intervals and optional settings. These settings include the multiplier (affects the length of the interval) and the shift (affects the starting point of the interval). For most time series, the start, length, and end of every observation is identifiable once the time interval is specified in this way. See “Time Variable” on page 106 for a description of each setting.

For the time variable, SAS Visual Forecasting detects the best default settings for the time interval, multiplier, and shift. They are determined by analysis of the time variable values. For input data that are already in time series form, you can often proceed with these default values. You can change them if you know the interval and disagree with the defaults. You can choose any time interval specification that is of lower frequency than the default setting. For transactional data, choose an interval specification in which most or all time periods encompass times at which transactions were recorded.
The examples in the following table show how the values that you specify for the interval, multiplier, and shift work together. For the interval name using SAS code, the `m.s` suffix indicates the multiplier and shift settings for that time interval. In the examples shown, YEAR2.7 indicates a time interval set to YEAR, multiplier set to 2, and shift set to 7.

**Table 5.4 Time Interval, Multiplier, and Shift Examples**

<table>
<thead>
<tr>
<th>Interval Name (in SAS code format)</th>
<th>Default Starting Point</th>
<th>Shift Period</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>YEARm.s</td>
<td>January 1</td>
<td>Months</td>
<td>YEAR2.7 specifies an interval of every two years with the first month of the year starting in July.</td>
</tr>
<tr>
<td>SEMIYEARm.s</td>
<td>January 1, July 1</td>
<td>Months</td>
<td>SEMIYEAR.3 specifies six-month intervals spanning March-August and September-February.</td>
</tr>
<tr>
<td>QTRm.s</td>
<td>January 1, April 1, July 1 October 1</td>
<td>Months</td>
<td>QTR.3 specifies three-month intervals starting on March 1, June 1, September 1, and December 1.</td>
</tr>
<tr>
<td>SEMIMONTHm.s</td>
<td>First and 16th of each month</td>
<td>Semimonthly periods</td>
<td>SEMIMONTH2.2 specifies intervals from the 16th of one month through the 15th of the next month.</td>
</tr>
<tr>
<td>MONTHm.s</td>
<td>First of each month</td>
<td>Months</td>
<td>MONTH2.2 specifies February-March, April-May, June-July, August-September, October-November, and December-January of the following year.</td>
</tr>
<tr>
<td>TENDAYm.s</td>
<td>First, 11th, and 21st of each month</td>
<td>Ten-day periods</td>
<td>TENDAY4.2 specifies four 10-day periods starting at the second 10-day period.</td>
</tr>
</tbody>
</table>
### Interval Name (in SAS code format)  
### Default Starting Point  
### Shift Period  
### Example

<table>
<thead>
<tr>
<th>Interval Name (in SAS code format)</th>
<th>Default Starting Point</th>
<th>Shift Period</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>WEEKm.s</td>
<td>Each Sunday</td>
<td>Days</td>
<td>WEEK6.3 specifies six-week intervals starting on Tuesdays.</td>
</tr>
<tr>
<td>DAYm.s</td>
<td>Each day</td>
<td>Days</td>
<td>DAY3 specifies three-day intervals starting on Sunday.</td>
</tr>
<tr>
<td>HOURm.s</td>
<td>Start of the day (midnight)</td>
<td>Hours</td>
<td>HOUR8.7 specifies eight-hour intervals starting at 6:00 a.m., 2:00 p.m., and 10:00 p.m.</td>
</tr>
</tbody>
</table>

### Demand Classification

Demand classification analyzes demand patterns to improve forecast accuracy. It uses analytical and statistical methods to classify demand patterns based on synchronized internal and external time series data. The process for demand classification involves classifying demand patterns, pattern grouping, and volume grouping.

For example, a grocery store might sell both regular types and holiday types of candy, which would have different demand patterns. You might segment some brands of regular candy as products with a long time span that sells all year round; and segment products like Valentines' day chocolates as a seasonal product with a short time span, which sells only around Valentines’ day. To generate a modeling strategy that produces accurate demand forecasts, products must be segmented appropriately, based on their demand patterns. Demand classification provides an automatic and data-driven way to handle segmentation and modeling.

Demand patterns in each time series are classified based on evaluating the demand life cycle, intermittency, and seasonality. The time series are analyzed and segmented using default attributes, which are based on the BY variables defined for the project.

SAS Visual Forecasting provides a Demand Classification pipeline template. This template includes a Demand Classification node that segments the time series in the project based on these demand classification patterns:

**SHORT**  
Time series with a short record of historical data. This could be a new series with only a few observations. The Naive (Moving Average) Forecasting pipeline is selected for this segment. Moving average is already selected as the naive model type.
LOW_VOLUME
Time series with low volumes. The Naive Forecasting pipeline is selected for this segment. Seasonal random walk is already selected as the naive model type.

INSEASON_INTERMITTENT
Short time span series with intermittent patterns. The Time Series Regression pipeline is selected for this segment.

INSEASON_NON_INTERMITTENT
Short time span series without intermittent patterns. The Time Series Regression pipeline is selected for this segment.

YEAR_ROUND_INTERMITTENT
Long time span series with intermittent patterns. The Auto-forecasting model (Intermittent) pipeline is selected for this segment. Only the IDM model is selected for inclusion.

YEAR_ROUND_SEASONAL
Long time span series with seasonal patterns. The Seasonal Forecasting pipeline is selected for this segment.

YEAR_ROUND_NON_SEASONAL
Long time span series without seasonal patterns. The Non-seasonal Forecasting pipeline is selected for this segment.

YEAR_ROUND_SEASONAL_INTERMITTENT
Long time span series with seasonal and intermittent patterns. The Temporal Aggregation Forecasting pipeline is selected for this segment. Moving average is already selected as the naive model type.

YEAR_ROUND_OTHER
Long time span series with no patterns that can be classified. The Naive (Moving Average) Forecasting pipeline is selected for this segment. Moving average is already selected as the naive model type.

OTHER
Time series that do not span long time periods and cannot be classified. The Naive (Moving Average) Forecasting pipeline is selected for this segment. Moving average is already selected as the naive model type.

RETIRED
Time series that are retired or are no longer active. The Retired Series model is selected for this segment.

After time series are moved into their corresponding demand classification segments, each segment is run using the modeling strategies appropriate to their demand classification. See “Customizing Each Segment” on page 60 for instructions to change or edit modeling strategies for any segment.

See Also

"Customizing Each Segment" on page 60
Understanding Hierarchy Reconciliation

When data is organized in a hierarchical fashion, there are often accounting constraints that link the data at different levels of the hierarchy. Typically, for any historical time period, the data in a parent node is either the sum or the average of the data of its child nodes. For example, the total sales of a product by a retail company are the sum of the sales of the same product in all stores that belong to the company. With forecast data, however, time series are often forecast independently at different levels.

As a result, the forecast values do not abide by the constraints that bind the original series. However, you can enforce these constraints by using an after-the-fact process known as reconciliation of hierarchical forecasts. In the hierarchical model, you can select from these reconciliation methods:

Top-down
   aggregates the data from the lowest levels in the forecast and then uses these values to generate the forecasts at the highest level. SAS Visual Forecasting then uses this forecast and the disaggregation method that you specified to reconcile the forecasts for lower levels in the hierarchy. The top-down method enables you to remove the excessive noise from the data at the lower levels of the hierarchy. However, you also might lose some components of the pattern (such as the seasonality) in the forecast.

You can also specify the type of loss function for top-down reconciliation:

- **Difference** - bases the loss function on the root mean square error (RMSE), which results in adjustments that are the (possibly weighted) mean difference of the aggregated child nodes and the parent node.

- **Proportions** - uses a loss function that results in reconciled forecasts that are the (possibly weighted) proportional disaggregation of the parent node.

Bottom-Up
   uses the data at the lowest level of the hierarchy to generate the forecasts. These forecasts are then used to reconcile the forecasts for the higher levels in the hierarchy.

   The bottom-up method enables you to see any patterns (such as seasonality) in the data. However, because you are using the lowest level of the hierarchy as a reference level for the forecasts, you can also have too much noise or randomness in the data. Also, these forecasts might fail because the data at the lowest level of the hierarchy can be sporadic or too sparse.

Middle-out
   aggregates the data from the lower levels and then uses these values to generate the forecasts for the middle level. SAS Visual Forecasting uses the forecasts at the middle level to reconcile the forecasts for both the higher and lower levels. Some hierarchies have more than one middle level, so you need to specify the level that you want to use.

The following example shows how the forecasts are generated for the hierarchy **Region > Product Category > Product Line > Product**, based on the reconciliation method that you choose. Aggregation forecasts for the higher levels in the hierarchy are created based on the aggregation statistic that you select. Disaggregation
forecasts for the lower levels in the hierarchy are created based on the disaggregation method that you select.

**Table 5.5 Reconciliation Methods**

<table>
<thead>
<tr>
<th>Hierarchy Level</th>
<th>Top Down</th>
<th>Middle Out – Product Category</th>
<th>Middle Out – Product Line</th>
<th>Bottom Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region</td>
<td>Forecast</td>
<td>Aggregation Forecast</td>
<td>Aggregation Forecast</td>
<td>Aggregation Forecast</td>
</tr>
<tr>
<td>Product Category</td>
<td>Disaggregation Forecast</td>
<td>Forecast</td>
<td>Aggregation Forecast</td>
<td>Aggregation Forecast</td>
</tr>
<tr>
<td>Product Line</td>
<td>Disaggregation Forecast</td>
<td>Disaggregation Forecast</td>
<td>Forecast</td>
<td>Aggregation Forecast</td>
</tr>
<tr>
<td>Product</td>
<td>Disaggregation Forecast</td>
<td>Disaggregation Forecast</td>
<td>Disaggregation Forecast</td>
<td>Forecast</td>
</tr>
</tbody>
</table>
Reference

Assign Variable Roles ......................................................... 106
  Time Variable ............................................................... 106
  Dependent Variable ...................................................... 109
  BY Variables ............................................................... 112

Time Interval Accumulation Settings .................................. 112

Project Settings ............................................................. 114
  Forecast Settings ......................................................... 114
  Reporting Settings ....................................................... 115
  Overrides Settings ....................................................... 115
  Output Library Settings ................................................. 115

Settings for Modeling Strategy Nodes ................................ 116
  Auto-forecasting Settings .............................................. 116
  Hierarchical Forecasting Settings ................................. 117
  Hierarchical Forecasting (Pluggable) Settings .................. 123
  Multistage Model Settings ............................................ 125
  Panel Series Neural Network Settings ............................ 131
  Stacked Model (NN + TS) Forecasting Settings .................. 139
  Seasonal Model Settings ............................................... 144
  Non-seasonal Model Settings ......................................... 145
  Temporal Aggregation Model Settings ............................. 145
  Regression for Time Series Settings ............................... 147

Derived Attributes .......................................................... 149
  Demand Classification Attributes ................................. 149
  Descriptive Statistics .................................................. 153
  Forecast Attributes ..................................................... 154

Model Selection Criteria .................................................. 158
  What Are Statistics of Fit? ............................................ 158
  Understanding the Formulas and Definitions .................... 158
  Descriptions of Model Selection Criteria ....................... 159

Weighted Model Comparison Selection Criteria .................... 162

Reserved Names for Variables and Models ......................... 163
  Reserved Variable Names ............................................. 163
  Reserved Model Names ................................................ 167

Promoting Projects and Templates from Previous Releases to
Model Studio 8.5 ............................................................. 168
  Promoting Projects and Templates from Model Studio 8.3 .... 168
  Promoting Projects and Templates from Model Studio 8.2 .... 169
  See Also ................................................................. 173
Assign Variable Roles

Use the Data tab to define the roles for the variables in your data source that are required for any forecasting project. The role assignments are committed after you run one or more pipelines and then open the Overrides tab. After the role assignments are committed, you can no longer change these assignments for this project.

The following fields are provided to provide specific settings for each role assignment.

Time Variable

A time variable is required for all forecasting projects. See “Understanding Time Intervals” on page 97 for more information about the other settings for the time variable.

Time variable
If a single candidate time variable is detected in your data source, it is automatically selected. If multiple candidates are detected, you must choose one.

Time interval
Select the time interval that you want to use to collect the data. For each time interval, you can specify the type (such as monthly or weekly), a multiplier, and a shift (the offset for the interval). You can specify a greater time interval than that found in the input data. Smaller intervals generate a large number of observations and should be avoided. Each time interval has its own Shift interval setting and some have specific starting subintervals. For example, by default, a weekly time interval starts on Sunday (Shift = 1), which you can change to another starting day of the week using the Shift setting.

Choose from the following options:

Second
specifies second intervals. The shift interval is in seconds.

Minute
specifies minute intervals. The shift interval is in minutes.
Hour
specifies hourly intervals. The shift interval is in hours.

Day
specifies daily intervals. The shift interval is in days.

Weekday
specifies daily intervals with weekend days included in the preceding weekday. The weekday interval is the same as the day interval, except that the weekend days are absorbed into the preceding weekday. The default weekend days are Saturday and Sunday. There are five weekday intervals in the default calendar week: Monday, Tuesday, Wednesday, Thursday, and the three-day period Friday, Saturday, and Sunday. The shift interval is in weekdays.

Weekend This button is enabled if you select Weekday. Click Weekend if you need to define the days of the week as the weekend. The defaults are Saturday and Sunday.

Week
specifies weekly intervals of seven days. The shift interval is days with the starting subinterval on Sunday.

The days of the week are numbered by the Shift setting as follows:

Table 6.1  Number Values for Days of the Week

<table>
<thead>
<tr>
<th>Value of the Shift</th>
<th>Day of the Week</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sunday</td>
</tr>
<tr>
<td>2</td>
<td>Monday</td>
</tr>
<tr>
<td>3</td>
<td>Tuesday</td>
</tr>
<tr>
<td>4</td>
<td>Wednesday</td>
</tr>
<tr>
<td>5</td>
<td>Thursday</td>
</tr>
<tr>
<td>6</td>
<td>Friday</td>
</tr>
<tr>
<td>7</td>
<td>Saturday</td>
</tr>
</tbody>
</table>

ISO 8601 week
specifies ISO 8601 weekly intervals of seven days. The shift interval is days with the starting subinterval on Monday.

Ten-day
specifies 10-day intervals. Each month consists of three periods. The first period is the 1st through the 10th day of the month. The second period is the 11th through the 20th day of the month. The third period is the 21st through the end of the month. The shift interval is 10-day periods.

Semimonth
specifies semimonthly intervals. SEMIMONTH breaks each month into two periods, starting on the first and 16th days. The shift period is in semimonthly intervals.
Month
specifies monthly intervals. The shift interval is in months.

Retail 4-4-5 month
specifies retail 4-4-5 monthly intervals. The 3rd, 6th, 9th, and 12th months are
five ISO 8601 weeks long with the exception that some 12th months contain leap
weeks. All other months are four ISO 8601 weeks long. R445MON intervals
begin with the 1st, 5th, 9th, 14th, 18th, 22nd, 27th, 31st, 35th, 40th, 44th, and
48th weeks of the ISO year. The shift interval is in retail 4-4-5 months.

Retail 4-5-4 month
specifies retail 4-5-4 monthly intervals. The 2nd, 5th, 8th, and 11th months are
five ISO 8601 weeks long. All other months are four ISO 8601 weeks long with
the exception that some 12th months contain leap weeks. R454MON intervals
begin with the 1st, 5th, 10th, 14th, 18th, 23rd, 27th, 31st, 36th, 40th, 44th, and
49th weeks of the ISO year. The shift interval is in retail 4-5-4 months.

Retail 5-4-4 month
specifies retail 5-4-4 monthly intervals. The 1st, 4th, 7th, and 10th months are
five ISO 8601 weeks long. All other months are four ISO 8601 weeks long with
the exception that some 12th months contain leap weeks. R544MON intervals
begin with the 1st, 6th, 10th, 14th, 19th, 23rd, 27th, 32nd, 36th, 40th, 45th, and
49th weeks of the ISO year. The shift interval is in retail 5-4-4 months.

Quarter
specifies quarterly intervals (every three months). The shift interval is in months.

Retail 4-4-5 quarter
specifies retail 4-4-5 quarterly intervals (every 13 ISO 8601 weeks). Some fourth
quarters will contain a leap week. The shift interval is in retail 4-4-5 months.

Retail 4-5-4 quarter
specifies retail 4-5-4 quarterly intervals (every 13 ISO 8601 weeks). Some fourth
quarters will contain a leap week. The shift interval is in retail 4-5-4 months.

Retail 5-4-4 quarter
specifies retail 5-4-4 quarterly intervals (every 13 ISO 8601 weeks). Some fourth
quarters will contain a leap week. The shift interval is in retail 5-4-4 months.

Semiyear
specifies intervals every six months. The shift interval is in months.

Year
specifies yearly intervals. The shift interval is in months.

ISO 8601 year
specifies ISO 8601 yearly intervals. The ISO 8601 year starts on the Monday on
or immediately preceding January 4 (or on January 4 if it is a Monday). Note that
it is possible for the ISO 8601 year to start in December of the preceding year.
Also, some ISO 8601 years contain a leap week. For more information about
ISO weeks, see Technical Committee ISO/TC 154 (Processes, Data Elements,
and Documents in Commerce, Industry, and Administration) (2004). The shift
interval is in ISO 8601 weeks (WEEKV).

Retail 4-4-5 year
is the same as YEARV except that the shift interval is in retail 4-4-5 months.

Retail 4-5-4 year
is the same as YEARV except that the shift interval is in retail 4-5-4 months

Retail 5-4-4 year
is the same as YEARV except that the shift interval is in retail 5-4-4 months
Multiplier
Specify an integer that is used to multiply the time interval for customizing the interval. For example, if time interval is set to MONTH and multiplier is 2, then the interval used by the project is every two months. This value can be any positive number.

Shift
Specify an integer to shift the starting point for the day that the interval begins. Specify the starting point for the interval. By default, this value is 1. A value greater than 1 shifts the start to a later point within the interval. The unit for the shift depends on the interval. For example, if shift is set to 4 and the time interval is YEAR, this shifts the start of the time interval to the fourth month. In this case, a year is from April 1 through March 31 of the following year. The starting point is dependent on the interval.

For a list of default starting points for each interval, see Table 5.4 on page 99.

Seasonal cycle length
Specify the length of a season in the units set by the time interval. This value is populated automatically if SAS Visual Forecasting can determine the seasonal cycle length from the time variable. However, you can specify a seasonal cycle length other than the default if you want to model a cycle in the data. For example, your data might contain a 13-week cycle, so you need to specify a 13-week seasonal cycle length in SAS Visual Forecasting.

---

Dependent Variable

If only a single candidate dependent variable is detected in your data source, it is automatically selected. If multiple candidates are detected, you must choose one.

Dependent variable
Specify the variable in your data set that you want to forecast. This variable must be numeric.

Hierarchy aggregation
Aggregation is the process of combining data from more than one time series to form a single series. Select the aggregation method that you want to use for all of the time series in each level of the hierarchy. The dependent variable has fewer aggregation options than independent variables. The following examples explain when you might want to use an aggregation method:

- Your data set contains the sales for a group of products and you want to know the total sales for a category.
- Your data contains the price of each product and you want to know the average price for a product line.

Select from one of the following options.

Sum of Values
aggregates the vector values based on the number of values.

\[ a = \bar{Q} \]

Average of Values
aggregates the vector values based on the average of their values.

\[ a = \bar{r} = \frac{1}{\bar{Q}N} \sum_{\bar{q}=1}^{\bar{Q}} r_{\bar{q}} \]
Missing values are ignored in the summation. If \( Q_N = 0 \), then \( a \) is set to missing.

**Note:** Average of values applies only to hierarchical modeling. The plot and table in Overrides shows a dynamic sum of values of the time series captured by the selected filters.

**Time interval accumulation**
Accumulation combines data within the same time interval into a summary value for that time interval.

**Average of Values**
accumulates the vector values based on the average of their values.

\[
a = \bar{r} = \frac{1}{Q_N} \sum_{q=1}^{Q} r_q
\]

Missing values are ignored in the summation. If \( Q_N = 0 \), then \( a \) is set to missing.

**Corrected Sum of Squares**
accumulates the vector values based on their corrected sum of squares.

\[
a = \sum_{q=1}^{Q} (r_q - \bar{r})^2
\]

Missing values are ignored in the summation. If \( Q_N = 0 \), then \( a \) is set to missing.

**Maximum of Values**
accumulates the vector values based on the maximum of their values.

\[
a = \max\{r_q\}_{q=1}^{Q}
\]

Missing values are ignored in the summation. If \( Q_N = 0 \), then \( a \) is set to missing.

**Minimum of Values**
accumulates the vector values based on the minimum of their values.

\[
a = \min\{r_q\}_{q=1}^{Q}
\]

Missing values are ignored in the summation. If \( Q_N = 0 \), then \( a \) is set to missing.

**Number of non-missing values**
accumulates the vector values based on the number of nonmissing values.

\[
a = Q_N
\]

**Number of Missing values**
accumulates the vector values based on the number of missing values.

\[
a = Q_{NMISS}
\]

**Standard Deviation of Values**
accumulates the vector values based on their standard deviation.

\[
a = \sqrt{\frac{1}{Q_N - 1} \sum_{q=1}^{Q} (r_q - \bar{r})^2}
\]

Missing values are ignored in the summation. If \( Q_N \leq 1 \), then \( a \) is set to missing.
**Sum of Values**
accumulates the vector values based on the summation of their values.

\[ a = \sum_{q=1}^{Q} r_q \]

Missing values are ignored in the summation. If \( Q_N = 0 \), then \( a \) is set to missing.

**Uncorrected Sum of Squares**
accumulates the vector values based on their uncorrected sum of squares.

\[ a = \sum_{q=1}^{Q} (r_q)^2 \]

Missing values are ignored in the summation. If \( Q_N = 0 \), then \( a \) is set to missing.

For more information, see “Time Interval Accumulation Settings” on page 112.

**Missing interpretation**
Once the data has been accumulated to form a time series, missing value interpretation is performed. If any time series contain missing values for the dependent or independent variables, you can specify how SAS Visual Forecasting should interpret these missing values. Choose from the following options:

- **First**
  Specifies that missing values are set to the first nonmissing value of all accumulated nonmissing values in the span of the series.

- **Last**
  Specifies that missing values are set to the last nonmissing value of all accumulated nonmissing values in the span of the series.

- **Maximum**
  Specifies that missing values are set to the maximum value of all accumulated nonmissing values in the span of the series.

- **Median**
  Specifies that missing values are set to the median value of all accumulated nonmissing values in the span of the series.

- **Minimum**
  Specifies that missing values are set to the minimum value of all accumulated nonmissing values in the span of the series.

- **Average**
  Specifies that missing values are set to the average value of all accumulated nonmissing values in the span of the series.

- **Missing**
  Specifies missing values to remain missing. Use this option if a missing value indicates an unknown value.

- **Next**
  Specifies that missing values are set to the next period’s accumulated nonmissing value. Missing values at the end of the accumulated series remain missing.

- **Previous**
  Specifies that missing values are set to the previous period’s accumulated nonmissing value. Missing values at the beginning of the accumulated series remain missing.
0
Specifies missing values to be set to zero. This setting is often used for transactional data, because no recorded data usually implies no activity.

BY Variables

Assign this role to one or more variables that should be used as the default attributes of the project. Use attributes to create filters to work with subsets of the project data. You can also order your BY variables for hierarchical forecasting. The default attributes should have discrete values, such as character data or numeric data with a limited set of values.

After you run a pipeline with the hierarchy defined, you can no longer remove variables from the selected list.

For more information, see “Assigning the Default Attributes” on page 8.

The order of the default attributes is also used for hierarchical forecasting. For more information, see “Defining the Hierarchy” on page 11.

Time Interval Accumulation Settings

Specify how you want the data to be accumulated within each time interval. Accumulation combines data within the same time interval into a summary value for that time period. Accumulation can be used in the following situations:

- converting a time series that has no fixed interval into a time series that has a fixed interval (such as hourly or monthly)
- converting a time series that has a fixed interval into a time series with a lower frequency time interval (such as hourly into daily)

Let \( R = \{r_q\}_{q=1}^Q \) be the data vector ordered by the time series occurrence in the data set with respect to the observation index. Let \( q = 1, \ldots, Q \) be the index that represents this ordering. Let \( Q_q \) be the number of nonmissing values and let \( Q_{NMISS} = Q - Q_N \) be the number of missing values in the data vector. Let

\[
\bar{r} = \frac{1}{Q_{NN}} \sum_{q=1}^Q r_q
\]

be the average value of the data vector with the missing values ignored.

The following example accumulates the observation series \( Z^{(N)} = \{z_i\}_{i=1}^N \) to the time series \( Y^{(T)} = \{y_t\}_{t=1}^T \), \( y_t = \text{Accumulate}(Z_t^{(T)}) \), for \( t = 1, \ldots, T \). In this situation, \( R = Z_t^{(T)} \) and \( Q = N_t^{(T)} \) for \( t = 1, \ldots, T \).

Let \( a = \text{Accumulate}(R) \) be this accumulated value for this data vector when the following accumulation methods are applied:

**Average of Values**
accumulates the vector values based on the average of their values.
\[ a = \bar{r} = \frac{1}{Q N} \sum_{q=1}^{Q} r_q \]

Missing values are ignored in the summation. If \( Q_N = 0 \), then \( a \) is set to missing.

**Corrected Sum of Squares**

accumulates the vector values based on their corrected sum of squares.

\[ a = \sum_{q=1}^{Q} (r_q - \bar{r})^2 \]

Missing values are ignored in the summation. If \( Q_N = 0 \), then \( a \) is set to missing.

**Maximum of Values**

accumulates the vector values based on the maximum of their values.

\[ a = \max\{ r_q \} \]

Missing values are ignored in the summation. If \( Q_N = 0 \), then \( a \) is set to missing.

**Minimum of Values**

accumulates the vector values based on the minimum of their values.

\[ a = \min\{ r_q \} \]

Missing values are ignored in the summation. If \( Q_N = 0 \), then \( a \) is set to missing.

**Number of non-missing values**

accumulates the vector values based on the number of nonmissing values.

\[ a = Q_N \]

**Number of Missing values**

accumulates the vector values based on the number of missing values.

\[ a = Q_{NMISS} \]

**Standard Deviation of Values**

accumulates the vector values based on their standard deviation.

\[ a = \sqrt{\frac{1}{Q N - 1} \sum_{q=1}^{Q} (r_q - \bar{r})^2} \]

Missing values are ignored in the summation. If \( Q_N \leq 1 \), then \( a \) is set to missing.

**Sum of Values**

accumulates the vector values based on the summation of their values.

\[ a = \sum_{q=1}^{Q} r_q \]

Missing values are ignored in the summation. If \( Q_N = 0 \), then \( a \) is set to missing.

**Uncorrected Sum of Squares**

accumulates the vector values based on their uncorrected sum of squares.

\[ a = \sum_{q=1}^{Q} (r_q)^2 \]
Missing values are ignored in the summation. If \( Q_N = 0 \), then \( a \) is set to missing.

---

**Project Settings**

Use the Project Settings dialog box to set properties that apply to the project as a whole. The Project Settings are available along the top of each tab by clicking \( \mathcal{O} \).

Review the values selected in Project Settings before you run any pipelines for your project.

**Forecast Settings**

If you have run any pipelines, they will need to be run again after changing these settings.

**Number of forecast periods (horizon)**

 Specify the number of periods to forecast (also called the horizon). For example, if the time interval for this project is set to WEEK, specify 12 to generate 12 weeks of forecasts for the dependent variable.

**Confidence limit**

 Specify the size of the confidence level for the forecasts. By default, this confidence level is 0.05, which is a 95% confidence limit.

**Champion selection criteria**

 Choose the statistics of fit analysis that is used to determine the champion pipeline for this project. The statistics of fit are statistical values that are used to evaluate how well a forecasting model performs by comparing the actual data to the predictions. For a given forecast model that has been fitted to the time series data, the model should be checked or evaluated to see how well it fits or forecasts the data.

 Each modeling node in a pipeline has its own model selection setting to determine the best model to use from within that node. The Champion selection criteria is used to select the best pipeline when more than one pipeline is run.

 For a full description of the statistics of fit options that are available, see "Model Selection Criteria" on page 158.

**Allow negative values for forecasts and overrides**

 Select this option to allow negative values for forecasts and overrides. If you clear this check box, then any negative values in the forecast model are set to 0 and any overrides with negative values are removed.

 This setting is turned off by default for time series projects. For external forecasts, this setting is turned on by default.

 If Missing interpretation for the dependent variable is set to 0, you could still see some negative values for the lower confidence limits (Forecast LCL) in the project time plots.

 Changes to this setting invalidate any existing overrides that have been created. Changes also invalidate the pipelines.
Note: To allow negative overrides, you must also set the Maximum decrease percentage on the Overrides tab to greater than 100%.

---

**Reporting Settings**

If you have run any pipelines, they will need to be run again after changing this setting.

**Forecast exception boundaries**

This range is used to identify forecast values that are problematic. Set the boundary below and above the historical average. Values that fall outside of the boundaries that you set are identified as problematic.

---

**Overrides Settings**

If you have already submitted any overrides, they will need to be resubmitted after you change any of these settings.

**Disaggregation method**

When you submit an override, SAS Visual Forecasting must reconcile the override with other forecasts in the same time period. Specify the default type of disaggregation method for reconciling overrides. Select one of the following methods:

- **Difference**
  
  reconciles the forecast based on equal split of the differences between aggregated forecast of the time series and their individual forecasts.

- **Proportions**
  
  reconciles the forecast based on the proportion of forecast contribution.

**Override filter warning threshold**

A large number of time series in an override can have adverse effects on performance. Use this setting to issue a warning message when an override contains an extremely high number of time series.

**Maximum percentage change of final forecast**

Specify the maximum percentage change that is allowed during reconciliation for final forecasts.

---

Note: Both values, Maximum increase, and Maximum decrease, must be zero or greater.

---

**Output Library Settings**

Browse through the libraries to specify the default location for storing project output data.
Settings for Modeling Strategy Nodes

Auto-forecasting Settings

Model Generation

You can change the following settings in the Options pane of the pipeline. For more information, see “Options Pane” on page 36.

For best results, make sure at least two of these models are selected. If none of the models are selected, an ESM model (ESMBEST) is used for all time series.

Include ESM models
Turn this setting on to include an exponential smoothing model (ESM) for diagnosis.

Include ARIMAX models
Turn this setting on to include an ARIMAX model for diagnosis.

Include IDM models
Turn this setting on to include an intermittent demand model (IDM) for diagnosis.

IDM Settings — This section is available when Include IDM models is enabled.

Sensitivity level for intermittency test
Specify an integer greater than one. This setting is used to determine whether a time series is intermittent. If the demand interval is equal to or greater than this number, then the series is assumed to be intermittent.

IDM method
Select one of the following models:

- Average: requests the extended sample autocorrelation function.
- Best: uses the single smoothing model to fit the average demand component.
- Croston: uses the two smoothing models to fit the demand interval component and the demand size component.

Include UCM models
Turn this setting on to include a UCM model for diagnosis.

Model Selection

Number of data points used in the holdout sample
Enter a positive integer to be used as the size of the holdout sample. The actual holdout sample will be the minimum between this value and the Percentage of data points used in the holdout sample. The default value is zero, which means no holdout sample is used.
Percentage of data points used in the holdout sample
Enter a value between 0 and 100 to specify the percentage of the sample that will be used for the holdout sample. The actual holdout sample will be the minimum between this value and the Number of data points used in the holdout sample. This option is displayed only if Number of data points used in the holdout sample is greater than zero.

Model selection criterion
Choose the statistics of fit to use for selecting the best model in this modeling strategy. For descriptions for each option, see “Descriptions of Model Selection Criteria” on page 159.

Hierarchical Forecasting Settings
You can change the following settings in the Options pane of the pipeline. For more information, see “Options Pane” on page 36.

Task Settings

Set the forecast task
Specify how to forecast the project. These tasks must be run sequentially. For example, the Diagnose task can be run independently, but the Fit task requires that you run the Diagnose task first.

Choose from one of the following tasks:

- **Diagnose**: performs model selection, estimates parameters of the selected model, and produces forecasts. This is the default.
- **Fit**: estimates parameters for the models that you select and then forecasts. No model selection is performed. You must run the Diagnose task successfully before running the Fit task.
- **Forecast**: forecasts using model parameters estimates. You must run the Diagnose task successfully before running the Forecast task.
- **Select**: performs model selection from the models that you select, estimates parameters of the selected model, and produces forecasts. You must run the Diagnose task successfully before running the Select task.
- **Update**: estimates parameters from the models that you select and then forecasts. No model selection is performed. Update differs from Fit in that the estimated parameters are used as starting values in the estimation. You must run the Diagnose task successfully before running the Update task.

Diagnostics Settings
Changes to these settings require that you rerun the pipeline.

Intermittency test
Turn this setting on to perform an intermittency test and use the IDM model for intermittent series
Sensitivity level for intermittency test
Specify an integer greater than one. This setting is used to determine whether a
time series is intermittent. If the demand interval is equal to or greater than this
number, then the series is assumed to be intermittent.

Seasonality test
Turn this setting on to perform seasonality testing for the time series

Sensitivity level for seasonality test
Specify the significance probability value to use in testing whether seasonality is
present in the time series. The value must be between 0 and 1. A smaller value
means that a stronger evidence of a seasonal pattern in the data is required
before seasonal models are used to forecast the time series.

Diagnose independent variable
Select from the following options:
- **Transform**: specifies that a transformation is applied to the time series
- **Trend**: specifies that trend analysis is performed on the time series based on
  the setting for minimum nonmissing values. Incorporation of a trend is
  checked only for smoothing, UCM, and ARIMA models. For the smoothing
  case, only simple smoothing is a non-trend model. For UCM, the absence of
  a slope component qualifies it as a non-trend model. For ARIMA, there must
  be no differencing of the dependent variable for the time series forecasting
  engine to consider it as a non-trend model.
- **Both**: specifies that transformation and trend analysis is performed for the
time series
- **None**: specifies that no transformation or trend analysis is performed for the
time series

Transformation
**Dependent variable transformation**
specify the type of functional transformation:
- **Auto**: Selects between logarithmic or no transformation as determined by the
  model selection criteria
- **Box-Cox**: Box-Cox transformation
- **Log**: Logarithmic transformation
- **Logistic**: Logistic transformation
- **None**: No transformations are processed on the time series
- **Square root**: Square-root transformation

Box-Cox parameter
Specify a number for the exponent, lambda ($\lambda$), which varies from -5 to 5,
exclusive. This setting is enabled only if **Functional transformation**
(dependent) is set to Box-Cox.

Forecast
Specify the forecast method when functional transformation is enabled.
Forecasts can be based on the mean or median. By default the mean value is
provided. This setting is disabled if **Dependent variable transformation** is set
to None.

Minimum number of observations required for a non-mean model
Specify a minimum value that a time series must meet to be fit using the models
in the selection list. Time series that do not meet this minimum value are forecast
as the mean of the observations in the series. This value must be greater than or equal to one. The default value is 2.

**Minimum number of observations required for a trend model**

Specify that a trend model is not fitted to any series with fewer nonmissing observations than the value specified. The value must be an integer that is greater than or equal to 1. The default value is 2. Trend models are not included for any series with fewer nonmissing observations than this value.

Incorporation of a trend is checked only for smoothing, UCM, and ARIMA models. For the smoothing case, only simple smoothing is a non-trend model. For UCM, the absence of a slope component qualifies it as a non-trend model. For ARIMA, there must be no differencing of the dependent variable for it to be considered as a non-trend model.

**Minimum number of seasons required for a seasonal model**

Seasonal models are not included for any series with fewer nonmissing observations than this value, multiplied by the seasonal length. Specify an integer greater than or equal to one.

**Model Generation Settings**

Changes to these settings require that you rerun the pipeline.

For best results, make sure at least two of these models are selected. If none of the models are selected, an ESM model (ESMBEST) is used for all time series.

**Include ARIMAX models**

Turn this setting on to include an ARIMAX model for diagnosis. With this setting on, you can set the control options for ARIMAX model parameter refinement. You can set the order in which events, inputs, or ARIMA components are included in the models. You can also set a significance level between 0 and 1.

The following settings are enabled when ARIMAX is turned on.

**Identification order for input variable and event coefficients**

Select one of the following options for the order in which input or ARIMA components are included in the model.

- Create two models, each of which uses a different identification method for model inclusion
- Identify coefficients for ARMA components before input variables and events
- Identify coefficients for input variables and events before ARMA components

**Refinement**

Specify ARIMA parameter refinement options. These options enable the refinement of insignificant parameters of the final model, identification of the factors to refine, and identification of the order of factors.

**Factor option**: Select from the available options to determine the order for diagnosing model components. For example, if you select ARMA:INPUT, ARMA coefficients are tested before input variable coefficients.

**Significance level**: Enter a number between 0 and 1 that specifies the cutoff value for refining all insignificant parameters

**Outlier Detection Settings**

Use the following for detecting outliers.

- **Specify the criterion for outlier detection** —
Select **Yes** to include detected outliers in a model if the model is successfully diagnosed.

Select **Maybe** to include detected outliers in a model if the model is successfully diagnosed and has a smaller criterion than the model without outliers.

Select **No** to ensure that no outlier detection is performed.

- **Specify the maximum number of outliers** — Provide a nonnegative integer that specifies the maximum number of outliers to include in a model. The actual number of outliers is the minimum number between this value and **Specify the maximum percentage of outliers**.

- **Specify the maximum percentage of outliers** — Provide a number between 0 and 100 that specifies the maximum number of outliers to include in a model as a percentage of the length of the dependent time series. The actual number of outliers is the minimum number between this value and **Specify the maximum number of outliers**.

- **Significance level** — Provide a value between 0 and 1 that specifies the cutoff value for outlier detection.

**Include ESM models**

Turn this setting on to include an ESM model for diagnosis.

With this setting on, specify how the best **ESM model** candidate is chosen, either seasonal (**BESTS**), nonseasonal (**BESTN**), or both (**BEST**).

**Include UCM models**

Turn this setting on to include an UCM model for diagnosis.

**Include external models**

Turn this setting on to include an external model to use for diagnosis.

- **Models from an external data source** - Enter the CAS library and table name for the external model, for example: mycas.externalmodel.

**Note:** If the external model uses a dummy data set, any independent variables declared in the dummy data set must use the same names as the independent variables in the project time series. Otherwise, the external model is not used.

**Number of levels required to use the system-generated ESM models (starting from the lowest level)**

Specify an integer ranging from 1 to the number of levels in the project hierarchy. A bigger number means the higher level in the hierarchy. The default is 0.

For example, if the hierarchy is `regionName ⇒ productLine ⇒ productName`, then the valid value could be one of the following:

- 0 — not applicable
- 1 — apply for productName
- 2 — apply for the productLine and the productName
- 3 — apply for the regionName, the productLine, and the productName

**Include combined models**

Specify whether to combine the selected models other than the external ones
Method for combination
Specify the method for determining the combination weights used in the weighted average of the candidate forecasts in the combination list. Choose from one of these options:

- **Akaike weights using AICC values** — computes the combination weights based on corrected AIC weights. By default, all AICC scored candidate forecasts are combined. Frequently, there is considerable disparity between the weights because of the exponential weighting scheme, so additional arguments are provided to affect the scaling and to cull low-scoring candidates from consideration for computational efficiency.

- **Average** — computes the simple average of the forecasts selected for combination

- **Equality constrained least squares** — computes the combination weights based on a constrained least squares problem to minimize the $\ell_2$ norm of the combined forecast residuals subject to the constraint that the weights sum to 1.

- **Equality constrained, non-negative least squares** — computes the combination weights based on a constrained least squares problem to minimize the $\ell_2$ norm of the combined forecast residuals subject to the constraints that the weights sum to 1 and be nonnegative.

- **Least absolute deviations** — computes the combination weights that result from the ordinary least squares problem to minimize the $\ell_2$ norm of the combined forecast residuals.

- **Non-negative least squares** — is equivalent to **Equality constrained, non-negative least squares** except that the resulting combination weights are not constrained to summing up to 1.

- **Ordinary least squares** — computes the combination weights that result from the ordinary least squares problem to minimize the $\ell_2$ norm of the combined forecast residuals.

- **Ranked weighting** — assigns weights by using the rank of the candidate forecasts when the combination is performed.

- **Scaled RMSE weights** — computes the combination weights based on the RMSE statistic of fit for the forecast contributors. The weights are normalized to sum to 1.

Statistic used for the encompassing test
Specify the encompassing test type. The encompassing test attempts to eliminate from consideration forecasts that fail to add significant information to the final forecast. Select one of the following values.

- **HLN** — uses the Harvey-Leybourne-Newbold (HLN) test to estimate pairwise encompassing between candidate forecasts

- **NONE** — performs no encompassing tests

- **OLS** — uses an OLS-based regression test to estimate pairwise encompassing between candidate forecasts

Encompassing test parameters
These fields are disabled if **Statistic used for the encompassing test** is set to **None**.

- **Significance level for the encompass test** — specifies the encompassing test significance level. The default value is 0.05.
Rank criterion — specifies the forecast combination criterion (statistic of fit) to be used when ranking forecast candidates in the context of the model combination. For descriptions for each option, see “Descriptions of Model Selection Criteria” on page 159.

Missing value interpretation
specifies a method for treating missing values in the forecast combination. In a given time slice across the combination ensemble, one or more combination contributors can have a missing value. This setting determines the treatment of those in the final combination for such time indices.

- **Rescale** — rescales the combination weights for the nonmissing contributors at each time index to sum to 1.
- **Missing** — generates a missing combined forecast at each time index with one or more missing contributors.

**Method to compute the prediction error variance**
specifies the method for computing the prediction error variance series. This series is used to compute the prediction standard error, which in turn is used to compute confidence bands on the combined forecast. Select from the following options.

- **DIAG** — computes the prediction error variance by assuming the forecast errors at time \( t \) are uncorrelated so that the simple diagonal form of \( \Sigma_t \) is used. This is the default method for computing prediction error variance.
- **ESTCORR** — computes the prediction error variance by using estimates of \( \rho_{ij,t} \), the sample cross-correlation between \( e_{i,t} \) and \( e_{j,t} \) over the time span \( t = 1,...,T \), where \( t \) denotes the last time index of the actual series \( y_t \). Of course, this option implies that the error series \( e_{i,t} \) and \( e_{j,t} \) are assumed to be jointly stationary.

**Maximum allowed percentage of missing in-sample data:**
specifies a threshold for the percentage of missing forecast values in the combination estimation region that is used to exclude a candidate forecast from consideration in the final combination. By default, no missing percentage test is performed on candidate forecasts. If specified, the admissible range is 1 to 100.

**Maximum allowed percentage of missing forecast values in the horizon**
specifies a threshold for the percentage of missing forecast values in the combination horizon used to exclude a candidate forecast from consideration in the final combination. By default, no horizon missing percentage test is performed on candidate forecasts. If specified, the admissible range is 1 to 100.

**Model Selection Settings**
Changes to these settings require that you rerun the pipeline.

**Number of data points used in the holdout sample**
Enter a positive integer to be used as the size of the holdout sample. The actual holdout sample will be the minimum between this value and the **Percentage of data points used in the holdout sample**. The default value is zero, which means no holdout sample is used.

**Percentage of data points used in the holdout sample**
Enter a value between 0 and 100 to specify the percentage of the sample that will be used for the holdout sample. The actual holdout sample will be the minimum between this value and the **Number of data points used in the**
holdout sample. This option is displayed only if Number of data points used in the holdout sample is greater than zero.

**Model selection criterion**
Choose the statistics of fit to use for selecting the best model in this modeling strategy. For descriptions for each option, see “Descriptions of Model Selection Criteria” on page 159.

### Reconciliation Settings
Setting the reconciliation level overrides the default setting selected when assigning the BY variable roles for the project. For more information, see “Assigning the Default Attributes” on page 8.

**Specify the reconciliation level**
Specify a positive integer that corresponds to the levels in your hierarchy, as follows:
- Set the value to 0 to perform top-down reconciliation.
- Specify an integer that matches the bottom level in the hierarchy to use bottom-up reconciliation. For example, if the hierarchy consists of two BY variables, specify 2 to perform bottom-up reconciliation.
- Specify an integer that corresponds to any middle level in the hierarchy to use middle-out reconciliation. For example, if the hierarchy consists of four BY variables, specify 2 to generate forecasts for that level and then reconcile forecasts for the upper and lower levels in the hierarchy.

If you leave this value blank, the reconciliation level set on the Data tab for the BY variables is used.

If the value provided does not match the hierarchy (for example, specifying 5 when the hierarchy is only 4 levels deep), then bottom-up reconciliation is performed.

For a description of these reconciliation methods, see “Understanding Hierarchy Reconciliation” on page 102.

**Disaggregation method during top-down disaggregation**
specifies the type of disaggregation method and type of loss function for top-down reconciliation. Select one of the following methods for top-down disaggregation:

- **Difference**
  bases the loss function on the root mean square error (RMSE). This results in adjustments that are the (possibly weighted) mean difference of the aggregated child nodes and the parent node.
- **Proportions**
  uses a loss function that results in reconciled forecasts that are the (possibly weighted) proportional disaggregation of the parent node.

### Hierarchical Forecasting (Pluggable) Settings
You can change the following settings in the Options pane of the pipeline. For more information, see “Options Pane” on page 36.
Set the forecast task

Specify how to forecast the project. These tasks must be run sequentially. For example, the Diagnose task can be run independently, but the Fit task requires that you run the Diagnose task first.

Choose from one of the following tasks:

- **Diagnose**: performs model selection, estimates parameters of the selected model, and produces forecasts. This is the default.
- **Fit**: estimates parameters for the models that you select and then forecasts. No model selection is performed. You must run the Diagnose task successfully before running the Fit task.
- **Forecast**: forecasts using model parameters estimates. You must run the Diagnose task successfully before running the Forecast task.
- **Select**: performs model selection from the models that you select, estimates parameters of the selected model, and produces forecasts. You must run the Diagnose task successfully before running the Select task.
- **Update**: estimates parameters from the models that you select and then forecasts. No model selection is performed. Update differs from Fit in that the estimated parameters are used as starting values in the estimation. You must run the Diagnose task successfully before running the Update task.

**Model Generation**

For best results, make sure at least two of these models are selected. If none of the models are selected, an ESM model (ESMBEST) is used for all time series.

**Include ESM models**

Turn this setting on to include an exponential smoothing model (ESM) for diagnosis.

**Include ARIMAX models**

Turn this setting on to include an ARIMAX model for diagnosis.

**Include IDM models**

Turn this setting on to include an intermittent demand model (IDM) for diagnosis.

- **Intermittency test** — This section is available when **Include IDM models** is enabled.
  - **Sensitivity level for intermittency test** — Specify an integer greater than one. This setting is used to determine whether a time series is intermittent. If the demand interval is equal to or greater than this number, then the series is assumed to be intermittent.

**Include UCM models**

Turn this setting on to include a UCM model for diagnosis.

**Include combined models**

Select **Yes** if you want to combine the selected models for diagnosis.

**Minimum number of observations required for a non-mean model**

Specify a minimum value that a time series must meet to be fit using the models in the selection list. Time series that do not meet this minimum value are forecast as the mean of the observations in the series. This value must be greater than or equal to one. The default value is 2.

**Minimum number of observations required for a trend model**

Specify that a trend model is not fitted to any series with fewer nonmissing observations than the value specified. The value must be an integer that is
greater than or equal to 1. The default value is 2. Trend models are not included for any series with fewer nonmissing observations than this value.

Incorporation of a trend is checked only for smoothing, UCM, and ARIMA models. For the smoothing case, only simple smoothing is a non-trend model. For UCM, the absence of a slope component qualifies it as a non-trend model. For ARIMA, there must be no differencing of the dependent variable for it to be considered as a non-trend model.

**Minimum number of seasons required for a seasonal model**
Seasonal models are not included for any series with fewer nonmissing observations than this value, multiplied by the seasonal length. Specify an integer greater than or equal to one.

**Number of data points used in the holdout sample**
Enter a positive integer to be used as the size of the holdout sample. The actual holdout sample will be the minimum between this value and the **Percentage of data points used in the holdout sample**. The default value is zero, which means no holdout sample is used.

**Percentage of data points used in the holdout sample**
Enter a value between 0 and 100 to specify the percentage of the sample that will be used for the holdout sample. The actual holdout sample will be the minimum between this value and the **Number of data points used in the holdout sample**. This option is displayed only if **Number of data points used in the holdout sample** is greater than zero.

**Model selection criterion**
Choose the statistics of fit to use for selecting the best model in this modeling strategy. For descriptions for each option, see “Descriptions of Model Selection Criteria” on page 159.

**Reconciliation level**
Specify a positive integer that corresponds to the levels in your hierarchy, as follows:

- Set the value to 0 to perform top-down reconciliation.
- Specify an integer that matches the bottom level in the hierarchy to use bottom-up reconciliation. For example, if the hierarchy consists of two BY variables, specify 2 to perform bottom-up reconciliation.
- Specify an integer that corresponds to any middle level in the hierarchy to use middle-out reconciliation. For example, if the hierarchy consists of four BY variables, specify 2 to generate forecasts for that level and then reconcile forecasts for the upper and lower levels in the hierarchy.

If you leave this value blank, the reconciliation level set on the Data tab for the BY variables is used.

For a description of these reconciliation methods, see “Understanding Hierarchy Reconciliation” on page 102.

---

**Multistage Model Settings**
You can change the following settings in the Options pane of the pipeline. For more information, see “Options Pane” on page 36.
Note: This modeling strategy requires a license for SAS Visual Data Mining and Machine Learning.

General Settings

**Highest level modeled in Stage 1**
Specifies an integer indicating the BY variable in the modeling hierarchy (for example, CATEGORY) to use as the high level for forecasts in Stage 1. The integer 0 indicates the top level node in the modeling hierarchy. If no value is specified for this parameter, no high-level forecasts in Stage 1 are generated.

**Lowest level modeled in Stage 1**
Specifies an integer indicating the BY variable in the modeling hierarchy (for example, PRODUCT) to use as the lowest level for forecasts in Stage 1. The integer 0 indicates the top level node in the modeling hierarchy. By default, this parameter value is the BY variable that is the second from the lowest in the hierarchy.

Feature Extraction Settings

Stage 1 and Stage 2
Select the model and settings for the two stages of forecasting.

**Feature extraction model**
Select either a **Regression** or **Neural network** model to extract the features for generating the forecasts.

For Stage 1, you can select **Time series** to generate forecasts without using feature extraction. **Time series** is not available for Stage 2.

**Dependent variable transformation**
Specify **Log** for a logarithmic transformation for the dependent variable or **None** for no transformation.

**Seasonal dummy variables**
Select this setting to generate seasonal dummy variables during feature extraction. The number of seasonal dummy variables corresponds to the **Seasonal cycle length** specified for the time variable.

**Time interval for creating seasonal dummy variables**
Enter a valid time interval value for creating seasonal dummy variables. If this value is left blank, the time interval specified for the time variable is used. For example, if the time variable for the project uses **Week** for the time interval, 52 seasonal dummy variables are generated. If you specify **Month**, then only 12 seasonal dummy variables are generated.

Enter one of the following specification values. Each specification shows the corresponding setting for the Time variable on the Data tab.
Table 6.2  Time interval specifications and corresponding Time variable intervals

<table>
<thead>
<tr>
<th>Time interval specification</th>
<th>Time variable setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>year</td>
<td>Year</td>
</tr>
<tr>
<td>yearv</td>
<td>ISO 8601</td>
</tr>
<tr>
<td>r445yr</td>
<td>Retail 4-4-5 year</td>
</tr>
<tr>
<td>r454yr</td>
<td>Retail 4-5-4 year</td>
</tr>
<tr>
<td>r544yr</td>
<td>Retail 5-4-4 year</td>
</tr>
<tr>
<td>semiyear</td>
<td>Semiyear</td>
</tr>
<tr>
<td>r445qtr</td>
<td>Retail 4-4-5 quarter</td>
</tr>
<tr>
<td>r454qtr</td>
<td>Retail 4-5-4 quarter</td>
</tr>
<tr>
<td>r544qtr</td>
<td>Retail 5-4-4 quarter</td>
</tr>
<tr>
<td>quarter</td>
<td>Quarter</td>
</tr>
<tr>
<td>month</td>
<td>Month</td>
</tr>
<tr>
<td>r445mon</td>
<td>Retail 4-4-5 month</td>
</tr>
<tr>
<td>r454mon</td>
<td>Retail 4-5-4 month</td>
</tr>
<tr>
<td>r544mon</td>
<td>Retail 5-4-4 month</td>
</tr>
<tr>
<td>semimonth</td>
<td>Semimonth</td>
</tr>
<tr>
<td>tenday</td>
<td>Ten-day</td>
</tr>
<tr>
<td>week</td>
<td>Week</td>
</tr>
<tr>
<td>weekv</td>
<td>ISO 8601 week</td>
</tr>
<tr>
<td>weekday</td>
<td>Weekday</td>
</tr>
<tr>
<td>day</td>
<td>Day</td>
</tr>
<tr>
<td>hour</td>
<td>Hour</td>
</tr>
<tr>
<td>minute</td>
<td>Minute</td>
</tr>
<tr>
<td>second</td>
<td>Second</td>
</tr>
</tbody>
</table>
Number of lags for the independent variables
Specify a positive integer for the number of independent variable lags to generate and include as independent variables. For example, setting the value to 3 computes three variables with lagged values for each independent variable defined in the project.

TIP  Setting this field equal to the dependent variable lags enables the neural network to better detect the interactions between the variables.

Regression Model Settings
Hierarchy level indicating BY variable to be used in Stage 1
Specify the level in the hierarchy for BY variables to use in the regression model during feature extraction in the first stage. Specify an integer between 0 and the number of BY variables assigned for this project. If not specified, or if this is set to 0, there will be no BY statement in the regression model.

For best results, this should be set to the higher levels in the hierarchy, for example, 0 or 1.

This setting is available only if Feature Extraction Model is set to Regression for Stage 1.

Hierarchy level indicating BY variable to be used in Stage 2
Specify the level in the hierarchy for BY variables to use in the regression model during feature extraction in the second stage. Specify an integer between 0 and the number of BY variables assigned for this project. If not specified, or if this is set to 0, there will be no BY statement in the regression model.

For best results, this should be set to the lower levels in the hierarchy, for example, 2 or 3.

This setting is available only if Feature Extraction Model is set to Regression for Stage 2.

Neural Network Settings
These settings are enabled only if Feature extraction model is set to Neural network. If you have Neural network enabled for both stages, these settings are used for both stages.

Dependent variable trend
Specify the method to create a dependent variable trend as an independent variable. You can choose a Linear trend or Damped trend. If you select None, no trend variable is created.

ESM Forecast of dependent variable
Specify Yes to use an ESM forecast of the dependent variable as an independent variable.

Number of lags for the dependent variable
Specify a positive integer for the number of dependent variable lags to generate and include as independent variables. Dependent variable lags are required for the neural network to learn the order of the time series.

For lags of missing values in the horizon, the previous forecasted values will be used to generate new lag values and extend the forecast recursively. For example, setting the value to 3 computes three independent variables with lagged values of the dependent variable.
Input standardization
Specify the method that is used to standardize the interval input variables. Select from these options:
- Midrange
- None
- Z-score

Number of hidden layers
Specify 0, 1, or 2 hidden layers to include in the neural network model. If the number of hidden layers is 0, a GLIM model is trained.

You are required to specify the number of neurons and the activation function for each hidden layer.

Layer 1 neurons
Specify an integer between 0 and 100 for the number of neurons in the first hidden layer. This is required if Number of hidden layers is greater than 0.

Layer 1 activation function
Specify the activation function for the first hidden layer. This is required if Number of hidden layers is greater than 0.
Select from these options:
- Exponential
- Identity
- Logistic
- Rectifier
- Sine
- Tanh (hyperbolic tangent)

Layer 2 neurons
Specify an integer between 0 and 100 for the number of neurons in the second layer. This is required if Number of hidden layers is 2.

Layer 2 activation function
Specify the activation function for the second hidden layer. This is required if Number of hidden layers is 2.
Select from these options:
- Exponential
- Identity
- Logistic
- Rectifier
- Sine
- Tanh (hyperbolic tangent)

Include direct connections between the input and output layers
Specify whether direct connections from nodes in the input layer to nodes in the output layer should be included in the neural network. By default, this is not selected. This setting is disabled if Number of hidden layers is 0.

Dependent variable standardization
Specify the method that is used to standardize the dependent variable. Select from these options:
- Midrange
- Std
- None

**Error function**
Specify the error function for the dependent variable output layer. Select one of the following options:

- **Gamma** — Selecting this value disables the **Activation function for the output layer**. The Exponential activation function is used with the Gamma error function.
- **Normal** — When there are no hidden layers, the normal error function is used.
- **Poisson** — Selecting this value disables the **Activation function for the output layer**. The Exponential activation function is used with the Poisson error function.

**Output layer activation function**
Specify the activation function to use on the output layer of the network. If **Error function** is not set to **Normal**, this setting is disabled and the **Exponential** function is used.

Specify the target layer activation function for interval targets. Select from these options:

- Identity
- Sine
- Tanh

**Random seed**
Specify a positive integer to use for generating random numbers to initialize the network.

**Maximum training iterations**
Specify the maximum number of training iterations within each try.

---

**Time Series Forecast Model Settings**
This setting is enabled only if **Feature extraction model** in **Stage 1** is set to **Time series**.

**Include ESM models**
Turn this setting on to include an exponential smoothing model (ESM) for diagnosis.

**Include ARIMAX models**
Turn this setting on to include an ARIMAX model for diagnosis.

**Include IDM models**
Turn this setting on to include an intermittent demand model (IDM) for diagnosis.

- **Intermittency test** — This section is available when **Include IDM models** is enabled.
  - **Sensitivity level for intermittency test** — Specify an integer greater than one. This setting is used to determine whether a time series is intermittent. If the demand interval is equal to or greater than this number, then the series is assumed to be intermittent.
Include UCM models
Turn this setting on to include a UCM model for diagnosis.

Include combined models
Select Yes if you want to combine the selected models for diagnosis.

Minimum number of observations required for a non-mean model
Specify a minimum value that a time series must meet to be fit using the models in the selection list. Time series that do not meet this minimum value are forecast as the mean of the observations in the series. This value must be greater than or equal to one. The default value is 2.

Minimum number of observations required for a trend model
Specify that a trend model is not fitted to any series with fewer nonmissing observations than the value specified. The value must be an integer that is greater than or equal to 1. The default value is 2. Trend models are not included for any series with fewer nonmissing observations than this value.

Incorporation of a trend is checked only for smoothing, UCM, and ARIMA models. For the smoothing case, only simple smoothing is a non-trend model. For UCM, the absence of a slope component qualifies it as a non-trend model. For ARIMA, there must be no differencing of the dependent variable for it to be considered as a non-trend model.

Minimum number of seasons required for a seasonal model
Seasonal models are not included for any series with fewer nonmissing observations than this value, multiplied by the seasonal length. Specify an integer greater than or equal to one.

Model selection criterion
Choose the statistics of fit to use for selecting the best model in this modeling strategy. For descriptions for each option, see “Descriptions of Model Selection Criteria” on page 159.

Model Selection Settings

Number of data points used in the holdout sample
Enter a positive integer to be used as the size of the holdout sample. The actual holdout sample will be the minimum between this value and the Percentage of data points used in the holdout sample. The default value is zero, which means no holdout sample is used.

Percentage of data points used in the holdout sample
Enter a value between 0 and 100 to specify the percentage of the sample that will be used for the holdout sample. The actual holdout sample will be the minimum between this value and the Number of data points used in the holdout sample. This option is displayed only if Number of data points used in the holdout sample is greater than zero.

Panel Series Neural Network Settings
You can change the following settings in the Options pane of the pipeline. For more information, see “Options Pane” on page 36.
Task Settings

Set the forecast task

Specify how to forecast the project. These tasks must be run sequentially. For example, the Diagnose task can be run independently, but the Fit task requires that you run the Diagnose task first.

- **Diagnose**: trains a neural network, computes the results, forecasts the time series, and sends the forecasts to output. This is the default.
- **Fit**: retrains a model with updated data. Parameters required for the fit task are copied from the most recently trained network when the Fit task is run. This includes variable names (inputs, nominals and target), activation functions, and the number of hidden layers and neurons. You must run the Diagnose task successfully before running the Fit task.
- **Forecast**: generates forecasts using input model data. You must run the Diagnose task successfully before running the Forecast task.
- **Update**: retrains a model with updated data. Hyperparameters and initial connection weights are copied from the most recently trained neural network. You must run the Diagnose task successfully before running the Update task.

Feature Generation Settings

To get a good model, enable some of the Feature Generation settings, which includes setting the number of dependent and independent variable lags. The generated features are included as independent variables when training the model.

**Number of lags for the dependent variable**

Specify a positive integer for the number of dependent variable lags to generate and include as independent variables. Dependent variable lags are required for the neural network to learn the order of the time series.

For lags of missing values in the horizon, the previous forecasted values will be used to generate new lag values and extend the forecast recursively. For example, setting the value to 3 computes three independent variables with lagged values of the dependent variable.

**Number of lags for the independent variables**

Specify a positive integer for the number of independent variable lags to generate and include as independent variables. For example, setting the value to 3 computes three variables with lagged values for each independent variable defined in the project.

**TIP** Setting this field equal to the dependent variable lags enables the neural network to better detect the interactions between the variables.

**Seasonal dummy variables**

Select this setting to generate seasonal dummy variables during feature extraction. The number of seasonal dummy variables corresponds to the Seasonal cycle length specified for the time variable.
**Time interval for creating seasonal dummy variables**

Enter a valid time interval value for creating seasonal dummy variables. If this value is left blank, the time interval specified for the time variable is used. For example, if the time variable for the project uses *Week* for the time interval, 52 seasonal dummy variables are generated. If you specify *Month*, then only 12 seasonal dummy variables are generated.

Enter one of the following specification values. Each specification shows the corresponding setting for the Time variable on the Data tab.

*Table 6.3  Time interval specifications and corresponding Time variable intervals*

<table>
<thead>
<tr>
<th>Time interval specification</th>
<th>Time variable setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>year</td>
<td>Year</td>
</tr>
<tr>
<td>yearv</td>
<td>ISO 8601</td>
</tr>
<tr>
<td>r445yr</td>
<td>Retail 4-4-5 year</td>
</tr>
<tr>
<td>r454yr</td>
<td>Retail 4-5-4 year</td>
</tr>
<tr>
<td>r544yr</td>
<td>Retail 5-4-4 year</td>
</tr>
<tr>
<td>semiyear</td>
<td>Semiyear</td>
</tr>
<tr>
<td>r445qtr</td>
<td>Retail 4-4-5 quarter</td>
</tr>
<tr>
<td>r454qtr</td>
<td>Retail 4-5-4 quarter</td>
</tr>
<tr>
<td>r544qtr</td>
<td>Retail 5-4-4 quarter</td>
</tr>
<tr>
<td>quarter</td>
<td>Quarter</td>
</tr>
<tr>
<td>month</td>
<td>Month</td>
</tr>
<tr>
<td>r445mon</td>
<td>Retail 4-4-5 month</td>
</tr>
<tr>
<td>r454mon</td>
<td>Retail 4-5-4 month</td>
</tr>
<tr>
<td>r544mon</td>
<td>Retail 5-4-4 month</td>
</tr>
<tr>
<td>semimonth</td>
<td>Semimonth</td>
</tr>
<tr>
<td>tenday</td>
<td>Ten-day</td>
</tr>
<tr>
<td>week</td>
<td>Week</td>
</tr>
<tr>
<td>weekv</td>
<td>ISO 8601 week</td>
</tr>
<tr>
<td>weekday</td>
<td>Weekday</td>
</tr>
</tbody>
</table>
### Time interval specification

<table>
<thead>
<tr>
<th>Time variable setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>day</td>
</tr>
<tr>
<td>hour</td>
</tr>
<tr>
<td>minute</td>
</tr>
<tr>
<td>second</td>
</tr>
</tbody>
</table>

### ESM Forecast of dependent variable

Specify **Yes** to use an ESM forecast of the dependent variable as an independent variable.

### Dependent variable trend

Specify the method to create a dependent variable trend as an independent variable. You can choose a **Linear trend** or **Damped trend**. If you select **None**, no trend variable is created.

### Model Generation Settings

#### Model Initialization Settings

- **Input standardization**
  Specify the method that is used to standardize the interval input variables. Select from these options:
  - **Midrange**
  - **None**
  - **Z-score**

- **Number of hidden layers**
  Specify 0, 1, or 2 hidden layers to include in the neural network model. If the number of hidden layers is 0, a GLIM model is trained.

  You are required to specify the number of neurons and the activation function for each hidden layer.

- **Layer 1 neurons**
  Specify an integer between 0 and 100 for the number of neurons in the first hidden layer. This is required if **Number of hidden layers** is greater than 0.

- **Layer 1 activation function**
  Specify the activation function for the first hidden layer. This is required if **Number of hidden layers** is greater than 0.

  Select from these options:
  - **Exponential**
  - **Identity**
  - **Logistic**
  - **Rectifier**
  - **Sine**
Layer 2 neurons
Specify an integer between 0 and 100 for the number of neurons in first second layer. This is required if **Number of hidden layers** is 2.

Layer 2 activation function
Specify the activation function for the second hidden layer. This is required if **Number of hidden layers** is 2.

Select from these options:
- Exponential
- Identity
- Logistic
- Rectifier
- Sine
- Tanh (hyperbolic tangent)

Include direct connections between the input and output layers
Specify whether direct connections from nodes in the input layer to nodes in the output layer should be included in the neural network. By default, this is not selected. This setting is disabled if **Number of hidden layers** is 0.

Dependent variable transformation
Specify **Log** for a logarithmic transformation for the dependent variable or **None** for no transformation.

Dependent variable standardization
Specify the method that is used to standardize the dependent variable. Select from these options:
- Midrange
- Std
- None

Error function
Specify the error function for the dependent variable output layer. Select one of the following options:
- **Gamma** — Selecting this value disables the **Activation function for the output layer**. The Exponential activation function is used with the Gamma error function.
- **Normal** — When there are no hidden layers, the normal error function is used.
- **Poisson** — Selecting this value disables the **Activation function for the output layer**. The Exponential activation function is used with the Poisson error function.

Output layer activation function
Specify the activation function to use on the output layer of the network. If **Error function** is not set to **Normal**, this setting is disabled and the **Exponential** function is used.

Specify the target layer activation function for interval targets. Select from these options:
- Identity
Neuron connection distribution
Specify the distribution of randomly generated initial neuron connection weights.
Select from these options:
- Cauchy
- MSRA
- Normal
- Uniform
- Xavier

Random seed
Specify a positive integer to use for generating random numbers to initialize the network.

Model Training Settings

Algorithm
Specify the optimization method used to train the neural network. Select one of the following options:
- LBFGS — limited memory version of Broyden-Fletcher-Goldfarb-Shanno (BFGS)
- SGD — Stochastic gradient descent

SGD Options
If you select the SGD algorithm, the following settings apply.

Learning rate
Specify the learning rate parameter for SGD optimization. The default value is 0.001.

Annealing rate
Specify the annealing rate parameter for SGD optimization. The default value is 0.000001.

SGD seed
Specify the random seed to use for the SGD algorithm.

Input layer dropout ratio
Specify the dropout ratio for the input layer when SGD optimization is used. The default value is 0. Enter a nonnegative number that is less than 1.

Hidden layer dropout ratio
Specify the dropout ratio for the hidden layers when SGD optimization is used. The default value is 0. Enter a nonnegative number that is less than 1.

Number of tries
Specify the number of times to train the network with different initial estimates for connection weights. The network with the smallest error is chosen as the optimal network.

Maximum training iterations
Specify the maximum number of training iterations within each try.
Maximum time (minutes)
Specify in minutes the maximum time allowed for each try. Training continues until the all tries have completed. Zero indicates not to use time to limit the training.

L1 regularization
Specify a positive number for the weight decay for L1 regularization. The default is 0.

L2 regularization
Specify a positive number for the weight decay for L2 regularization. The default is 0.1.

Enable early stopping
Select this option to stop training when the model begins to overfit. The training stops after a number of consecutive iterations without improvement in the holdout region. Set the number of consecutive iterations in Stagnation limit for early stopping.

Stagnation limit for early stopping
Specify the number of consecutive iterations without improvement in validation error before stopping the optimization. Specifying 0 has the same effect as deselecting Enable early stopping.
This field is available only when Enable early stopping is selected.

Autotune Settings
Enable Autotune — Turn this setting on to enable autotuning of neural network parameters. Autotuning is applied only when running the Diagnose task.

Note: Autotuning can significantly increase the amount of time required to train the model.

Note:
If you enable autotune, this modeling strategy fails if Number of tries is set to a number greater than 1.

Hidden layer tuning
Specify whether to autotune the number of hidden layers. If you enable this setting for autotuning, specify the initial value and the lower and upper bounds for the number of hidden layers. This setting overrides Number of hidden layers in the Model Initialization settings.

Hidden neuron tuning
Specify whether to autotune the number of neurons in each hidden layer. Specify the initial value and the lower and upper bounds for the number of neurons. This setting overrides Number of hidden layers in the Model Initialization settings.
This setting is enabled if Hidden layer tuning is enabled.

L1 regularization
Specify whether to autotune the weight decay for the L1 regularization parameter. If you enable this setting for autotuning, specify the initial value and the lower and upper bounds for the weight decay. This setting overrides L1 regularization in Model Training settings.
L2 regularization
Specify whether to autotune the weight decay for the L2 regularization parameter. If you enable this setting for autotuning, specify the initial value and the lower and upper bounds for the weight decay. This setting overrides L2 regularization in Model Training settings.

Learning rate
Specify whether to autotune the learning rate for the hidden layers. If you enable this setting for autotuning, specify the initial value and the lower and upper bounds for the learning rate. This setting overrides Learning Rate in the Model Training settings.

This setting is disabled if the Algorithm under Model Training is set to LBFGS.

Annealing rate
Specify whether to autotune the annealing rate for the hidden layers. If you enable this setting for autotuning, specify the initial value and the lower and upper bounds for the learning rate. This setting overrides Annealing Rate in the Model Training settings.

This setting is disabled if the Algorithm under Model Training is set to LBFGS.

Maximum autotuning iterations
Specify the maximum number of iterations for autotuning. Autotuning continues until all iterations are completed or the Maximum autotuning time (minutes) is reached, whichever is first.

Maximum autotuning time (minutes)
Specify the maximum time allowed for autotuning in minutes. Autotuning continues until this time is reached or the Maximum autotuning iterations are completed, whichever is first.

Autotune seed
Specify the random seed for the autotuning.

Autotune objective function
Specify the objective function to optimize when tuning parameters. You can select one of these options:

- ASE — Average squared error
- MAE — Mean absolute error
- MSE — Mean squared error
- MSLE — Mean squared logarithmic error
- RMAE — Root mean absolute error
- RMSE — Root mean square error
- RMSLE — Root mean squared logarithmic error

Model Validation Settings
The holdout sample is the minimum calculated value between these two settings.

Number of data points used in the holdout sample
Enter a positive integer to be used as the size of the holdout sample. The actual holdout sample will be the minimum between this value and the Percentage of data points used in the holdout sample. The default value is zero, which means no holdout sample is used.
**Percentage of data points used in the holdout sample**

Enter a value between 0 and 100 to specify the percentage of the sample that will be used for the holdout sample. The actual holdout sample will be the minimum between this value and the **Number of data points used in the holdout sample**. This option is displayed only if **Number of data points used in the holdout sample** is greater than zero.

---

**Stacked Model (NN + TS) Forecasting Settings**

You can change the following settings in the Options pane of the pipeline. For more information, see “Options Pane” on page 36.

**Note:** This modeling strategy requires a license for SAS Visual Data Mining and Machine Learning.

---

**Feature Generation Settings**

**Number of lags for the dependent variable**

Specify a positive integer for the number of dependent variable lags to generate and include as independent variables. Dependent variable lags are required for the neural network to learn the order of the time series.

For lags of missing values in the horizon, the previous forecasted values will be used to generate new lag values and extend the forecast recursively. For example, setting the value to 3 computes three independent variables with lagged values of the dependent variable.

**Number of lags for the independent variables**

Specify a positive integer for the number of independent variable lags to generate and include as independent variables. For example, setting the value to 3 computes three variables with lagged values for each independent variable defined in the project.

**TIP** Setting this field equal to the dependent variable lags enables the neural network to better detect the interactions between the variables.

**Seasonal dummy variables**

Select this setting to generate seasonal dummy variables during feature extraction. The number of seasonal dummy variables corresponds to the **Seasonal cycle length** specified for the time variable.

**Time interval for creating seasonal dummy variables**

Enter a valid time interval value for creating seasonal dummy variables. If this value is left blank, the time interval specified for the time variable is used. For example, if the time variable for the project uses `Week` for the time interval, 52 seasonal dummy variables are generated. If you specify `Month`, then only 12 seasonal dummy variables are generated.

Enter one of the following specification values. Each specification shows the corresponding setting for the Time variable on the Data tab.
<table>
<thead>
<tr>
<th>Time interval specification</th>
<th>Time variable setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>year</td>
<td>Year</td>
</tr>
<tr>
<td>yearv</td>
<td>ISO 8601</td>
</tr>
<tr>
<td>r445yr</td>
<td>Retail 4-4-5 year</td>
</tr>
<tr>
<td>r454yr</td>
<td>Retail 4-5-4 year</td>
</tr>
<tr>
<td>r544yr</td>
<td>Retail 5-4-4 year</td>
</tr>
<tr>
<td>semiyear</td>
<td>Semiyear</td>
</tr>
<tr>
<td>r445qtr</td>
<td>Retail 4-4-5 quarter</td>
</tr>
<tr>
<td>r454qtr</td>
<td>Retail 4-5-4 quarter</td>
</tr>
<tr>
<td>r544qtr</td>
<td>Retail 5-4-4 quarter</td>
</tr>
<tr>
<td>quarter</td>
<td>Quarter</td>
</tr>
<tr>
<td>month</td>
<td>Month</td>
</tr>
<tr>
<td>r445mon</td>
<td>Retail 4-4-5 month</td>
</tr>
<tr>
<td>r454mon</td>
<td>Retail 4-5-4 month</td>
</tr>
<tr>
<td>r544mon</td>
<td>Retail 5-4-4 month</td>
</tr>
<tr>
<td>semimonth</td>
<td>Semimonth</td>
</tr>
<tr>
<td>tenday</td>
<td>Ten-day</td>
</tr>
<tr>
<td>week</td>
<td>Week</td>
</tr>
<tr>
<td>weekv</td>
<td>ISO 8601 week</td>
</tr>
<tr>
<td>weekday</td>
<td>Weekday</td>
</tr>
<tr>
<td>day</td>
<td>Day</td>
</tr>
<tr>
<td>hour</td>
<td>Hour</td>
</tr>
<tr>
<td>minute</td>
<td>Minute</td>
</tr>
<tr>
<td>second</td>
<td>Second</td>
</tr>
</tbody>
</table>
ESM Forecast of dependent variable
Specify Yes to use an ESM forecast of the dependent variable as an
independent variable.

Dependent variable trend
Specify the method to create a dependent variable trend as an independent
variable. You can choose a Linear trend or Damped trend. If you select None,
no trend variable is created.

Model Generation Settings

Model Initialization

Input standardization
Specify the method that is used to standardize the interval input variables. Select
from these options:
- Midrange
- None
- Z-score

Number of hidden layers
Specify 0, 1, or 2 hidden layers to include in the neural network model. If the
number of hidden layers is 0, a GLIM model is trained.

You are required to specify the number of neurons and the activation function for
each hidden layer.

Layer 1 neurons
Specify an integer between 0 and 100 for the number of neurons in the first
hidden layer. This is required if Number of hidden layers is greater than 0.

Layer 1 activation function
Specify the activation function for the first hidden layer. This is required if
Number of hidden layers is greater than 0.

Select from these options:
- Exponential
- Identity
- Logistic
- Rectifier
- Sine
- Tanh (hyperbolic tangent)

Layer 2 neurons
Specify an integer between 0 and 100 for the number of neurons in first second
layer. This is required if Number of hidden layers is 2.

Layer 2 activation function
Specify the activation function for the second hidden layer. This is required if
Number of hidden layers is 2.

Select from these options:
- Exponential
Include direct connections between the input and output layers
Specify whether direct connections from nodes in the input layer to nodes in the output layer should be included in the neural network. By default, this is not selected. This setting is disabled if Number of hidden layers is 0.

Dependent variable transformation
Specify Log for a logarithmic transformation for the dependent variable or None for no transformation.

Dependent variable standardization
Specify the method that is used to standardize the dependent variable. Select from these options:
- Midrange
- Std
- None

Error function
Specify the error function for the dependent variable output layer. Select one of the following options:
- Gamma — Selecting this value disables the Activation function for the output layer. The Exponential activation function is used with the Gamma error function.
- Normal — When there are no hidden layers, the normal error function is used.
- Poisson — Selecting this value disables the Activation function for the output layer. The Exponential activation function is used with the Poisson error function.

Output layer activation function
Specify the activation function to use on the output layer of the network. If Error function is not set to Normal, this setting is disabled and the Exponential function is used.

Specify the target layer activation function for interval targets. Select from these options:
- Identity
- Sine
- Tanh

Neuron connection distribution
Specify the distribution of randomly generated initial neuron connection weights. Select from these options:
- Cauchy
- MSRA
- Normal
**Settings for Modeling Strategy Nodes**

- **Uniform**
- **Xavier**

**Random seed**
Specify a positive integer to use for generating random numbers to initialize the network.

**Model Training Settings**

**Algorithm**
Specify the optimization method used to train the neural network. Select one of the following options:

- **LBFGS** — limited memory version of Broyden-Fletcher-Goldfarb-Shanno (BFGS)
- **SGD** — Stochastic gradient descent

**SGD Options**
If you select the SGD algorithm, the following settings apply.

- **Learning rate**
  Specify the learning rate parameter for SGD optimization. The default value is 0.001.

- **Annealing rate**
  Specify the annealing rate parameter for SGD optimization. The default value is 0.000001.

- **SGD seed**
  Specify the random seed to use for the SGD algorithm.

- **Input layer dropout ratio**
  Specify the dropout ratio for the input layer when SGD optimization is used. The default value is 0. Enter a nonnegative number that is less than 1.

- **Hidden layer dropout ratio**
  Specify the dropout ratio for the hidden layers when SGD optimization is used. The default value is 0. Enter a nonnegative number that is less than 1.

**Number of tries**
Specify the number of times to train the network with different initial estimates for connection weights. The network with the smallest error is chosen as the optimal network.

**Maximum training iterations**
Specify the maximum number of training iterations within each try.

**Maximum time (minutes)**
Specify in minutes the maximum time allowed for each try. Training continues until the all tries have completed. Zero indicates not to use time to limit the training.

**L1 regularization**
Specify a positive number for the weight decay for L1 regularization. The default is 0.

**L2 regularization**
Specify a positive number for the weight decay for L2 regularization. The default is 0.1.
Enable early stopping
Select this option to stop training when the model begins to overfit. The training stops after a number of consecutive iterations without improvement in the holdout region. Set the number of consecutive iterations in Stagnation limit for early stopping.

Stagnation limit for early stopping
Specify the number of consecutive iterations without improvement in validation error before stopping the optimization. Specifying 0 has the same effect as deselecting Enable early stopping.

This field is available only when Enable early stopping is selected.

Model Selection Options

Model selection criterion for the second stage model
Specify the model selection criteria for the second stage. For a description of each option, see “Descriptions of Model Selection Criteria” on page 159.

Seasonal Model Settings

You can change the following settings in the Options pane of the pipeline. For more information, see “Options Pane” on page 36.

Model Generation

For best results, make sure at least two of these models are selected. If none of the models are selected, an ESM model (ESMBEST) is used for all time series.

Include ESM models
Turn this setting on to include an exponential smoothing model (ESM) for diagnosis.

Include ARIMAX models
Turn this setting on to include an ARIMAX model for diagnosis.

Include UCM models
Turn this setting on to include a UCM model for diagnosis.

Model Selection

Number of data points used in the holdout sample
Enter a positive integer to be used as the size of the holdout sample. The actual holdout sample will be the minimum between this value and the Percentage of data points used in the holdout sample. The default value is zero, which means no holdout sample is used.

Percentage of data points used in the holdout sample
Enter a value between 0 and 100 to specify the percentage of the sample that will be used for the holdout sample. The actual holdout sample will be the minimum between this value and the Number of data points used in the holdout sample. This option is displayed only if Number of data points used in the holdout sample is greater than zero.
Model selection criterion
Choose the statistics of fit to use for selecting the best model in this modeling strategy. For descriptions for each option, see “Descriptions of Model Selection Criteria” on page 159.

Non-seasonal Model Settings
You can change the following settings in the Options pane of the pipeline. For more information, see “Options Pane” on page 36.

Model Generation
For best results, make sure at least two of these models are selected. If none of the models are selected, an ESM model (ESMBEST) is used for all time series.

- Include ESM models
  Turn this setting on to include an exponential smoothing model (ESM) for diagnosis.

- Include ARIMAX models
  Turn this setting on to include an ARIMAX model for diagnosis.

- Include UCM models
  Turn this setting on to include a UCM model for diagnosis.

Model Selection

Number of data points used in the holdout sample
Enter a positive integer to be used as the size of the holdout sample. The actual holdout sample will be the minimum between this value and the Percentage of data points used in the holdout sample. The default value is zero, which means no holdout sample is used.

Percentage of data points used in the holdout sample
Enter a value between 0 and 100 to specify the percentage of the sample that will be used for the holdout sample. The actual holdout sample will be the minimum between this value and the Number of data points used in the holdout sample. This option is displayed only if Number of data points used in the holdout sample is greater than zero.

Model selection criterion
Choose the statistics of fit to use for selecting the best model in this modeling strategy. For descriptions for each option, see “Descriptions of Model Selection Criteria” on page 159.

Temporal Aggregation Model Settings
You can change the following settings in the Options pane of the pipeline. For more information, see “Options Pane” on page 36.
High Frequency Settings

**Sensitivity level for intermittency test**
Specify an integer greater than one. This setting is used to determine whether a time series is intermittent. If the demand interval is equal to or greater than this number, then the series is assumed to be intermittent.

**IDM method**
Select one of the following models:
- **Average**: requests the extended sample autocorrelation function.
- **Best**: uses the single smoothing model to fit the average demand component.
- **Croston**: uses the two smoothing models to fit the demand interval component and the demand size component.

Low Frequency Settings

**Model Generation**
For best results, make sure at least two of these models are selected. If none of the models are selected, an ESM model (ESM BEST) is used for all time series.

**Include ESM models**
Turn this setting on to include an exponential smoothing model (ESM) for diagnosis.

**Include ARIMAX models**
Turn this setting on to include an ARIMAX model for diagnosis.

**Include UCM models**
Turn this setting on to include a UCM model for diagnosis.

**Model Selection**

**Number of data points used in the holdout sample**
Enter a positive integer to be used as the size of the holdout sample. The actual holdout sample will be the minimum between this value and the **Percentage of data points used in the holdout sample**. The default value is zero, which means no holdout sample is used.

**Percentage of data points used in the holdout sample**
Enter a value between 0 and 100 to specify the percentage of the sample that will be used for the holdout sample. The actual holdout sample will be the minimum between this value and the **Number of data points used in the holdout sample**. This option is displayed only if **Number of data points used in the holdout sample** is greater than zero.

**Model selection criterion**
Choose the statistics of fit to use for selecting the best model in this modeling strategy. For descriptions for each option, see "Descriptions of Model Selection Criteria" on page 159.
Regression for Time Series Settings

You can change the following settings in the Options pane of the pipeline. For more information, see “Options Pane” on page 36.

Feature Extraction

**Observation index**
Select this option to include the observation index to capture linear trend.

**Observation index squared**
Select this option to include the square of the observation index to capture second-order effects.

**Observation index cubed**
Select this option to include the cube of the observation index to capture third-order effects.

**Seasonal dummy variables**
Select this setting to generate seasonal dummy variables during feature extraction. The number of seasonal dummy variables corresponds to the **Seasonal cycle length** specified for the time variable.

**Time interval for creating seasonal dummy variables**
Enter a valid time interval value for creating seasonal dummy variables. If this value is left blank, the time interval specified for the time variable is used. For example, if the time variable for the project uses **Week** for the time interval, 52 seasonal dummy variables are generated. If you specify **Month**, then only 12 seasonal dummy variables are generated.

Enter one of the following specification values. Each specification shows the corresponding setting for the Time variable on the Data tab.

*Table 6.5  Time interval specifications and corresponding Time variable intervals*

<table>
<thead>
<tr>
<th>Time interval specification</th>
<th>Time variable setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>year</td>
<td>Year</td>
</tr>
<tr>
<td>yearv</td>
<td>ISO 8601</td>
</tr>
<tr>
<td>r445yr</td>
<td>Retail 4-4-5 year</td>
</tr>
<tr>
<td>r454yr</td>
<td>Retail 4-5-4 year</td>
</tr>
<tr>
<td>r544yr</td>
<td>Retail 5-4-4 year</td>
</tr>
<tr>
<td>semiyear</td>
<td>Semiyear</td>
</tr>
<tr>
<td>r445qtr</td>
<td>Retail 4-4-5 quarter</td>
</tr>
<tr>
<td>r454qtr</td>
<td>Retail 4-5-4 quarter</td>
</tr>
</tbody>
</table>
### Time interval specification

<table>
<thead>
<tr>
<th>Time interval specification</th>
<th>Time variable setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>r544qtr</td>
<td>Retail 5-4-4 quarter</td>
</tr>
<tr>
<td>quarter</td>
<td>Quarter</td>
</tr>
<tr>
<td>month</td>
<td>Month</td>
</tr>
<tr>
<td>r445mon</td>
<td>Retail 4-4-5 month</td>
</tr>
<tr>
<td>r454mon</td>
<td>Retail 4-5-4 month</td>
</tr>
<tr>
<td>r544mon</td>
<td>Retail 5-4-4 month</td>
</tr>
<tr>
<td>semimonth</td>
<td>Semimonth</td>
</tr>
<tr>
<td>tenday</td>
<td>Ten-day</td>
</tr>
<tr>
<td>week</td>
<td>Week</td>
</tr>
<tr>
<td>weekv</td>
<td>ISO 8601 week</td>
</tr>
<tr>
<td>weekday</td>
<td>Weekday</td>
</tr>
<tr>
<td>day</td>
<td>Day</td>
</tr>
<tr>
<td>hour</td>
<td>Hour</td>
</tr>
<tr>
<td>minute</td>
<td>Minute</td>
</tr>
<tr>
<td>second</td>
<td>Second</td>
</tr>
</tbody>
</table>

### Model Selection

**Number of data points used in the holdout sample**

Enter a positive integer to be used as the size of the holdout sample. The actual holdout sample will be the minimum between this value and the **Percentage of data points used in the holdout sample**. The default value is zero, which means no holdout sample is used.

**Percentage of data points used in the holdout sample**

Enter a value between 0 and 100 to specify the percentage of the sample that will be used for the holdout sample. The actual holdout sample will be the minimum between this value and the **Number of data points used in the holdout sample**. This option is displayed only if **Number of data points used in the holdout sample** is greater than zero.

**Model selection criterion**

Choose the statistics of fit to use for selecting the best model in this modeling strategy. For descriptions for each option, see “**Descriptions of Model Selection Criteria**” on page 159.
Derived Attributes

SAS Visual Forecasting generates attributes for each project. These attributes listed are included in all forecasting projects. The Demand Classification and Descriptive Statistics attributes are generated when the project is created and are available in the Filters pane for all viewers. The Statistics of Fit and Forecast Attributes are available in the Filters pane only for the Forecast viewer and Overrides tab.

Demand Classification Attributes

SAS Visual Forecasting analyzes patterns in the time series for a project and classifies these patterns to generate attributes. Demand classification attributes represent characteristics of times series such as intermittence, volume, volatility, and seasonality. See “Demand Classification” on page 100 for complete descriptions of these demand patterns.

All demand classification attributes have discrete values. Use a check box to select one or more values for creating filters.

In time series projects, you can use these attributes to create filters in all of the viewers. See “Viewers for Time Series in Your Project” on page 77 for more information. The filters are listed in the Filters pane by the label.

Note: For external forecast projects, demand classification attributes are listed on the Data tab. However, they are not available for the project in the Filters pane for the Time series viewer, Forecast viewer, or Overrides.
<table>
<thead>
<tr>
<th>Label</th>
<th>Variable Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume Volatility Class</td>
<td>_VOLUME_VOLATILITY_CLAS S</td>
<td>values that identify different ranges of volume and volatility in the time series. Time series that exhibit high volatility are difficult to forecast accurately and might need special attention to identify any inputs that might generate better results.</td>
</tr>
</tbody>
</table>

In the Filters pane, you can select from the following patterns that have been detected in the time series.

- HIGH_VOLUME_HIGH_VOLATILITY
- LOW_VOLUME_HIGH_VOLATILITY
- HIGH_VOLUME_LOW_VOLATILITY
- LOW_VOLUME_LOW_VOLATILITY
- RETIRED
- SHORT
- INTERMITTENT
- OTHER
### Demand Class

**_DEMAND_CLASS**

A specific set of classifications based on patterns detected in the time series. These classifications are used to segment the time series in the Demand Classification pipeline. See "Demand Classification" on page 100 for more information.

Select from the following possible values for the _DEMAND_CLASS attribute. Some of these values might not show in the Filters pane if no time series meet the criteria for that attribute.

- **SHORT**
  - Time series with a short record of historical data.
  - This could be a new series with only a few observations.

- **LOW_VOLUME**
  - Time series with low volumes

- **INSEASON_INTERMITTENT**
  - Short time span series with intermittent patterns

- **INSEASON_NON_INTERMITTENT**
  - Short time span series without intermittent patterns

- **YEAR_ROUND_INTERMITTENT**
  - Long time span series with intermittent patterns

- **YEAR_ROUND_SEASONAL**
  - Long time span series with seasonal patterns

- **YEAR_ROUND_NON_SEASONAL**
  - Long time span series without seasonal patterns

- **YEAR_ROUND_SEASONAL_INTERMITTENT**
  - Long time span series with seasonal and intermittent patterns

- **YEAR_ROUND_OTHER**
  - Long time span series with no patterns that can be classified

- **OTHER**
  - Time series that do not span long time periods and cannot be classified

- **RETIRED**
  - Time series that are retired or are no longer active

### Seasonal

**_SEASONAL**

Indicates whether the time series is seasonal. Seasonality is determined by a significance probability of 0.01 or less. In the Filters pane, select Y to show only seasonal series or N to show only series that are not seasonal.

Select ND for time series for which seasonal patterns could not be determined.

### Intermittent

**_INTERMITTENT**

Indicates whether the time series is intermittent. In the Filters pane, select Y to show series that indicate intermittent patterns or N for series that are not intermittent.
<table>
<thead>
<tr>
<th>Label</th>
<th>Variable Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seasonal</td>
<td>_SEASONAL_INTERMITTENT</td>
<td>indicates whether the time series is seasonal intermittent. In the Filters pane, select Y to show series that are both seasonal and intermittent or N for series that are not a combination of seasonal and intermittent. Select ND for time series for which seasonal intermittent patterns could not be determined.</td>
</tr>
<tr>
<td>Retired</td>
<td>_RETIRED</td>
<td>indicates whether the time series is retired. In the Filters pane, select Y to show only series that are retired or N for series that are not retired.</td>
</tr>
<tr>
<td>Short</td>
<td>_SHORT</td>
<td>indicates whether the time series is a short period. In the Filters pane, select Y for series that indicate a short pattern or N for series that are not short.</td>
</tr>
<tr>
<td>Volume</td>
<td>_VOLUME</td>
<td>the volume of the time series. In the Filters pane, select from LOW, MEDIUM, or HIGH.</td>
</tr>
<tr>
<td>Volatility</td>
<td>_VOLATILITY</td>
<td>the volatility of the time series. Usually, series with high volatility are harder to automatically forecast than series with low volatility. In the Filters pane, select either LOW or HIGH.</td>
</tr>
<tr>
<td>Demand Span</td>
<td>_DEMAND_SPAN</td>
<td>the length of the time series. In the Filters pane, select any of the following options.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>YEAR_ROUND</strong> for time series with values that spread throughout the year, like sales for basic goods</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>INSEASON</strong> for time series that occur only during certain seasons, such as sales for seasonal goods</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>ND</strong> for time series with values that spread throughout the year, like sales for basic goods</td>
</tr>
<tr>
<td>Trailing Zero</td>
<td>_TRAILING_ZERO_LENGTH</td>
<td>the number of trailing zeros in the time series. In the Filters pane, select from any detected numeric range.</td>
</tr>
<tr>
<td>Length</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum</td>
<td>_MAXIMAL_CYCLE_LENGTH</td>
<td>the maximum value of the demand cycle lengths detected from the time series. In the Filters pane, select from the values detected from time series.</td>
</tr>
<tr>
<td>Cycle Length</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volume</td>
<td>_VOLUME_MEASURE</td>
<td>the actual volume measurement value of the time series. In the Filters pane, select from a numeric range, starting with the lowest detected volume measure in the time series to the highest.</td>
</tr>
</tbody>
</table>
Table 6.6 OUTSUM Statistics

<table>
<thead>
<tr>
<th>Label</th>
<th>Variable Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demand Interval</td>
<td>_DEMAND_INTERVAL_MEASURE</td>
<td>the median value of the demand intervals. This value is used for detecting if the series is intermittent or not. Select from the numeric values detected in the time series. Note: This attribute is listed in the Filters pane as Demand Interval Measure.</td>
</tr>
<tr>
<td>Volatility Measure</td>
<td>_VOLATILITY_MEASURE</td>
<td>the actual volatility measurement value of the time series. In the Filters pane, select from a numeric range, starting with the lowest detected volatility measure in the time series to the highest.</td>
</tr>
</tbody>
</table>
| Status                 | _STATUS_     | a value that indicates whether analysis for each time series was successful. You can select one or more of the following conditions.  
  - 0 - Analysis was successful  
  - 3000 - Accumulation failed  
  - 4000 - Missing value interpretation failed  
  - 6000 - Series is all missing  
  - 9000 - Descriptive statistics could not be computed |

Descriptive Statistics

Each project contains a data set of summary statistics for the dependent variable as determined by the value of the forecast horizon. These summary statistics are stored in the OUTSUM data set. OUTSUM is also displayed in the results for each modeling strategy.

You can use these attributes to create filters in all of the viewers. See “Viewers for Time Series in Your Project” on page 77 for more information. The filters are listed in the Filters pane by the label.
<table>
<thead>
<tr>
<th>Label</th>
<th>Variable Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time Series Length</td>
<td>NOBS</td>
<td>the number of observations, including observations with missing values, in each input time series</td>
</tr>
<tr>
<td>Number of Nonmissing Values</td>
<td>N</td>
<td>the number of observations in each time series that do not have missing values in each input time series</td>
</tr>
<tr>
<td>Number of Missing Values</td>
<td>NMISS</td>
<td>the number of observations in each time series that have missing values. In the Filters pane, you can select among the range of missing values that have been collected from the time series.</td>
</tr>
<tr>
<td>Minimum Value of Series</td>
<td>MIN</td>
<td>the smallest value of observations in each time series. In the Filters pane, you can select among the range of minimum values that have been collected from the time series.</td>
</tr>
<tr>
<td>Maximum Value of Series</td>
<td>MAX</td>
<td>the largest value of observations in each time series. In the Filters pane, you can select among the range of maximum values that have been collected from the time series.</td>
</tr>
<tr>
<td>Mean Value of Series</td>
<td>MEAN</td>
<td>the mean value of observations in each time series. In the Filters pane, you can select among the range of mean values that have been collected from the time series.</td>
</tr>
<tr>
<td>Standard Deviation of Series</td>
<td>STDDEV</td>
<td>the standard deviation of observations in each time series. In the Filters pane, you can select among the range of standard deviation values that have been collected from the time series.</td>
</tr>
<tr>
<td>Condition Code for Series</td>
<td><em>STATUS</em></td>
<td>a value that indicates whether analysis for each time series was successful. You can select one or more of the following conditions.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 0 - Analysis was successful</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 3000 - Accumulation failed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 4000 - Missing value interpretation failed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 6000 - Series is all missing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 9000 - Descriptive statistics could not be computed</td>
</tr>
</tbody>
</table>

**Forecast Attributes**

You can select filters of time series in the Forecast viewer based on forecast attributes. These attributes are generated by each modeling strategy in the pipeline and are not available in the Time series viewer. These summary statistics are stored.
in the OUTMODELINFO data set. OUTMODELINFO is also displayed in the results for each modeling strategy.

Forecast attributes includes the following attribute types.

**Forecast Attributes for Selected Model**

These attributes contain detailed information about the selected forecast model. The following describes each attribute by variable name and variable label. The filters are listed in the Filters pane by the label.

---

**Note:** Some of these attributes are not available in every modeling strategy.

---

**Table 6.7 Forecast Attributes from OUTMODELINFO**

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Variable Label</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>MODEL</em></td>
<td>Model</td>
<td>the name of the model selected for this time series.</td>
</tr>
<tr>
<td><em>MODELTYPE</em></td>
<td>Model Family</td>
<td>the type of model used (ESM, ARIMA, UCM, or IDM)</td>
</tr>
<tr>
<td><em>DEPTRANS</em></td>
<td>Dependent Variable Transform</td>
<td>the type of dependent variable transformation that was used for this time series.</td>
</tr>
<tr>
<td><em>SEASONAL</em></td>
<td>Seasonal model</td>
<td>indication of seasonal model used for this time series. A value of 1 indicates a seasonal model was used.</td>
</tr>
<tr>
<td><em>TREND</em></td>
<td>Trend Model</td>
<td>indication of trend model used for this time series. A value of 1 indicates a trend model was used.</td>
</tr>
<tr>
<td><em>INPUTS</em></td>
<td>Inputs present</td>
<td>the number of input variables used in the model</td>
</tr>
<tr>
<td><em>EVENTS</em></td>
<td>Events Present</td>
<td>indication of whether events were used in the model. A value of 1 indicates events are present.</td>
</tr>
<tr>
<td><em>OUTLIERS</em></td>
<td>Outliers Present</td>
<td>the number of outliers discovered in the model</td>
</tr>
<tr>
<td><em>SOURCE</em></td>
<td>Model Source</td>
<td>the named source of the model</td>
</tr>
<tr>
<td>Variable Name</td>
<td>Variable Label</td>
<td>Description</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------</td>
<td>-------------</td>
</tr>
<tr>
<td><em>STATUS</em></td>
<td>Model Status</td>
<td>the execution status of the model. The following status codes can be present.</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>The forecast completed successfully.</td>
</tr>
<tr>
<td></td>
<td>3000</td>
<td>Model selection could not be completed. Forecast values are set to missing.</td>
</tr>
<tr>
<td></td>
<td>3001</td>
<td>Model selection could not be completed and NOALTLIST prohibits use of default exponential smoothing. Forecast values are set to missing.</td>
</tr>
<tr>
<td></td>
<td>3002</td>
<td>The forecast was completed subject to qualification that one or more input variables were omitted from the selected model. This can occur only in the context of ARIMAX or UCM models.</td>
</tr>
<tr>
<td></td>
<td>3003</td>
<td>The desired model could not be forecast. The forecast reverted to the default exponential smoothing model.</td>
</tr>
<tr>
<td></td>
<td>3004</td>
<td>The attempt to forecast the desired model produced an arithmetic exception. The forecast is generated by CATCH(ESM) processing.</td>
</tr>
<tr>
<td></td>
<td>3005</td>
<td>The attempt to forecast the desired model produced an arithmetic exception. The forecast is generated by CATCH(RW) processing.</td>
</tr>
<tr>
<td></td>
<td>3006</td>
<td>The attempt to forecast the desired model produced an arithmetic exception. The forecast is generated by CATCH(MISSING) processing.</td>
</tr>
<tr>
<td></td>
<td>3007</td>
<td>The mean value forecast is generated as a result of the MINOBS criterion.</td>
</tr>
<tr>
<td></td>
<td>3008</td>
<td>There were insufficient nonmissing observations in the variable to be forecast. A missing value forecast is produced.</td>
</tr>
<tr>
<td></td>
<td>3009</td>
<td>There were insufficient nonzero observations in the variable to be forecast. A zero-valued forecast is produced.</td>
</tr>
</tbody>
</table>

Statistics of Fit Attributes

You can filter the Forecast viewer based on a selected range for one or more statistics of fit. The following general statistics of fit are available for filtering. You can also filter using statistics of fit that are used for model selection. See “Descriptions of Model Selection Criteria” on page 159 for a complete list of these attributes.
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Variable Label</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Number of Observations</td>
<td>the number of the qualifying subset of observations in the specified model region (either FIT region or FORECAST region) that are potentially available for the computation of fit statistics</td>
</tr>
<tr>
<td>NOBS</td>
<td>Number of Observations Used</td>
<td>the actual number of nonmissing observations from the qualifying subset of observations in the specified model region that were actually used to compute fit statistics</td>
</tr>
<tr>
<td>NMISSA</td>
<td>Number of Missing Actual Values</td>
<td></td>
</tr>
<tr>
<td>NMISSP</td>
<td>Number of Missing Predicted Values</td>
<td></td>
</tr>
<tr>
<td>NPARMS</td>
<td>Number of Model Parameters</td>
<td></td>
</tr>
<tr>
<td><em>REGION</em></td>
<td>Region</td>
<td>specifies the time region in which the statistics were collected. Select from the following.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FIT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>indicates that fit statistics were calculated over the fit region</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FORECAST</td>
</tr>
<tr>
<td></td>
<td></td>
<td>indicates that fit statistics were calculated over the forecast region</td>
</tr>
<tr>
<td><em>SELECT</em></td>
<td>Selection List</td>
<td>name of model selection list to which <em>MODEL</em> belongs</td>
</tr>
<tr>
<td><em>MODEL</em></td>
<td>Model</td>
<td></td>
</tr>
<tr>
<td>DFE</td>
<td>Degrees of Freedom Error</td>
<td></td>
</tr>
<tr>
<td>SST</td>
<td>Corrected Total Sum of Squares</td>
<td></td>
</tr>
<tr>
<td>TSS</td>
<td>Total Sum of Squares</td>
<td></td>
</tr>
</tbody>
</table>
Model Selection Criteria

What Are Statistics of Fit?

The statistics of fit are statistical values that are used to evaluate how well a forecasting model performs by comparing the actual data to the predictions. For a given forecast model that has been fitted to the time series data, the model should be checked or evaluated to see how well it fits or forecasts the data. Commonly used statistics of fit are Root Mean Square Error (RMSE), Mean Absolute Percent Error (MAPE), Akaike Information Criteria (AIC), and many others. The statistics of fit can be computed from the model residuals or the prediction errors.

When the full range of data is used to both fit and evaluate the model, this is referred to as in-sample evaluation. When the most recent data is excluded for parameter estimation (holdout) and this holdout sample is used for evaluation, this is referred to as holdout sample evaluation. Holdout sample analysis is similar to the training and testing of neural networks. A portion of the data is withheld from training (fit) and the withheld data (holdout) is used to test performance.

When a particular statistic of fit is used for forecast model selection, it is referred to as the model selection criterion. For example, if the MAPE (an often recommended choice) is used as a model selection criterion, the forecast model with the smallest MAPE in the evaluation region (in-sample or holdout-sample) is chosen as the best model.

When a particular statistic of fit is used to judge how well the forecasting process is predicting the future, it is referred to as the performance statistic.

Understanding the Formulas and Definitions

To understand the formulas and definitions for the statistics of fit, review the following information:

- $n$ is the number of nonmissing observations.
- $k$ is the number of fitted parameters in the model.
- $APE = 100 \times \frac{|y_t - \hat{y}_t|}{y_t}$ is the absolute percent error.
- $ASPE = 100 \times \frac{|y_t - \hat{y}_t|}{0.5(y_t + \hat{y}_t)}$ is the absolute symmetric percent error.
- $APPE = 100 \times \frac{|y_t - \hat{y}_t|}{|\hat{y}_t|}$ is the absolute predictive percent error.
- $RAE = \frac{(y_t - \hat{y}_t)}{(y_t - y_{t-1})}$ is relative absolute error.
Descriptions of Model Selection Criteria

The definitions and formulas for the statistics of fit that are available in SAS Visual Forecasting are described below. You can use statistics of fit to measure how well different models fit the data. The statistics of fit for the various forecasting models can be printed or stored in a data set. Choose from one of the following options.

**Adjusted R-square (ADJRSQ)**

The adjusted $R^2$ statistic, $1 - \frac{1}{(n-k)}(1 - R^2)$

**Akaike information Corrected criterion (AICC)**

Akaike’s information criterion with an empirical correction for small sample sizes, $AIC + \frac{2k(k + 1)}{n - k - 1}$

**Akaike information criterion (AIC)**

Akaike’s information criterion, $n \ln(\frac{SSE}{n}) + 2k$

**Amemiya’s adjusted R-square (AADJRSQ)**

Amemiya’s adjusted $R^2$, $1 - \frac{1}{n-k}(1 - R^2)$

**Amemiya’s prediction criterion (APC)**

Amemiya’s prediction criterion, $\frac{1}{n}SST(\frac{n+k}{n-k})(1 - R^2) = (\frac{n+k}{n-k})\frac{1}{n}SSE$

**Geometric mean absolute error percent of standard deviation (GMAPES)**

The geometric mean of the absolute error as a percentage of the standard deviation

**Geometric mean percent error (GMAPE)**

The geometric mean percent error

**Geometric mean predictive percent error (GMAPPE)**

The geometric mean absolute predictive percent prediction error

**Geometric mean relative absolute error (GMRAE)**

The geometric mean of the relative absolute errors

**Geometric mean symmetric percent error (GMASPE)**

The geometric mean of the absolute symmetric percent errors

**Maximum Absolute Error Percent of Standard Deviation (MAXAPES)**

The maximum of the absolute error as a percentage of the standard deviation

**Maximum error (MAXERR)**

The largest prediction error

**Maximum percent error (MAXPE)**

The largest percent prediction error, $100\max(\frac{y_i - \hat{y}_i}{y_i})$. The summation ignores observations where $y_i = 0$.

**Maximum predictive percent error (MAXPPE)**

The maximum of the predictive percent errors

**Maximum relative error (MAXRE)**

The maximum of the relative errors
Maximum symmetric percent error (MAXSPE)
The maximum of the symmetric percent errors

Mean absolute error (MAE)
The mean absolute prediction error, \( \frac{1}{n} \sum_{t=1}^{n} | y_t - \hat{y}_t | \).

Mean absolute error percent of standard deviation (MAPES)
The mean of the absolute error as a percentage of the standard deviation.

Mean absolute percent error (MAPE)
The mean of the absolute percent errors.

Mean absolute predictive symmetric percent error (MAPPE)
The mean of the absolute symmetric predictive percent error.

Mean absolute scaled error (MASE)
The mean of the absolute scaled errors, where \( MAE_{in-sample, naive} \) is the mean absolute error produced by a naive forecast calculated on the in-sample data:
\[
\text{MASE} = \frac{\text{MAE}}{\text{MAE}_{in-sample, naive}}
\]

For a non-seasonal series, the naïve forecast is generated by using the actual value from previous record. For a seasonal series, the naïve forecast is generated by using the actual value from previous season.

Mean error (ME)
The mean prediction error, \( \frac{1}{n} \sum_{t=1}^{n} (y_t - \hat{y}_t) \)

Mean percent error (MPE)
The mean percent prediction error, \( \frac{1}{n} \sum_{t=1}^{n} \frac{(y_t - \hat{y}_t)}{y_t} \). The summation ignores observations where \( y_t = 0 \).

Mean predictive percent error (MPPE)
The mean of the predictive percent error.

Mean relative absolute error (MRAE)
The mean of the relative absolute errors.

Mean relative error (MRE)
The mean of the relative errors.

Mean squared error (MSE)
The mean squared prediction error calculated from the one-step-ahead forecasts, \( MSE = \frac{1}{n} \sum_{t=1}^{n} (y_t - \hat{y}_t)^2 \). This formula enables you to evaluate small holdout samples.

Mean symmetric percent error (MSPE)
The mean of the symmetric percent errors.

Mean absolute symmetric percent error (SMAPE)
The symmetric mean of the absolute percent error.

Median absolute error percent of standard deviation (MDAPES)
The median of the absolute error as a percentage of the standard deviation.

Median percent error (MDAPE)
The median of the percent errors.

Median predictive percent error (MDAPPE)
The median of the predictive percent errors.
Median relative absolute error (MDRAE)
The median of the relative absolute errors.

Median symmetric percent error (MDASPE)
The median of the symmetric percent errors.

Minimum Absolute Error Percent of Standard Deviation (MINAPES)
The minimum of the absolute error as a percentage of the standard deviation.

Minimum error (MINERR)
The smallest prediction error.

Minimum percent error (MINPE)
The smallest percent prediction error, $100\min\left(\frac{y_t - \hat{y}_t}{y_t}\right)$. The summation ignores observations where $y_t = 0$.

Minimum predictive percent error (MINPPE)
The smallest predictive percent error.

Minimum relative error (MINRE)
The smallest relative error.

Minimum symmetric percent error (MINSPE)
The smallest symmetric percent error.

R-square (RSQUARE)
The $R^2$ statistic, $R^2 = 1 - \frac{SSE}{SST}$. If the model fits the series badly, the model sum of square error, $SSE$, might be larger than $SST$ and the $R^2$ statistic will be negative.

Random walk R-square (RWRSQ)
The random walk $R^2$ statistic (Harvey’s $R^2$ statistic using the random walk model for comparison), $1 - \left(\frac{n - 1}{n}\right) \frac{SSE}{RW_{SSE}}$, where $RW_{SSE} = \sum_{t=2}^{n} (y_t - y_{t-1} - \mu)^2$, and
\[
\mu = \frac{1}{n-1} \sum_{t=2}^{n} (y_t - y_{t-1}).
\]

Root mean squared error (RMSE)
The root mean square error, $\sqrt{MSE}$.

Schwarz Bayesian information criterion (SBC)
Schwarz Bayesian information criterion, $n \ast \ln\left(\frac{SSE}{n}\right) + k \ast \ln(n)$.

Sum of square error (SSE)
The sum of the squared prediction errors. $SSE = \sum_{t=1}^{n} (y_t - \hat{y}_t)^2$, where $\hat{y}$ is the one-step predicted value.

Unbiased mean squared error (UMSE)
The unbiased mean squared error.

Unbiased root mean squared error (URMSE)
The unbiased root mean squared error.
Weighted Model Comparison Selection Criteria

Set the selection criteria for choosing the champion pipeline on the Pipeline Comparison tab and for choosing the champion model within a pipeline.

- Use the Project Settings to set the selection criteria used for pipeline comparison.
- The Model Comparison node has a property to set the selection criteria for model comparison within the pipeline.

Note: For projects using the External Forecast pipeline, if the values are missing from the ACTUAL column, the weighted metrics will have missing values in the Pipeline Comparison results. See “Working with External Forecast Projects” on page 27 for more information about the composition of external forecast projects.

To understand the absolute percent error or APE, where \( n \) is the number of nonmissing observations and \( k \) is the number of fitted parameters in the model, the following formula applies:

\[
APE = 100 \times \left| \frac{y_t - \hat{y}_t}{y_t} \right|
\]

Note: For the following statistics of fit, the final computation is based on weighted measurements from each time series, where more weight is given to time series with a higher average of the dependent variable.

**Weighted mean absolute percent error (WMAPE)**

The weighted mean of the absolute percent errors.

**Weighted mean absolute error (WMAE)**

The weighted mean absolute prediction error, \( \frac{1}{n} \sum_{t=1}^{n} |y_t - \hat{y}_t| \).

**Weighted mean absolute scaled error (WMASE)**

The weighted mean of the absolute scaled errors, where \( MAE_{in\text{-sample}, \text{naive}} \) is the mean absolute error produced by a naive forecast calculated on the in-sample data:

\[
\frac{MAE}{MAE_{in\text{-sample}, \text{naive}}}
\]

For a non-seasonal series, the naïve forecast is generated by using the actual value from previous record. For a seasonal series, the naïve forecast is generated by using the actual value from previous season.

**Weighted root mean squared error (WRMSE)**

The weighted root mean square error, \( \sqrt{MSE} \).

Note: For the following statistics of fit, \( a \) is the actual value.
**Weighted absolute percent error (WAPE)**

The weighted APE is computed as \( \frac{\Sigma(a \times APE)}{\Sigma a} \).

**Weighted absolute scaled error (WASE)**

WASE is the weighted ASE for all time series. The WASE comparison metric is computed as \( \frac{\Sigma(a \times ASE)}{\Sigma a} \).

---

**Reserved Names for Variables and Models**

**Reserved Variable Names**

For each project, SAS Visual Forecasting creates several output data sets. The variable names in your input data set cannot match any of the variable names in these output data sets. The variable names in your input data set also cannot start with an underscore. If you try to assign a variable to a role and the variable name matches either of these conditions, then an error message appears.

The following table lists alphabetically the variables that are used by SAS Visual Forecasting.

*Table 6.9  Reserved Variable Names*

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>_VariableName</td>
<td>Any variable name that begins with an underscore</td>
</tr>
<tr>
<td>AADJRSQ</td>
<td>Amemiya’s adjusted R-square</td>
</tr>
<tr>
<td>ACTUAL</td>
<td>Dependent series value</td>
</tr>
<tr>
<td>ADJRSQ</td>
<td>Adjusted R-square</td>
</tr>
<tr>
<td>AGGCHILDPREDICT</td>
<td>Aggregated prediction of child nodes</td>
</tr>
<tr>
<td>AIC</td>
<td>Akaike Information Criterion</td>
</tr>
<tr>
<td>AICC</td>
<td>Finite sample corrected Akaike Information Criterion</td>
</tr>
<tr>
<td>APC</td>
<td>Amemiya’s Prediction Criterion</td>
</tr>
<tr>
<td>DFE</td>
<td>Degrees of freedom error</td>
</tr>
<tr>
<td>END</td>
<td>Ending value of the time variable</td>
</tr>
<tr>
<td>Variable Name</td>
<td>Description</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------</td>
</tr>
<tr>
<td>ENDOBS</td>
<td>Number of the last observation in the data</td>
</tr>
<tr>
<td>ERROR</td>
<td>Prediction errors</td>
</tr>
<tr>
<td>FINALPREDICT</td>
<td>Predicted value for the parent node</td>
</tr>
<tr>
<td>GMAPE</td>
<td>Geometric mean percent error</td>
</tr>
<tr>
<td>GMAPPE</td>
<td>Geometric mean predictive percent error</td>
</tr>
<tr>
<td>GMAPES</td>
<td>Geometric mean absolute error percent of standard deviation</td>
</tr>
<tr>
<td>GMASPE</td>
<td>Geometric mean symmetric percent error</td>
</tr>
<tr>
<td>GMRAE</td>
<td>Geometric mean relative absolute error</td>
</tr>
<tr>
<td>ISRECONCILED</td>
<td>If the node is reconciled, then this variable is 1; if the node is not reconciled, then this variable is 0.</td>
</tr>
<tr>
<td>LEAF</td>
<td>Keyword used in model generation</td>
</tr>
<tr>
<td>LLOCK</td>
<td>Lock level for lower bound on the forecast</td>
</tr>
<tr>
<td>LOWBROVR</td>
<td>Lower confidence limits before override reconciliation</td>
</tr>
<tr>
<td>LOWER</td>
<td>Lower confidence limits</td>
</tr>
<tr>
<td>LOWERBD</td>
<td>Lower bound on the forecast</td>
</tr>
<tr>
<td>MAE</td>
<td>Mean absolute error</td>
</tr>
<tr>
<td>MAPE</td>
<td>Mean absolute percent error</td>
</tr>
<tr>
<td>MAPPE</td>
<td>Symmetric mean absolute predictive percent error</td>
</tr>
<tr>
<td>MASE</td>
<td>Mean absolute scaled error</td>
</tr>
<tr>
<td>MAX</td>
<td>Maximum value</td>
</tr>
<tr>
<td>MAXAPES</td>
<td>Maximum absolute error percent of standard deviation</td>
</tr>
<tr>
<td>MAXERR</td>
<td>Maximum error</td>
</tr>
<tr>
<td>Variable Name</td>
<td>Description</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------</td>
</tr>
<tr>
<td>MAXPE</td>
<td>Maximum percent error</td>
</tr>
<tr>
<td>MAXPPE</td>
<td>Maximum predictive percent error</td>
</tr>
<tr>
<td>MAXRE</td>
<td>Maximum relative error</td>
</tr>
<tr>
<td>MAXSPE</td>
<td>Maximum symmetric percent error</td>
</tr>
<tr>
<td>MDAPE</td>
<td>Median percent error</td>
</tr>
<tr>
<td>MDAPES</td>
<td>Median absolute error percent of standard deviation</td>
</tr>
<tr>
<td>MDAPPE</td>
<td>Median predictive percent error</td>
</tr>
<tr>
<td>MDASPE</td>
<td>Median symmetric percent error</td>
</tr>
<tr>
<td>MDRAE</td>
<td>Median relative absolute error</td>
</tr>
<tr>
<td>ME</td>
<td>Mean error</td>
</tr>
<tr>
<td>MEAN</td>
<td>Mean value</td>
</tr>
<tr>
<td>MIN</td>
<td>Minimum value</td>
</tr>
<tr>
<td>MINAPES</td>
<td>Minimum absolute error percent of standard deviation</td>
</tr>
<tr>
<td>MINERR</td>
<td>Minimum error</td>
</tr>
<tr>
<td>MINPE</td>
<td>Minimum percent error</td>
</tr>
<tr>
<td>MINPPE</td>
<td>Minimum predictive percent error</td>
</tr>
<tr>
<td>MINRE</td>
<td>Minimum relative error</td>
</tr>
<tr>
<td>MINSPE</td>
<td>Minimum symmetric percent error</td>
</tr>
<tr>
<td>MPE</td>
<td>Mean percent error</td>
</tr>
<tr>
<td>MPPE</td>
<td>Mean predictive percent error</td>
</tr>
<tr>
<td>MRAE</td>
<td>Mean relative absolute error</td>
</tr>
<tr>
<td>MRE</td>
<td>Mean relative error</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean square error</td>
</tr>
<tr>
<td>Variable Name</td>
<td>Description</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------</td>
</tr>
<tr>
<td>MSPE</td>
<td>Mean symmetric percent error</td>
</tr>
<tr>
<td>N</td>
<td>Number of nonmissing observations or number of variance products</td>
</tr>
<tr>
<td>NAME</td>
<td>Variable name</td>
</tr>
<tr>
<td>NMISS</td>
<td>Number of missing observations</td>
</tr>
<tr>
<td>NMISSA</td>
<td>Number of missing actuals</td>
</tr>
<tr>
<td>NMISSP</td>
<td>Number of missing predicted</td>
</tr>
<tr>
<td>NOBS</td>
<td>Number of observations</td>
</tr>
<tr>
<td>NONMISSCHLD</td>
<td>Number of nonmissing children in the current AGGBY group</td>
</tr>
<tr>
<td>NOTE</td>
<td>Text that user specifies for a series</td>
</tr>
<tr>
<td>NPARMS</td>
<td>Number of model parameters</td>
</tr>
<tr>
<td>OLOCK</td>
<td>Lock level for equality constraint</td>
</tr>
<tr>
<td>OVERRIDE</td>
<td>Equality constraint on the forecast</td>
</tr>
<tr>
<td>PREBFOVR</td>
<td>Predicted values before override reconciliation</td>
</tr>
<tr>
<td>PREDICT</td>
<td>Predicted values</td>
</tr>
<tr>
<td>RECDIFF</td>
<td>Reconciliation difference</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root mean square error</td>
</tr>
<tr>
<td>RSQUARE</td>
<td>R-square</td>
</tr>
<tr>
<td>RWRSQ</td>
<td>Random walk R-square</td>
</tr>
<tr>
<td>SBC</td>
<td>Schwarz Bayesian information criterion</td>
</tr>
<tr>
<td>SMAPE</td>
<td>Symmetric mean absolute percent error</td>
</tr>
<tr>
<td>SSE</td>
<td>Sum of squares error</td>
</tr>
<tr>
<td>SST</td>
<td>Corrected total sum of squares</td>
</tr>
<tr>
<td>START</td>
<td>Beginning value of the time variable</td>
</tr>
<tr>
<td>Variable Name</td>
<td>Description</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------------------------------------------------</td>
</tr>
<tr>
<td>STARTOBS</td>
<td>Number of the first observation</td>
</tr>
<tr>
<td>STD</td>
<td>Prediction standard errors</td>
</tr>
<tr>
<td>STDBFOVR</td>
<td>Standard deviation before override reconciliation</td>
</tr>
<tr>
<td>STDDEV</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SUM</td>
<td>Summation value</td>
</tr>
<tr>
<td>TOP</td>
<td>Keyword used in model generation</td>
</tr>
<tr>
<td>TSS</td>
<td>Total sum of squares</td>
</tr>
<tr>
<td>ULOCK</td>
<td>Lock level for the upper bound on the forecast</td>
</tr>
<tr>
<td>UMSE</td>
<td>Unbiased mean square error</td>
</tr>
<tr>
<td>UNLOCK</td>
<td>For locked overrides, the value of this variable is 0. For unlocked overrides, the value is 1.</td>
</tr>
<tr>
<td>UPPEFOVR</td>
<td>Upper confidence limits before override reconciliation</td>
</tr>
<tr>
<td>UPPER</td>
<td>Upper confidence limits</td>
</tr>
<tr>
<td>UPPERBD</td>
<td>Upper bound on the forecast</td>
</tr>
<tr>
<td>URMSE</td>
<td>Unbiased root mean square error</td>
</tr>
<tr>
<td>XML</td>
<td>User preferences for a series in XML format</td>
</tr>
</tbody>
</table>

Reserved Model Names

SAS Visual Forecasting uses the following keywords in the names of the automatically generated models:
- LEAF
- TOP
- HPF
These words are reserved for SAS Visual Forecasting models. If the name of the model that you are creating contains these keywords, then a warning message appears, and you are prompted to select a different name.

**Note:** If you are copying a system-generated model that contains one of these keywords, then the copied name can contain one of these keywords. For example, if you copy a model named LEAF, then SAS Visual Forecasting names the copied model LEAFCOPYn, where n is the lowest available integer value. Because this name is unique, SAS Visual Forecasting allows this name.

---

## Promoting Projects and Templates from Previous Releases to Model Studio 8.5

Model Studio provides menu options for administrators to import projects and templates from another system using the current release of Model Studio. Projects and templates from a previous release of Model Studio must use promotion as described here.

**Note:** After upgrading Model Studio to the latest release, each project must be individually upgraded by the owner to take advantage of the new features. See “Upgrade Considerations” in *SAS Visual Forecasting: Overview* for more information.

---

## Promoting Projects and Templates from Model Studio 8.3

An administrator can use SAS Environment Manager to promote projects and templates from previous releases of Model Studio to the current release. Before promoting a project, make sure the source data for the project is already loaded to CAS on the target server. For SAS Visual Forecasting, this includes any data sources for additional attributes or custom event definitions.

### Exporting Projects and Templates

The instructions for exporting Model Studio projects and templates using SAS Environment Manager can be found in *Export Content* in *SAS Viya 3.4 Administration / Content Management*.

You can find the projects and templates in SAS Environment Manager by clicking in the left side toolbar. On the left side of the Content page, navigate through to locate the project or template that you want to export.
Importing Projects and Templates

The instructions for importing Model Studio projects and templates using SAS Environment Manager can be found in Import SAS Viya Resources in SAS Viya 3.4 Administration / Promotion (Import and Export).

Note: Before importing any projects, make sure any custom pipeline or node templates used by the project are promoted first.

---

Promoting Projects and Templates from Model Studio 8.2

Overview

An administrator must use the transfer command line interface (CLI) to promote projects and templates from Model Studio 8.2 to the current release. You must first apply the latest software update on the Model Studio 8.2 server. See SAS Note 62339 to obtain information about this hot fix.

Before promoting a project, make sure the source data for the project is already loaded to the CAS server on the target server. For SAS Visual Forecasting, this includes any data sources for additional attributes or custom event definitions.

For general information about the transfer CLI, see "Promotion: How to Import (Command-Line Interface)" in SAS Viya Administration: Promotion (Import and Export).

Before you start the promotion process, you must first get the unique ID for the resource that you need to promote.

Getting the ID for a Project

To promote a project to another server, you must first get the unique ID for the project. Follow these steps to get the ID for the project that you want to export.

1. From the Model Studio project list, select the project that you want to export to the target server.
2. Click and select Batch API.
3. Copy the project ID from the Batch Code dialog box.

   - If SAS is selected for Download Type, the project ID is in a macro at the end of the code. For example:

     \%let projectId = fef185c3-c0c2-46fe-bf73-417784c74bba;

   - If Python is selected for Download Type, the project ID is listed with other arguments near the top of the code. For example:

     parser.add_argument('--projectId', default='fef185c3-c0c2-46fe-bf73-417784c74bba')
Getting the ID for a Template

Follow these steps to get the ID for a pipeline or node template that you want to export from Model Studio 8.2.

1. Click at the top left corner of SAS Home and select Manage Environment. SAS Environment Manager is opened.
2. Click on the left side.
3. Open Users ⇒ username ⇒ My Folder ⇒ Analytics Toolbox
   For username, provide the user ID for the person who created the template. Custom templates created by the user are listed on the left.
4. Select the template that you want to export.
   The properties of the selected template are displayed to the right. Each will have a URI value, which includes a path followed by the ID for the template, for example:
   
   **Pipeline template**
   `/analyticsComponents/templates/84271a16-a22b-4b0a-b01d-100f806758c4`
   
   The ID for this pipeline template is `84271a16-a22b-4b0a-b01d-100f806758c4`.
   
   **Node template**
   `/analyticsComponents/templates/e26bb3e0-cc68-4153-b42b-a3df6ab1ccc8`
   
   The ID for this pipeline template is `e26bb3e0-cc68-4153-b42b-a3df6ab1ccc8`.

Promoting Projects and Templates

In the steps below, the following conventions are used.

**source server**
refers to the Model Studio 8.2 installation from which the project is exported. These steps use the example `test01.example.org`.

**target server**
refers to the Model Studio 8.5 installation to which the project is imported. These steps use the example `test02.example.org`.

You need a directory on each server to store the JSON transfer file for the project or template. The examples below use the path `/tmp/projects`.

After you have the project or template ID, follow these steps to promote the resource to the target server using the transfer CLI for SAS Viya. You must have administrator access for both installations.

1. From a command prompt, issue an SSH connection to the source server for the project.
   
   `ssh root@test01.example.org`
   
2. Change to the `/opt/sas/viya/home/bin` directory.
3 Follow the instructions in “Create at Least One Profile” in SAS Viya Administration: Using the Command-Line Interfaces to create a profile for the source server. Use these responses when prompted by the profile initialization:

<table>
<thead>
<tr>
<th>Specification</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service Endpoint&gt;</td>
<td>Specify the URL for the source server that contains the project that you want to export. For example: <a href="http://test01.example.org">http://test01.example.org</a></td>
</tr>
<tr>
<td>Output type (text</td>
<td>json</td>
</tr>
<tr>
<td>Enable ANSI colored output (y/n)?</td>
<td>Specify y.</td>
</tr>
</tbody>
</table>

4 Sign in to the server as described in “Use a Profile to Sign In” in SAS Viya Administration: Using the Command-Line Interfaces.

5 Export the project specifying the path and ID of the project or template.

```
./sas-admin transfer export --resource-uri path/ID
```

where path is one of the following:

- project
  /analyticsGateway/projects/
- pipeline template
  /analyticsGateway/pipelineTemplates/
- node template
  /analyticsGateway/componentTemplates/

The following example shows the output when a project is successfully exported.

```
[root@test01 bin]# ./sas-admin transfer export --resource-uri /analyticsGateway/projects/3fdde180-be4c-4bc9-9385-f1db9d0a3069
running ........ completed
{
  "items": [
    {
      "message": "Transfer object processing complete.",
      "name": "project",
      "state": "completed"
    }
  ]
}

Package created: 4458628a-47b1-4eaa-846a-b7ff50d7c811
```

The output of this command includes the ID for the transfer package. You need this ID for the next step.

6 Download the transfer package as a JSON file. Specify a valid path and give the package file a name. The ID for the package from the previous step must be specified for the --id argument. In the following example, the download path and filename for the transfer file is /opt/projects/sales01.json.

```
./sas-admin transfer download -file /tmp/projects/sales01.json --id 4458628a-47b1-4eaa-846a-b7ff50d7c811
```
7 Copy the transfer file to the target server, specifying a valid target path.
scp /tmp/projects/sales01.json root@test02.example.org:/tmp/projects/sales01.json

8 Start a new SSH session with the target server.
ssh root@test02.example.org

9 Change to the /opt/sas/viya/home/bin directory.

10 Follow the instructions in “Command-Line Interface: Preliminary Instructions” in SAS Viya Administration: Using the Command-Line Interfaces to create a profile and sign in to the target server. Use these responses when prompted by the profile initialization:

Service Endpoint>
Specify the URL for the Model Studio environment where you want to import the project. For example:
http://test02.example.org

Output type (text|json|fulljson)>
Specify json.

Enable ANSI colored output (y/n)?
Specify y.

11 Sign in to the server as described in “Use a Profile to Sign In” in SAS Viya Administration: Using the Command-Line Interfaces.

12 Upload the transfer file, for example:
./sas-admin transfer upload --file /opt/projects/sales01.json

```
{
  "id": "09206455-906b-4ae6-805b-6eb7c69f4f20",
  "name": "Export"
}
```

The output of this command includes the ID for the transfer package. You need this ID for the next step.

13 Import the transfer file using the package ID output from the previous step.
./sas-admin transfer import --request "{"\"packageUrl\":\"/transfer/packages/09206455-906b-4ae6-805b-6eb7c69f4f20\"}"

```
routing .........................completed
Summary results of import:

Total tasks: 5  Succeeded: 5  Failed: 0  Completed with errors: 0
Completed with warnings: 0
```

14 Log on to the target server and open Model Studio.
Verify that you can access the project or template.
See Also

*SAS Viya 3.4 Administration: Promotion (Import and Export)*

Troubleshooting

Here are some issues that you might find using SAS Visual Forecasting, along with steps to take to work around the problem.

Pipeline Fails to Run for Project with User-Defined Formats

Description

If you create a project using a data set that uses formats created with the FORMAT procedure, running the pipeline fails.

Solution

There are some steps that you need to take to make the format library available to SAS Visual Forecasting and CAS.

1. Make sure the formats are in a path accessible by CAS, such as `cas_path/default/formats/`.

2. Add the format library to CAS. You can use the CAS `addFmtLib` action, for example:

   ```
cas.addfmtlib="myformats1.sashdat myuserformats2.sashdat"
   
   For more information about this action, see *addFmtLib Action* in *SAS 9.4 and SAS Viya 3.4 Programming Documentation*.
   ```

3. Set the `FMTSEARCH=` system option to include the format library, for example:

   ```
cas casauto fmtsearch=(myformats1 myuserformats2);
   
   For more information about this system option, see *FMTSEARCH= System Option* in *SAS 9.4 and SAS Viya 3.4 Programming Documentation*.
   ```

4. Restart the CAS server.

For more information about user-defined formats in CAS, see these resources.

- Manage Your User-Defined Formats in *SAS 9.4 and SAS Viya 3.4 Programming Documentation*
- *SAS Cloud Analytic Services: User-Defined Formats*
- Add a Global Format Library to a Session’s Format Search List in *SAS 9.4 and SAS Viya 3.4 Programming Documentation*
Cannot Access the Bottom of the Overrides Tab

Description
You open the Overrides tab but the content at the bottom is cut off and you cannot scroll to access it.

Solution
Resize the browser window. This forces the vertical scroll bar to appear so that you can scroll to the bottom.

Cannot Open Project When Source Data Tables Are Not Loaded

Description
You attempt to open a project and receive this error message:

The project cannot be opened because the required data tables for the project are not loaded. Contact your system administrator.

This issue often occurs for projects that are shared as read-only. If the CAS server has been restarted, the source data tables cannot be reloaded for projects that are read-only.

Solution
The project owner can remove the read-only setting from the Model Studio project page and then open the project. The source data tables are reloaded and the read-only setting can then be applied back to the project.

Pipeline Fails with Hierarchical Forecasting Model

Description
You run a pipeline after setting the forecast task to a value other than Diagnose. The pipeline run fails. The forecast tasks must be run sequentially. Each task requires the preceding task to be run first. These settings are described in “Task Settings” on page 117.
Solution
Create a new pipeline and start running the Diagnose task first, followed by the other tasks in the order provided.

Modeling Strategies Take a Long Time to Complete

Description
There are several factors that can influence the performance of modeling strategies during a pipeline run. In some circumstances, if there are a lot of users running pipelines at the same time, node execution time can be affected. By default, a maximum of five users can run the modeling strategies simultaneously. For any additional users, the modeling strategies are placed in Pending state until a run completes and another node can be started.

Solution
To change the number of users that can concurrently run modeling strategies, the administrator can update the maximumConcurrentNodeExecution property following these steps.

1 Click in the upper left corner and select Manage Environment.
   This opens SAS Environment Manager.
2 In the left side toolbar, click .
   The Configuration page opens.
3 On the left side, click the View drop-down list and select All services.
4 In the list of services, select Advanced Analytics Flows service. You can find it faster by entering the service name in the Filter field.
   On the right side of the page is a list of instances used to configure the service.
5 Scroll down to the sas.analytics.flows instance and click .
6 In the Edit sas.analytics.flows Configuration dialog box, scroll down to Maximum Concurrent Nodes.
   This is set to 5 by default. If you set this too high, many of the platform services can become highly stressed and fail. Start with setting this slightly higher, perhaps 6 or 7. If you do not see a significant impact on node execution time, you can repeat these steps to raise it a little higher.
7 Click Save.
Saved Model Filter Conflicts with Other Model Attributes

Description

Saving filters based on model-specific attributes can cause conflicts with different models, in the same pipeline or other pipelines for a project.

For example, the user can run the Retired Forecasting model in one pipeline and save a filter using the Model forecast attribute value, `customModel`. If the user then goes to the Forecast viewer for the Time Series Regression modeling strategy, simply selecting the Model value, `DIAG1_ARIMAX1`, invokes the saved filter for `customModel` from the retired modeling strategy. The new Model selection cannot be saved. This situation can also occur when working across neural network modeling strategies.

This should occur only when the Model value selects all of the time series in the Forecast viewer for both models.

Solution

If you need to save filters based on model-specific attributes, consider using those filters in a separate project from the other pipelines. For a list of model-specific attributes, see "Forecast Attributes for Selected Model" on page 155.

Error Generated When Closing Output Data

Description

When viewing the results for a Model Comparison or any modeling node in the pipeline, you can click View Output Data on the Output Data tab to load the output table generated by the model. If you click Close before the table finishes loading, you might get the following error:

An error occurred. Please contact your system administrator.

The details of the error could start with this message:

```
line: 1
col: 3185
TypeError: Cannot read property 'setText' of null
```

Solution

This error does not occur every time users click Close while the output data table is loading. When it does occur, the error can be ignored. To prevent the error from appearing again, wait until the output data finishes loading before closing it.
Error Generated When Changing the Activate Attribute Setting

Description

The _SEG_ attribute is used when importing attributes to a project to segment the data. After the attributes are imported, you can generate an error in the project by following these steps.

1. Follow the steps in "Importing Attributes to Your Project" on page 15. To generate this error, the table that you import must include the _seg_ variable. When the table is imported, it is selected under Attributes in the left pane.

2. Without navigating anywhere else in the project, select the _seg_ attribute in the middle pane.

3. In the right pane, set Activate attribute to No.

4. Now set Activate attribute to Yes.

The following error message is displayed:

The "If-Match" header value for the request is invalid.

After dismissing the error, the right pane remains in a busy state.

Solution

Exit from the project and then reopen the project. Select the _seg_ attribute from the imported attributes. The right pane is no longer in busy state. You can change the Activate attribute setting without generating the error again.

After Project Upgrade, Pipeline Fails

Description

Pipelines for some projects that are upgraded from SAS Visual Forecasting 8.2 might fail. This failure has occurred for projects with Champion selection criteria set to Weighted absolute scaled error (WASE) in the project settings.

Solution

Choose another setting for Champion selection criteria.
After Project Upgrade, Segmented Pipeline Fails

Description
After upgrading a SAS Visual Forecasting 8.4 project to 8.5, if you select a segmented pipeline and run it, it fails. In some cases, the Merge Segments node shows this error:

Error: There are no output tables found from segments.

If you open the modeling node, all segments show completed. If you open the pipeline for a segment, it shows that it is out of date.

Segmented pipelines for SAS Visual Forecasting 8.3 projects do not fail when upgraded to 8.5.

Solution
Duplicate the segmented pipeline and run the duplicate. After it runs successfully, you can delete the original pipeline.

To duplicate a pipeline, click ⚡ by the tab title and select Duplicate.

Data Is Misplaced When Downloaded from Forecast Viewer

Description
When working with the Forecast viewer, if you click ⬇️ to download the envelope plot data to your local machine, the lower and upper confidence limits for the forecast horizon are misplaced. They are placed at the beginning of the historical record. They belong in the rows associated with the forecast horizon dates.

This error occurs in the following situations:
- The project has no default attributes.
- The project has default attributes and you have selected one or more time series in the right pane.

Solution
Move the data from the rows at the beginning of the historical record to the rows in the forecast horizon, which is at the bottom of the data. Be sure to place the data in the correct columns that correspond to the lower confidence limit and upper confidence limit.