Data Flow Elements

**SQL Server 2008 R2**

[Other Versions](javascript:;)

SQL Server Integration Services provides three different types of data flow components: sources, transformations, and destinations. Sources extract data from data stores such as tables and views in relational databases, files, and Analysis Services databases. Transformations modify, summarize, and clean data. Destinations load data into data stores or create in-memory datasets.

Additionally, Integration Services provides paths that connect the output of one component to the input of another component. Paths define the sequence of components, and let you add annotations to the data flow or view the source of the column.

The following diagram shows a data flow that has a source, a transformation with one input and one output, and a destination. The diagram includes the inputs, outputs, and error outputs in addition to the input, output, and external columns.

[Sources](javascript:void(0))

In Integration Services, a source is the data flow component that makes data from different external data sources available to the other components in the data flow.

The source for a data flow typically has one regular output. The regular output contains output columns, which are columns the source adds to the data flow.

The regular output references external columns. An external column is a column in the source. For example, the MakeFlag column in the Product table of the AdventureWorks2008R2 database is an external column that can be added to the regular output. Metadata for external columns includes such information as the name, data type, and length of the source column.

An error output for a source contains the same columns as the regular output, and also contains two additional columns that provide information about errors. The Integration Services object model does not restrict the number of regular outputs and error outputs that sources can have. Most of the sources that Integration Services includes, except the Script component, have one regular output, and many of the sources have one error output. Custom sources can be coded to implement multiple regular outputs and error outputs.

All the output columns are available as input columns to the next data flow component in the data flow.

For more information, see [Integration Services Sources](https://technet.microsoft.com/en-us/library/ms141093(v=sql.105).aspx).

[Transformations](javascript:void(0))

The capabilities of transformations vary broadly. Transformations can perform tasks such as updating, summarizing, cleaning, merging, and distributing data.

The inputs and outputs of a transformation define the columns of incoming and outgoing data. Depending on the operation performed on the data, some transformations have a single input and multiple outputs, while other transformations have multiple inputs and a single output. Transformations can also include error outputs, which provide information about the error that occurred, together with the data that failed: for example, string data that could not be converted to an integer data type. The Integration Services object model does not restrict the number of inputs, regular outputs, and error outputs that transformations can contain. You can create custom transformations that implement any combination of multiple inputs, regular outputs, and error outputs.

The input of a transformation is defined as one or more input columns. Some Integration Services transformations can also refer to external columns as input. For example, the input to the OLE DB Command transformation includes external columns. An output column is a column that the transformation adds to the data flow. Both regular outputs and error outputs contain output columns. These output columns in turn act as input columns to the next component in the data flow, either another transformation or a destination.

For more information, see [Integration Services Transformations](https://technet.microsoft.com/en-us/library/ms141713(v=sql.105).aspx).

[Destinations](javascript:void(0))

A destination is the data flow component that writes the data from a data flow to a specific data store, or creates an in-memory dataset.

An Integration Services destination must have at least one input. The input contains input columns, which come from another data flow component. The input columns are mapped to columns in the destination.

Many destinations also have one error output. The error output for a destination contains output columns, which typically contain information about errors that occur when writing data to the destination data store. Errors occur for many different reasons. For example, a column may contain a null value, whereas the destination column cannot be set to null.

The Integration Services object model does not restrict the number of regular inputs and error outputs that destinations can have, and you can create custom destinations that implement multiple inputs and error outputs.

For more information, see [Integration Services Destinations](https://technet.microsoft.com/en-us/library/ms141089(v=sql.105).aspx).

[External Metadata](javascript:void(0))

When you create a data flow in a package using SSIS Designer, the metadata from the sources and destinations is copied to the external columns on sources and destinations, serving as a snapshot of the schema. When Integration Services validates the package, SSIS Designer compares this snapshot against the schema of the source or destination, and posts errors and warnings, depending on the changes.

The Integration Services project provides an offline mode. When you work offline no connections are made to the sources or destinations the package uses, and the metadata of external columns is not updated.

[Inputs and Outputs](javascript:void(0))

Sources have outputs, destinations have inputs, and transformations have both inputs and outputs. Additionally, many data flow components can be configured to use an error output.

Inputs

Destinations and transformations have inputs. An input contains one or more input columns, which can refer to external columns if the data flow component has been configured to use them. Inputs can be configured to monitor and control the flow of data: for example, you can specify if the component should fail in response to an error, ignore errors, or redirect error rows to the error output. You can also assign a description to the input or update the input name. In SSIS Designer, inputs are configured by using the **Advanced Editor** dialog box. For more information about the **Advanced Editor,**see [Integration Services User Interface](https://technet.microsoft.com/en-us/library/ms141730(v=sql.105).aspx).

Outputs

Sources and transformations always have outputs. An output contains one or more output columns, which can refer to external columns if the data flow component has been configured to use them. Outputs can be configured to provide information useful to downstream processing of the data. For example, you can indicate whether the output is sorted. You can also provide a description for the output, or update the output name. In SSIS Designer, outputs are configured by using the **Advanced Editor** dialog box.

Error Outputs

Sources, destinations, and transformations can include error outputs. You can specify how the data flow component responds to errors in each input or column by using the **Configure Error Output** dialog box. If an error or data truncation occurs at run time and the data flow component is configured to redirect rows, the data rows with the error are sent to the error output. By default, an error output contains the output columns and two error columns: ErrorCode and ErrorColumn. The output columns contain the data from the row that failed, ErrorCode provides the error code, and ErrorColumn identifies the failing column.

For more information, see [Handling Errors in the Data Flow](https://technet.microsoft.com/en-us/library/ms141679(v=sql.105).aspx).

Columns

Inputs, outputs, and error outputs are collections of columns. Each column is configurable and depending on the column type—input, output, or external—Integration Services provides different properties for the column. Integration Services provides three different ways of setting column properties: programmatically, by using component-specific dialog boxes, or by using the **Advanced Editor** dialog box.

[Paths](javascript:void(0))

Paths connect data flow components. In SSIS Designer, you can view and modify the path properties, view the output metadata for the path start point, and attach data viewers to a path.

For more information, see [Integration Services Paths](https://technet.microsoft.com/en-us/library/ms139791(v=sql.105).aspx) and [Debugging Data Flow](https://technet.microsoft.com/en-us/library/ms137944(v=sql.105).aspx).

[Configuration of Data Flow Components](javascript:void(0))

Data flow components can be configured at the component level; at the input, output, and error output levels; and at the column level.

* At the component level, you set properties that are common to all components, and you set the custom properties of the component.
* At the input, output, and error output levels, you set the common properties of inputs, outputs, and the error output. If the component supports multiple outputs, you can add outputs.
* At the column level, you set the properties that are common to all columns, in addition to any custom properties that the component provides for columns. If the component supports the addition of output columns, you can add columns to outputs.

You can set properties through SSIS Designer or programmatically. In SSIS Designer, you can set element properties using the custom dialog boxes provided for each element type, or by using the Properties window or the **Advanced Editor** dialog box.

For more information about how to set properties by using SSIS Designer, see [How to: Set the Properties of a Data Flow Component](https://technet.microsoft.com/en-us/library/ms140097(v=sql.105).aspx).