Diving Into Expressions

Introduction
Expressions can be used to help with a number of data preparation and analytic tasks such as cleansing string data, applying conditional logic, and mathematically transforming numeric values, just to name a few. Most commonly, expressions are created in the Formula tool but, by learning to write expressions, you are harnessing the power of not just one but 13 other tools in Designer that can use expressions as part of their configurations. For the purposes of this lesson and other lessons in this course, we will focus on writing expressions with the Formula tool.

An input dataset contains information on counties in the United States, their populations over time, and their area in square miles. Create expressions to modify values in the column [Name] and calculate the population density for each county. Drag a Formula tool from the Preparation tool palette onto the Canvas and connect it to the Input Data tool.

Create or Overwrite a Column
The Formula tool can be used to create new columns of data or modify values in an existing column. Select an option below.

Add a New Column
Create a new column called [Density] to calculate the population density of each county. Use the Drop Down menu to select "Add Column".

New columns of data must be named. Create a new column called [Density].

By default, new columns of data are assigned a datatype of V_Wstring. For this new column, a numeric datatype is more appropriate. Use the DropDown to select Double from the list of options.

Modify an Existing Column
Use the Drop Down to select an existing column to modify: [Name].

When modifying an existing column, its data type and size cannot be changed in the Formula tool's configuration.

Expression Editor
All tools capable of leveraging expressions have an expression editor. Use the expression editor to insert variables, functions, and syntax.

Inserting Variables
Variables, or columns, can be inserted into the expression editor in a couple of different ways.

Click to access a list of available Columns and Constants. Then, click [POP2010] to insert it into the Expression Editor.

Columns and constants can also be inserted with keystrokes. After typing an open bracket into the expression editor, a list of available columns and constants appears. Click or use the up and down arrows to select a value from this list.

Inserting Functions
The Formula tool contains a library of all the functions that can be applied in an expression. The function library organizes functions into categories.

Like variables, functions can be inserted into the Expression Editor in a few ways.
Click to access the function library and expand the category "String".

Hover on each function for a brief description of what it does. Select the function that converts all letters to uppercase by clicking it to insert it into the expression editor.

Functions can also be inserted by typing directly into the expression editor, allowing you to search and filter functions quickly. Select a function by clicking it or pressing "Enter" to insert it into the expression editor.

Expression Syntax
Syntax highlighting and the Formula tool's Data Preview allow you to assess whether an expression is complete or correct. Functions, variables, operators and syntax are highlighted in different colors. When an expression is not correctly formatted, values in an expression may not be colored, and the Data Preview will display an error in red text. When an expression is formatted correctly, the Data Preview will display the expected output for the value in the first row of the input dataset.

Add or Remove Expressions
Create multiple expressions in a single Formula tool. To add another expression in a sequence, click the "plus sign". Expressions are executed in the order in which they are displayed. To adjust the order of expressions, drag an expression editor into place. To delete an expression from the list, click the trash can icon.