Datasets

Contents

Copyright GE Digital

© 2023 General Electric Company.

GE, the GE Monogram, and Predix are either registered trademarks or trademarks of All other trademarks are the property of their respective owners.

This document may contain Confidential/Proprietary information of and/or its suppliers or vendors. Distribution or reproduction is prohibited without permission.

THIS DOCUMENT AND ITS CONTENTS ARE PROVIDED "AS IS," WITH NO REPRESENTATION OR WARRANTIES OF ANY KIND, WHETHER EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO WARRANTIES OF DESIGN, MERCHANTABILITY, OR FITNESS FOR A PARTICULAR PURPOSE. ALL OTHER LIABILITY ARISING FROM RELIANCE UPON ANY INFORMATION CONTAINED HEREIN IS EXPRESSLY DISCLAIMED.

Access to and use of the software described in this document is conditioned on acceptance of the End User License Agreement and compliance with its terms.

Chapter

1

Overview

Topics:

Overview of Datasets

A dataset is a fixed set of information that is stored as a Catalog item.

More Details

Datasets contain information that can be viewed and modified in a grid. While queries return the most current information in the database, a dataset displays the same information until it is modified manually. For example, you might use a dataset to store information from an external system so that the data can referenced or used to create graphs in GE Digital APM.

When a dataset is created, a data type is defined for each column to indicate the type of data that can be stored in cells in that column. When you create or modify a dataset, if you attempt to enter data that conflicts with the data type defined for that column, the data will not be retained. The following types of data can be stored in a dataset:

- **Character :** Stores alphanumeric data. Cells in columns of this data type will appear as text boxes in which you can enter letters, numbers, and symbols.
- **Numeric :** Stores numeric data. Cells in columns of this data type will appear as text boxes in which you can enter numbers.
- **Date :** Stores date and time data in the format MM/DD/YYYY. Cells in columns of this data type will display the ⊞ button, which can be selected to access a calendar from which a date can be selected. You can also enter dates directly into cells.
- **Logical :** Stores the value True or False. Cells in columns of this data type will appear as text boxes in which you can enter the value True or False.

You can create new datasets manually within GE Digital APM , or by importing a Microsoft Excel file (.xlsx), an XML file (.xml), or a text file (.txt or .csv).

After you create a dataset, you can access it, modify it, use it to create graphs, and export it as a Microsoft Excel file, an XML file, or a text file.

About Creating Datasets by Importing Files

You can create a new dataset in GE Digital APM by importing a Microsoft Excel file (.xlsx), an XML file (.xml), or a text file (.txt or .csv).

When importing a dataset, if you specify a Data Type for a column and that column contains values that are not supported by the selected Data Type, the entire column will be excluded from the dataset.

When importing a file, in the **Modify Columns** section, if you select the check box in the **Unique Value** column for a given column and the file that you are importing has the same value in multiple cells within that column, only the first row with a unique value will be imported into the dataset.

About Importing an Excel File

Microsoft Excel files with the file extension .xlsx can be imported to create a dataset. In a Microsoft Excel file, data is organized within grids, and can contain multiple worksheets. When importing a Microsoft Excel file, a dataset can be created only from a single worksheet. If you want to import multiple worksheets from a Microsoft Excel file, you must create a new dataset for each worksheet.

Note: You cannot create a dataset by importing binary Microsoft Excel files with the file extension .xls. You must convert the binary Microsoft Excel files to the OpenXML format with the file extension .xlsx to import the files.

About Importing an XML File

Extensible Markup Language (XML) files that have the file extension .xml can be imported to create a dataset. XML files can be created in text editors, such as Notepad and Microsoft Word, but can also be opened and viewed in an Internet browser.

When importing an XML file, the following data types are supported:

- string: Used for columns that contain alphanumeric values. Columns of this type will be imported using the Character data type.
- double: Used for columns that contain numeric values and may include a - (negative sign) and decimal. Columns of this type will be imported using the Numeric data type.
- dateTime: Used for columns that contain date and time values. Date values should use the format YYYY-MM-DD, and time values should use the 24-hour time format THH:MM:SS (e.g., T14:30:23). Columns of this type will be imported using the Date data type.
- boolean: Used for columns that contain logical data. Data in these columns should be limited to the following values: True, 1, False, and 0. Columns of this type will be imported as a Logical data type.

Each row of the dataset is contained with a Table tag. The attributes within the Table tag correspond to the column names defined via the xs:attribute tags.

About Importing a Text File

Text files with the extension .txt or .csv can be imported to create a dataset. These files contain lines of data that are delimited, or separated, by punctuation marks, with rows of data separated by line breaks.

GE Digital APM accepts the following characters as delimiters:

- , (comma).
- ; (semicolon).
- | (pipe character).

To maintain the structure of the dataset, delimiters must be used throughout the file, even for cells that do not contain values. To identify an empty cell, you would need to include two consecutive delimiters (e.g., A,B,,D,E).

About Dataset Modification

After you access a dataset, or before you save when creating a dataset manually in GE Digital APM or creating a dataset by importing a file, you can modify the dataset. It is important to understand GE Digital APM's capabilities and limitations when modifying datasets.

Details

On the page displaying a dataset, the dataset is displayed as a grid. You can specify column names only during the creation of a dataset; once the dataset is created, the number of columns in the dataset and the names of those columns cannot be altered.

If you need to modify the information contained in a dataset, you can do so in the following ways:

- Modify the data in a cell.
- Add rows.
- Delete rows.
- Move a column by selecting it and dragging it to the desired position.

After a dataset exists in GE Digital APM, if you need to modify the column structure or the data type of a column, you can perform the following steps:

- Export the dataset.
- Perform the necessary modifications.
- Import the updated file to create a new dataset.

Access the Dataset Page

Procedure

In the **Applications** menu, navigate to the **TOOLS** section, and then select **Datasets**.
The **Dataset** page appears.

Note: Only Datasets created or accessed from the **Dataset** page appear in this list. The 25 most recently accessed Datasets appear in the list.

Create a Dataset in GE Digital APM

About This Task

The following instructions describe how to create a dataset manually within GE Digital APM. You can also create a dataset by importing a file.

Procedure

1. Access the **Dataset** page.
2. In the upper-right corner of the page, select **Create New**, and then select **Manual**.

   The **New Dataset** page appears, displaying the **Modify Columns** section.

3. To create a new column, select ✛.

   A new row appears in the grid.
4. In the new row, in the box in the **Column Name** column, enter a name for the column.
5. In the same row, in the box in the **Data Type** column, select the type of data that will be allowed in the column.
6. In the same row, if you want each cell in the column to contain a unique value, select the check box in the **Unique Value** column.
7. To add another column, repeat steps 3 through 6. When you are finished adding columns, select **Next**.

   The **Preview** section appears.

   Note: Each column name must be unique. If each column name is not unique, when you select **Next**, an error message will appear, and you will have to return to the **Modify Columns** section to modify the column names.
8. In each cell in which you wish to enter data, select the cell, and then enter the data.

   Note:

   - At any time, you can create a new row by selecting ✛.
   - The value that you enter in a cell must correspond to the data type that you selected for the column in which it appears.
9. Select **Finish**.

   A new page appears, displaying the dataset.

10. Save the dataset.

Create a Dataset by Importing a File

About This Task

The following instructions describe how to create a dataset by importing a file. You can also create a dataset manually within GE Digital APM .

Procedure

1. Access the **Dataset** page.
2. In the upper-right corner of the page, select **Create New**, and then, depending on the type of file that you want to import, select **Excel**, **XML**, or **Text**.

   A new page appears, displaying the **Select File** section. Depending on the type of file that you are importing, to continue the operation, proceed to the appropriate drop-down section below.
3. To import an Excel file:

   a) Select **Choose File**, and then select the file that you want to import.
   b) Select **Next**.
   c) In the **Worksheet Name** box, select the worksheet that you want to import.
   d) If the first row in the worksheet should be used as column names in the dataset, select the **First row contains header definition** check box.
   e) If you want to import values only from certain cells in the worksheet, select the **Import specific range values** check box, and then, in the Start and End boxes, enter the values that define the range that you want (e.g., A2 to E20). At any time, you can select **See Preview** to access a preview of the dataset based on your current selections.
   f) Select **Next**.

   The **Modify Columns** section appears, displaying a grid.
   g) As needed, in the boxes in the **Column Name** column, modify the column names. You can only modify these column names if, in step 6, you did not select the **First row contains header definition** check box.
   h) As needed, in the boxes in the **Data Type** column, select the type of data that will be allowed in the column.
   i) For any columns in which you want each cell in that column to contain a unique value, in the row representing that column, select the check box in the **Unique Value** column.
   j) Select **Next**.
   The **Preview** section appears, displaying a preview of the dataset.
   k) Select **Finish**.
   A new page appears, displaying the dataset.
   l) Save the dataset.
4. To import an XML file:

   a) Select **Choose File**, and then select the file that you want to import.
   b) Select **Next**.

   The **Modify Columns** section appears, displaying a grid.
   c) As needed, in the boxes in the **Data Type** column, select the type of data that will be allowed in the column.
   d) For any columns in which you want each cell in that column to contain a unique value, in the row representing that column, select the check box in the **Unique Value** column.

e) Select **Next**.
   The **Preview** section appears, displaying a preview of the dataset.

f) Select **Finish**.
   A new page appears, displaying the dataset.

g) Save the dataset.

5. To import a text file:

   a) Select **Choose File**, and then select the file that you want to import.

   b) Select **Next**.

   c) In the **Column Delimiter** box, select the delimiter that separates values within the selected file.

   d) If the first row in the worksheet should be used as column names in the dataset, select the **First row contains header definition** check box. At any time, you can select **See Preview** to view a preview of the dataset based on your current selections.

   e) Select **Next**.

   The **Modify Columns** section appears, displaying a grid.

   f) As needed, in the boxes in the **Column Name** column, modify the column names. You can only modify these column names if, in step 6, you did not select the **First row contains header definition** check box.

   g) As needed, in the boxes in the **Data Type** column, select the type of data that will be allowed in the column.

   h) For any columns in which you want each cell in that column to contain a unique value, in the row representing that column, select the check box in the **Unique Value** column.

   i) Select **Next**.
      The **Preview** section appears, displaying a preview of the dataset.

   j) Select **Finish**.
      A new page appears, displaying the dataset.

   k) Save the dataset.

Access a Dataset

Procedure

1. Access the **Dataset** page.

2. To access a dataset via the grid on the **Dataset** page:

   a. If you want to filter the datasets displayed in the grid, select ▼, and then, in the boxes that appear in the grid, enter information in the box in any column to limit the results displayed to datasets whose column values contain the information entered.

   b. In the grid, select the dataset that you want to access.
      A new page appears, displaying the selected dataset.

   -or-

   To access a dataset via the Catalog:

   a. In the upper-right corner of the **Dataset** page, select **Browse**.
      The **Select a dataset from the catalog** window appears.

   b. On the left side of the window, navigate to the folder that contains the dataset that you want to view.
      On the right side of the window, in the grid, the datasets in the selected folder appear.

   c. In the grid, select the dataset that you want to access, and then select **Open**.

A new page appears, displaying the selected dataset.

Modify a Dataset

Procedure

1. Access a dataset.

   - To delete a row, in the row that you want to delete, select **Delete**.
   - To add a row to the dataset, select ╋ .
   - To modify the value in any cell below the column headings, select the cell, and then modify the value.

   Note: Any value within a cell that has been modified since the last interim save will be surrounded by a green box. At any time, you can select ↺ to revert the dataset to the condition it had before the last interim save. At any time, you can select 💾 to perform an interim save, which will not overwrite the previously saved Catalog item.

2. Save the dataset.

Save a Dataset

Before You Begin

Saving a dataset is the final step of the process when you create a dataset in GE Digital APM , create a dataset by importing a file, or modify a dataset. To save a dataset, you should follow one of those procedures, and then complete the steps described in this topic.

Procedure

1. If you are creating a dataset:

   a) In the upper-right corner of the page, select 💾.
   b) On the left side of the window, navigate to the folder in which you want to save the dataset.
   c) On the bottom of the window, enter values in the **Name**, **Caption**, and **Description** boxes, and then select **Save**.
   d) Select **OK**.
      The dataset is saved as a Catalog item in the selected folder.

2. If you are modifying a dataset and want to overwrite that dataset's previously saved data:

   a) In the upper-right corner of the page, select 💾.
      The dataset is saved.

3. If you are modifying a dataset and want to save the modified dataset as a new Catalog item:

   a) In the upper-right corner of the page, select 💾.
   b) On the left side of the window, navigate to the folder in which you want to save the dataset.
   c) On the bottom of the window, enter values in the **Name**, **Caption**, and **Description** boxes, and then select **Save**.
   d) Select **OK**.
      The dataset is saved as a Catalog item in the selected folder.

Create a Graph from a Dataset

Procedure

1. Access a dataset.

2. In the upper-right corner of the page, select ⋮ , and then select **Create Graph**.

   A new page appears, displaying the dataset in the **Data Source** workspace.

3. Based on the type of graph that you want to create, complete the necessary steps.

   The graph is created, and appears on the **Graph Viewer** page.

Export a Dataset

Procedure

1. Access a dataset.

2. In the upper-right corner of the page, select ⋮ , and then select **Export to a File**.

   The **Export to a File** window appears.

3. In the **Please provide a File Name** box, enter a file name for the file that will be created.

4. In the drop-down list box, select the type of file that will be created, and then select **Export**.

   The specified file is created, and can be viewed, modified, and saved using a program that supports the specified file type.

Chapter

2

Reference

Topics:

- Dataset URLs

Dataset URLs

There is one URL route associated with datasets: **dataset**. The following table describes the various paths that build on the route, and the elements that you can specify for each.

Element	Description	Accepted Value(s)	Notes
#dataset/**Catalog Item Key**	Specifies the Catalog Item Key of the dataset that you want to open in the Dataset tool.	Numeric Catalog Item Key	None
#dataset?path=**Catalog Path \Dataset name**	Specifies the path and name of the dataset that you want to open in the Dataset tool.	Catalog item path	None

Example: Dataset URLs

Example URL	Destination
#dataset/**3223198**	Opens the dataset with the Catalog Item Key of 3223198 in the Dataset tool.
#dataset?path=**Public\Training Content\APM Framework \Datasets\APM Dataset**	Opens the dataset named 'APM Dataset' in the Dataset tool that is found in the specified Catalog folder.