Importing Data Into a Repository

In this section:

- Introduction
- Preparing to Import from Excel
- Adding the Data Repository Tool
- Configuring the Data Repository Tool
- Running the Tool

Introduction

The most common method for populating a repository is to use the **Generate Parameterized Messages** wizard when creating a test suite or virtual asset, but you can also structure and populate a data repository with Excel, WSDL, or schema files using the Data Repository Tool. The Data Repository Tool specifies the data to import, how to structure it, and which repository data source should be include with the imported data.

Use Case Clarifications

- Data Repository Tool functionality is not applicable to SQL data sets. For details on how to create a virtual asset from a database recording, see Creating SQL Responders from a Database Recording.
- The Data Repository Tool is not intended to create virtual assets that are parameterized from a data repository. For details on how to create parameterized message responders from traffic, see Creating Parameterized Message Responders from Traffic.

Preparing to Import from Excel

Additional preparation is required to successfully import data from an Excel file.

Defining the Structure

In order for the Data Repository Tool to recognize data record type hierarchies, at least one worksheet should have a column with a header labeled `<child_sheet_name> dsref*`, which establishes the rows in the worksheet as parent nodes. The `<child_sheet_name>` value should refer to the name of a sheet that contains child rows of data. The child sheet (sheet referenced with `<child_sheet_name>`) should contain a column with a header labeled `ParentIndex`. The cells in this column are numerical values that correspond to rows in the parent sheet. See the Example.

The dsref* and ParentIndex columns in your worksheets correspond to the **Join Column** and **Parent Join Column** fields in the Data Repository Tool configuration (see Configuring Nodes).

Example

In the following example, the Categories sheet contains a dsref* column that references the Lessons sheet, indicating that the rows of data in Lessons are children of the rows in Categories:

	A	B
1	Category	Lessons dsref*
2	test automation	
3	service virtualization	
4	environment management	
5	static analysis	
6	unit testing	
Categories		

The Lessons sheet contains a ParentIndex column, the values of which correspond to rows in the Categories sheet. The Data Repository Tool ignores the column header row, so the enumeration starts with row 2 in the Categories sheet:

	A	B	C	D

1	ID	Lesson	Schedule	ParentIndex
2	AT-1	Introduction to Automated Testing	Mondays	1
3	AT-2	Intermediate Automated Testing	Wednesdays	1
4	AT-3	Advanced Automated Testing	Fridays	1
5	SV-1	Introduction to Service Virtualization	Mondays	2
6	SV-2	Intermediate Service Virtualization	Wednesdays	2
7	SV-3	Advanced Service Virtualization	Fridays	2
8	EM-1	Introduction to Environment Management	Mondays	3
9	EM-2	Intermediate Environment Management	Wednesdays	3
10	EM-3	Advanced Environment Management	Fridays	3
11	SA-1	Static Analysis for Beginners	Mondays	4
12	SA-2	Intermediate Static Analysis Techniques	Wednesdays	4
13	SA-3	Advanced Static Analysis Techniques	Fridays	4
14	UT-1	Unit Testing for Beginners	Mondays	5
15	UT-2	Intermediate Unit Testing Techniques	Wednesdays	5
16	UT-3	Advanced Unit Testing Techniques	Fridays	5
Lessons				

When this example file is processed by the Data Repository Tool, the following structure will be applied:

- test automation
  - AT-1, Introduction to Automated Testing, Mondays
  - AT-2, Intermediated Automated Testing, Wednesdays
  - AT-3, Advanced Automated Testing, Fridays
- service virtualization
  - SV-1, Introduction to Service Virtualization, Mondays
  - SV-2, Intermediate Service Virtualization, Wednesdays
  - SV-3, Advanced Service Virtualization, Fridays
- environment management
  - EM-1, Introduction to Environment Management, Mondays
  - EM-2, Intermediate Environment Management, Wednesdays
  - EM-3, Advanced Environment Management, Fridays
- static analysis
  - SA-1, Introduction to Static Analysis, Mondays
  - SA-2, Intermediate Static Analysis, Wednesdays
  - SA-3, Advanced Static Analysis, Fridays
- unit testing
  - UT-1, Introduction to Unit Testing, Mondays
  - UT-2, Intermediate Unit Testing, Wednesdays
  - UT-3, Advanced Unit Testing, Fridays

See

Parameterizing Arrays of Varying Size in SOAtest or Parameterizing Arrays of Varying Size in Virtualize

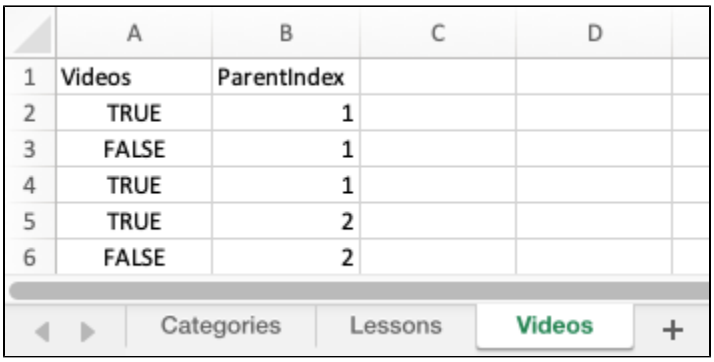
for additional information about dsref* and ParentIndex columns.

Support for Primitive Lists

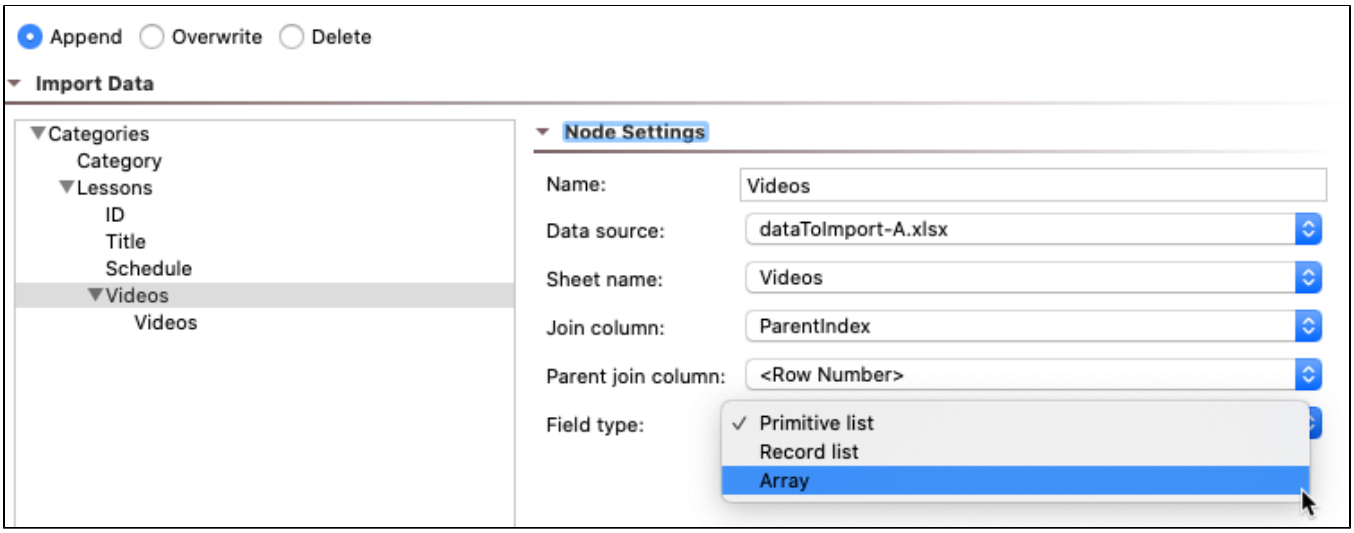
You can configure the Data Repository Tool to import columns as primitive lists, record lists, and arrays using the **Field type** setting (see Configuring the Data Repository Tool). The **Field type** setting becomes available when the following conditions apply:

- The child worksheet contains only one column.
- The name of the child worksheet is the same as the first column in the sheet.

In the following example, the **Field type** option will appear for the Videos node when configuring the Data Repository Tool to import the Excel file:



During the tool configuration step, you will be able to choose a field type for the Videos node:



Null and Exclude Values

The following strings are reserved values that are interpreted as "null" ("nil" for XML traffic):

- **[parasoft_null]**
- **[null]**

The following strings are reserved values that exclude the value from the generated message, regardless of the message type:

- **[parasoft_exclude]**
- **[exclude]**

For details on how [parasoft_exclude] works with URL parameter data source correlations, see Matching on Absent/Empty Fields and Parameters.

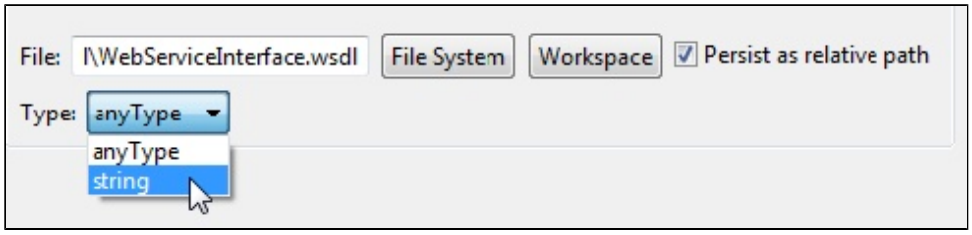
Adding the Data Repository Tool

1. Ensure that the Data Repository Server you want to import data into is running.
2. Create a new.tst or provisioning action file with an empty test or action suitein a new project or existing project. See Adding Projects, .tst files, and Test Suites or Adding Projects, Virtual Assets, and Responder Suites.
3. Add a new Repository Data Source to the suite. See Creating a Repository Data Source.
4. Open the data source and specify the target data set for importing the data. You can browse existing repositories on the server in the Repository view (see Views in Virtualize and Views in SOAtest). If the specified repository name or data set does not exist, it will be created at runtime.
5. Right-click the **Test** or **Action Suite** node and choose **Add New> Test** or **Action**.and choose **Data Repository Tool.**
6. Click **Next**. and configure the data and structure initialization options. See Configuring Initialization Options.
7. Click **Finish** and configure the Data Repository Tool.

If you initialized the data and the structure based on an Excel file, the workbook will be imported as an Excel data source and the tool will be configured according to the data. Review the tool configuration and modify the settings as necessary before running the tool (see Configuring the Data Repository Tool).

Configuring Initialization Options

You can configure the following initialization settings when creating a Data Repository Tool. The settings in this screen determine how to initialize the imported data structure.

Initialize data from	• Choose **None** to manually initialize the data. See Configuring the Data Repository Tool. • Choose **Excel** to initialize data from an Excel file. Refer to the Preparing to Import from Excel section for information about preparing the data.
Initialize structure from	• Choose **None** to manually structure the data. • Choose **Excel** to use the structure from an Excel file to structure the data. Refer to the Preparing to Import from Excel section for information about preparing the data. • Choose **WSDL** to use the structure from a WSDL definition file to structure the data. • Choose **Schema** to use the structure from a schema file to structure the data. When specifying WSDLs and schemas, specify the definition file and choose the data type definition from the Type menu. 
Join column	If initializing from an Excel file, specify the name of the column that indicates relationships across Excel sheets in the Join column field. By default, the Join column is the ParentIndex column, but you can also point to another column that contains values for referencing the parent row. See Defining the Structure additional information.

Configuring the Data Repository Tool

Verify that the Target Repository is set to the Repository Data Source that you created (see Adding the Data Repository Tool).

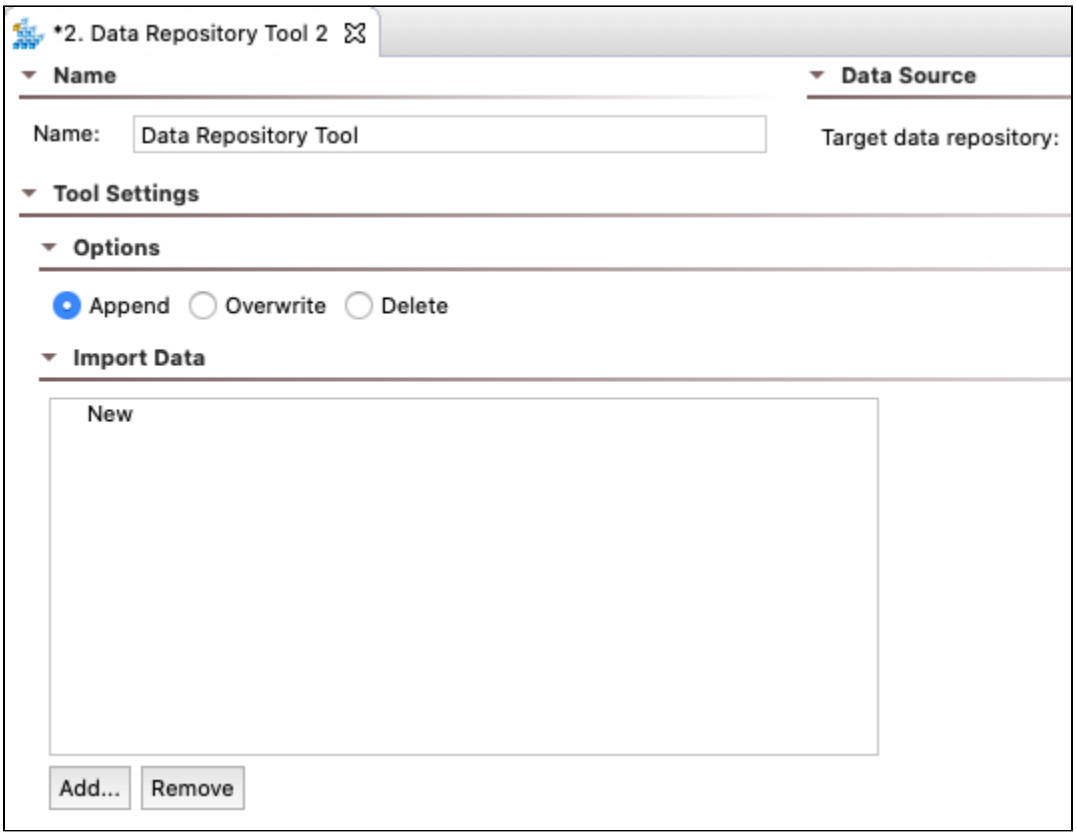
Data Processing Options

In the **Tool Settings > Options** section, enable a mode for processing data:

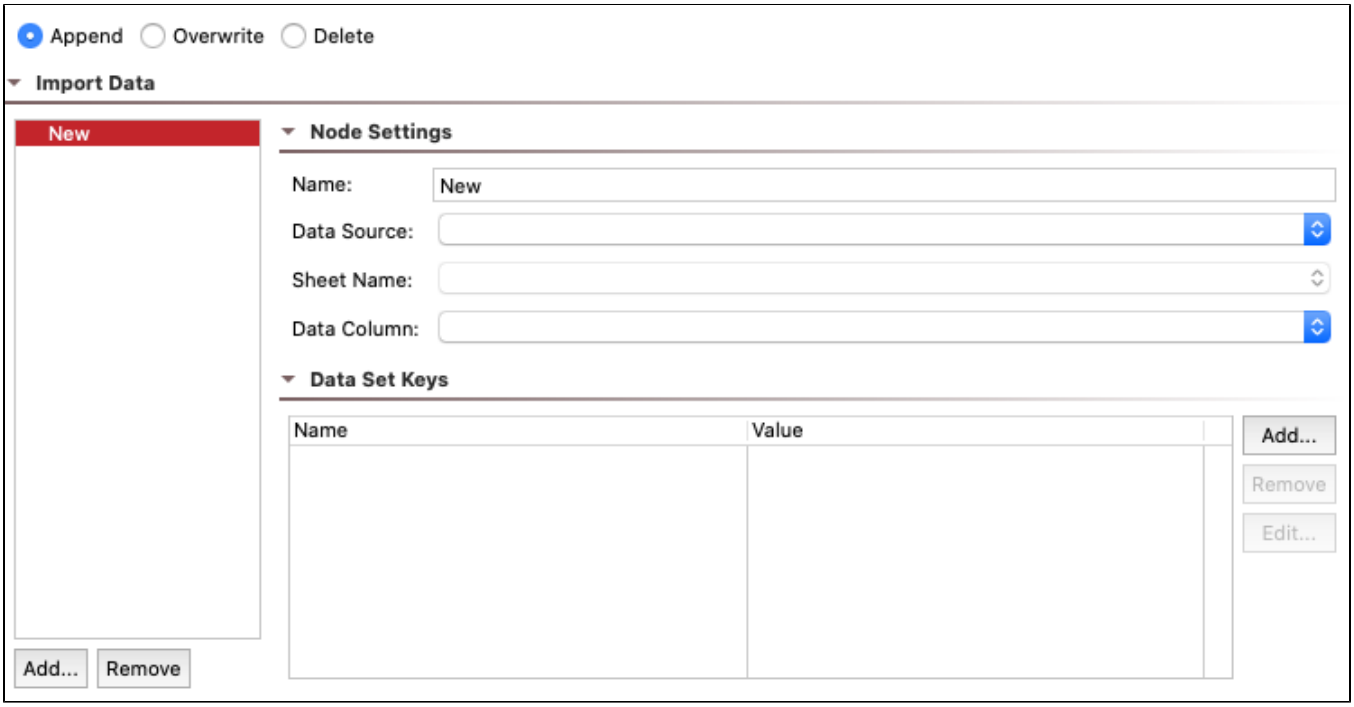
- **Append**: Adds records to the data an existing set. Existing records will not be altered.
- **Overwrite**: If matching records are detected in overwrite mode, they will be replaced.
- **Delete**: Removes matching records.

Configuring the Data Structure

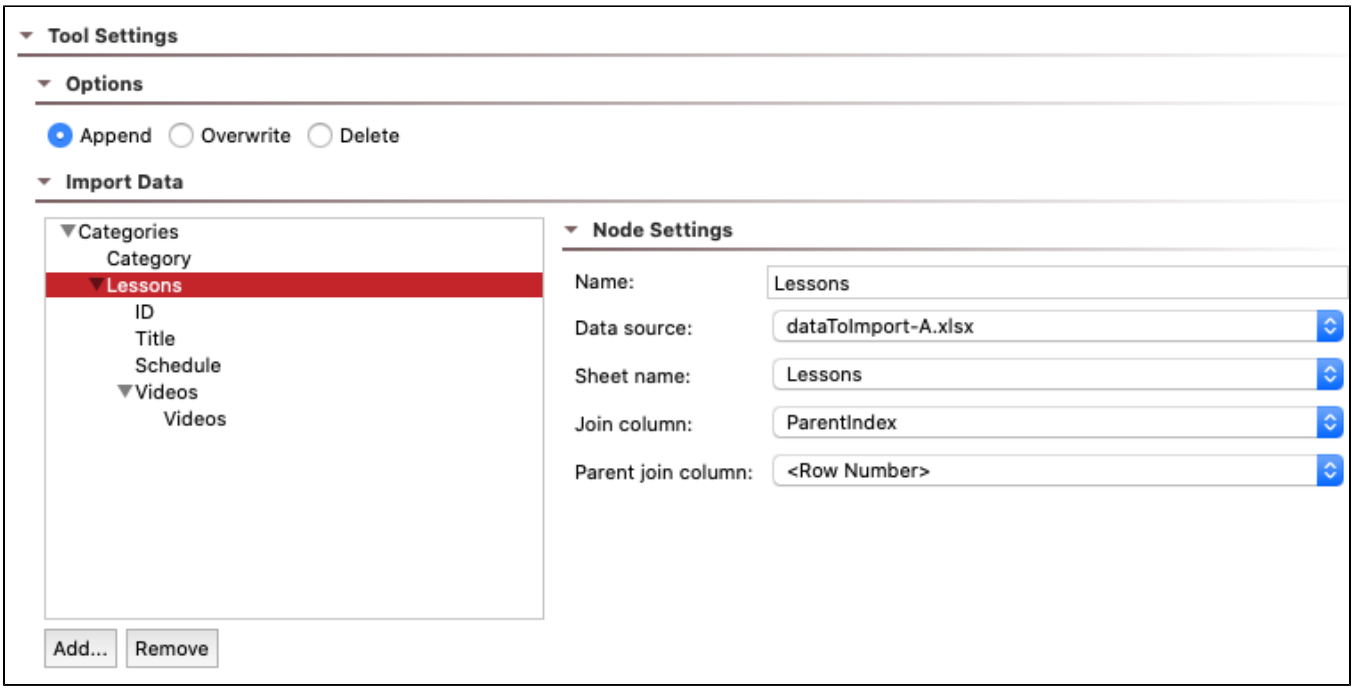
If you created the Data Repository Tool without initializing data or a structure, a default New node will appear in the panel.



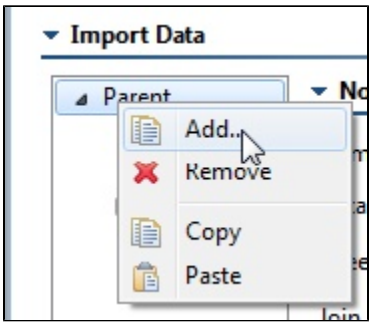
Click on the node to access it's configuration options (see Configuring Nodes).



If you initialized the structure from Excel during tool creation, the structure will already be applied. You can choose a node from the panel and configure as necessary (see Configuring Nodes).

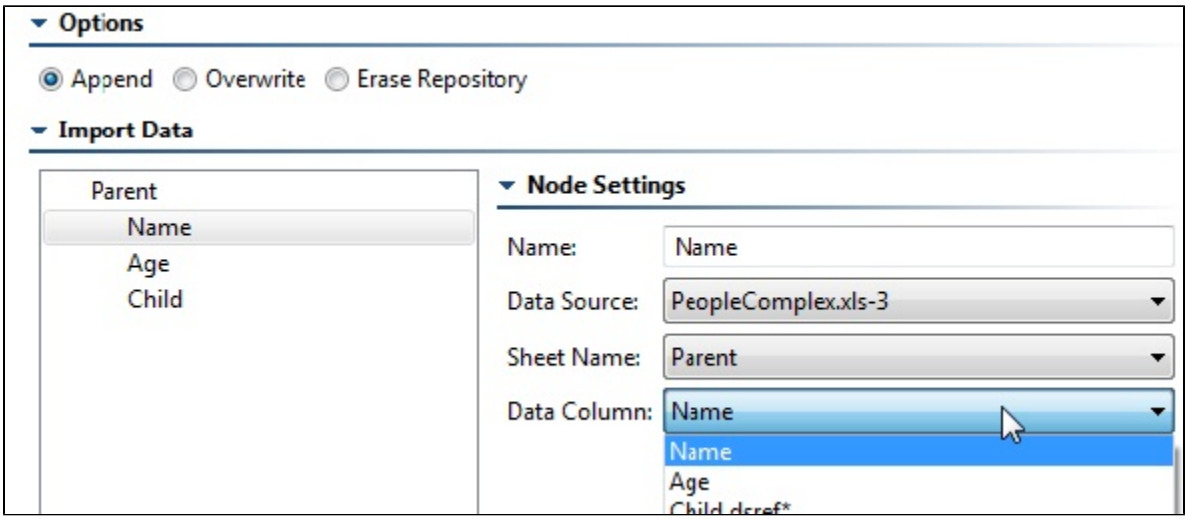


You can click **Add...** and **Remove** in the Import Data panel to manually add nodes to the data tree. You can also right-click on existing nodes to remove the node or add children.



Configuring Nodes

Each node in the data tree represents a part of the source data you are importing. If you chose to initialize data based on one of the options in the initialization screen (Configuring Initialization Options), this section will be configured automatically, but you can modify the following settings as necessary prior to running the tool.

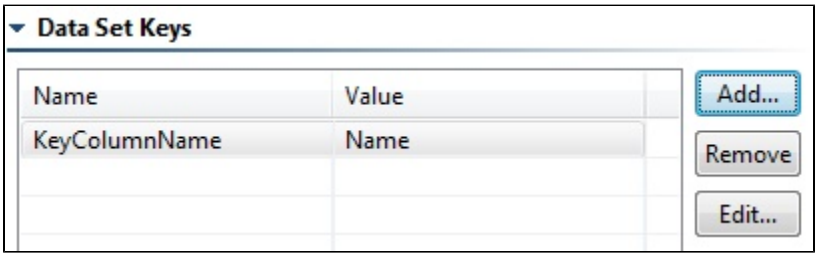


The available settings depend on the nodes' relationships:

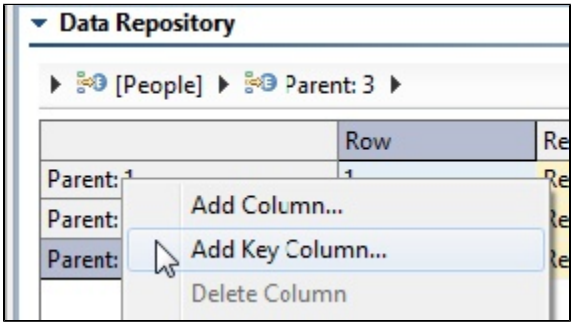
Name	Specifies the name for the node.
Data Source	Specifies the Excel workbook used to initialize the data (see Configuring Initialization Options). If the **None** option was enabled, this field will be empty and data will need to be added manually using the data source editor (see Viewing and Modifying the Repository Structure and Contents).
Sheet Name	Specifies a sheet name when importing from Excel.
Join Column	Specifies the name of the column used for indicating relationships across Excel sheets. By default, the ParentIndex column is used (see Defining the Structure). If the value in a record's Join Column matches the value in a Parent Join Column record, then that record becomes a child of the parent record. This setting is not available for column-level nodes. This setting cannot be configured for worksheet-level nodes at the root of the tree.
Parent Join Column	Specifies the column in the parent node used for indicating relationships across Excel sheets. By default, the value is set to **<Row Number>**, which specifies that the column correlates to the parent row number. If the value in a record's Join Column matches the value in a Parent Join Column record, then that record becomes a child of the parent record. This setting is not available for column-level nodes. This setting cannot be configured for worksheet-level nodes at the root of the tree.
Data Column	Specifies name of the column containing the data.
Field Type	Specifies which type the data should be imported as. You can specify one of the following types: • Record List: Series of complex records that are hierarchical and have multiple fields/columns. • Primitive List (see Support fro Primitive Lists): A list of primitive types. • Array: An array of primitive types. See Viewing and Modifying the Repository Structure and Contents for additional information about types.

Specifying Data Set Keys

Data set keys allow you to specify key columns that will be used for Virtualize responder correlations. These columns can later be selected in the Data Source Correlations tab.



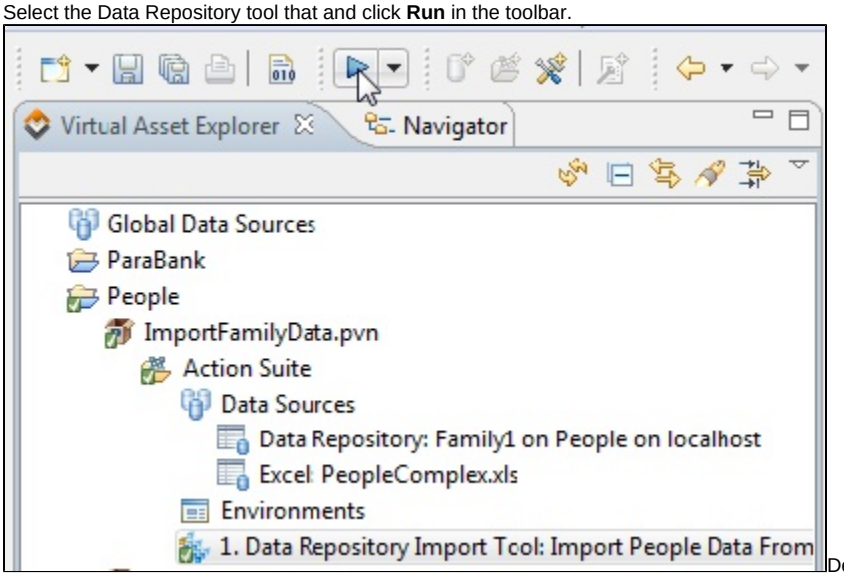
If you don't specify these key columns now, you can add them later in the Data Editor.



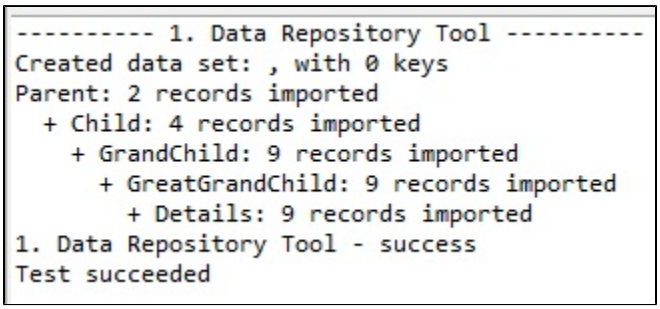
See Responder Correlation Tab for additional information.

Running the Tool

Select the Data Repository tool that and click **Run** in the toolbar.



Details on the import will be reported in the Console view.



Any problems with the import will be reported in the progress view. For instance, the following message might occur if you are trying to import data into a Repository Server that is not currently running.