Paper S107-2011

Moving between an XML document and SAS using the SAS XML Mapper

Margo Law, South Dakota State University, Brookings, SD

Abstract:

Moving between an XML document and SAS using the SAS XMLMapper. Starting with an XML document containing multiple tables this presentation will demonstrate how to use the SAS XML Mapper to move the multiple tables into SAS, either individually or all at once. Additional items that will also will be demonstrated and discussed include the difference between an XML library and a SAS library and moving between the two, and moving SAS data sets back out to an XML document, either individually or multiple data sets at once, along with the benefits and difficulties of doing so.

Introduction:

With XML files that contain a large number of different tables and/or a lot of data it can be difficult to navigate through all of the information, especially when trying to locate particular information. The SAS XML Mapper allows for all of the information within the XML file to be separated out into individual data sets, which then, if desired, can be moved into SAS, where the information can then be more easily located and managed.

Moving the XML document into the XML Mapper

- Beginning with an XML document:
  - Open XML Mapper

- Opening the XML file:
  - File → open XML → (locate the desired xml file and click open)
    (information will appear in the upper left window and on the bottom window, under the XML source tab, of the XML mapper)

- Auto generating the XML map schema:
  - Tools → AutoMap using XML (or Control+M)
    (information will appear in the upper right window of the XML mapper)

    Clicking on a data set name (right window) and the table view tab (bottom window) enables you to view all the data that is in that table; clicking on the contents tab (bottom window) enables you to view the information about all the variables contained within that table.

Clicking on the + next to the data set name (right window) expands the table and lists all of the variables found within that table. The blue spheres represent numeric variables, red pyramids represent character variables. When clicking on a specific variable this row will be highlighted under the contents tab (bottom window) or the column will be highlighted under the table view tab (bottom window).

Suppressing specific variable information within a table

- Click on the desired variable within the table expansion (right window).
  - Dates (for example) could be converted to text by clicking on the format tab along the top right side, selecting character (under the type dropdown box) and string (under the datatype dropdown box), then clearing all formats and informats (the clear button on the right).
  - Character variables could be made blank by clicking the condition tab along the top right side and the condition box.
  - Numeric variables could be changed to all -3 (for example) by clicking on the enumeration tab along the top right, typing in -3 into the box on the right side, clicking 'add', then changing the default value to -3.

Saving the XML map for all tables at one time

- Click on AUTO_GEN (in the right window at the very top of all the tables)

  (this name could be changed by clicking on the properties tab along the top right and changing the Name line)

- File → Save XML Map As → (place in a desired location and provide a name)

Saving an XML map for one table at a time

- Click on the table name (right window)
- File→Save XML Map As→(place in a desired location and provide a name)

Moving the tables from the SAS XML Mapper (XML document) into SAS

```
/* importing the entire xml document to multiple SAS data sets in an
   xml library */

filename xmlfile 'I:\sugar_test_data.xml';
      /* the original xml document */
```

```
filename map 'I:\wholexml.map';
     /* wholexml.map is the xml map that was saved in the xml mapper */

libname xmlfile xml xmlmap=map access=readonly;
     /* xmlfile is the name of the xml library

         xml after the library name is required

         access=readonly allows for viewing the tables that are in the
         xml library (this is optional, but if it is not included the
         tables in the xml library will not be able to be opened and
         viewed). */


/*************************************************************************/

/* importing one table from the xml document to a SAS data set in an
xml library */


filename xmlfile 'I:\sugar_test_data.xml';
            /* the original xml document */

filename map 'I:\onetable.map';
     /* onetable.map is the xml map that was saved in the xml mapper */

libname xmlfile xml xmlmap=map access=readonly;
           /* xmlfile is the name of the XML library

              xml after the library name is required

       access=readonly allows for viewing the table that is in the xml
       library (this is optional, but if it is not included the tables
       in the xml library will not be able to be opened and viewed). */
```

Moving from an XML library to a SAS library

```
/* copying everything in the XML library (xmlfile) into a SAS library
                                                          (myfiles) */

libname myfiles 'I:\myfiles';
      /* myfiles is the SAS library */


proc copy in=xmlfile out=myfiles;
run;
```

When moving into SAS, all of the tables from within the XML document are arranged in alphabetical order within the SAS library.

```
/*****************************************************************/

/* Moving one table within the XML library (xmlfile) into a SAS library
                                                          (myfiles) */

Data myfiles.person;
 Set xmlfile.person;
Run;
```

Moving the SAS data sets into an XML document

```
/* moving all the SAS data sets from the SAS library myfiles into one
   XML document */


libname xmlfile xml 'I:\output_file.xml';
       /* output_file is the name of the XML document */

proc copy in=myfiles out=xmlfile;
run;
```

Within SAS all the tables are arranged in alphabetical order, and therefore the corresponding XML document will have the tables arranged in alphabetical order.

Conclusion:

Moving an XML file into the SAS XML Mapper, then from the XML Mapper into SAS, allows for viewing all of the information contained within the XML file and locating and working with specific information much quicker and easier. However, currently there is no way of maintaining the order of the tables from the original XML document when moving into SAS or from within SAS back out to an XML document. There is also no way of moving directly from an XML document into the SAS XML Mapper and back out to an XML document without moving the tables into SAS.

References:

'Help' tab within the XML Mapper

- 'Help Topics'
  - Provides examples and step by step instructions for certain tasks.
- 'Web Resources and Updates'
  - A link to support.sas.com, the XML LIBNAME Engine page.

Contact Information:
Margo Law
South Dakota State University
325 Wecota Hall
Brookings, SD 57007
605-688-4608 (office)
605-688-4220 (fax)
Margo.law@sdstate.edu