## Contents

Exporting an XML Document Using a Customized Tagset .............................. 60

### PART 2  LIBNAME Statement Syntax 67

**Chapter 7 • Introduction to the XMLV2 LIBNAME Statements** ............................. 69
  Using the LIBNAME Statement ................................................. 69
  Comparing the XMLV2 and XML Engines .................................... 69

**Chapter 8 • LIBNAME Statement for the XMLV2 Engine** .................................. 73
  Dictionary .............................................................................. 73

### PART 3  XMLMap File Syntax 89

**Chapter 9 • Introduction to XMLMap Syntax** .................................................. 91
  Using XMLMap Syntax ............................................................. 91
  Comparing the XMLMap Syntax ................................................. 91

**Chapter 10 • XMLMap Syntax Version 2.1** ..................................................... 95
  Dictionary .............................................................................. 95

**Chapter 11 • Using SAS XML Mapper to Generate and Update an XMLMap** ........... 113
  What Is SAS XML Mapper? ....................................................... 113
  Using the Windows ................................................................. 114
  Using the Menu Bar ............................................................... 114
  Using the Toolbar ................................................................. 114
  How Do I Get SAS XML Mapper? ................................................ 115
  To Start SAS XML Mapper ..................................................... 115

### PART 4  Appendixes 117

**Appendix 1 • Usage Not Supported for the XMLV2 Engine** ........................... 119
  Exporting an XML Document for Use by Oracle .............................. 119
  Exporting Numeric Values ....................................................... 121
  Exporting an XML Document in CDISC ODM Markup ...................... 126
  Importing an XML Document with Numeric Values ......................... 127
  Importing an XML Document Created by Microsoft Access ................ 128
  Importing Concatenated XML Documents ..................................... 132
  Importing a CDISC ODM Document .......................................... 134

**Appendix 2 • Example CDISC ODM Document** ............................................. 137
  Recommended Reading .................................................................. 143
  Glossary ................................................................................. 145
  Index ....................................................................................... 151
What’s New in the SAS 9.4 XMLV2 Engine

Overview

The XMLV2 LIBNAME statement supports new options and provides more functionality in z/OS environments. Several documentation enhancements have been made.

New XMLV2 LIBNAME Statement Options

In SAS 9.4M6, the following LIBNAME statement options are new:

• **CHARMULTIPLIER**= option XMLV2 Only expands column (variable) lengths by a multiplier value. These column (variable) lengths are specified in an XMLMap.

• **DERIVECHARMULTIPLIER**= option XMLV2 Only expands column (variable) lengths by a default multiplier value that is based on the session encoding. These column (variable) lengths are specified in an XMLMap.

In SAS 9.4M4, the new **PREFIXATTRIBUTES**= option XMLV2 Only specifies whether the element name is concatenated to the attribute name when generating each XMLMap COLUMN element.

Documentation Enhancements

In SAS 9.4M6, the XMLV2 engine is supported throughout the documentation, except when a feature is supported by the XML engine only. In the syntax documentation, each language element is labeled with a superscript such as XMLV2 Only or XML Only or XMLV2 and XML. Examples that are supported for the XML engine only and not supported for the XMLV2 engine are now in an appendix: Appendix 1, “Usage Not Supported for the XMLV2 Engine,” on page 119. For more information, see “Comparing the XMLV2 and XML Engines” on page 69.

In SAS 9.4M6, information about the encoding of imported XML documents is added in “Transferring an XML Document across Environments” on page 6.

In SAS 9.4M5, documentation for SAS Viya is merged with SAS 9.4M5 documentation.
Operating Environment Support

In SAS 9.4, XMLV2 engine functionality for the z/OS environment changed from preproduction to production. The engine is production in all SAS 9.4 operating environments.
Part 1

Usage

Chapter 1
  Getting Started with the XMLV2 Engine  3

Chapter 2
  Importing XML Documents  9

Chapter 3
  Exporting XML Documents  15

Chapter 4
  Importing XML Documents Using an XMLMap  19

Chapter 5
  Exporting XML Documents Using an XMLMap  55

Chapter 6
  Understanding and Using Tagsets for the XMLV2 Engine  59
## Chapter 1

### Getting Started with the XMLV2 Engine

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>What Does the XMLV2 Engine Do?</td>
<td>3</td>
</tr>
<tr>
<td>Understanding How the XMLV2 Engine Works</td>
<td>4</td>
</tr>
<tr>
<td>Assigning a Libref</td>
<td>4</td>
</tr>
<tr>
<td>Importing an XML Document</td>
<td>4</td>
</tr>
<tr>
<td>Exporting an XML Document</td>
<td>5</td>
</tr>
<tr>
<td>SAS Processing Supported by the XMLV2 Engine</td>
<td>5</td>
</tr>
<tr>
<td>Transferring an XML Document across Environments</td>
<td>6</td>
</tr>
<tr>
<td>Frequently Asked Questions</td>
<td>6</td>
</tr>
<tr>
<td>Is the XMLV2 Engine a DOM or SAX Application?</td>
<td>6</td>
</tr>
<tr>
<td>Does the XMLV2 Engine Validate an XML Document?</td>
<td>6</td>
</tr>
<tr>
<td>What Is the Difference between Using the XMLV2 Engine and Using ODS MARKUP?</td>
<td>7</td>
</tr>
<tr>
<td>Why Do I Get Errors When Importing XML Documents Not Created with SAS?</td>
<td>7</td>
</tr>
<tr>
<td>What Are the XMLV2 and XML Engines?</td>
<td>7</td>
</tr>
<tr>
<td>Accessibility Features of the XMLV2 Engine</td>
<td>7</td>
</tr>
</tbody>
</table>

### What Does the XMLV2 Engine Do?

The XMLV2 LIBNAME engine processes an XML document. The engine can do the following:

- export (write to an output location) an XML document from a SAS data set of type DATA by translating the SAS proprietary file format to XML markup. The output XML document can then be:
  - used by a product that processes XML documents.
  - moved to another host for the engine to process by translating the XML markup back to a SAS data set.
  - import (read from an input location) an external XML document. The input XML document is translated to a SAS data set.
Understanding How the XMLV2 Engine Works

Assigning a Libref

The XMLV2 LIBNAME engine works much like other SAS engines. That is, you execute a LIBNAME statement to assign a libref and specify an engine. You use that libref throughout the SAS session where a libref is valid.

A libref for the XMLV2 engine can be associated with either a specific XML document or the physical location of a SAS library in a directory-based environment. (The older XML engine can associate a libref with a specific document, but not a library.) When you use the libref, SAS either translates the data in a SAS data set into XML markup, or translates the XML markup into SAS format.

Importing an XML Document

To import an XML document as a SAS data set, the following LIBNAME statement assigns a libref to a specific XML document and specifies the XMLV2 engine:

```
libname myxml xmlv2 'C:\Example\Students.xml';
```

Executing the DATASETS procedure shows that SAS interprets the XML document as a SAS data set:

```
proc datasets library=myxml;
run;
quit;
```

Output 1.1  DATASETS Procedure Output for MyXml Library

<table>
<thead>
<tr>
<th>The SAS System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directory</td>
</tr>
<tr>
<td>Libref</td>
</tr>
<tr>
<td>Engine</td>
</tr>
<tr>
<td>Physical Name</td>
</tr>
<tr>
<td>XML Type</td>
</tr>
<tr>
<td>Version</td>
</tr>
</tbody>
</table>

The PRINT procedure results in the following output:

```
proc print data=myxml.Students;
run;
```
Exporting an XML Document

To export an XML document from a SAS data set, the LIBNAME statement for the XMLV2 engine assigns a libref to the XML document to be created.

In the following code, the first LIBNAME statement assigns the libref MyFiles to the SAS library that contains the SAS data set Singers. The second LIBNAME statement assigns the libref MyXml to the physical location of the XML document that is to be exported from MyFiles.Singers:

libname myfiles 'C:\Example\';
libname myxml xmlv2 'C:\Output\Singers.xml';

Executing these statements creates the XML document named Singers.XML:

data myxml.Singers;
  set myfiles.Singers;
run;

Output 1.3 XML Document Singers.XML

```xml
<?xml version="1.0" encoding="utf-8" ?>
<TABLE>
  <SINGERS>
    <FirstName>Tom</FirstName>
    <Age>62</Age>
  </SINGERS>
  <SINGERS>
    <FirstName>Willie</FirstName>
    <Age>70</Age>
  </SINGERS>
  <SINGERS>
    <FirstName>Randy</FirstName>
    <Age>43</Age>
  </SINGERS>
</TABLE>
```

SAS Processing Supported by the XMLV2 Engine

The XMLV2 engine supports the following processing:

- The engine supports input (read) and output (create) processing. The engine does not support update processing.
The engine is a sequential access engine in that it processes data one record after the other. The engine starts at the beginning of the file and continues in sequence to the end of the file. The engine does not provide random (direct) access, which is required for some SAS applications and features. For example, you cannot use the SORT procedure or ORDER BY in the SQL procedure. If you request processing that requires random access, a message in the SAS log notifies you that the processing is not valid for sequential access. If this message occurs, put the XML data into a temporary SAS data set before you continue.

Transferring an XML Document across Environments

When you transfer an XML document across environments (for example, using FTP), you must be aware of the document's content to determine the appropriate transfer mode. If the document contains an encoding attribute in the XML declaration or if a byte-order mark precedes the XML declaration, transfer the file in binary mode. If the document contains neither criteria and you are transferring the document across similar hosts, transfer the file in text mode.

When you import an XML document that was created in a different character encoding, be aware of the possibility of transcoding errors. If an XML document or XMLMap does not specify an ENCODING= attribute in the XML declaration, then the engine attempts to identify the encoding from a byte-order mark. This behavior is consistent with World Wide Web Consortium (W3C) specifications. If neither an ENCODING= attribute nor a byte-order mark is found, the default is UTF-8 encoding. If the document's encoding is not compatible with UTF-8, a transcoding error could occur.

When you export an XML document using the XMLV2 engine, by default, the XML document contains an encoding attribute in the XML declaration from the SAS data set's encoding. Here is an example:

```xml
<?xml version="1.0" encoding="utf-8" ?>
```

You can override the SAS data set's encoding when you export the XML document by specifying the XMLENCODING= LIBNAME statement option.

Frequently Asked Questions

Is the XMLV2 Engine a DOM or SAX Application?

The XMLV2 engine uses a Simple API for XML (SAX) model, not a Document Object Model (DOM). SAX does not provide a random-access lookup to the document's contents. It scans the document sequentially and presents each item to the application one item at a time.

Does the XMLV2 Engine Validate an XML Document?

The XML engine does not validate an input XML document. The engine assumes that the data passed to it is in valid, well-formed XML markup. Because the engine does not use a DTD (Document Type Definition) or SCHEMA, there is nothing to validate.
against. Successful processing without log messages does not ensure that the XML markup is valid. You must check the data to ensure that it is complete.

**What Is the Difference between Using the XMLV2 Engine and Using ODS MARKUP?**

The XMLV2 engine creates and reads XML documents. ODS MARKUP creates, but does not read, XML documents. Typically, you use the engine to transport data, and you use the ODS MARKUP destination to create XML from SAS output.

**Why Do I Get Errors When Importing XML Documents Not Created with SAS?**

The XMLV2 engine reads files that conform to the markup types supported in the XMLTYPE= LIBNAME statement option. Attempting to import free-form XML documents that do not conform to the specifications required by the supported markup types can generate errors. To successfully import files that do not conform to the XMLTYPE= markup types, you can create a separate XML document, called an XMLMap. The XMLMap syntax tells the engine how to interpret the XML markup into a SAS data set or data sets, variables (columns), and observations (rows). See Chapter 4, “Importing XML Documents Using an XMLMap,” on page 19.

**What Are the XMLV2 and XML Engines?**

SAS provides two versions of XML LIBNAME engine functionality by supporting the two engine names **XMLV2** and **XML** in the LIBNAME statement. See “Comparing the XMLV2 and XML Engines” on page 69.

---

**Accessibility Features of the XMLV2 Engine**

The XMLV2 LIBNAME engine is a command-based product. For this release, no features were added to address accessibility, but the product might be compliant to accessibility standards because it does not have a graphical user interface, and all of its features are available to anyone who can type or otherwise produce a command. If you have specific questions about the accessibility of SAS products, send them to accessibility@sas.com or call SAS Technical Support.
Chapter 2
Importing XML Documents

Understanding How to Import an XML Document

Importing an XML document is the process of reading an external XML document as a SAS data set. The XMLV2 engine translates the input XML document to the SAS proprietary file format.

To import an XML document, you execute the LIBNAME statement for the XMLV2 engine in order to assign a libref to the physical location of an existing XML document. Then, you execute SAS code to access the XML document as a SAS data set.

Importing an XML Document Using the GENERIC Markup Type

This example imports the following XML document, which conforms to the physical structure for the GENERIC markup type. For information about the required physical structure, see “Understanding the Required Physical Structure for an XML Document to Be Imported Using the GENERIC Markup Type” on page 20.

```xml
<?xml version="1.0" encoding="windows-1252" ?>
<TABLE>
  <CLASS>
    <Name> Alfred </Name>
    <Gender> M </Gender>
    <Age> 14 </Age>
    <Height> 69 </Height>
    <Weight> 112.5 </Weight>
  </CLASS>
  <CLASS>
    <Name> Alice </Name>
  </CLASS>
</TABLE>
```
The following SAS program translates the XML markup to SAS proprietary format:

```sas
libname trans xmlv2 'C:\Example\class.xml'; */1*/
libname myfiles 'C:\Output\'; /*2*/

data myfiles.class; /*3*/
   set trans.class;
run;
```

1 The first LIBNAME statement assigns the libref Trans to the physical location of the XML document (complete pathname, filename, and file extension) and specifies the XMLV2 engine. By default, the XMLV2 engine expects GENERIC markup.

2 The second LIBNAME statement assigns the libref MyFiles to the physical location of the SAS library that will store the resulting SAS data set. The V9 engine is the default.

3 The DATA step reads the XML document and writes its content in SAS proprietary format.

Issuing the following PRINT procedure produces the output for the data set that was translated from the XML document:

```sas
proc print data=myfiles.class;
run;
```
Importing an XML Document with Non-Escaped Character Data

W3C specifications (section 4.6, Predefined Entities) state that for character data, certain characters such as the ampersand (&) and the apostrophe (’) must be escaped using character references or strings like & and ’. For example, to allow attribute values to contain both single and double quotation marks, the apostrophe or single-quotation character (’) can be represented as &apos; and the double-quotation character (“) as &quot;.

To import an XML document that contains non-escaped characters, you can specify the LIBNAME statement option XMLPROCESS=PERMIT. When that option is specified, the XMLV2 engine accepts non-escaped characters like the apostrophe, double quotation marks, and the ampersand in character data. (Non-escaped angle brackets are not supported.)
Note: Use XMLPROCESS=PERMIT cautiously. If an XML document consists of non-escaped characters, the content is not standard XML construction. The option is provided for convenience, not to encourage invalid XML markup.

This example imports the following XML document named Strings.XML, which contains non-escaped character data:

```xml
<?xml version="1.0" ?>
<STRINGS>
  <CHARS>
    <status>proper escape sequence</status>
    <ampersand>&amp;</ampersand>
    <squote>&apos;</squote>
    <dquote>&quot;</dquote>
    <greater>&gt;</greater>
  </CHARS>
  <CHARS>
    <status>unescaped character in CDATA</status>
    <ampersand><![CDATA[Abbott & Costello]]></ampersand>
    <squote><![CDATA[Logan's Run]]></squote>
    <dquote><![CDATA[As Benjamin Franklin advised, "Well done is better than well said."]]></dquote>
    <greater><![CDATA[x > y]]></greater>
  </CHARS>
  <CHARS>
    <status>unescaped character in string</status>
    <ampersand>Dunn & Bradstreet</ampersand>
    <squote>Isn't this silly?</squote>
    <dquote>Quoth the raven, "Nevermore!"</dquote>
    <greater> & </greater>
  </CHARS>
</CHARS>
</STRINGS>
```

The example code below shows the default behavior, which is XMLPROCESS=CONFORM. Under the default, the XMLV2 engine expects XML markup to conform to W3C specifications. The following LIBNAME statement results in errors, and the MyFiles libref is not assigned.

```sas
libname myfiles xmlv2 'C:\Example\strings.xml';
```

Log 2.1  SAS Log Output

```
1 libname myfiles xmlv2 'C:\Example\strings.xml';
ERROR: There is an illegal character in the entity name. occurred at or near line 19, column 24
ERROR: XML parsing error. Please verify that the XML content is well-formed.
ERROR: Error in the LIBNAME statement.
```

Specifying the LIBNAME statement option XMLPROCESS=PERMIT enables the XMLV2 engine to import the XML document:

```sas
libname myfiles xmlv2 'C:\Example\strings.xml' xmlprocess=permit;
proc print data=myfiles.chars;
run;
```
Output 2.2  PRINT Procedure Output for MyFiles.Chars

<table>
<thead>
<tr>
<th>Obs</th>
<th>status</th>
<th>ampersand</th>
<th>squote</th>
<th>dquote</th>
<th>greater</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>proper escape sequence</td>
<td>&amp;</td>
<td>'</td>
<td>&quot;</td>
<td>&gt;</td>
</tr>
<tr>
<td>2</td>
<td>unescaped character in CDATA</td>
<td>Abbott &amp; Costello</td>
<td>Logans Run</td>
<td>As Benjamin Franklin advised, &quot;Well done is better than well said.&quot;</td>
<td>x &gt; y</td>
</tr>
<tr>
<td>3</td>
<td>unescaped character in string</td>
<td>Dunn &amp; Bradstreet</td>
<td>Isn't this silly?</td>
<td>Quoth the raven, &quot;Nevermore!&quot;</td>
<td>&gt;</td>
</tr>
</tbody>
</table>
Chapter 3
Exporting XML Documents

Understanding How to Export an XML Document
Exporting an XML Document That Contains SAS Dates, Times, and Datetimes
Exporting an XML Document with Separate Metadata

Understanding How to Export an XML Document

Exporting an XML document is the process of writing a SAS data set of type DATA to an output XML document. The XMLV2 LIBNAME engine exports an XML document by translating SAS proprietary format to XML markup.

To export an XML document, you execute the LIBNAME statement for the engine in order to assign a libref to the physical location of an XML document to be created. Then you execute SAS code that produces output such as a DATA step or the COPY procedure.

Exporting an XML Document That Contains SAS Dates, Times, and Datetimes

This example exports an XML document from a SAS data set that contains datetime, date, and time values. The XML document is generated for the GENERIC markup type.

First, the following SAS program creates a simple SAS data set and prints the contents of the data set. The variables generate their values from the DATETIME(), DATE(), and TIME() functions.

```sas
data test;
  Var1=DATETIME();
  format Var1 datetime.;
  Var2=DATE();
  format Var2 date9.;
  Var3=TIME();
  format Var3 timeampm.;
  proc print data=test;
  run;
```
Output 3.1 PRINT Procedure Output for Work.Test Containing SAS Datetime, Date, and Time

The following code exports an XML document for the GENERIC markup type that includes the SAS datetime, date, and time information:

libname trans xmlv2 'C:\Output\Test.xml' xmltype=generic; /*

data trans.test; /*
   set work.test;
run;

The LIBNAME statement assigns the libref Trans to the physical location of the file that will store the exported XML document and specifies the XMLV2 engine. The physical location includes the complete pathname, filename, and file extension. XMLTYPE= specifies the GENERIC markup type, which is the default.

The DATA step reads the SAS data set Work.Test and writes its content in XML markup to the specified XML document.

Here is the resulting XML document.

Output 3.2 XML Document Using GENERIC Markup

<?xml version="1.0" encoding="utf-8" ?>
<TABLE>
   <TEST>
      <Var1>2017-10-10T13:36:32</Var1>
      <Var2>2017-10-10</Var2>
      <Var3>13:36:32</Var3>
   </TEST>
</TABLE>

Exporting an XML Document with Separate Metadata

This example exports an XML document from a SAS data set and specifies a separate file to contain metadata-related information. The example illustrates using the XMLMETA= option and XMLSCHEMA= option and uses a SAS data set from the Sashelp library.

First, here is the CONTENTS procedure output for the SAS data set Sashelp.Snacks:
The following SAS program exports an XML document from the SAS data set Sashelp.Snacks:

```sas
filename myxsd 'C:\Output\snacks.xsd'; /* 1 */
libname output xmlv2 'C:\Output\snacks.xml'
  xmlmeta=schemadata xmlschema=myxsd; /* 2 */
data output.snacks; /* 3 */
  set sashelp.snacks;
run;
```

1. TheFILENAME statement assigns the fileref MyXsd to the physical location of the separate external file that will contain the metadata-related information.

2. TheLIBNAME statement assigns the libref Output to the physical location of the file that will store the exported XML document and specifies theXMLV2 engine. Here are the engine options:
   - XMLMETA=SCHEMADATA specifies to include both data content and metadata-related information in the exported markup.
   - XMLSCHEMA= specifies the fileref that is assigned, in the previousFILENAME statement, to the separate external file that will contain the metadata-related information.

3. TheDATA step reads the SAS data set Sashelp.Snacks and writes its data content inXML markup to the XML document Snacks.XML. TheDATA step also writes the metadata information to the separate external file Snacks.XSD.

Part of the resulting XML document is shown inOutput 3.4 on page 18. The separate metadata information is shown inOutput 3.5 on page 18.
Output 3.4  XML Document Snacks.XML

```xml
<?xml version="1.0" encoding="utf-8" ?>
<TABLE xmlns:xs="http://www.w3.org/2001/XMLSchema"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:noNamespaceSchemaLocation="snacks.xsd">
    <SNACKS>
        <QtySold>0</QtySold>
        <Price>1.99</Price>
        <Advertised>0</Advertised>
        <Holiday>0</Holiday>
        <Date>2002-01-01</Date>
        <Product>Baked potato chips</Product>
    </SNACKS>
    <SNACKS>
        <QtySold>0</QtySold>
        <Price>1.99</Price>
        <Advertised>0</Advertised>
        <Holiday>0</Holiday>
        <Date>2002-01-02</Date>
        <Product>Baked potato chips</Product>
    </SNACKS>
</TABLE>
```

Output 3.5  Separate Metadata Information Snacks.XSD

```xml
<?xml version="1.0" encoding="utf-8" ?>
<x:schema xmlns:x="http://www.w3.org/2001/XMLSchema">
    <x:element name="TABLE">
        <x:complexType>
            <x:sequence>
                <x:element ref="SNACKS" minOccurs="0" maxOccurs="unbounded" />
            </x:sequence>
        </x:complexType>
    </x:element>
    <x:element name="SNACKS">
        <x:complexType>
            <x:sequence>
                <x:element name="QtySold" minOccurs="0" type="xs:double" />
                <x:element name="Price" minOccurs="0" type="xs:double" />
                <x:element name="Advertised" minOccurs="0" type="xs:double" />
                <x:element name="Holiday" minOccurs="0" type="xs:double" />
                <x:element name="Date" minOccurs="0" type="xs:date" />
                <x:element name="Product" minOccurs="0">
                    <x:simpleType>
                        <x:restriction base="xs:string">
                            <x:maxLength value="40" />
                        </x:restriction>
                    </x:simpleType>
                </x:element>
            </x:sequence>
        </x:complexType>
    </x:element>
</x:schema>
```
Chapter 4
Importing XML Documents
Using an XMLMap

Why Use an XMLMap When Importing? .................................................. 19
Understanding the Required Physical Structure for an XML Document to Be Imported Using the GENERIC Markup Type .................. 20
  What Is the Required Physical Structure? ........................................... 20
  Why Is a Specific Physical Structure Required? .............................. 21
  Handling XML Documents That Are Not in the Required Physical Structure . . 23
Using an XMLMap to Import an XML Document as One SAS Data Set ........ 23
Using an XMLMap to Import an XML Document as Multiple SAS Data Sets . . . 26
Importing Hierarchical Data as Related Data Sets ............................... 29
Including a Key Field with Generated Numeric Keys ............................. 33
Determining the Observation Boundary to Avoid Concatenated Data ........... 36
Using ISO 8601 SAS Informats and Formats to Import Dates .................. 39
Using ISO 8601 SAS Informats and Formats to Import Time Values with a Time Zone ............................................................. 40
Referencing a Fileref Using the URL Access Method ............................ 43
Specifying a Location Path on the PATH Element ............................... 44
Including Namespace Elements in an XMLMap .................................. 47
Using the AUTOMAP= Option to Generate an XMLMap ........................ 49
Avoiding Truncation Errors When Importing in a Multi-Byte Encoding ...... 53

Why Use an XMLMap When Importing?

The XMLV2 engine imports only XML documents that conform to the markup types supported in the XMLTYPE= option. Attempting to import free-form XML documents that do not conform to the specifications required by the supported markup types will generate errors. To successfully import files that do not conform to the XMLTYPE= markup types, you can create a separate XML document, called an XMLMap.

If your XML document does not import successfully, you can use an XMLMap to interpret the XML markup. An XMLMap contains specific XMLMap syntax (which is in itself XML markup). The XMLMap syntax tells the XMLV2 engine how to interpret
the XML markup into a SAS data set or data sets, variables (columns), and observations (rows).

As an alternative to you creating an XMLMap by coding XMLMap syntax, the SAS XML Mapper can generate XMLMap syntax. SAS XML Mapper removes the tedium of creating and modifying an XMLMap by providing a GUI that generates the appropriate XML elements for you. SAS XML Mapper analyzes the structure of an XML document or an XML schema, and generates basic syntax for the XMLMap. See Chapter 11, “Using SAS XML Mapper to Generate and Update an XMLMap,” on page 113.

After the XMLMap is created, use the XMLMAP= option in the LIBNAME statement to specify the file.

The AUTOMAP= option in the LIBNAME statement provides another alternative to creating an XMLMap. AUTOMAP= specifies to automatically generate an XMLMap file to import an XML document. See the AUTOMAP= option on page 76.

---

Understanding the Required Physical Structure for an XML Document to Be Imported Using the GENERIC Markup Type

What Is the Required Physical Structure?

In order for an XML document to be successfully imported, the requirements for well-formed XML must translate as follows:

- The root-enclosing element (top-level node) of an XML document is the document container. For SAS, it is like the SAS library
- The nested elements (repeating element instances) that occur within the container begin with the second-level instance tag.
- The repeating element instances must represent a rectangular organization. For a SAS data set, they determine the observation boundary that becomes a collection of rows with a constant set of columns.

Here is an example of an XML document that illustrates the physical structure that is required:

```
<?xml version="1.0" encoding="windows-1252" ?>
<LIBRARY> <!-- 1 -->
  <STUDENTS> <!-- 2 -->
    <ID> 0755 </ID>
    <NAME> Brad Martin </NAME>
    <ADDRESS> 1611 Glengreen </ADDRESS>
    <CITY> Huntingville </CITY>
    <STATE> Texas </STATE>
  </STUDENTS>
  <STUDENTS> <!-- 3 -->
    <ID> 1522 </ID>
    <NAME> Zac Harvell </NAME>
    <ADDRESS> 11900 Glenda </ADDRESS>
    <CITY> Houston </CITY>
    <STATE> Texas </STATE>
  </STUDENTS>
</LIBRARY>
```
When the previous XML document is imported, the following happens:

1. The XMLV2 engine recognizes <LIBRARY> as the root-enclosing element.
2. The engine goes to the second-level instance tag, which is <STUDENTS>, and translates it as the data set name. The engine begins scanning the elements that are nested (contained) between the <STUDENTS> start tag and the </STUDENTS> end tag, looking for variables.
3. Because the instance tags <ID>, <NAME>, <ADDRESS>, <CITY>, and <STATE> are contained within the <STUDENTS> start tag and </STUDENTS> end tag, the XMLV2 engine interprets them as variables. The individual instance tag names become the data set variable names. The repeating element instances are translated into a collection of rows with a constant set of columns.

These statements result in the following SAS output:

```
libname test xmlv2 'C:\Example\students.xml';
proc print data=test.students;
run;
```

**Output 4.1 PRINT Procedure Output for Test.Students**

<table>
<thead>
<tr>
<th>Obs</th>
<th>ID</th>
<th>NAME</th>
<th>ADDRESS</th>
<th>CITY</th>
<th>STATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>755</td>
<td>Brad Martin</td>
<td>1611 Glenn</td>
<td>Huntsville</td>
<td>Texas</td>
</tr>
<tr>
<td>2</td>
<td>1522</td>
<td>Zac Harvell</td>
<td>11900 Glenda</td>
<td>Houston</td>
<td>Texas</td>
</tr>
</tbody>
</table>

**Why Is a Specific Physical Structure Required?**

Well-formed XML is determined by structure, not content. Although the XMLV2 engine can assume that the XML document is valid, well-formed XML, the engine cannot assume that the root element encloses only instances of a single node element (that is, only a single data set). Therefore, the XMLV2 engine has to account for the possibility of multiple nodes (that is, multiple SAS data sets).

For example, when the following correctly structured XML document is imported, it is recognized as containing two SAS data sets: HighTemp and LowTemp.

```xml
<?xml version="1.0" encoding="windows-1252" ?>
<CLIMATE> <!-- 1 -->
  <HIGHTEMP> <!-- 2 -->
    <PLACE> Libya </PLACE>
    <DATE> 1922-09-13 </DATE>
    <DEGREE-F> 136 </DEGREE-F>
    <DEGREE-C> 58 </DEGREE-C>
  </HIGHTEMP>
  .  more instances of <HIGHTEMP>
</CLIMATE>
```
When the previous XML document is imported, the following happens:

1. The XMLV2 engine recognizes the first instance tag <CLIMATE> as the root-enclosing element, which is the container for the document.

2. Starting with the second-level instance tag, which is <HIGHTEMP>, the XMLV2 engine uses the repeating element instances as a collection of rows with a constant set of columns.

3. When the second-level instance tag changes, the XMLV2 engine interprets that change as a different SAS data set.

The result is two SAS data sets: HighTemp and LowTemp. Both happen to have the same variables but different data.

To ensure that an import result is what you expect, use the DATASETS procedure. For example, these SAS statements result in the following:

```sas
libname climate xmlv2 'C:\Example\climate.xml';
proc datasets library=climate;
quit;
```

**Output 4.2** DATASETS Procedure Output for Climate Library

| The SAS System |
| --- | --- | --- |
| **Directory** |  |  |
| Libref | CLIMATE |  |
| Engine | XMLV2 |  |
| Physical Name | C:\Example\climate.xml |  |
| XMLType | SAS XML Generic |  |
| **Version** |  |  |

<table>
<thead>
<tr>
<th>#</th>
<th>Name</th>
<th>Member Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HIGHTEMP</td>
<td>DATA</td>
</tr>
<tr>
<td>2</td>
<td>LOWTEMP</td>
<td>DATA</td>
</tr>
</tbody>
</table>
Handling XML Documents That Are Not in the Required Physical Structure

If your XML document is not in the required physical structure, you can tell the XMLV2 engine how to interpret the XML markup to successfully import the document. See “Why Use an XMLMap When Importing?” on page 19.

Using an XMLMap to Import an XML Document as One SAS Data Set

This example explains how to create and use an XMLMap in order to tell the XMLV2 engine how to map XML markup to a SAS data set, variables, and observations.

Here is the XML document Nhl.xml to be imported. Although simply constructed and relatively easy for you to read, it does not import successfully because its XML markup is not in the required physical structure:

```xml
<?xml version="1.0" encoding="iso-8859-1" ?>
<NHL>
 <CONFERENCE> Eastern
  <DIVISION> Southeast
   <TEAM name="Thrashers" abbrev="ATL" />
   <TEAM name="Hurricanes" abbrev="CAR" />
   <TEAM name="Panthers" abbrev="FLA" />
   <TEAM name="Lightning" abbrev="TB" />
   <TEAM name="Capitals" abbrev="WSH" />
  </DIVISION>
 </CONFERENCE>

 <CONFERENCE> Western
  <DIVISION> Pacific
   <TEAM name="Stars" abbrev="DAL" />
   <TEAM name="Kings" abbrev="LA" />
   <TEAM name="Ducks" abbrev="ANA" />
   <TEAM name="Coyotes" abbrev="PHX" />
   <TEAM name="Sharks" abbrev="SJ" />
  </DIVISION>
 </CONFERENCE>
</NHL>
```

To successfully import the XML document, an XMLMap is needed. After familiarizing yourself with the data to be imported, you can code the XMLMap syntax so that the data is successfully imported. Here is the XMLMap used to import the XML document, with notations for the data investigation:

```xml
<?xml version="1.0" ?>
<SXLEMAP version="2.1">
 <TABLE name="TEAMS">
  <TABLE-PATH syntax="XPath">
   /NHL/CONFERENCE/DIVISION/TEAM
  </TABLE-PATH>
 </TABLE>
</SXLEMAP>
```
The previous XMLMap syntax defines how to translate the XML markup as explained below using the following data investigation steps:

1 Locate and identify distinct tables of information.

You want a SAS data set (table) that contains some of the teams of the National Hockey League. Because that is the only information contained in the XML document, you can define a single data set named Teams in the XMLMap. (Note that other XML documents might contain more than one table of related information. Importing multiple tables is supported by the XMLMap syntax as shown in “Using an XMLMap to Import an XML Document as Multiple SAS Data Sets” on page 26.)

2 Identify the SAS data set observation boundary, which translates into a collection of rows with a constant set of columns.

In the XML document, information about individual teams occurs in a <TEAM> tag located with <CONFERENCE> and <DIVISION> enclosures. You want a new observation generated each time a TEAM element is read.

3 Collect column definitions for each table.
For this XML document, the data content form is mixed. Some data occurs as XML PCDATA (for example, CONFERENCE), and other data is contained in attribute-value pairs (for example, NAME). Data types are all string values. The constructed observation will also include the team NAME and ABBREV. A length of 30 characters is sufficient for the NAME, and three characters is enough for the ABBREV field contents.

4 Add foreign keys or required external context.

You want to include information about the league orientation for the teams. Also, you want to extract CONFERENCE and DIVISION data.

*Note:* The retain= attribute in the column definition forces retention of processed data values after an observation is written to the output data set. Because the foreign key fields occur outside the observation boundary (that is, they are more sparsely populated in the hierarchical XML data than in the SAS observation), their values for additional rows need to be retained as they are encountered.

5 Define a location path for each variable definition.

The PATH element identifies a position in the XML document from which to extract data for each column. Element-parsed character data is treated differently than attribute values. There is no conditional selection criteria involved.

The following SAS statements import the XML document Nhl.xml:

```sas
filename nhl 'C:\Example\Nhl.xml'; /*1*/
filename map 'C:\Example\Nhl.map'; /*2*/
libname nhl xmlv2 xmlmap=map; /*3*/
proc print data=nhl.teams; /*4*/
run;
```

1 The first FILENAME statement assigns the file reference Nhl to the physical location (complete pathname, filename, and file extension) of the XML document named Nhl.xml.

2 The second FILENAME statement assigns the file reference Map to the physical location of the XMLMap named Nhl.map.

3 The LIBNAME statement uses the file reference Nhl to reference the XML document. It specifies the XMLV2 engine and uses the file reference Map to reference the XMLMap.

4 The PRINT procedure produces output, verifying that the import was successful.
The following code is not necessary for this example, but it creates the data set on disk to use in “Using an XMLMap to Export an XML Document with a Hierarchical Structure” on page 55.

```
libname myfiles 'C:\MyFiles\';
data myfiles.teams;
  set nhl.teams;
run;
```

### Using an XMLMap to Import an XML Document as Multiple SAS Data Sets

This example explains how to create and use an XMLMap to import an XML document as two SAS data sets. The example uses the XML document `RSS.xml` below. `RSS.xml` does not import successfully without an XMLMap because its XML markup is a nonstandard structure. If you attempt to import the XML document without an XMLMap, an error is written to the log that the XML data is not in a format supported natively by the XML LIBNAME engine. The XML content in this example uses a simplified set of RSS (rich site summary) tags.

```xml
<?xml version="1.0" encoding="UTF-8" ?>
<rss version="2.0">
  <channel>
    <title>My RSS Channel</title>
    <description>This is a simplified example of an RSS feed.</description>
    <link>http://www.example.com/main.html</link>
    <language>en-us</language>
    <item>
      <title>My news item</title>
    </item>
  </channel>
</rss>
```
The XML document can be successfully imported by creating an XMLMap that defines the XML elements as the columns and other attributes of one or more data sets. The following XMLMap, which is named RSS.map, defines two SAS data sets. The Channel data set contains content information. The Item data set contains individual news stories.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<SXLEMAP name="SXLEMap" version="2.1"> <!-- 1 -->
  <TABLE name="CHANNEL"> <!-- 2 -->
    <TABLE-PATH syntax="XPath">/rss/channel</TABLE-PATH> <!-- 3 -->
    <COLUMN name="Channel_Title"> <!-- 4 -->
      <PATH syntax="XPath">/rss/channel/title</PATH>
      <TYPE>character</TYPE>
      <DATATYPE>string</DATATYPE>
      <LENGTH>200</LENGTH>
    </COLUMN>
    <COLUMN name="Channel_URL">
      <PATH syntax="XPath">/rss/channel/link</PATH>
      <DESCRIPTION>Channel link</DESCRIPTION>
      <TYPE>character</TYPE>
      <DATATYPE>string</DATATYPE>
      <LENGTH>200</LENGTH>
    </COLUMN>
    <COLUMN name="Channel_Description">
      <PATH syntax="XPath">/rss/channel/description</PATH>
      <TYPE>character</TYPE>
      <DATATYPE>string</DATATYPE>
      <LENGTH>1024</LENGTH>
    </COLUMN>
  </TABLE>
  <TABLE description="Individual news stories" name="ITEM"> <!-- 5 -->
    <TABLE-PATH syntax="XPath">/rss/channel/item</TABLE-PATH> <!-- 6 -->
    <COLUMN name="Title">
      <PATH syntax="XPath">/rss/channel/item/title</PATH>
      <TYPE>character</TYPE>
      <DATATYPE>string</DATATYPE>
      <LENGTH>200</LENGTH>
    </COLUMN>
  </TABLE>
</SXLEMAP>
```
SXLEMAP is the root-enclosing element for SAS data set definitions.

This TABLE element defines the Channel data set.

This TABLE-PATH element specifies the location path in the XML document to collect variables for the Channel data set.

This COLUMN element defines the Channel_Title variable in the Channel data set. The XPath construction specifies where to find the current tag and where to access data from the named element. The next two COLUMN elements define the Channel_URL and Channel_Description variables.

This TABLE element defines the Item data set.

This TABLE-PATH element specifies the location path in the XML document to collect variables for the Item data set.

This COLUMN element defines the Title variable in the Item data set. The next two COLUMN elements define the URL and Description variables.

The following SAS statements import the XML document RSS.xml and specify the XMLMap named RSS.map. The DATASETS procedure verifies the import results.

```
filename myrss 'C:\Example\rss.xml';
filename mymap 'C:\Example\rss.map';
libname myrss xmlv2 xmlmap=mymap access=readonly;
proc datasets library=myrss;
run;
quit;
```
Importing Hierarchical Data as Related Data Sets

XML documents often contain hierarchical data in that the data is structured into different levels like a company organization chart. Hierarchical structures are one-to-many relationships. Top items having one or more items below it (for example, customer to orders).

This example explains how to define an XMLMap in order to import an XML document as two data sets that have related information.

Here is the XML document Pharmacy.xml. The file contains hierarchical data with related entities in the form of individual customers and their prescriptions. Each customer can have one or multiple prescriptions. Notice that PRESCRIPTION elements are nested within each <PERSON> start tag and </PERSON> end tag:

```xml
<?xml version="1.0" ?>
<PHARMACY>
  <PERSON>
    <NAME>Brad Martin</NAME>
    <STREET>11900 Glenda Court</STREET>
    <CITY>Austin</CITY>
    <PRESCRIPTION>
      <NUMBER>1234</NUMBER>
      <DRUG>Tetracycline</DRUG>
    </PRESCRIPTION>
    <PRESCRIPTION>
      <NUMBER>1245</NUMBER>
      <DRUG>Lomotil</DRUG>
    </PRESCRIPTION>
  </PERSON>
  <PERSON>
    ... (additional PERSON elements)
  </PERSON>
</PHARMACY>
```
To import separate data sets, one describing the customers and the other containing prescription information, a relation between each customer and associated prescriptions must be designated. The relationship determines which prescriptions belong to each customer.

An XMLMap defines how to translate the XML markup into two SAS data sets. The Person data set imports the name and address of each customer, and the Prescription data set imports the customer's name, prescription number, and drug. Notations in the XMLMap syntax are explained below.

Note: The XMLMap was generated by using SAS XML Mapper.

<?xml version="1.0" encoding="windows-1252"?>
<!-- ############################################################ -->
<!-- 2017-10-21T11:57:46 -->
<!-- SAS XML Libname Engine Map -->
<!-- Generated by XML Mapper, 904500.0.0.20170816190000_v940m5 -->
<!-- ############################################################ -->
<!-- ### Validation report ### -->
<!-- Column (NAME) in table (PRESCRIPTION) has an XPath outside the scope of the table path. The contents of this column may not correspond to other row values and/or may be missing entirely. -->
<!-- XMLMap validation completed successfully. -->
<!-- ############################################################ -->
<SXLEMAP name="Pharmacy" version="2.1"> <!-- 2 -->
<NAMESPACES count="0"/>

<TABLE description="PERSON" name="PERSON"> <!-- 3 -->
<TABLE-PATH syntax="XPath">/PHARMACY/PERSON</TABLE-PATH>
<COLUMN name="NAME"> <!-- 4 -->
(PATH syntax="XPath">/PHARMACY/PERSON/NAME</PATH>
<TYPE>character</TYPE>
<DATATYPE>string</DATATYPE>
<LENGTH>11</LENGTH>
</COLUMN>

<COLUMN name="STREET"> <!-- 4 -->
(PATH syntax="XPath">/PHARMACY/PERSON/STREET</PATH>
<TYPE>character</TYPE>
<DATATYPE>string</DATATYPE>
<LENGTH>18</LENGTH>
</COLUMN>
</TABLE>
</SXLEMAP>
If the SAS XML Mapper application issues a warning, a good practice is to print the data set and check for errors. In this case, no error occurs.

SXLEMAP is the root-enclosing element for the two SAS data set definitions.

First TABLE element defines the Person data set.

COLUMN elements contain the attributes for the Name, Street, and City variables in the Person data set.

Second TABLE element defines the Prescription data set.

COLUMN element contains the attributes for the Name variable in the Prescription data set. Specifying the retain="YES" attribute causes the name to be held for each observation until it is replaced by a different value. (The retain= attribute is like the SAS DATA step RETAIN statement, which causes a variable to retain its value from one iteration of the DATA step to the next.)

COLUMN elements contain the attributes for the Number and Drug variables in the Prescription data set.

The following SAS statements import the XML document and specify the XMLMap:
filename pharm 'C:\Example\Pharmacy.xml';
filename map 'C:\Example\Pharmacy.map';
libname pharm xmlv2 xmlmap=map;

The DATASETS procedure verifies that SAS interprets the XML document Pharmacy.xml as two SAS data sets: Pharm.Person and Pharm.Prescription.

proc datasets library=pharm;
quit;

**Output 4.5** DATASETS Procedure Output for Pharm Library

<table>
<thead>
<tr>
<th>The SAS System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directory</td>
</tr>
<tr>
<td>Libref</td>
</tr>
<tr>
<td>Engine</td>
</tr>
<tr>
<td>Physical Name</td>
</tr>
<tr>
<td>XML Type</td>
</tr>
<tr>
<td>Version</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>#</th>
<th>Name</th>
<th>Member Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PERSON</td>
<td>DATA</td>
</tr>
<tr>
<td>2</td>
<td>PRESCRIPTION</td>
<td>DATA</td>
</tr>
</tbody>
</table>

Here is the PRINT procedure output for both of the imported SAS data sets.

**Output 4.6** PRINT Procedure Output for Pharm.Person

<table>
<thead>
<tr>
<th>The SAS System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obs</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
</tbody>
</table>
Including a Key Field with Generated Numeric Keys

This example imports the XML document Pharmacy.xml, which contains hierarchical data and is used in the example “Importing Hierarchical Data as Related Data Sets” on page 29. This example continues with the XMLMap by adding a key field with generated numeric key values to provide a relationship between the two data sets. (A key field holds unique data to identify that record from the other records. For example, account number, product code, and customer name are typical key fields.)

To generate key field values, use the class="ORDINAL" attribute in the COLUMN element to create a counter variable. A counter variable keeps track of the number of times the location path, which is specified by the INCREMENT-PATH element, is encountered. The counter variable increments its count by 1 each time the location path is matched. (The counter variable is similar to the _N_ automatic variable in DATA step processing in that it counts the number of observations being read into a SAS data set.)

Note: When using a counter variable to create a key field for related data sets, you must specify the same location paths for both TABLE elements. Otherwise, the results will not match. Each table must have the same generated key for like-named data elements.

The following XMLMap imports Pharmacy.xml document as two SAS data sets that have related information and also creates a key field that holds generated numeric key values:

```xml
<?xml version="1.0" encoding="windows-1252"?>
<!-- #-------------------------------------------------------------------------- -->
<!-- 2017-10-23T11:00:08 -->
<!-- SAS XML Libname Engine Map -->
<!-- Generated by XML Mapper, 904500.0.0.20170816190000_940m5 -->
<!-- # Validation report -->
<!-- XLMap validation completed successfully. -->
<!-- #-------------------------------------------------------------------------- -->
</SXLMAP name="PharmacyOrdinal" version="2.1">
```
<TABLE description="PERSON" name="PERSON">
  <TABLE-PATH syntax="XPath">/PHARMACY/PERSON</TABLE-PATH> <!-- 1 -->
  <COLUMN class="ORDINAL" name="KEY" retain="YES"> <!-- 2 -->
    <INCREMENT-PATH beginend="BEGIN" syntax="XPath">/PHARMACY/PERSON</INCREMENT-PATH>
    <TYPE>numeric</TYPE>
    <DATATYPE>integer</DATATYPE>
    <FORMAT width="3">Z</FORMAT>
  </COLUMN>
  <COLUMN name="NAME">
    <PATH syntax="XPath">/PHARMACY/PERSON/NAME</PATH>
    <TYPE>character</TYPE>
    <DATATYPE>string</DATATYPE>
    <LENGTH>11</LENGTH>
  </COLUMN>
  <COLUMN name="STREET">
    <PATH syntax="XPath">/PHARMACY/PERSON/STREET</PATH>
    <TYPE>character</TYPE>
    <DATATYPE>string</DATATYPE>
    <LENGTH>18</LENGTH>
  </COLUMN>
  <COLUMN name="CITY">
    <PATH syntax="XPath">/PHARMACY/PERSON/CITY</PATH>
    <TYPE>character</TYPE>
    <DATATYPE>string</DATATYPE>
    <LENGTH>6</LENGTH>
  </COLUMN>
</TABLE>

<!-- 3 -->
<TABLE description="PRESCRIPTION" name="PRESCRIPTION">
  <TABLE-PATH syntax="XPath">/PHARMACY/PERSON/PRESCRIPTION</TABLE-PATH> <!-- 4 -->
  <COLUMN class="ORDINAL" name="KEY" retain="YES"> <!-- 5 -->
    <INCREMENT-PATH beginend="BEGIN" syntax="XPath">/PHARMACY/PERSON</INCREMENT-PATH>
    <TYPE>numeric</TYPE>
    <DATATYPE>integer</DATATYPE>
    <FORMAT width="3">Z</FORMAT>
  </COLUMN>
  <COLUMN name="NUMBER">
    <PATH syntax="XPath">/PHARMACY/PERSON/PRESCRIPTION/NUMBER</PATH>
    <TYPE>numeric</TYPE>
    <DATATYPE>integer</DATATYPE>
  </COLUMN>
  <COLUMN name="DRUG">
    <PATH syntax="XPath">/PHARMACY/PERSON/PRESCRIPTION/DRUG</PATH>
    <TYPE>character</TYPE>
  </COLUMN>
</TABLE>
In the TABLE element that defines the Person data set, the TABLE-PATH element identifies the observation boundary for the data set. The location path generates a new observation each time a PERSON element is read.

For the Person data set, the COLUMN element for the Key variable contains the class="ORDINAL" attribute as well as the INCREMENT-PATH element. The XMLV2 engine follows this process to generate the key field values for the Person data set:

1. When the XMLV2 engine encounters the <PERSON> start tag, it reads the value into the input buffer, and then increments the value for the Key variable by 1.
2. The XMLV2 engine continues reading values into the input buffer until it encounters the </PERSON> end tag, at which time it writes the completed input buffer to the SAS data set as one observation.
3. The process is repeated for each <PERSON> start tag (from INCREMENT-PATH) and </PERSON> end tag (from TABLE-PATH) sequence.
4. The result is four variables and two observations.

In the TABLE element that defines the Prescription data set, the TABLE-PATH element identifies the observation boundary for the data set. The location path generates a new observation each time a PRESCRIPTION element is read.

For the Prescription data set, the COLUMN element for the Key variable contains the class="ORDINAL" attribute as well as the INCREMENT-PATH element. The XMLV2 engine follows this process to generate the key field values for the Prescription data set:

1. When the XMLV2 engine encounters the <PERSON> start tag, it reads the value into the input buffer, and then increments the value for the Key variable by 1.
2. The XMLV2 engine continues reading values into the input buffer until it encounters the </PRESCRIPTION> end tag, at which time it writes the completed input buffer to the SAS data set as one observation. Because the location paths for the counter variables must be the same for both TABLE elements, the behavior of the XMLV2 engine for the Prescription data set Key variable is the same as the Person data set Key variable. Although the XMLV2 engine tracks the occurrence of a PERSON tag as a key for both counter variables, the observations are derived from different TABLE-PATH locations.
3. The process is repeated for each <PERSON> start tag (from INCREMENT-PATH) and </PRESCRIPTION> end tag (from TABLE-PATH) sequence.
4. The result is three variables and three observations.

The following SAS statements import the XML document:

```sas
filename pharm 'C:\Example\Pharmacy.xml';
filename map 'C:\Example\PharmacyOrdinal.map';
libname pharm xmlv2 xmlmap=map;
```
Here is the PRINT procedure output for both of the imported SAS data sets with a numeric key:

**Output 4.8**  PRINT Procedure Output for Pharm.Person

```
The SAS System

<table>
<thead>
<tr>
<th>Obs</th>
<th>KEY</th>
<th>NAME</th>
<th>STREET</th>
<th>CITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>001</td>
<td>Brad Martin</td>
<td>11900 Glenda Court</td>
<td>Austin</td>
</tr>
<tr>
<td>2</td>
<td>002</td>
<td>Jim Spano</td>
<td>1611 Glendreen</td>
<td>Austin</td>
</tr>
</tbody>
</table>
```

**Output 4.9**  PRINT Procedure Output for Pharm.Prescription

```
The SAS System

<table>
<thead>
<tr>
<th>Obs</th>
<th>KEY</th>
<th>NUMBER</th>
<th>DRUG</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>001</td>
<td>1234</td>
<td>Tetracycline</td>
</tr>
<tr>
<td>2</td>
<td>001</td>
<td>1245</td>
<td>Lomotil</td>
</tr>
<tr>
<td>3</td>
<td>002</td>
<td>1268</td>
<td>Nexium</td>
</tr>
</tbody>
</table>
```

---

**Determining the Observation Boundary to Avoid Concatenated Data**

This example illustrates how to determine the observation boundary so that an import results in separate observations and not concatenated data.

The observation boundary translates into a collection of rows with a constant set of columns. Using an XMLMap, you determine the observation boundary with the TABLE-PATH element by specifying a location path. The end tag for the location path determines when data is written to the SAS data set as an observation.

Identifying the observation boundary can be tricky due to sequences of start-tag and end-tag pairing. If you do not identify the appropriate observation boundary, the result could be a concatenated data string instead of separate observations. This example illustrates pairing situations that can cause unwanted results.

For the following XML document, an XMLMap is necessary to import the file successfully. Without an XMLMap, the XMLV2 engine would import a data set named Ford with variables Row0, Model0, Year0, Row1, Model1, Year1, and so on.

```
<?xml version="1.0" ?>
<VEHICLES>
  <FORD>
    <ROW>
```

---
<Model>Mustang</Model>  
<Year>1965</Year>  
</ROW>  
<ROW>  
<Model>Explorer</Model>  
<Year>1982</Year>  
</ROW>  
<ROW>  
<Model>Taurus</Model>  
<Year>1998</Year>  
</ROW>  
<ROW>  
<Model>F150</Model>  
<Year>2000</Year>  
</ROW>  
</FORD>  
</VEHICLES>

Looking at the above XML document, there are three sequences of element start tags and end tags: VEHICLES, FORD, and ROW. If you specify the following table location path and column locations paths, the XMLV2 engine processes the XML document as follows:

<TABLE-PATH syntax="XPath"> /VEHICLES/FORD </TABLE-PATH>  
<PATH syntax="XPath"> /VEHICLES/FORD/ROW/Model </PATH>  
<PATH syntax="XPath"> /VEHICLES/FORD/ROW/Year </PATH>

1. The XMLV2 engine reads the XML markup until it encounters the <FORD> start tag, because FORD is the last element specified in the table location path. 
2. The XMLV2 engine clears the input buffer and scans subsequent elements for variables based on the column location paths. As a value for each variable is encountered, it is read into the input buffer. For example, after reading the first ROW element, the input buffer contains the values Mustang and 1965.
3. The XMLV2 engine continues reading values into the input buffer until it encounters the </FORD> end tag, at which time it writes the completed input buffer to the SAS data set as an observation.
4. The end result is one observation, which is not what you want.

To get separate observations, you must change the table location path so that the XMLV2 engine writes separate observations to the SAS data set. Here are the correct location paths and the process that the engine would follow:

<TABLE-PATH syntax="XPath"> /VEHICLES/FORD/ROW </TABLE-PATH>  
<PATH syntax="XPath"> /VEHICLES/FORD/ROW/Model </PATH>  
<PATH syntax="XPath"> /VEHICLES/FORD/ROW/Year </PATH>

1. The XMLV2 engine reads the XML markup until it encounters the <ROW> start tag, because ROW is the last element specified in the table location path. 
2. The XMLV2 engine clears the input buffer and scans subsequent elements for variables based on the column location paths. As a value for each variable is encountered, it is read into the input buffer.
3. The XMLV2 engine continues reading values into the input buffer until it encounters the </ROW> end tag, at which time it writes the completed input buffer to the SAS data set as an observation. That is, one observation is written to the SAS data set that contains the values Mustang and 1965.
4. The process is repeated for each `<ROW>` start-tag and `</ROW>` end-tag sequence.

5. The result is four observations.

Here is the complete XMLMap syntax:

```xml
<?xml version="1.0" ?>
<SXLEMAP version="2.1" name="path" description="XMLMap for path">  
<TABLE name="FORD">  
<TABLE-PATH syntax="XPath"> /VEHICLES/FORD/ROW </TABLE-PATH>  
<COLUMN name="Model">  
<DATATYPE> string </DATATYPE>  
<LENGTH> 20 </LENGTH>  
<TYPE> character </TYPE>  
<PATH syntax="XPath"> /VEHICLES/FORD/ROW/Model </PATH>  
</COLUMN>  
<COLUMN name="Year">  
<DATATYPE> string </DATATYPE>  
<LENGTH> 4 </LENGTH>  
<TYPE> character </TYPE>  
<PATH syntax="XPath"> /VEHICLES/FORD/ROW/Year </PATH>  
</COLUMN>  
</TABLE>  
</SXLEMAP>
```

The following SAS statements import the XML document and specify the XMLMap. The PRINT procedure verifies the results.

```sas
filename path 'C:\Example\path.xml';
filename map 'C:\Example\path.map';
libname path xmlv2 xmlmap=map;
proc print data=path.ford noobs;
run;
```

**Output 4.10**  PRINT Procedure Output Showing Path.Ford Data Set

<table>
<thead>
<tr>
<th>Model</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mustang</td>
<td>1965</td>
</tr>
<tr>
<td>Explorer</td>
<td>1982</td>
</tr>
<tr>
<td>Taurus</td>
<td>1998</td>
</tr>
<tr>
<td>F-150</td>
<td>2000</td>
</tr>
</tbody>
</table>
Using ISO 8601 SAS Informats and Formats to Import Dates

This simple example illustrates importing an XML document that contains date values in both the basic format and the extended format. The XMLMap uses the FORMAT and INFORMAT elements to specify the appropriate SAS format and SAS informat in order to represent the dates according to ISO 8601 standards.

Here is the XML document:

```xml
<?xml version="1.0" ?>
<Root>
  <ISODATE>
    <BASIC>19450508</BASIC>
    <EXTENDED>1945-05-08</EXTENDED>
  </ISODATE>
</Root>
```

The following XMLMap imports the XML document using the SAS informats and formats to read and write the date values:

```xml
<?xml version="1.0" encoding="windows-1252"?>
<!-- ############################################################ -->
<!-- 2017-10-23T13:33:48 -->
<!-- SAS XML Libname Engine Map -->
<!-- Generated by XML Mapper, 904500.0.0.20170816190000_v940m5 -->
<!-- ############################################################ -->
<!-- ### Validation report                                   ### -->
<!-- ############################################################ -->
<!-- XMLMap validation completed successfully. -->
<!-- ############################################################ -->
<XMLmap description="Reading a Basic and Extended format ISO date field" name="ISOdate" version="2.1">
  <NAMESPACES count="0"/>
  <!-- ############################################################ -->
  <TABLE name="ISODATE">
    <TABLE-PATH syntax="XPath">/Root/ISODATE</TABLE-PATH>
    <COLUMN name="BASIC">
      <PATH syntax="XPath">/Root/ISODATE/BASIC</PATH>
      <TYPE>numeric</TYPE>
      <DATATYPE>date</DATATYPE>
      <FORMAT width="10">E8601DA</FORMAT>
      <INFORMAT width="8">B8601DA</INFORMAT>
    </COLUMN>
    <COLUMN name="EXTENDED">
      <PATH syntax="XPath">/Root/ISODATE/EXTENDED</PATH>
      <TYPE>numeric</TYPE>
      <DATATYPE>date</DATATYPE>
    </COLUMN>
  </TABLE>
</XMLmap>
```
The following explains the XMLMap syntax that imports the date values:

1. For the Basic variable, the FORMAT element specifies the E8601DA SAS format, which writes data values in the extended format `yyyy-mm-dd`.

2. For the Basic variable, the INFORMAT element specifies the B8601DA SAS informat, which reads date values into a variable in the basic format `yyyymmdd`.
   
   Note: As recommended, when you read values into a variable with a basic format SAS informat, this example writes the values with the corresponding extended format SAS format.

3. For the Extended variable, the FORMAT element specifies the E8601DA SAS format, which writes data values in the extended format `yyyy-mm-dd`.

4. For the Extended variable, the INFORMAT element specifies the E8601DA SAS informat, which reads date values into a variable in the basic format `yyyymmdd`.

The following SAS statements import the XML document and display PRINT procedure output:

```sas
filename dates 'C:\Example\isodate.xml';
filename map 'C:\Example\isodate.map';
libname dates xmlv2 xmlmap=map;
proc print data=dates.isodate;
run;
```

Output 4.11 PRINT Procedure Output for Imported Data Set Dates.IsoDate

![The SAS System](image)

<table>
<thead>
<tr>
<th>Obs</th>
<th>BASIC</th>
<th>EXTENDED</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1945-05-08</td>
<td>1945-05-08</td>
</tr>
</tbody>
</table>

Using ISO 8601 SAS Informats and Formats to Import Time Values with a Time Zone

This example illustrates importing an XML document that contains time values in various forms. The XMLMap uses the FORMAT and INFORMAT elements to specify the appropriate SAS formats and SAS informats in order to represent the times appropriately.

Here is an XML document that contains a variety of time values:
Using ISO 8601 SAS Informs and Formats to Import Time Values with a Time Zone

The following XMLMap imports the XML document using the SAS informs and formats to read and write the time values:

```xml
<?xml version="1.0" encoding="UTF-8"?>
<!- - ############################################################################# - - >
<!- - SAS XML Libname Engine Map -->
<!- - Generated by XML Mapper, 903000.1.0.20101208190000_v930 -->
<!- - ############################################################################# - - >
<!- - Validation report # -->
<!- - XMLMap validation completed successfully. -->
<!- - ############################################################################# - - >
<SXLEMAP name="ISOtime" version="2.1">
  <NAMESPACES count="0"/>
  <TABLE name="TIME">
    <TABLE-PATH syntax="XPath">/Root/TIME</TABLE-PATH>
    <COLUMN name="LOCAL">
      <PATH syntax="XPath">/Root/TIME/LOCAL</PATH>
      <TYPE>numeric</TYPE>
      <DATATYPE>time</DATATYPE>
      <FORMAT width="8">E8601TM</FORMAT>
      <INFORMAT width="8">E8601TM</INFORMAT>
    </COLUMN>
    <COLUMN name="LOCALZONE">
      <PATH syntax="XPath">/Root/TIME/LOCAL</PATH>
      <TYPE>numeric</TYPE>
      <DATATYPE>time</DATATYPE>
      <FORMAT width="14">E8601LZ</FORMAT>
      <INFORMAT width="8">E8601TM</INFORMAT>
    </COLUMN>
    <COLUMN name="UTC">
      <PATH syntax="XPath">/Root/TIME/UTC</PATH>
      <TYPE>numeric</TYPE>
      <DATATYPE>time</DATATYPE>
      <FORMAT width="9">E8601TZ</FORMAT>
      <INFORMAT width="9">E8601TZ</INFORMAT>
    </COLUMN>
    <COLUMN name="OFFSET">
```
```
The following explains the XMLMap syntax that imports the time values:

1. For the Local variable, the INFORMAT and FORMAT elements specify the E8601TM SAS informat and format, which reads and writes time values in the extended format \( hh:mm:ss.ffffff \). Because there is no time zone indicator, the context of the value is local time.

2. For the Localzone variable, which reads the same value as the Local variable, the INFORMAT element specifies the E8601TM SAS informat, which reads time values in the extended format \( hh:mm:ss.ffffff \). Because there is no time zone indicator, the context of the value is local time.

The FORMAT element, however, specifies the E8601LZ SAS format, which writes time values in the extended format \( hh:mm:ss+|-hh:mm \). The E8601LZ format appends the UTC offset to the value as determined by the local, current SAS session. Using the E8601LZ format enables you to provide a time notation in order to eliminate the ambiguity of local time.

Note: Even with the time notation, it is recommended that you do not mix time-based values.

3. For the UTC variable, the INFORMAT and FORMAT elements specify the E8601TZ SAS informat and format, which reads and writes time values in the extended format \( hh:mm:ss+|-hh:mm \). Because there is a time zone indicator, the value is assumed to be expressed in UTC. No adjustment or conversion is made to the value.

4. For the Offset variable, the INFORMAT and FORMAT elements specify the E8601TZ SAS informat and format, which reads and writes time values in the extended format \( hh:mm:ss+|-hh:mm \). A time zone offset is present. When the time value is read into the variable using the time zone-sensitive SAS informat, the value is adjusted to UTC as requested via the time zone indicator. The time zone context is not stored with the value. When the time value is written using the time zone sensitive SAS format, the value is expressed as UTC with a zero offset value and is not adjusted to or from local time.

The following SAS statements import the XML document and display the PRINT procedure output:

```sas
filename timzn 'C:\Example\Time.xml';
filename map 'C:\Example\Time.map';
libname timzn xmlv2 xmlmap=map;
proc print data=timzn.time;
run;
```
Referencing a Fileref Using the URL Access Method

Using several methods, the XMLV2 engine can access an XML document that is referenced by a fileref. When using the URL access method to reference a fileref, you should also specify an XMLMap. Specifying an XMLMap causes the XMLV2 engine to process the XML document with a single pass of the file, rather than a double pass, which is what happens when you do not specify an XMLMap.

This example illustrates how to access an XML document by referencing a fileref and using the URL access method:

```sas
filename mynl url 'http://www.a.com/NHL.xml'; /*1*/
filename mymap 'C:\Example\NHL.map'; /*2*/
libname mynl xmlv2 xmlmap=mymap; /*3*/
proc print data=mynl.TEAMS; /*4*/
run;
```

1. The first FILENAME statement assigns the fileref MyNhl to the XML document by using the URL access method.
2. The second FILENAME statement assigns the fileref MyMap to the physical location of the XMLMap NHL.map.
3. The LIBNAME statement uses the fileref MyNhl to reference the XML document, specifies the XMLV2 engine, and uses the fileref MyMap to reference the XMLMap.
4. PROC PRINT produces output, verifying that the import was successful.
Specifying a Location Path on the PATH Element

The XMLMap PATH element supports several XPath forms to specify a location path. The location path tells the XMLV2 engine where in the XML document to locate and access a specific tag for the current variable. In addition, the location path tells the XMLV2 engine to perform a function, which is determined by the XPath form, to retrieve the value for the variable.

This example imports an XML document and illustrates each of the supported XPath forms, which include three element forms and two attribute forms.

Here is the XML document NhlShort.xml to be imported:

```xml
<?xml version="1.0" encoding="iso-8859-1" ?>
<NHL>
  <CONFERENCE> Eastern
    <DIVISION> Southeast
      <TEAM founded="1999" abbrev="ATL"> Thrashers </TEAM>
      <TEAM founded="1997" abbrev="CAR"> Hurricanes </TEAM>
      <TEAM founded="1993" abbrev="FLA"> Panthers </TEAM>
      <TEAM founded="1992" abbrev="TB"> Lightning </TEAM>
      <TEAM founded="1974" abbrev="WSH"> Capitals </TEAM>
    </DIVISION>
  </CONFERENCE>
</NHL>
```

Here is the XMLMap Nhl1.map used to import the XML document, with notations for each XPath form on the PATH element:

```xml
<?xml version="1.0" ?>
<SXLEMAP version="2.1">
  <TABLE name="TEAMS">
    <Obs NAME ABBREV CONFERENCE DIVISION>
      <1 Thrashers ATL Eastern Southeast</1>
      <2 Hurricanes CAR Eastern Southeast</2>
      <3 Panthers FLA Eastern Southeast</3>
      <4 Lightning TB Eastern Southeast</4>
      <5 Capitals WSH Eastern Southeast</5>
      <6 Stars DAL Western Pacific</6>
      <7 Kings LA Western Pacific</7>
      <8 Ducks ANA Western Pacific</8>
      <9 Coyotes PHX Western Pacific</9>
      <10 Sharks SJ Western Pacific</10>
    </Obs>
  </TABLE>
</SXLEMAP>
```
The Abbrev variable uses the attribute form that selects values from a specific attribute. The engine scans the XML markup until it finds the TEAM element. The engine retrieves the value from the abbrev attribute, which results in each team abbreviation.
The Founded variable uses the attribute form that conditionally selects from a specific attribute based on the value of another attribute. The engine scans the XML markup until it finds the TEAM element. The engine retrieves the value from the founded= attribute where the value of the abbrev= attribute is ATL, which results in the value 1999. The two attributes must be for the same element.

The Conference variable uses the element form that selects PCDATA from a named element. The engine scans the XML markup until it finds the CONFERENCE element. The engine retrieves the value between the <CONFERENCE> start tag and the </CONFERENCE> end tag, which results in the value Eastern.

The Team variable uses the element form that conditionally selects PCDATA from a named element. The engine scans the XML markup until it finds the TEAM element where the value of the founded= attribute is 1993. The engine retrieves the value between the <TEAM> start tag and the </TEAM> end tag, which results in the value Panthers.

The Team5 variable uses the element form that conditionally selects PCDATA from a named element based on a specific occurrence of the element. The position function tells the engine to scan the XML markup until it finds the fifth occurrence of the TEAM element. The engine retrieves the value between the <TEAM> start tag and the </TEAM> end tag, which results in the value Capitals.

The following SAS statements import the XML document NhlShort.xml and specify the XMLMap named Nhl1.map. The PRINT procedure shows the resulting variables with selected values:

```sas
filename nhl 'C:\Example\NhlShort.xml';
filename map 'C:\Example\Nhl1.map';
libname nhl xmlv2 xmlmap=map;
proc print data=nhl.teams noobs;
run;
```

**Output 4.14 PRINT Procedure Output Showing Resulting Variables with Selected Values**

<table>
<thead>
<tr>
<th>ABBREV</th>
<th>FOUNDED</th>
<th>CONFERENCE</th>
<th>TEAM</th>
<th>TEAM5</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATL</td>
<td>1999</td>
<td>Eastern</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAR</td>
<td></td>
<td>Eastern</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FLA</td>
<td></td>
<td>Eastern</td>
<td>Panthers</td>
<td></td>
</tr>
<tr>
<td>TB</td>
<td></td>
<td>Eastern</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WSH</td>
<td></td>
<td>Eastern</td>
<td>Capitals</td>
<td></td>
</tr>
</tbody>
</table>
Including Namespace Elements in an XMLMap

This example illustrates the XMLMap namespace elements. The XMLMap namespace elements enable you to import an XML document with like-named elements that are qualified with XML namespaces. The XMLMap namespace elements maintain XML namespaces from the imported XML document to export an XML document with namespaces from the SAS data set.

Here is an XML document named NSSample.xml to be imported. The XML document contains three XML namespaces. The namespaces distinguish ADDRESS elements by qualifying them with references to unique URIs. The ADDRESS elements are highlighted below in the first PERSON repeating element:

```xml
<?xml version="1.0" encoding="UTF-8"?>
<PEOPLE xmlns:HOME="http://sample.url.org/home"
xmlns:IP="http://sample.url.org/ip"
xmlns:WORK="http://sample.url.org/work">
  <PERSON>
    <NAME>Joe Smith</NAME>
    <HOME:ADDRESS>1234 Elm Street</HOME:ADDRESS>
    <HOME:PHONE>999-555-0011</HOME:PHONE>
    <WORK:ADDRESS>2001 Office Drive, Box 101</WORK:ADDRESS>
    <WORK:PHONE>999-555-0101</WORK:PHONE>
    <IP:ADDRESS>192.168.1.1</IP:ADDRESS>
  </PERSON>
  <PERSON>
    <NAME>Jane Jones</NAME>
    <HOME:ADDRESS>9876 Main Street</HOME:ADDRESS>
    <HOME:PHONE>999-555-0022</HOME:PHONE>
    <WORK:ADDRESS>2001 Office Drive, Box 102</WORK:ADDRESS>
    <WORK:PHONE>999-555-0102</WORK:PHONE>
    <IP:ADDRESS>172.16.1.2</IP:ADDRESS>
  </PERSON>
  <PERSON>
    <NAME>Pat Perkinson</NAME>
    <HOME:ADDRESS>1395 Half Way</HOME:ADDRESS>
    <HOME:PHONE>999-555-0033</HOME:PHONE>
    <WORK:ADDRESS>2001 Office Drive, Box 103</WORK:ADDRESS>
    <WORK:PHONE>999-555-0103</WORK:PHONE>
    <IP:ADDRESS>10.0.1.3</IP:ADDRESS>
  </PERSON>
</PEOPLE>
```

Here is the XMLMap that was used to import the XML document. Notations describe the namespace elements.

```xml
<SXLEMAP name="Namespace" version="2.1">
  <NAMESPACES count="3"> <!-- 1 -->
    <NS id="1" prefix="HOME">http://sample.url.org/home</NS> <!-- 2 -->
    <NS id="2" prefix="IP">http://sample.url.org/ip</NS>
    <NS id="3" prefix="WORK">http://sample.url.org/work</NS>
  </NAMESPACES>
</SXLEMAP>
```
1. A NAMESPACES element contains NS elements for defining XML namespaces. The count= attribute specifies that there are three defined XML namespaces.

2. Three NS elements define the XML namespaces by referencing unique URIs. The id= attribute specifies the identification numbers 1, 2, and 3 for the three XML namespaces. The prefix= attribute assigns the names HOME, WORK, and IP to the referenced URIs.
3. The XMLMap TABLE element contains the data set definition for the PERSON repeating element.

4. XMLMap COLUMN elements contain variable definitions for each nested element within PERSON, which includes NAME, ADDRESS, PHONE, ADDRESS1, PHONE1, and ADDRESS2.

5. In the PATH element for each COLUMN element, the type of syntax is specified as XPathENR (XPath with Embedded Namespace Reference). This type indicates that the syntax is not compliant with the XPath specification. In addition, the identification number is included in the location path preceding the element that is being defined. The identification number is enclosed in braces. For example, this is the PATH element for the ADDRESS element: `<PATH syntax="XPathENR">
PEOPLE/PERSON/{1}ADDRESS</PATH>`.

The following SAS statements import the XML document and specify an XMLMap named NSSample.map. The PRINT procedure shows the resulting SAS data set:

```sas
filename ns 'C:\Example\NSSample.xml';
filename nsmap 'C:\Example\NSSample.map';

libname ns xmlv2 xmlmap=nsmap;

proc print data=ns.person noobs;
run;
```

Output 4.15  PRINT Procedure Output for NS.Person

![The SAS System Table]

Using the AUTOMAP= Option to Generate an XMLMap

This example illustrates how to import an XML document using the AUTOMAP= option to automatically generate an XMLMap file. By specifying the AUTOMAP= option in the LIBNAME statement, SAS analyzes the structure of the specified XML document and generates XMLMap syntax that describes how to interpret the XML markup into a SAS data set or data sets, variables (columns), and observations (rows). The AUTOMAP= option is supported by the XMLV2 engine only.

Here is the XML document Nhl.xml to be imported. If you try to import the document without an XMLMap, an error indicates that the data is not in a supported format.

```xml
<?xml version="1.0" encoding="iso-8859-1" ?>
```
The following SAS statements import the XML document Nhl.xml:

```sas
filename nhl 'C:\Example\Nhl.xml'; /*1*/
filename map 'C:\Output\NhlGenerate.map'; /*2*/
libname nhl xmlv2 automap=replace xmlmap=map; /*3*/
proc print data=nhl.team; /*4*/
run;
```

1. The first FILENAME statement assigns the file reference Nhl to the physical location (complete pathname, filename, and file extension) of the XML document named Nhl.xml to be imported.

2. The second FILENAME statement assigns the file reference Map to the physical location of the XMLMap named NhlGenerate.map to be generated.

3. The LIBNAME statement includes the following arguments:
   - The LIBNAME statement assigns the library reference Nhl, which matches the file reference that is assigned in the first FILENAME statement. Because the library reference and file reference match, the physical location of the XML document to be imported does not have to be specified in the LIBNAME statement.
   - The XMLV2 engine is specified.
   - The AUTOMAP=REPLACE option requests an XMLMap file to be generated and to overwrite the filename, if it exists. The default setting PREFIXATTRIBUTES=YES specifies that the element name is concatenated to the attribute name in each generated XMLMap /COLUMN element, which defines the SAS variable name.
   - The XMLMAP= option specifies the file reference map, which matches the file reference that is assigned in the second FILENAME statement. The file reference is associated with the physical location of the XMLMap to be generated.

4. PROC PRINT produces output, verifying that the import was successful.
Here is the generated XMLMap NhlGenerate.map:

```xml
<?xml version="1.0" encoding="UTF-8"?>
<!-- #------------------------------------------------------------ -->
<!-- 2012-09-21T10:30:48 -->
<!-- SAS XML Libname Engine Map -->
<!-- Generated by XML Mapper, 904000.4.0.20120905190000_v940 -->
<!-- #------------------------------------------------------------ -->
<!-- Validation report # -->
<!-- XMLMap validation completed successfully. -->
<!-- #------------------------------------------------------------ -->
<?xml version="1.0" encoding="UTF-8"?>
<NAME> name="AUTO_GBN" version="2.1">
  <NAMESPACES count="0"/>
  <!-- #------------------------------------------------------------ -->
  <TABLE description="NHL" name="NHL">
    <TABLE-PATH syntax="XPath">/NHL</TABLE-PATH>
    <COLUMN class="ORDINAL" name="NHL_ORDINAL">
      <INCREMENT-PATH beginend="BEGIN" syntax="XPath">/NHL</INCREMENT-PATH>
      <TYPE>numeric</TYPE>
      <DATATYPE>integer</DATATYPE>
    </COLUMN>
  </TABLE>
  <!-- #------------------------------------------------------------ -->
  <TABLE description="CONFERENCE" name="CONFERENCE">
    <TABLE-PATH syntax="XPath">/NHL/CONFERENCE</TABLE-PATH>
    <COLUMN class="ORDINAL" name="NHL_ORDINAL">
      <INCREMENT-PATH beginend="BEGIN" syntax="XPath">/NHL</INCREMENT-PATH>
      <TYPE>numeric</TYPE>
      <DATATYPE>integer</DATATYPE>
    </COLUMN>
    <COLUMN class="ORDINAL" name="CONFERENCE_ORDINAL">
      <INCREMENT-PATH beginend="BEGIN" syntax="XPath">/NHL/CONFERENCE</INCREMENT-PATH>
      <TYPE>numeric</TYPE>
      <DATATYPE>integer</DATATYPE>
    </COLUMN>
    <COLUMN name="CONFERENCE">
      <PATH syntax="XPath">/NHL/CONFERENCE</PATH>
      <TYPE>character</TYPE>
      <DATATYPE>string</DATATYPE>
      <LENGTH>12</LENGTH>
    </COLUMN>
  </TABLE>
</NAME>
<!-- #------------------------------------------------------------ -->
```
<TABLE description="DIVISION" name="DIVISION">
  <TABLE-PATH syntax="XPath">/NHL/CONFERENCE/DIVISION</TABLE-PATH>

  <COLUMN class="ORDINAL" name="CONFERENCE_ORDINAL">
    <INCREMENT-PATH beginend="BEGIN" syntax="XPath">/NHL/CONFERENCE</INCREMENT-PATH>
    <TYPE>numeric</TYPE>
    <DATATYPE>integer</DATATYPE>
  </COLUMN>

  <COLUMN class="ORDINAL" name="DIVISION_ORDINAL">
    <INCREMENT-PATH beginend="BEGIN" syntax="XPath">/NHL/CONFERENCE/DIVISION</INCREMENT-PATH>
    <TYPE>numeric</TYPE>
    <DATATYPE>integer</DATATYPE>
  </COLUMN>

  <COLUMN name="DIVISION">
    <PATH syntax="XPath">/NHL/CONFERENCE/DIVISION</PATH>
    <TYPE>character</TYPE>
    <DATATYPE>string</DATATYPE>
    <LENGTH>28</LENGTH>
  </COLUMN>
</TABLE>

<!-- ############################################################ -->

<TABLE description="TEAM" name="TEAM">
  <TABLE-PATH syntax="XPath">/NHL/CONFERENCE/DIVISION/TEAM</TABLE-PATH>

  <COLUMN class="ORDINAL" name="DIVISION_ORDINAL">
    <INCREMENT-PATH beginend="BEGIN" syntax="XPath">/NHL/CONFERENCE/DIVISION</INCREMENT-PATH>
    <TYPE>numeric</TYPE>
    <DATATYPE>integer</DATATYPE>
  </COLUMN>

  <COLUMN class="ORDINAL" name="TEAM_ORDINAL">
    <INCREMENT-PATH beginend="BEGIN" syntax="XPath">/NHL/CONFERENCE/DIVISION/TEAM</INCREMENT-PATH>
    <TYPE>numeric</TYPE>
    <DATATYPE>integer</DATATYPE>
  </COLUMN>

  <COLUMN name="TEAM_name">
    <PATH syntax="XPath">/NHL/CONFERENCE/DIVISION/TEAM/@name</PATH>
    <TYPE>character</TYPE>
    <DATATYPE>string</DATATYPE>
    <LENGTH>10</LENGTH>
  </COLUMN>

  <COLUMN name="TEAM_abbrev">
    <PATH syntax="XPath">/NHL/CONFERENCE/DIVISION/TEAM/@abbrev</PATH>
    <TYPE>character</TYPE>
    <DATATYPE>string</DATATYPE>
    <LENGTH>3</LENGTH>
  </COLUMN>
</TABLE>
Here is the PRINT procedure output for Nhl.Team

**Output 4.16  PRINT Procedure Output for Nhl.Team**

<table>
<thead>
<tr>
<th>Obs</th>
<th>DIVISION_ORDINAL</th>
<th>TEAM_ORDINAL</th>
<th>TEAM_name</th>
<th>TEAM_abbrev</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>Thrashers</td>
<td>ATL</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>Hurricanes</td>
<td>CAR</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>3</td>
<td>Panthers</td>
<td>FLA</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>4</td>
<td>Lightning</td>
<td>TB</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>5</td>
<td>Capitals</td>
<td>WSH</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>6</td>
<td>Stars</td>
<td>DAL</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>7</td>
<td>Kings</td>
<td>LA</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>8</td>
<td>Ducks</td>
<td>ANA</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>9</td>
<td>Coyotes</td>
<td>PHX</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>10</td>
<td>Sharks</td>
<td>SJ</td>
</tr>
</tbody>
</table>

---

**Avoiding Truncation Errors When Importing in a Multi-Byte Encoding**

To avoid truncation in a double-byte character set (DBCS) or a multi-byte character set (MBCS), use the CHARMULTIPLIER= or DERIVECHARMULTIPLIER= option to expand the column (variable) length of a generated XMLMap. The issue is that when an XMLMap is generated or created by SAS XML Mapper, the LENGTH element specifies the number of characters in a column. SAS XML Mapper does not have knowledge of your SAS session encoding. The XMLV2 engine does use the SAS session encoding, and it can use the correct number of bytes to represent a multi-byte character. For example, a character might be represented in the wlatin1 encoding as one byte, but in UTF-8, it is represented as two bytes. The XMLMap assigns a length of 1 to that character. The XMLV2 engine assigns a length of 2 to that character. Because of this difference, when your SAS session encoding uses more than one byte to represent a character, truncation could occur if the XMLMap column length does not accommodate the larger character size.

The example uses NhlGenerate.map, which is generated in “Using the AUTOMAP= Option to Generate an XMLMap” on page 49. (The XML content does not include multi-byte characters. The XML and XMLMap are used to show the expansion behavior only.)

```sas
filename nhl 'C:\Example\Nhl.xml';
filename map 'C:\Output\NhlGenerate.map';
libname nhl xmlv2 xmlmap=map;
```
Here is a portion of the PROC CONTENTS output, showing the column lengths without expansion:

```
  proc contents data=nhl.team;
  run;
```

The following example code uses the `CHARMULTIPLIER=` option to expand column lengths:

```
  libname nhl xmlv2 xmlmap=map charmultiplier=2.5;
  proc contents data=nhl.team;
  run;
```

Here is a portion of the PROC CONTENTS output, showing expanded column lengths for the two character variables:
Chapter 5
Exporting XML Documents Using an XMLMap

Why Use an XMLMap when Exporting?

To export an XML document that was imported using an XMLMap, you can use the same XMLMap or create a new one. The XMLMap syntax tells the XMLV2 engine how to map a SAS data set into a specific XML document structure.

To export an XML document using an XMLMap, specify the XMLV2 engine in the LIBNAME statement, and use the XMLMAP= option to specify the file. Using an XMLMap for exporting XML data is supported by the XMLV2 engine only, and not the older XML engine.

See the XMLMap syntax on page 95 for the version numbers that are supported for export.

Using an XMLMap to Export an XML Document with a Hierarchical Structure

This example modifies an existing XMLMap to export a SAS data set to a different XML document structure. For the original XMLMap and data set, see “Using an XMLMap to Import an XML Document as One SAS Data Set” on page 23.

Here is the Nhl.Teams data set to be exported:
If the data were exported without an XMLMap, the structure of the resulting XML document would be rectangular and consist of a TEAMS element for each observation in the SAS data set. For example:

```xml
<?xml version="1.0" encoding="windows-1252" ?>
<TABLE>
  <TEAMS>
    <NAME>Thrashers</NAME>
    <ABBREV>ATL</ABBREV>
    <CONFERENCE>Eastern</CONFERENCE>
    <DIVISION>Southeast</DIVISION>
  </TEAMS>

  <TEAMS>
    <NAME>Hurricanes</NAME>
    <ABBREV>CAR</ABBREV>
    <CONFERENCE>Eastern</CONFERENCE>
    <DIVISION>Southeast</DIVISION>
  </TEAMS>

  .
  .
  .
</TABLE>
```

To export the SAS data set as an XML document that structures data hierarchically by division within each conference, an XMLMap is required. The only change to the existing XMLMap is to include the OUTPUT element. Notations in the XMLMap syntax are explained.

```xml
<?xml version="1.0" ?>
<SXLEMAP version="2.1"> <!--

1

-->
<OUTPUT> <!--

2

--> 
<HEADING>
  <ATTRIBUTE name="description"
 value="Teams of the National Hockey League" />
</HEADING>
<TABLEREF name="TEAMS" />
</OUTPUT>

<TABLE name="TEAMS">
<TABLE-PATH syntax="XPath">/NHL/CONFERENCE/DIVISION/TEAM</TABLE-PATH>

<COLUMN name="NAME">
(PATH syntax="XPath">/NHL/CONFERENCE/DIVISION/TEAM/uname</PATH>
<TYPE>character</TYPE>
<DATATYPE>STRING</DATATYPE>
<LENGTH>30</LENGTH>
</COLUMN>

<COLUMN name="ABBREV">
(PATH syntax="XPath">/NHL/CONFERENCE/DIVISION/TEAM/abbrev</PATH>
<TYPE>character</TYPE>
<DATATYPE>STRING</DATATYPE>
<LENGTH>3</LENGTH>
</COLUMN>

<COLUMN name="CONFERENCE" retain="YES">
(PATH syntax="XPath">/NHL/CONFERENCE</PATH>
<TYPE>character</TYPE>
<DATATYPE>STRING</DATATYPE>
<LENGTH>10</LENGTH>
</COLUMN>

<COLUMN name="DIVISION" retain="YES">
(PATH syntax="XPath">/NHL/CONFERENCE/DIVISION</PATH>
<TYPE>character</TYPE>
<DATATYPE>STRING</DATATYPE>
<LENGTH>10</LENGTH>
</COLUMN>
</TABLE>
</SXLEMAP>

1 See the XMLMap syntax on page 95 for the version numbers that are supported for export.

2 To use an XMLMap to export the SAS data set as an XML document, you must include the OUTPUT element in the XMLMap. The OUTPUT element can contain the optional HEADING element.

3 The optional ATTRIBUTE element defines additional XML file attribute information. Each ATTRIBUTE element specifies a name/value pair. To be usable XML, the specification must be a valid XML attribute name and value. Each ATTRIBUTE element is included as an attribute of the root element of the exported document in the form name=value. See the result in the exported XML below.

4 The TABLEREF element references the name of the table to be exported. The TABLEREF element must reference the name of a TABLE element in the XMLMap.

The following SAS statements export the SAS data set named Myfiles.Teams to an XML document named NhlOut.XML, using an XMLMap named NhlExport.MAP:

```
libname myfiles 'C:\MyFiles\';
filename out 'C:\Output
hlout.xml';
```
libname out xmlv2 xmltype=xmlmap xmlmap='C:\Example\nhlexport.map';
data out.TEAMS;
   set myfiles.teams;
run;

Notice in the code above that TEAMS is uppercase to match the casing in the XMLMap. Here is the resulting XML document:

```xml
<?xml version="1.0" encoding="utf-8" ?>
<!-- SAS XML Libname Engine (SAS92XML)
   SAS XMLMap Generated Output
   Version V.03.03M010122017
   Created 2017-10-13T15:54:55
-->

<NHL description="Teams of the National Hockey League">
   <CONFERENCE>Eastern
      <DIVISION>Southeast
         <TEAM name="Thrashers" abbrev="ATL" />
         <TEAM name="Hurricanes" abbrev="CAR" />
         <TEAM name="Panthers" abbrev="FLA" />
         <TEAM name="Lightning" abbrev="TB" />
         <TEAM name="Capitals" abbrev="WSH" />
      </DIVISION>
   </CONFERENCE>
   <CONFERENCE>Western
      <DIVISION>Pacific
         <TEAM name="Stars" abbrev="DAL" />
         <TEAM name="Kings" abbrev="LA" />
         <TEAM name="Ducks" abbrev="ANA" />
         <TEAM name="Coyotes" abbrev="PHX" />
         <TEAM name="Sharks" abbrev="SJ" />
      </DIVISION>
   </CONFERENCE>
</NHL>
```
Chapter 6
Understanding and Using Tagsets for the XMLV2 Engine

What Is a Tagset?
A tagset specifies instructions for generating a markup language from your SAS data set. The resulting output contains embedded instructions defining layout and some content. SAS provides tagsets for a variety of markup languages, including the XML markup language.

Creating Customized Tagsets
In addition to using the tagsets provided by SAS, you can modify the SAS tagsets, and you can create your own tagsets. To create a tagset, use the TEMPLATE procedure to define the tagset definition. For information about defining customized tagsets, see the TEMPLATE procedure in SAS Output Delivery System: Procedures Guide.

CAUTION:
Use customized tagsets with caution. If you are unfamiliar with XML output, do not specify different tagsets. If you alter the tagset when exporting an XML document, and then attempt to import the XML document generated by that altered tagset, the engine might not be able to translate the XML markup back to SAS proprietary format.
Exporting an XML Document Using a Customized Tagset

Example Overview

This example defines a customized tagset, and then uses the tagset with the XMLV2 engine to export an XML document with customized tags.

Define Customized Tagset Using TEMPLATE Procedure

The following TEMPLATE procedure defines a customized tagset named Tagsets.Custom.

You can use the following code as a template to define your own customized tagsets. For example, to create your own customized tagset, only the EmitMeta, EmitRow, and EmitCol events would require minor modifications. Submit the following PROC TEMPLATE code in a SAS session:

```sas
proc template;
    /* +------------------------------------------------+*/
    /* |                                                | */
    /* | +------------------------------------------------+ */
    define tagset tagsets.custom ;
    notes "SAS XMLV2 engine output event model(interface)";
    indent = 3;
    map = '<>&"''';
    mapsub = '/</>/&/"'/';
    /* +------------------------------------------------+*/
    /* |                                                | */
    /* | +------------------------------------------------+ */
    define event XMLversion;
    put  '<?xml version="1.0"';
    putq ' encoding=' ENCODING;
    put  ' ?>' CR;
    break;
    end;

    define event XMLcomment;
    put  '<!-- ' CR;
    put  '     ' TEXT CR;
    put  '  -->' CR;
    break;
    end;
```

60 Chapter 6 • Understanding and Using Tagsets for the XMLV2 Engine
define event initialize;
set $LIBRARYNAME 'LIBRARY' ;
set $TABLENAME 'DATASET' ;
set $COLTAG 'column' ;
set $META 'FULL' ;

eval $is_engine 1;
eval $is_procprint 0;
eval $is_OUTBOARD 1;
end;

/* +------------------------------------------------+
|                                                |
|+------------------------------------------------+ */

define event doc;
start:
    trigger initialize;
    trigger XMLversion;
    break;
finish:
    break;
end;

define event doc_head;
start:
    break;
finish:
    break;
end;

define event doc_body;
start:
    break;
finish:
    break;
end;

define event proc;
start:
    break / if frame_name ; /* set by ODS statement use */
eval $is_OUTBOARD 0 ; /* default for non-engine */
do / if cmp(XMLCONTROL, "OUTBOARD"); /* only the engine sets this */
eval $is_OUTBOARD 1 ;
else ;
    eval $is_OUTBOARD 0 ;
done ;
break;
finish:
    break;
end;
define event leaf;
start:
  /*
   * PROC PRINT
   * data set reference is in the value and label fields
   * and NOT in the output_label field
   */
eval $is_engine   0; /* NOT ENGINE */
break / if ^cmp("Print", name);
eval $is_procprint 1; /* PROC PRINT */
eval $regex prxpath("/(.+)/");
eval $match prxmatch($regex, value);
set $TABLENAME prxposn($regex, 1, value);
break;
finish:
  break;
end;

define event output;
start:
  break / if $is_procprint ;
eval $is_engine   0; /* NOT ENGINE */
set $TABLENAME name / if name; /* TABLE VIEWER */
break;
finish:
  break;
end;

define event table;
start:
  unset $col_names;
  unset $col_types;
  unset $col_width;
eval $index      1;
eval $index_max  0;
set $TABLENAME name / if name; /* LIBNAME ENGINE */
set $META XMLMETADATA / if XMLMETADATA ; /* LIBNAME ENGINE */
set $SCHEMA XMLSCHEMA / if XMLSCHEMA ; /* LIBNAME ENGINE */
break;
finish:
  break;
end;

define event colspecs;
start:
  break / if cmp(XMLMETADATA, "NONE");
finish:
  break / if cmp(XMLMETADATA, "NONE");
end;

define event colgroup;
start:
  break / if cmp(XMLMETADATA, "NONE");
finish:
  break / if cmp(XMLMETADATA, "NONE");
end;

define event colspec_entry;
start:
  break / if ^$is_engine and $index eq 1 and cmp(name, "Obs");
  eval $index_max $index_max+1;
  set $col_names[] name;
  set $col_types[] type;
  set $col_width[] width;
  break;
finish:
  break;
end;

define event table_head;
start:
  break;
finish:
  break;
end;

define event table_body;
start:
  trigger EmitMeta ;
  break;
finish:
  trigger EmitMeta ;
  break;
end;

define event row;
start:
  break / if !cmp(SECTION, "body");
  break / if cmp(XMLMETADATA, "ONLY");
  eval $index 1;
  unset $col_values;
  break;
finish:
  break / if !cmp(SECTION, "body");
  break / if cmp(XMLMETADATA, "ONLY");
  trigger EmitRow ;
break;
end;

define event data;
start:
  break / if !cmp(SECTION, "body");
  do / if $is_engine ;
    break / if !cmp(XMLCONTROL, "Data");
  else ;
    break / if !cmp(HTMLCLASS, "Data");
  done ;
  break / if cmp(XMLMETADATA, "ONLY");
  set $name $col_names[$index];
  do / if exists(MISSING);
    eval $is_MISSING 1;
    eval $value_MISSING MISSING;
    set $col_values[$name] " ";
  else ;
    eval $is_MISSING 0;
    set $col_values[$name] VALUE;
  done;
  break;
finish:
  break / if !cmp(SECTION, "body");
  do / if $is_engine ;
    break / if !cmp(XMLCONTROL, "Data");
  else ;
    break / if !cmp(HTMLCLASS, "Data");
  done ;
  break / if cmp(XMLMETADATA, "ONLY");
  set $name $col_names[$index];
  eval $index $index+1;
  break;
end;

/* +------------------------------------------------+
|                                                |
| at this point, we just take over XML output.  |
| EmitRow() is triggered each time the data is   |
| loaded into the $col_values array.            |
|                                                |
| we can output anything we desire from here...  |
|                                                |
| +------------------------------------------------+ */
define event EmitMeta; /* */
start:
  put ' <' $LIBRARYNAME ' >' CR ;
  put ' <!-- List of available columns' CR ;
  eval $index 1;
  iterate $col_names ;
  do /while _value_;
The EmitMeta event generates an XML comment that contains a list of the variables from the SAS data set. The event contains an example of iteration for a list variable, which processes all of the variables in the SAS data set. For more information about iteration, see the ITERATE statement in the TEMPLATE procedure DEFINE EVENT statement in SAS Output Delivery System: Procedures Guide.

The EmitRow event creates XML output from the three SAS data set observations. The EmitRow event names specific variables to process, which are Name, Height, and Weight.

The EmitCol event creates generic-looking XML for each processed variable.
The DATA step creates a data set named Work.Class that consists of only three observations.

The FILENAME statement assigns the fileref XmlOut to the physical location of the file that will store the exported XML document (complete pathname, filename, and file extension).

The LIBNAME statement uses the fileref to reference the XML document and specifies the XMLV2 engine. The TAGSET= option specifies the customized tagset named Tagsets.Custom.

The DATA step reads the data set Work.Class and writes its content to the MyClass.XML document in the format that is defined by the customized tagset.

Here is the resulting XML document:

Output 6.1  Exported XML Document Using Customized Tagset

```xml
<?xml version="1.0" encoding="windows-1252" ?>
<LIBRARY>
   <!-- List of available columns
       1 Name
       2 Sex
       3 Age
       4 Height
       5 Weight
   -->
   <STUDENT>
      <Name>Alfred</Name>
      <Height>69</Height>
      <Weight>112.5</Weight>
   </STUDENT>
   <STUDENT>
      <Name>Alice</Name>
      <Height>56.5</Height>
      <Weight>84</Weight>
   </STUDENT>
   <STUDENT>
      <Name>Barbara</Name>
      <Height>65.3</Height>
      <Weight>98</Weight>
   </STUDENT>
</LIBRARY>
```
Part 2

LIBNAME Statement Syntax

Chapter 7
Introduction to the XMLV2 LIBNAME Statements .................. 69

Chapter 8
LIBNAME Statement for the XMLV2 Engine ......................... 73
Chapter 7

Introduction to the XMLV2 LIBNAME Statements

Using the LIBNAME Statement

For the XMLV2 engine, the LIBNAME statement associates a SAS libref with either a SAS library that stores XML documents or a specific XML document to import or export.


Comparing the XMLV2 and XML Engines

Overview of XMLV2 and XML Engine Differences

SAS provides two versions of functionality by implementing two engine names in the LIBNAME statement:

- The XMLV2 engine supports current functionality and includes enhancements that have been made since SAS 9.2. (If you are running in a SAS 9.2 session, use the alias XML92.)
- The XML engine is a compatibility engine for legacy programs to support SAS 9.1.3 and earlier functionality.

The differences between the two engines are explained in the following topics.
**XML Compliance**

The XMLV2 engine is XML compliant, which means the XML markup must be well-formed and in valid construction that is in compliance with the W3C specifications. XML compliance could affect the following situations:

- XML documents that can be imported with the older XML engine might not pass the more strict parsing rules in the XMLV2 engine. For example, XML compliance means that the markup is case sensitive. Opening and closing tags must be written in the same case, such as `<BODY> ...</BODY>` and `<Message>...</Message>`. For the XMLV2 engine, the tag `<Letter>` is different from the tag `<letter>`. Attribute names are also case sensitive, and the attribute value must be enclosed in quotation marks, such as `<Note date="09/24/1975">`.

- XMLMap files that are accepted by the older XML engine might not work with the XMLV2 engine. For example, XML compliance means that the markup and the supported XPath syntax are case sensitive. In addition, the XMLMap markup must follow specific XMLMap rules. Tag names must be uppercase. Element attributes must be lowercase. An example is `<SXLemap version="2.1">`.

**XMLMap Files**

The XMLV2 engine supports XMLMap files starting with XMLMap version 1.2. The older XML engine supports all XMLMap files starting with XMLMap version 1.0.

The documented XMLMap syntax version is 2.1. See Chapter 9, “Introduction to XMLMap Syntax,” on page 91.

The XMLV2 engine supports using XMLMap for exporting and also supports XML namespaces.

**LIBNAME Statement Functionality Enhancements for XMLV2**

The XMLV2 engine provides the following LIBNAME statement functionality:

- The ability to assign a libref to a SAS library rather than assigning the libref to a specific XML document.
- Using the GENERIC markup type, you can export a single XML document from multiple SAS data sets.
- Additional options and option values are available for the XMLV2 engine. Some legacy options or option values are not supported for the XMLV2 engine. The following topic covers these differences.

**XMLV2 and XML Engine LIBNAME Statement Options**

The following table lists the available LIBNAME statement options. The ■ symbol indicates whether the option is available for each engine.
### Table 7.1 LIBNAME Statement Options

<table>
<thead>
<tr>
<th>Task</th>
<th>Option</th>
<th>XML</th>
<th>XMLV2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automatically generate an XMLMap file to import an XML document</td>
<td>AUTOMAP= on page 76</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Expand XMLMap column (variable) lengths by a multiplier value</td>
<td>CHARMULTIPLIER= on page 77</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Expand XMLMap column (variable) lengths by a default multiplier value that is based on the session encoding</td>
<td>DERIVECHARMULTIPLIER= on page 77</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Determine whether SAS formats are used with the GENERIC markup type</td>
<td>FORMATACTIVE= on page 77</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Determine whether SAS formats are used with the CDISCODM markup type</td>
<td>FORMATACTIVE= on page 77</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Specify the libref to create a format catalog with the CDISCODM markup type</td>
<td>FORMATLIBRARY= on page 78</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Replace existing format entries in the format catalog with the CDISCODM markup type</td>
<td>FORMATNOREPLACE= on page 78</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Indent nested elements in exported XML document</td>
<td>INDENT= on page 79</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Specify the character set to use for the output file</td>
<td>ODSCHARSET= on page 79</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Control the generation of a record separator</td>
<td>ODSRECSEP= on page 79</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Specify the translation table to use for the output file</td>
<td>ODSTRANTAB= on page 80</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Task</td>
<td>Option</td>
<td>XML</td>
<td>XMLV2</td>
</tr>
<tr>
<td>---------------------------------------------------------------------</td>
<td>-------------------------------</td>
<td>-----</td>
<td>-------</td>
</tr>
<tr>
<td>Specifies whether the element name is concatenated to the attribute name when generating each XMLMap COLUMN element</td>
<td>PREFIXATTRIBUTES= on page 80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Override the default tagset</td>
<td>TAGSET= on page 80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Import concatenated XML documents</td>
<td>XMLCONCATENATE= on page 81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specify the tag format to contain SAS variable information</td>
<td>XMLDATAFORM= on page 81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control the results of numeric values</td>
<td>XMLDOUBLE= on page 81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Override the SAS data set's encoding for the output file</td>
<td>XMLENCODING= on page 82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specify a fileref for the XML document</td>
<td>XMLFILEREF= on page 83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specify an XMLMap</td>
<td>XMLMAP= on page 83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Determine whether metadata-related information is included</td>
<td>XMLMETA= on page 84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Determine whether to process nonconforming character data</td>
<td>XMLPROCESS= on page 85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specify an external file to contain exported metadata-related information</td>
<td>XMLSCHEMA= on page 85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specify the XML markup type</td>
<td>XMLTYPE= on page 86</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Chapter 8
LIBNAME Statement for the XMLV2 Engine

Dictionary

LIBNAME Statement for the XMLV2 and XML Engines

Processes an XML document.

Valid in: Anywhere
Category: Data Access

Syntax

LIBNAME libref engine 'SAS-library | XML-document-path' <option(s)>;

Summary of Optional Arguments

AUTOMAP=REPLACE | REUSE XMLV2 Only
specifies to automatically generate an XMLMap file to import an XML document.

CHARMULTIPLIER=value XMLV2 Only
expands column (variable) lengths by a multiplier value. These are column (variable) lengths specified in an XMLMap.

DERIVECHARMULTIPLIER=YES | NO XMLV2 Only
expands column (variable) lengths by a default multiplier value that is based on the session encoding. These are column (variable) lengths specified in an XMLMap.

FORMATACTIVE=NO | YES XMLV2 and XML with differences
determines whether SAS formats are used.

FORMATLIBRARY=libref XML Only
specifies the libref of an existing SAS library in which to create the format catalog.

FORMATNOREPLACE=NO | YES XML Only
specifies whether to replace existing format entries in the format catalog search path in cases where an existing format entry has the same name as a format that is being created by the XML engine when it converts a CDISC ODM CodeList element.

**INDENT=integer** XMLV2 and XML

specifies the number of columns to indent each nested element in the exported XML document.

**ODSCHARSET=character-set** XMLV2 and XML

specifies the character set to use for the output file.

**ODSRECSEP=DEFAULT | NONE | YES** XML Only

controls the generation of a record separator that marks the end of a line in the output XML document.

**ODSTRANTAB=table-name** XMLV2 and XML

specifies the translation table to use for the output file.

**PREFIXATTRIBUTES=YES | NO** XMLV2 Only

specifies whether the element name is concatenated to the attribute name when generating each XMLMap COLUMN element, which defines the SAS variable name.

**TAGSET=tagset-name** XMLV2 and XML

specifies the name of a tagset to override the default tagset that is used by the markup type that is specified with XMLTYPE=.

**XMLCONCATENATE=NO | YES** XMLV2 and XML

specifies whether the file to be imported contains multiple, concatenated XML documents.

**XMLDATAFORM=ELEMENT | ATTRIBUTE** XMLV2 and XML

specifies whether the tag for the element to contain SAS variable information (name and data) is in open element or enclosed attribute format.

**XMLDOUBLE=DISPLAY | INTERNAL**

determines how the engine processes character data that does not conform to W3C specifications.

**XMLENCODING='encoding-value'** XMLV2 and XML

overrides the SAS data set’s encoding for the output file.

**XMLFILEREF=fileref** XMLV2 and XML

is the SAS name that is associated with the physical location of the XML document to be exported or imported.

**XMLMAP=fileref | 'XMLMap-location'** XMLV2 and XML

specifies an XML document that contains specific XMLMap syntax.

**XMLMETA=DATA | SCHEMADATA | SCHEMA** XMLV2 and XML

specifies whether to include metadata-related information in the exported markup, or specifies whether to import metadata-related information that is included in the input XML document.

**XMLPROCESS=CONFORM | PERMIT** XMLV2 and XML

determines how the engine processes character data that does not conform to W3C specifications.

**XMLSCHEMA=fileref | 'external-file'** XMLV2 and XML

specifies an external file to contain metadata-related information.

**XMLTYPE=GENERIC | CDISCODM | MSACCESS | ORACLE | XMLMAP**

specifies the XML markup type.
**Required Arguments**

**libref**

is a valid SAS name that serves as a shortcut name to associate with the physical location of the XML document. The name must conform to the rules for SAS names. A libref cannot exceed eight characters.

**engine**

is one of two supported engines that import and export an XML document. The syntax documentation is labeled with a superscript such as `XMLV2 Only` or `XML Only` or `XMLV2 and XML`.

**XMLV2**

specifies the XMLV2 engine, which accesses the engine functionality of SAS 9.2 and later.

<table>
<thead>
<tr>
<th>Alias</th>
<th>XML92</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restriction</td>
<td>When exporting an XML document using the XMLV2 engine, you can specify up to 19 SAS data sets to export.</td>
</tr>
</tbody>
</table>

**XML**

specifies the XML engine, which accesses the engine functionality in SAS 9.1.3 and earlier.

**Note**  
For more information about the two engine names, see “Comparing the XMLV2 and XML Engines” on page 69.

'SAS-library | XML-document-path'

is the physical location of the XML document for export or import. Enclose the physical location in single or double quotation marks.

**SAS-library**  
XMLV2 Only

is the pathname for a collection of one or more files that are recognized by SAS and that are referenced and stored as a unit. Here is an example: `'C:\My Documents\XML'`.

**XML-document-path**  
XMLV2 and XML

includes the pathname, filename, and file extension. Here is an example: `'C:\My Documents\XML\myfile.xml'`.

**Operating Environment Information**

For details about specifying the physical location of files, see the SAS documentation for your operating environment.

**Interactions**

You can use the FILENAME statement in order to assign a fileref to be associated with the physical location of the XML document to be exported or imported. If the fileref matches the libref, you do not need to specify the physical location of the XML document in the LIBNAME statement. For example, the following code reads from the XML document Fred.XML:

```sas
filename bedrock 'C:\XMLdata\fred.xml';
libname bedrock xmlv2;
proc print data=bedrock.fred;
run;
```
To specify a fileref for the XML document that does not match the
libref, you can use the **XMLFILEREF=** option on page 83.

---

**Optional Arguments**

**AUTOMAP=REPLACE | REUSE**  XMLV2 Only

specifies to automatically generate an XMLMap file to import an XML document. The XMLMap file contains specific syntax that describes how to interpret the XML markup into a SAS data set or data sets, variables (columns), and observations (rows). XMLMap syntax is generated by analyzing the structure of the specified XML document. To automatically generate the XMLMap file, you must specify an existing XML document and the physical location for the output XMLMap file.

**REPLACE**

overwrites an existing XMLMap file.

If an XMLMap file exists at the specified physical location, the generated XMLMap file overwrites the existing one. If an XMLMap file does not exist at the specified physical location, the generated XMLMap file is written to the specified pathname and filename.

**REUSE**

does not overwrite an existing XMLMap file.

If an XMLMap file exists at the specified physical location, the existing XMLMap file is used. If an XMLMap file does not exist at the specified physical location, the generated XMLMap file is written to the specified pathname and filename.

**Restrictions**

Use this option when importing only.

The AUTOMAP= option is not supported in partitioned data sets (PDS) on z/OS.

**Requirements**

You must specify the physical location of an existing XML document with either the complete pathname, filename, and file extension, or with a file reference that is associated with the physical location for the DISK or TEMP device type only. The XML document must exist on disk. The AUTOMAP= option does not support accessing an XML document by using access methods such as FTP, SFTP, URL, or WebDAV.

You must include the **XMLMAP=** option to specify the physical location of the generated XMLMap file with either the complete pathname, filename, and file extension, or with a file reference that is associated with the physical location for the DISK or TEMP device type only. The AUTOMAP= option does not support accessing an XMLMap file by using access methods such as FTP, SFTP, URL, or WebDAV.

**Interaction**

To specify whether the element name is concatenated to the attribute name when generating each XMLMap COLUMN element, which defines the SAS variable name, use the **PREFIXATTRIBUTES=** option. By default, element names are attached to the resulting SAS variable names.

**Tips**

The functionality to automatically generate an XMLMap file is also available with SAS XML Mapper. The AUTOMAP= option
provides the functionality to create and use the XMLMap with a single LIBNAME statement.

You can assign the generated XMLMap file to a temporary location, which is deleted at the end of the SAS session. To do this, associate a file reference with a physical location and specify the TEMP device type.

Example

“Using the AUTOMAP= Option to Generate an XMLMap” on page 49

**CHARMULTIPLIER=value** XMLV2 Only

expands column (variable) lengths by a multiplier value. These are column (variable) lengths specified in an XMLMap. Use CHARMULTIPLIER= to avoid truncation when importing XML data that contains double-byte character set (DBCS) or multi-byte character set (MBCS) data. An XMLMap that is generated or created by SAS XML Mapper could specify column lengths that do not accommodate the larger character size.

For the multiplier value, specify a real number greater than or equal to 1.0 and less than or equal to 5.0. The engine multiplies the column lengths that are specified in an XMLMap by this number. For information about DBCS and MBCS issues, see *SAS National Language Support (NLS): Reference Guide*.

Example

“Avoiding Truncation Errors When Importing in a Multi-Byte Encoding” on page 53

**DERIVECHARMULTIPLIER=YES | NO** XMLV2 Only

expands column (variable) lengths by a default multiplier value that is based on the session encoding. These are column (variable) lengths specified in an XMLMap. Use DERIVECHARMULTIPLIER= YES to avoid truncation when importing XML data that contains DBCS or MBCS data. An XMLMap that is generated or created by SAS XML Mapper could specify column lengths that do not accommodate the larger character size.

The engine multiplies the column lengths that are specified in an XMLMap by a default multiplier value that is based on the current session encoding. Use DERIVECHARMULTIPLIER= YES rather than CHARMULTIPLIER= when you are uncertain of the session encoding. For information about DBCS and MBCS issues, see *SAS National Language Support (NLS): Reference Guide*.

Default

**NO**

Interaction

If CHARMULTIPLIER= is also specified, then DERIVECHARMULTIPLIER= is ignored.

**FORMATACTIVE=NO | YES** XMLV2 and XML with differences

determines whether SAS formats are used. Behavior differs between the two supported markup types as follows.

For the CDISCODM markup type: XML Only

FORMATACTIVE= specifies whether CDISC ODM CodeList elements, which contain instructions for transcoding display data in a CDISC ODM document, are converted to SAS formats, and vice versa.

**NO**

causes formatting controls to be suppressed for both importing and exporting.
YES
when importing, converts the CDISC ODM CodeList elements to the corresponding SAS formats, registers the SAS formats on the referenced variables, and stores the created SAS formats in the format catalog.

when exporting, converts the SAS formats to the corresponding CDISC ODM CodeList elements.

Tips
By default, the format catalog is created in the Work library. If you want to store the catalog in a permanent library, use the FORMATLIBRARY= option on page 78.

When the format catalog is updated, the default behavior is that any new SAS formats that are created by converting CDISC ODM CodeList elements will overwrite any existing SAS formats that have the same name. To prevent existing SAS formats from being overwritten, specify FORMATNOREPLACE=YES.

Example
“Exporting an XML Document in CDISC ODM Markup” on page 126

For the GENERIC markup type; XMLV2 and XML
FORMATACTIVE= specifies whether output values are affected by SAS formats.

NO
writes the actual data value to the XML markup.

YES
causes the XML markup to contain the formatted data value.

Restriction
For the GENERIC markup type, if you export a SAS data set with formatted data values, and then you try to import the XML document back into the existing SAS data set, the import might fail. Exporting a SAS data set with formatted data values can result in different variables or different variable attributes.

Default
NO

Restriction
Use this option for the CDISCODM and GENERIC markup types only

FORMATLIBRARY=libref XML Only
specifies the libref of an existing SAS library in which to create the format catalog.

Restrictions
Use this option when importing an XML document only.

Use this option only for the CDISCODM markup type with FORMATACTIVE=YES.

FORMATNOREPLACE=NO | YES XML Only
specifies whether to replace existing format entries in the format catalog search path in cases where an existing format entry has the same name as a format that is being created by the XML engine when it converts a CDISC ODM CodeList element.

NO
does not replace formats that have the same name.

YES
replaces formats that have the same name.
**Default** | NO  

**Restrictions** | Use this option when importing an XML document only. Use this option for the CDISCODM markup type only.

**INDENT=integer** XMLV2 and XML

specifies the number of columns to indent each nested element in the exported XML document. The value can be from 0 (which specifies no indentation) through 32. This specification is cosmetic and is ignored by an XML-enabled browser.

**Default** | 3

**Restriction** | Use this option when exporting an XML document only.

**ODSCHARSET=character-set** XMLV2 and XML

specifies the character set to use for the output file. A character set includes letters, logograms, digits, punctuation, symbols, and control characters that are used for display and printing. An example of a character set is ISO-8859-1.

**Restriction** | Use this option when exporting an XML document only.

**Requirement** | Use this option with caution. If you are unfamiliar with character sets, encoding methods, or translation tables, do not use this option without proper technical advice.

**Tip** | The combination of the character set and translation table (encoding method) results in the file's encoding.


**ODSRECSEP=DEFAULT | NONE | YES** XML Only

controls the generation of a record separator that marks the end of a line in the output XML document.

**DEFAULT**

enables the XML engine to determine whether to generate a record separator based on the operating environment where you run the SAS job.

The use of a record separator varies by operating environment.

**Tip** | If you do not transfer XML documents across environments, use the default behavior.

**NONE**

specifies to not generate a record separator.

The XML engine uses the logical record length of the file that you are writing to and writes one line of XML markup at a time to the output file.

**Requirement** | The logical record length of the file that you are writing to must be at least as long as the longest line that is produced. If the logical record length of the file is not long enough, then the markup might wrap to another line at an inappropriate place.
## Interaction

Transferring an XML document that does not contain a record separator can be a problem. For example, FTP needs a record separator to transfer data properly in ASCII (text) mode.

**YES**

specifies to generate a record separator.

**Default**

The XML engine determines whether to generate a record separator based on the operating environment where you run the SAS job.

**Restriction**

Use this option when exporting an XML document only.

**Interaction**

Most transfer utilities interpret the record separator as a carriage return sequence. For example, using FTP in ASCII (text) mode to transfer an XML document that contains a record separator results in properly constructed line breaks for the target environment.

### ODSTRANTAB=

**XMLV2 and XML**

specifies the translation table to use for the output file. The translation table (encoding method) is a set of rules that are used to map characters in a character set to numeric values. An example of a translation table is one that converts characters from EBCDIC to ASCII-ISO. The `table-name` can be any translation table that SAS provides or any user-defined translation table. The value must be the name of a SAS catalog entry in either the Sasuser.Profile catalog or the Sashelp.Host catalog.

**Restriction**

Use this option when exporting an XML document only.

**Requirement**

Use this option with caution. If you are unfamiliar with character sets, encoding methods, or translation tables, do not use this option without proper technical advice.

**Tip**

The combination of the character set and translation table results in the file's encoding.

**See**


### PREFIXATTRIBUTES= YES | NO

**XMLV2 Only**

specifies whether the element name is concatenated to the attribute name when generating each XMLMap COLUMN element, which defines the SAS variable name. If you specify PREFIXATTRIBUTES=NO, the element name is not attached to the resulting SAS variable name.

**Default**

YES

**Restrictions**

Use this option when importing an XML document only.

Use this option with the AUTOMAP= option on page 76.

**Example**

“Using the AUTOMAP= Option to Generate an XMLMap” on page 49

### TAGSET=

**XMLV2 and XML**

specifies the name of a tagset to override the default tagset that is used by the markup type that is specified with XMLTYPE=. To change the tags that are produced, you can create a customized tagset and specify it with the TAGSET=...
XMLCONCATENATE=NO | YES  
XMLV2 and XML

specifies whether the file to be imported contains multiple, concatenated XML documents. Importing multiple, concatenated XML documents can be useful (for example, if an application is producing a complete document per query or response as in a web form). For the XMLV2 engine, this option is not necessary and is ignored.

<table>
<thead>
<tr>
<th>Alias</th>
<th>XMLCONCAT=</th>
</tr>
</thead>
<tbody>
<tr>
<td>Default</td>
<td>NO</td>
</tr>
</tbody>
</table>

Restriction  Use this option when importing an XML document only.

Requirement Use XMLCONCATENATE=YES cautiously. If an XML document consists of concatenated XML documents, the content is not standard XML construction. The option is provided for convenience, not to encourage invalid XML markup.

Example  “Importing Concatenated XML Documents” on page 132

XMLDATAFORM=ELEMENT | ATTRIBUTE  
XMLV2 and XML

specifies whether the tag for the element to contain SAS variable information (name and data) is in open element or enclosed attribute format. For example, if the variable name is PRICE and the value of one observation is 1.98, the generated output for ELEMENT is `<PRICE> 1.98 </PRICE>` and for ATTRIBUTE is `<COLUMN name="PRICE" value="1.98" />`.

| Default       | ELEMENT |

Restrictions  Use this option when exporting an XML document only.

Use this option for the GENERIC markup type only.

XMLDOUBLE=DISPLAY | INTERNAL  
XMLV2 and XML

controls the results of importing or exporting numeric values.

DISPLAY  when exporting, the engine retrieves the stored value for the numeric variable, determines an appropriate display for the value in a readable form, and writes the
display value to the XML document. The display value is affected by the engine and whether a format is assigned.

- The XML engine uses an assigned format. The maximum value is 16 digits. For example, if a numeric variable has an assigned format width that is 20 digits, such as BEST20., the engine truncates the exported value. If there is not an assigned format, the engine displays the value using BEST10.
- The XMLV2 engine ignores any assigned format and displays the value using BEST16.

When importing, the engine retrieves PCDATA (parsed character data) from the named element in the XML document and converts the data into numeric variable content.

Alias FORMAT

INTERNAL XML import, XMLV2 and XML export

when exporting, the engine retrieves the stored value for the numeric variable and writes the raw value to a generated attribute value pair (of the form rawvalue="value"). SAS uses the base64 encoding of a portable machine representation. (The base64 encoding method converts binary data into ASCII text and vice versa and is similar to the MIME format.)

When importing, the engine retrieves the stored value from the rawvalue= attribute from the named element in the XML document. It converts that value into numeric variable content. The PCDATA content of the element is ignored. When importing, XMLDOUBLE=INTERNAL is not supported for the XMLV2 engine.

Alias PRECISION

Tip Typically, you use XMLDOUBLE=INTERNAL to import or export an XML document when content is more important than readability.

Default DISPLAY

Restriction You can specify the XMLDOUBLE= option for the GENERIC markup type only.

Examples “Exporting Numeric Values” on page 121

“Importing an XML Document with Numeric Values” on page 127

XMLENCODING='encoding-value' XMLV2 and XML

overrides the SAS data set's encoding for the output file. If an encoding value contains a hyphen, enclose the value in quotation marks.

Restriction Use this option when exporting an XML document only.

Requirement Use this option with caution. If you are unfamiliar with character sets, encoding methods, or translation tables, do not use this option without proper technical advice.

Tips When transferring an XML document across environments (for example, using FTP), you must be aware of the document's content to determine the appropriate transfer mode. If the document contains an encoding attribute in the XML declaration, or if a byte-order mark (BOM) precedes the XML declaration, transfer the XML document
in binary mode. If the document contains neither of these and you are transferring the document across similar environments, transfer the XML document in text mode.

The combination of the character set and translation table (encoding method) results in the file's encoding.


XMLFILEREF=fileref** XMLV2 and XML**
is the SAS name that is associated with the physical location of the XML document to be exported or imported. To assign the fileref, use the FILENAME statement. The engine can access any data referenced by a fileref. For example, the following code writes to the XML document Wilma.XML:

```
libname source 'C:\Example';
filename cartoon 'C:\XMLdata\wilma.xml';
libname bedrock xmlv2 xmlfileref=cartoon;
data bedrock.wilma;
set source.wilma;
run;
```

Tip When using the URL access method to reference a fileref that is assigned to an XML document, you should also specify an XMLMap. Specifying an XMLMap causes the engine to process the XML document with a single pass. Whether you need to specify an XMLMap depends on your web server. For an example, see “Referencing a Fileref Using the URL Access Method” on page 43.

**XMLMAP=fileref | 'XMLMap-location'** **XMLV2 and XML**
specifies an XML document that contains specific XMLMap syntax. The syntax tells the engine how to interpret the XML markup for importing or exporting. When you import with the AUTOMAP=YES option, specify the fileref or location of the XMLMap to be generated. When you export, or when you import without the AUTOMAP=YES option, specify the fileref or location of the existing XMLMap.

fileref

is the SAS name that is associated with the physical location of the XMLMap. To assign a fileref, use the FILENAME statement.

Tip To assign a fileref to an XMLMap using the URL access method, your web server might require that the file extension be .xml instead of .map.

'XMLMap-location'
is the physical location of the XMLMap.

Include the complete pathname and the filename. It is suggested that you use the file extension .map. Enclose the physical name in single or double quotation marks.

For example, the following statements import an XML document named My.XML and specify the XMLMap named My.MAP, which contains specific XMLMap syntax. The engine interprets the XML document as a SAS data set named Test.My. In this example, XMLMAP= is used as an option in the LIBNAME statement:

```
libname test xmlv2 'C:\Example\Nhl.xml' xmlmap='C:\Example\Nhl.map';
proc print data=test.teams;
run;
```
Restriction  See the XMLMap syntax on page 95 for the version numbers that are supported for import and export by the XMLV2 engine, and for import only by the XML engine.

Interactions  If you specify an XMLMap, you must either specify XMLTYPE=XMLMAP or do not specify a markup type. If you explicitly specify a markup type other than XMLMAP (such as XMLTYPE=GENERIC), an error occurs.

If you specify the AUTOMAP= option, you must specify the XMLMAP= option.

See  Chapter 9, “Introduction to XMLMap Syntax,” on page 91

Examples  Chapter 4, “Importing XML Documents Using an XMLMap,” on page 19

Chapter 5, “Exporting XML Documents Using an XMLMap,” on page 55

**XMLMETA=DATA | SCHEMADATA | SCHEMA**

XMLV2 and XML.

specifies whether to include metadata-related information in the exported markup, or specifies whether to import metadata-related information that is included in the input XML document. Metadata-related information is metadata that describes the characteristics (types, lengths, levels, and so on) of columns within the table markup. Including the metadata-related information can be useful when exporting an XML document from a SAS data set to process on an external product.

**DATA**

ignores metadata-related information. DATA includes only data content in the exported markup and imports only data content in the input XML document.

**SCHEMADATA**

includes both data content and metadata-related information in the exported markup and imports both data content and metadata-related information in the input XML document.

**SCHEMA**

ignores data content. SCHEMA includes only metadata-related information in the exported markup and imports only metadata-related information in the input XML document.

Default  DATA

Restriction  Use this option for the GENERIC and MSACCESS markup types only.

Interaction  If XMLMETA=SCHEMADATA and XMLSCHEMA= is specified, the data is written to the physical location of the XML document specified in the LIBNAME statement. Separate metadata-related information is written to the physical location specified with XMLSCHEMA=. If XMLSCHEMA= is not specified, the metadata-related information is embedded with the data content in the XML document.

Tip  Prior to SAS 9, the functionality for the XMLMETA= option used the keyword XMLSCHEMA=. SAS 9 changed the option keyword XMLSCHEMA= to XMLMETA=. SAS 9.1 added new functionality using the XMLSCHEMA= option.
XMLPROCESS=CONFORM | PERMIT XMLV2 and XML

determines how the engine processes character data that does not conform to W3C specifications.

CONFORM

requires that the XML conform to W3C specifications. W3C specifications state that for character data, certain characters such as the left angle bracket (<), the ampersand (&), and the apostrophe (') must be escaped using character references or strings like &amp; . For example, to allow attribute values to contain both single and double quotation marks, the apostrophe or single quotation mark character (') can be represented as &apos; and the double quotation mark character (") can be represented as &quot;.

PERMIT

permits character data that does not conform to W3C specifications to be accepted. That is, in character data, non-escaped characters such as the apostrophe, double quotation marks, and the ampersand are accepted.

Restrictions

Non-escaped angle brackets in character data are not accepted.

Use XMLPROCESS=PERMIT cautiously. If an XML document consists of non-escaped characters, the content is not standard XML construction. The option is provided for convenience, not to encourage invalid XML markup.

Default

CONFORM

Example

“Importing an XML Document with Non-Escaped Character Data” on page 11

XMLSCHEMA=fileref | 'external-file' XMLV2 and XML

specifies an external file to contain metadata-related information.

fileref

is the SAS name that is associated with the physical location of the output file. To assign a fileref, use the FILENAME statement.

'external-file'

is the physical location of the file to contain the metadata-related information. Include the complete pathname and the filename. Enclose the physical name in single or double quotation marks.

Restrictions

Use this option when exporting an XML document only.

Use this option only for the GENERIC and MSACCESS markup types with XMLMETA=SCHEMADATA.

Interaction

If XMLMETA=SCHEMADATA and XMLSCHEMA= is specified, the data is written to the physical location of the XML document specified in the LIBNAME statement. Separate metadata-related information is written to the physical location specified with XMLSCHEMA= . If XMLSCHEMA= is not specified, the metadata-
related information is embedded with the data content in the XML document.

<table>
<thead>
<tr>
<th>Example</th>
<th>“Exporting an XML Document with Separate Metadata” on page 16</th>
</tr>
</thead>
</table>

**XMLTYPE=GENERIC | CDISCODM | MSACCESS | ORACLE | XMLMAP**

specifies the XML markup type.

**GENERIC** XMLV2 and XML

is a simple, well-formed XML markup type. The XML document consists of a root (enclosing) element and repeating element instances. GENERIC determines a variable's attributes from the data content.

**Requirement** When importing, the GENERIC markup type requires a specific physical structure.

**See** “Understanding the Required Physical Structure for an XML Document to Be Imported Using the GENERIC Markup Type” on page 20

**Examples**

| “Exporting an XML Document That Contains SAS Dates, Times, and Datetimes” on page 15 |
| “Exporting Numeric Values” on page 121 |
| “Importing an XML Document Using the GENERIC Markup Type” on page 9 |

**CDISCODM** XML Only

is the XML markup type for the markup standards that are defined in the Operational Data Model (ODM) that was created by the Clinical Data Interchange Standards Consortium (CDISC). The XML engine supports the ODM 1.2 schema specification. ODM supports the electronic acquisition, exchange, and archiving of clinical trials data and metadata for medical and biopharmaceutical product development.

**Tip** Use the FORMATACTIVE=, FORMATNOREPLACE=, and FORMATLIBRARY= options to specify how display data are read and stored in the target environment.

**Examples**

| “Importing a CDISC ODM Document” on page 134 |
| “Exporting an XML Document in CDISC ODM Markup” on page 126 |

**MSACCESS** XML Only

is the XML markup type for the markup standards supported for a Microsoft Access database (.mdb). If the Microsoft Access file contains metadata-related information, then you must specify MSACCESS rather than the default GENERIC markup type. If there is an embedded XML schema, specifying MSACCESS and the XMLMETA=SCHEMADATA option causes a variable's attributes to be obtained from the embedded schema. If there is not an embedded schema, MSACCESS uses default values for attributes.

| Example | “Importing an XML Document Created by Microsoft Access” on page 128 |
**ORACLE** XML Only

is the XML markup type for the markup standards equivalent to the Oracle 8i XML implementation. The number of columns to indent each nested element is one, and the enclosing element tag for the contents of the SAS data set is ROWSET.

Example  “Exporting an XML Document for Use by Oracle” on page 119

**XMLMAP** XMLV2 Only

specifies that XML markup is determined by an XMLMap, which is an XML document that you create that contains specific XMLMap syntax.

The XMLMap syntax tells the engine how to map the SAS data into a specific XML document structure. To specify the XMLMap in the LIBNAME statement, use the XMLMAP= option on page 83.

Restriction  Exporting an XML document that is controlled by an XMLMap is limited to a single SAS data set.

Example  “Using an XMLMap to Export an XML Document with a Hierarchical Structure” on page 55

**Default** GENERIC

**Tip** You can control the markup by specifying options such as INDENT=, XMLDATAFORM=, XMLMETA= (when applicable), and TAGSET=.
Part 3

XMLMap File Syntax

Chapter 9
Introduction to XMLMap Syntax .................................................. 91

Chapter 10
XMLMap Syntax Version 2.1 .......................................................... 95

Chapter 11
Using SAS XML Mapper to Generate and Update an XMLMap .... 113
Chapter 9
Introduction to XMLMap Syntax

Using XMLMap Syntax

The XML elements for the XMLMap syntax for version 2.1 are explained in this chapter. The elements are listed in the order in which you would typically include them in an XMLMap. That is:

- The first element in the XMLMap is the SXLEMAP element, which is the primary (root) enclosing element that contains the definition for the generated output file. See “SXLEMAP Element” on page 95.
- The namespace elements define XML namespaces, which distinguish element and attribute names by qualifying them with Uniform Resource Identifier (URIs). See “Elements for XML Namespaces” on page 96.
- If you use an XMLMap for exporting, you must include the exporting elements. See “Elements forExporting” on page 98.
- The table elements define the SAS data set. See “Elements for Tables” on page 99.
- The column elements define the variables for the SAS data set. See “Elements for Columns” on page 102.

CAUTION:

The XMLMap markup, as XML itself, is case sensitive. The tag names must be uppercase, and the element attributes must be lowercase. Here is an example: <SXLEMAP version="2.1">. In addition, the supported XPath syntax is case sensitive as well.

Comparing the XMLMap Syntax

The following table lists the available XMLMap syntax. The ■ symbol indicates whether the syntax is available for importing or exporting, and whether the syntax is available for each engine.
### Table 9.1 XMLMap Syntax

<table>
<thead>
<tr>
<th>Syntax</th>
<th>Description</th>
<th>Import</th>
<th>Export</th>
<th>XML</th>
<th>XMLV2</th>
</tr>
</thead>
<tbody>
<tr>
<td>SXLEMAP on page 95</td>
<td>Primary (root) enclosing element</td>
<td>■</td>
<td>■</td>
<td>■</td>
<td>■</td>
</tr>
<tr>
<td>NAMESPACES on page 96</td>
<td>Contains elements for defining XML namespaces</td>
<td>■</td>
<td>■</td>
<td></td>
<td>■</td>
</tr>
<tr>
<td>NS on page 97</td>
<td>Defines an XML namespace by referencing a unique URI</td>
<td>■</td>
<td>■</td>
<td></td>
<td>■</td>
</tr>
<tr>
<td>OUTPUT on page 98</td>
<td>Contains elements for exporting a SAS data set as an XML document</td>
<td>■</td>
<td>■</td>
<td></td>
<td>■</td>
</tr>
<tr>
<td>HEADING on page 98</td>
<td>Contains ATTRIBUTE elements for exporting</td>
<td>■</td>
<td>■</td>
<td></td>
<td>■</td>
</tr>
<tr>
<td>ATTRIBUTE on page 98</td>
<td>Contains file attribute information for exporting</td>
<td>■</td>
<td>■</td>
<td></td>
<td>■</td>
</tr>
<tr>
<td>TABLEREF on page 98</td>
<td>Specifies the name of the table for exporting</td>
<td>■</td>
<td>■</td>
<td></td>
<td>■</td>
</tr>
<tr>
<td>TABLE on page 99</td>
<td>Contains a data set definition</td>
<td>■</td>
<td>■</td>
<td>■</td>
<td>■</td>
</tr>
<tr>
<td>TABLE-PATH on page 99</td>
<td>Specifies a location path for variables</td>
<td>■</td>
<td>■</td>
<td>■</td>
<td>■</td>
</tr>
<tr>
<td>TABLE-END-PATH on page 100</td>
<td>Specifies a location path to stop processing</td>
<td>■</td>
<td>■</td>
<td>■</td>
<td>■</td>
</tr>
<tr>
<td>TABLE-DESCRIPTION on page 102</td>
<td>Specifies a SAS data set description</td>
<td>■</td>
<td>■</td>
<td>■</td>
<td>■</td>
</tr>
<tr>
<td>COLUMN name= on page 103</td>
<td>Specifies the variable name</td>
<td>■</td>
<td>■</td>
<td>■</td>
<td>■</td>
</tr>
<tr>
<td>COLUMN retain= on page 103</td>
<td>Determines the contents of the input buffer</td>
<td>■</td>
<td>■</td>
<td>■</td>
<td>■</td>
</tr>
<tr>
<td>COLUMN class= on page 103</td>
<td>Determines the type of variable</td>
<td>■</td>
<td>■</td>
<td>■</td>
<td>■</td>
</tr>
<tr>
<td>TYPE on page 104</td>
<td>Specifies the SAS data type for the variable</td>
<td>■</td>
<td>■</td>
<td>■</td>
<td>■</td>
</tr>
<tr>
<td>DATATYPE on page 104</td>
<td>Specifies the type of data being read</td>
<td>■</td>
<td>■</td>
<td>■</td>
<td>■</td>
</tr>
<tr>
<td>DEFAULT on page 105</td>
<td>Specifies a default value for a missing value</td>
<td>■</td>
<td>■</td>
<td>■</td>
<td>■</td>
</tr>
<tr>
<td>Syntax</td>
<td>Description</td>
<td>Import</td>
<td>Export</td>
<td>XML</td>
<td>XMLV2</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------</td>
<td>--------</td>
<td>-----</td>
<td>-------</td>
</tr>
<tr>
<td>ENUM on page 105</td>
<td>Contains a list of valid values for the variable</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>FORMAT on page 106</td>
<td>Specifies a SAS format for the variable</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>INFORMAT on page 106</td>
<td>Specifies a SAS informat for the variable</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>DESCRIPTION on page 106</td>
<td>Specifies a description for the variable</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>LENGTH on page 106</td>
<td>Determines the maximum field storage length for a character variable</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>PATH on page 107</td>
<td>Specifies a location path for the current variable</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>INCREMENT-PATH on page 109</td>
<td>Specifies a location path for incrementing the accumulated value for a counter variable</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>RESET-PATH on page 110</td>
<td>Specifies a location path for resetting the accumulated value for a counter variable to zero</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>DECREMENT-PATH on page 111</td>
<td>Specifies a location path to decrement the accumulated value for the counter variable by 1</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>
Chapter 10
XMLMap Syntax Version 2.1

Dictionary

SXLEMAP Element

Is the primary (root) enclosing element that contains the definition for the generated output file. The element provides the XML well-formed constraint for the definition.

Restriction: When importing an XML document, the definition can define more than one output SAS data set. When exporting an XML document from a SAS data set, the definition can define only one output XML document.

Requirement: The SXLEMAP element is required.

Syntax

SXLEMAP version="number" name="XMLMap" description="description"

Attributes

version="number"

specifies the version of the XMLMap syntax. The documented XMLMap syntax version is 2.1 and must be specified to obtain full functionality.

Default The default version is the first version of XMLMap syntax. It is retained for compatibility with prior releases of the XMLMap syntax. It is recommended that you update existing XMLMaps to version 2.1.

Restrictions For import, the XMLV2 engine supports XMLMap syntax versions 1.2, 1.9, and 2.1. For export, the XMLV2 engine supports version 2.1, and the XML92 engine (alias for XMLV2) supports version 1.9.
For import, the XML engine supports XMLMap syntax versions 1.0, 1.1, and 1.2. The XML engine does not support using an XMLMap for export.

Tip
To update an XMLMap to version 2.1, load the existing XMLMap into SAS 9.3 or later XML Mapper, and then save the XMLMap. For information about SAS XML Mapper, see Chapter 11, “Using SAS XML Mapper to Generate and Update an XMLMap,” on page 113.

name="XMLMap"
  is an optional attribute that specifies the filename of the XMLMap.

description="description"
  is an optional attribute that specifies a description of the XMLMap. The description can be up to 256 characters.

Details
In the example below, the SXLEMAP element specifies all three attributes and contains two TABLE elements.

```xml
<?xml version="1.0" ?>
<SXLEMAP version="2.1" name="Myxmlmap" description="sample XMLMap">
  <TABLE name="test1">
    .
    .
    .
  </TABLE>
  <TABLE name="test2">
    .
    .
    .
  </TABLE>
</SXLEMAP>
```

**Elements for XML Namespaces**
Define XML namespaces.

**Syntax**

```xml
<NAMESPACE count="number">
  <NS id="number" prefix="name">
```

**Elements**

```xml
NAMESPACE count="number" XMLV2 Only
  is an optional element that contains one or more NS elements for defining XML namespaces. Here is an example: <NAMESPACE count="2">

XMLMap namespace elements enable you to import an XML document with like-named elements that are qualified with XML namespaces. In addition, XMLMap
namespace elements maintain XML namespaces from the imported XML document to export an XML document from the SAS data set.

An XML namespace is a W3C specification that distinguishes element and attribute names by qualifying them with Uniform Resource Identifiers (URIs). For example, if an XML document contains a CUSTOMER element and a PRODUCT element, and both elements contain a nested ID element, XML namespaces make each nested ID element unique.

count="number"

specifies the number of defined XML namespaces.

Requirements

The count= attribute is required. The specified value must match the total number of NS elements.

Example

“Including Namespace Elements in an XMLMap” on page 47

NS id="number" <prefix="name"> XMLV2 Only

is an optional element that defines an XML namespace by referencing a unique URI. The URI is a string of characters that identifies a resource on the internet. The URI is treated by an XML parser as simply a string. Specifying the URI does not require that it be used to retrieve information. The most common URI is the Uniform Resource Locator (URL), which identifies an internet domain address. The use of URIs as namespace names must follow the same rules as the W3C specification for XML namespaces. Here is an example: <NS id="1" prefix="freq"> http://www.hurricanefrequency.com </NS>.

Note: It is recommended that you do not use non-escaped characters in a URI.

id="number"

specifies an identification number for the XML namespace.

Requirements

In the variable definition, the identification number must be included in the location path preceding the element that is being defined. See “PATH syntax=type” XMLV2 and XML” on page 107.

prefix="name"

specifies a qualified name that is associated with the referenced URI. The prefix is used with each element or attribute to indicate to which XML namespace it belongs. Prefix names must follow the same rules as the W3C specification for element names.

Requirements

The referenced URI must be unique.

The total number of NS elements must match the specified value in the NAMESPACES count= attribute.

Tip

It is recommended that you do not use non-escaped characters in a URI.

Example

“Including Namespace Elements in an XMLMap” on page 47
Elements for Exporting

Export an XML document from a SAS data set by using the XMLMap that was created to import the XML document.

**Restriction:** The engine supports exporting from one SAS data set only.

### Syntax

```
OUTPUT
HEADING
ATTRIBUTE name="name" value="value"
TABLEREF name="name"
```

### Elements

**OUTPUT** XMLV2 Only  
is an optional element that contains zero or more HEADING elements and one TABLEREF element for exporting a SAS data set as an XML document.

**Requirement**  
If you specify version 1.9 or 2.1 in an XMLMap to export a SAS data set as an XML document, you must include the OUTPUT element in the XMLMap.

**Example**  
“Using an XMLMap to Export an XML Document with a Hierarchical Structure” on page 55

**HEADING** XMLV2 Only  
is an optional element that contains zero or more ATTRIBUTE elements.

**ATTRIBUTE** name="attribute-name" value="attribute-value" XMLV2 Only  
is an optional element that contains additional file attribute information for the exported XML document, such as a schema reference or other general attributes. The specified name-value pairs are added as attributes to the first generated element in the exported XML document. Here is an example: `<NHL description="Teams of the National Hockey League">`.

- **name="attribute-name"**  
  specifies a name for a file attribute. In the example code above, it is `name="description"`.

- **value="attribute-value"**  
  specifies a value for the attribute. In the example code above, it is `value="Teams of the National Hockey League"`.

**TABLEREF** name="table-name" XMLV2 Only  
specifies the name of the table in the XMLMap to be exported.

- **name="table-name"**  
  is a valid SAS name that is unique in the XMLMap definition.

**Requirements**  
If you specify the OUTPUT element, you must specify one (and no more than one) TABLEREF element.

The specified name must match a TABLE element name= attribute.
Elements for Tables

Define the SAS data set.

Syntax

TABLE description="description" name="data-set-name"
TABLE-PATH syntax="type"
TABLE-END-PATH syntax="type" beginend="BEGIN | END"
TABLE-DESCRIPTION

Elements

TABLE description="description" name="data-set-name" XMLV2 and XML

is an element that contains a data set definition. Here is an example: <TABLE name="channel">

description="description"

is an optional attribute that specifies a description of the SAS data set. The description can be up to 256 characters.

name="data-set-name"

specifies the name for the SAS data set. The name must be unique in the XMLMap, and the name must be a valid SAS name, which can be up to 32 characters.

Requirement The name= attribute is required.

Requirement The TABLE element is required.

Interaction The TABLE element can contain one or more of the following elements: TABLE-PATH, TABLE-END-PATH, TABLE-DESCRIPTION, and COLUMN.

TABLE-PATH syntax="type" XMLV2 and XML

specifies a location path that tells the XML engine where in the XML document to locate and access specific elements in order to collect variables for the SAS data set. The location path defines the repeating element instances in the XML document, which is the SAS data set observation boundary. The observation boundary is translated into a collection of rows with a constant set of columns.

For example, in “Using an XMLMap to Import an XML Document as Multiple SAS Data Sets” on page 26, consider this TABLE-PATH element in RSS.map for the Item data set:

<TABLE-PATH syntax="XPath"> /rss/channel/item </TABLE-PATH>

The TABLE-PATH causes the following to occur

1. The XML engine reads the XML markup until it encounters the <ITEM> start tag.

2. The XML engine clears the input buffer, sets the contents to MISSING (by default), and scans elements for variable names based on the COLUMN element definitions. As values are encountered, they are read into the input buffer. (Note
that whether the XML engine resets to MISSING is determined by the
DEFAULT element as well as the COLUMN element retain= attribute.)

3. When the </ITEM> end tag is encountered, the XML engine writes the
completed input buffer to the SAS data set as a SAS observation.

4. The process is repeated for each <ITEM> start-tag and </ITEM> end-tag
sequence until the end-of-file is encountered in the input stream or until the
TABLE-END-PATH (if specified) is achieved, which results in six observations.

**syntax=\"type\"**
is an optional attribute that specifies the type of syntax in the location path. The
syntax is valid XPath construction in compliance with the W3C specifications.
Here is an example: **syntax=\"XPath\"**.

**Default**
XPath

**Requirements**
The value must be XPath or XPathENR.

If an XML namespace is defined with the NAMESPACES
element, you must specify the type of syntax as XPathENR
(XPath with Embedded Namespace Reference). This is because
the syntax is different from the XPath specification. Here is an
example: **syntax=\"XPathENR\"**.

**CAUTION:**
Specifying the table location path, which is the observation boundary, can
be complicated due to start-tag and end-tag pairing. The table location path
determines which end tag causes the XML engine to write the completed input
buffer to the SAS data set. If you do not identify the appropriate end tag, the
result could be concatenated data instead of separate observations, or an
unexpected set of columns. See “Determining the Observation Boundary to
Avoid Concatenated Data” on page 36.

**Requirements**
The TABLE-PATH element is required.

If an XML namespace is defined with the NAMESPACES element,
you must include the identification number in the location path
preceding the element that is being defined. The identification
number is enclosed in braces. Here is an example: **<TABLE-
PATH syntax=\"XPathENR\">/Table/{1}Hurricane</TABLE-
PATH>**.

The XPath construction is a formal specification that puts a path
description similar to UNIX on each element of the XML structure.
Note that XPath syntax is case sensitive. For example, if an element
tag name is uppercase, it must be uppercase in the location path. If it
is lowercase, it must be lowercase. All location paths must begin
with the root-enclosing element (denoted by a slash '/') or with the
"any parent" variant (denoted by double slashes ///). Other W3C
documented forms are not currently supported.

**TABLE-END-PATH syntax=\"type\" beginend=\"BEGIN | END\"** XMLV2 and XML
is an optional, optimization element that saves resources by stopping the processing
of the XML document before the end of the file. The location path tells the XML
element where in the XML document to locate and access a specific element in order
to stop processing the XML document.
For example, in “Using an XMLMap to Import an XML Document as Multiple SAS Data Sets” on page 26, here is the TABLE-PATH location path in RSS.map for the Channel data set:

```xml
<TABLE-PATH syntax="XPath"> /rss/channel </TABLE-PATH>
```

The XML document RSS.xml has only one `<CHANNEL>` start tag and one `</CHANNEL>` end tag, so processing could conclude after the `</CHANNEL>` end tag. The engine does not store new data in the input buffer after it encounters the first `<ITEM>` start tag because the remaining elements no longer qualify. However, the engine continues to process the entire XML document. To improve efficiency for a long XML document, you could add a TABLE-END-PATH element. For example, if you add the following specification for the Channel data set, the XML engine stops processing when the first `<ITEM>` start tag is encountered:

```xml
<TABLE-END-PATH syntax="XPath" beginend="BEGIN"> /rss/channel/item </TABLE-END-PATH>
```

With both the TABLE-PATH and TABLE-END-PATH specifications, the XML engine processes only the highlighted data in the RSS.xml document for the Channel data set rather than the entire XML document.

```xml
<?xml version="1.0" encoding="UTF-8" ?>
<rss version="2.0">
  <channel>
    <title>My RSS Channel</title>
    <description>This is a simplified example of an RSS feed.</description>
    <link>http://www.example.com/main.html</link>
    <language>en-us</language>
    <item>
      <title>My news item</title>
      <description>This is a detailed summary of my news item.</description>
      <link>http://www.example.com/blog/post/1</link>
    </item>
    <item>
      <title>Another news item</title>
      <description>This description is shorter.</description>
      <link>http://www.example.com/blog/post/2</link>
    </item>
  </channel>
</rss>
```

The `syntax="type"` is an optional attribute that specifies the type of syntax in the location path. The syntax is valid XPath construction in compliance with the W3C specifications. The XPath form supported by the XML engine allows elements and attributes to be individually selected for exclusion in the generated SAS data set. Here is an example: `syntax="XPath"`.

<table>
<thead>
<tr>
<th>Default</th>
<th>XPath</th>
</tr>
</thead>
<tbody>
<tr>
<td>Requirements</td>
<td>The value must be XPath or XPathENR.</td>
</tr>
</tbody>
</table>

If an XML namespace is defined with the NAMESPACES element, you must specify the type of syntax as XPathENR (XPath with Embedded Namespace Reference). This is because the syntax is different from the XPath specification. Here is an example: `syntax="XPathENR"`. 
**Beginend**="BEGIN | END"

is an optional attribute that specifies to stop processing when either the element start tag is encountered or the element end tag is encountered.

**Default** BEGIN

**Default**

Processing continues until the last end tag in the XML document.

**Requirements**

If an XML namespace is defined with the NAMESPAACES element, you must include the identification number in the location path preceding the element that is being defined. The identification number is enclosed in braces. Here is an example:

```xml
<TABLE-END-PATH syntax="XPathENR">/Table/{1}Hurricane</TABLE-END-PATH>.
```

The XPath construction is a formal specification that puts a path description similar to UNIX on each element of the XML structure. Note that XPath syntax is case sensitive. For example, if an element tag name is uppercase, it must be uppercase in the location path. If it is lowercase, it must be lowercase. All location paths must begin with the root-enclosing element (denoted by a slash '/') or with the "any parent" variant (denoted by double slashes '///'). Other W3C documented forms are not currently supported.

**Interaction**

The TABLE-END-PATH element does not affect the observation boundary; that is determined with the TABLE-PATH element.

**Tip**

Specifying a location to stop processing is useful for an XML document that is hierarchical, but generally not appropriate for repeating instance data.

**Example**

“Using an XMLMap to Import an XML Document as Multiple SAS Data Sets” on page 26

**TABLE-DESCRIPTION**

XMLV2 and XML

is an optional element that specifies a description for the SAS data set, which can be up to 256 characters. Here is an example: `<TABLE-DESCRIPTION> Data Set contains TV channel information </TABLE-DESCRIPTION>.

---

**Elements for Columns**

Define the variables for the SAS data set.

**Syntax**

COLUMN name="name" retain="NO | YES" class="ORDINAL"

TYPE

DATATYPE

DEFAULT

ENUM

FORMAT width="w" ndec="d"

INFORMAT width="w" ndec="d"
DESCRIPTION
LENGTH
PATH syntax="type"
INCREMENT-PATH syntax="type" beginend="BEGIN | END"
RESET-PATH syntax="type" beginend="BEGIN | END"
DECREMENT-PATH syntax="type" beginend="BEGIN | END"

Elements

COLUMN name="name" retain="NO | YES" class="ORDINAL"

is an element that contains a variable definition. Here is an example: <COLUMN name="Title">

name="name" XMLV2 and XML

specifies the name for the variable. The name must be a valid SAS name, which can be up to 32 characters.

Requirement  The name= attribute is required.

retain="NO | YES" XMLV2 and XML

is an optional attribute that determines the contents of the input buffer at the beginning of each observation.

NO  sets the value for the beginning of each observation either to MISSING or to the value of the DEFAULT element if specified.

YES  keeps the current value until it is replaced by a new, nonmissing value. Specifying YES is much like the RETAIN statement in DATA step processing. It forces the retention of processed values after an observation is written to the output SAS data set.

Default  NO

Example  “Importing Hierarchical Data as Related Data Sets” on page 29

class="ORDINAL" XMLV2 Only

is an optional attribute that determines the type of variable.

ORDINAL

specifies that the variable is a numeric counter variable that keeps track of the number of times the location path, which is specified by the INCREMENT-PATH element or the DECREMENT-PATH element, is encountered. (This is similar to the _N_ automatic variable in DATA step processing.) The counter variable increments or decrements its count by 1 each time the location path is encountered. Counter variables can be useful for identifying individual occurrences of like-named data elements or for counting observations.

Restriction  When exporting an XML document, variables with class="ORDINAL" are not included in the output XML document.

Requirements  You must use the INCREMENT-PATH element or the DECREMENT-PATH element. The PATH element is not allowed.
The TYPE element must specify the SAS data type as numeric, and the DATATYPE element must specify the type of data as integer.

Example

“Including a Key Field with Generated Numeric Keys” on page 33

Requirement

At least one COLUMN element is required.

Interaction

COLUMN can contain one or more of the following elements that describe the variable attributes: DATATYPE, DEFAULT, ENUM, FORMAT, INFORMAT, DESCRIPTION, LENGTH, TYPE, PATH, INCREMENT-PATH, DECREMENT-PATH, and RESET-PATH.

**TYPE** XMLV2 and XML

specifies the SAS data type (character or numeric) for the variable, which is how SAS stores the data. For example, `<TYPE> numeric </TYPE>` specifies that the SAS data type for the variable is numeric.

Requirement

The TYPE element is required.

Tips

To assign a floating-point type, use

```
<TYPE> float </TYPE>
```

To apply output formatting in SAS, use the FORMAT element.

To control output data type conversion in input, use the INFORMAT element. Here is an example: `<INFORMAT> datetime </INFORMAT>`.

**DATATYPE** XMLV2 and XML

specifies the type of data being read from the XML document for the variable. For example, `<DATATYPE> string </DATATYPE>` specifies that the data contains alphanumeric characters.

The type of data specification can be

- **string**
  
  specifies that the data contains alphanumeric characters and does not contain numbers used for calculations.

- **integer**
  
  specifies that the data contains whole numbers used for calculations.

- **double**
  
  specifies that the data contains floating-point numbers.

- **datetime**
  
  specifies that the input represents a valid datetime value, which is either
  
  - in the form of the XML specification ISO 8601 format. The default form is: `yyyy-mm-ddThh:mm:ss.ffffff`.
  
  - in a form for which a SAS informat (either supplied by SAS or user-written) properly translates the input into a valid SAS datetime value. See also the INFORMAT element on page 106.
date
specifies that the input represents a valid date value, which is either
• in the form of the XML specification ISO 8601 format. The default form is: yyyy-mm-dd.
• in a form for which a SAS informat (either supplied by SAS or user-written) properly translates the input into a valid SAS date value. See also the INFORMAT element on page 106.

time
specifies that the input represents a valid time value, which is either
• in the form of the XML specification ISO 8601 format. The default form is: hh:mm:ss.ffffff.
• in a form for which a SAS informat (either supplied by SAS or user-written) properly translates the input into a valid SAS date value. See also the INFORMAT element on page 106.

Restriction The values for previous versions of XMLMap syntax are not accepted by versions 1.9 and 2.1.

Requirement The DATATYPE element is required.

DEFAULT XMLV2 and XML
is an optional element that specifies a default value for a missing value for the variable. Use the DEFAULT element to assign a nonmissing value to missing data. For example, <DEFAULT> single </DEFAULT> assigns the value single when a missing value occurs.

Default By default, the XML engine sets a missing value to MISSING.

ENUM XMLV2 and XML
is an optional element that contains a list of valid values for the variable. The ENUM element can contain one or more VALUE elements to list the values. By using ENUM, values in the XML document are verified against the list of values. If a value is not valid, it is either set to MISSING (by default) or set to the value specified by the DEFAULT element. Note that a value specified for DEFAULT must be one of the ENUM values in order to be valid.

    <COLUMN name="filing_status">
        ...
        ...
        ...
        <DEFAULT> single </DEFAULT>
        ...
        ...
        ...
        <ENUM>
            <VALUE> single </VALUE>
            <VALUE> married filing joint return </VALUE>
            <VALUE> married filing separate return </VALUE>
            <VALUE> head of household </VALUE>
            <VALUE> qualifying widow(er) </VALUE>
        </ENUM>
    </COLUMN>
FORMAT width="w" ndec="d" XMLV2 and XML

is an optional element that specifies a SAS format for the variable. A format name can be up to 31 characters for a character format and 32 characters for a numeric format. A SAS format is an instruction that SAS uses to write values. You use formats to control the written appearance of values. Do not include a period (.) as part of the format name. Specify a width and length as attributes, not as part of the format name.

For a list of the SAS formats, including the ISO 8601 SAS formats, see SAS Formats and Informs: Reference.

width="w"

is an optional attribute that specifies a format width, which for most formats is the number of columns in the output data.

ndec="d"

is an optional attribute that specifies a decimal scaling factor for numeric formats.

Here is an example:

```xml
<FORMAT> E8601DA </FORMAT>
<FORMAT width="8"> best </FORMAT>
<FORMAT width="8" ndec="2"> dollar </FORMAT>
```

INFORMAT width="w" ndec="d" XMLV2 and XML

is an optional element that specifies a SAS informat for the variable. An informat name can be up to 30 characters for a character informat and 31 characters for a numeric informat. A SAS informat is an instruction that SAS uses to read values into a variable (that is, to store the values). Do not include a period (.) as part of the informat name. Specify a width and length as attributes, not as part of the informat name.

For a list of the SAS informats, including the ISO 8601 SAS informats, see SAS Formats and Informs: Reference.

Here is an example:

```xml
<INFORMAT> E8601DA </INFORMAT>
<INFORMAT width="8"> best </INFORMAT>
<INFORMAT width="8" ndec="2"> dollar </INFORMAT>
```

width="w"

is an optional attribute that specifies an informat width, which for most informats is the number of columns in the input data.

ndec="d"

is an optional attribute that specifies a decimal scaling factor for numeric informats. SAS divides the input data by 10 to the power of this value.

DESCRIPTION XMLV2 and XML

is an optional element that specifies a description for the variable, which can be up to 256 characters. The following example shows that the description is assigned as the variable label.

```xml
<DESCRIPTION> Story link </DESCRIPTION>
```

LENGTH XMLV2 and XML

is the maximum field storage length from the XML data for a character variable. The value refers to the number of characters used to store each of the variable's values in the SAS data set. The value can be 1 to 32,767. During the input process, a
maximum length of characters is read from the XML document and transferred to the observation buffer. Here is an example: `<LENGTH> 200 </LENGTH>.

**Restriction**  
LENGTH is not valid for numeric data.

**Requirement**  
For data that is defined as a STRING data type, the LENGTH element is required.

**Tips**  
You can use LENGTH to truncate a long field. Multi-byte character strings that are longer than the specified length are truncated on a character boundary, not on a byte boundary.

When importing a table with multi-byte character set (MBCS) data, you might need to expand the number of bytes specified in the LENGTH element. Use the CHARMULTIPLIER= or DERIVECHARMULTIPLIER= LIBNAME statement options to ensure that the SAS data set column accommodates the larger character size.

**PATH syntax=**

XMLV2 and XML

specifies a location path that tells the XML engine where in the XML document to locate and access a specific tag for the current variable. In addition, the location path tells the XML engine to perform a function, which is determined by the location path form, to retrieve the value for the variable. The XPath forms that are supported allow elements and attributes to be individually included in the generated SAS data set.

**syntax=**

is an attribute that specifies the type of syntax used in the location path. The syntax is valid XPath construction in compliance with the W3C specifications. The XPath form supported by the XML engine allows elements and attributes to be individually included in the generated SAS data set.

**Default**  
XPath

**Requirements**  
The value must be XPath or XPathENR.

If an XML namespace is defined with the NAMESPACES element, you must specify the type of syntax as XPathENR (XPath with Embedded Namespace Reference). This is because the syntax is different from the XPath specification. Here is an example: `syntax="XPathENR"`.

To specify the PATH location path, use one of the following forms:

**CAUTION:**

These forms are the only XPath forms that the XML engine supports. If you use any other valid W3C form, the results will be unpredictable.

**element-form**

selects PCDATA (parsed character data) from a named element. The following element forms enable you to select from a named element, conditionally select from a named element based on a specific attribute value, or conditionally select from a named element based on a specific occurrence of the element using the position function:

```
<PATH> /LEVEL/ITEM </PATH>
<PATH> /LEVEL/ITEM[@attr="value"] </PATH>
<PATH> /LEVEL/ITEM[position()=n] | [n] </PATH>
```
The following examples illustrate the element forms. For more information about the examples, see “Specifying a Location Path on the PATH Element” on page 44.

- The following location path tells the XML engine to scan the XML markup until it finds the CONFERENCE element. The XML engine retrieves the value between the <CONFERENCE> start tag and the </CONFERENCE> end tag.

```
<Path> /NHL/CONFERENCE </Path>
```

- The following location path tells the XML engine to scan the XML markup until it finds the TEAM element where the value of the founded= attribute is 1993. The XML engine retrieves the value between the <TEAM> start tag and the </TEAM> end tag.

```
<Path> /NHL/CONFERENCE/DIVISION/TEAM[@founded="1993"] </Path>
```

- The following location path uses the position function to tell the XML engine to scan the XML markup until it finds the fifth occurrence of the TEAM element. The XML engine retrieves the value between the <TEAM> start tag and the </TEAM> end tag.

```
<Path> /NHL/CONFERENCE/DIVISION/TEAM[position()=5] </Path>
```

You can use the following shorter version for the position function:

```
<Path> /NHL/CONFERENCE/DIVISION/TEAM[5] </Path>
```

### attribute-form

selects values from an attribute. The following attribute forms enable you to select from a specific attribute or conditionally select from a specific attribute based on the value of another attribute:

```
<Path> /LEVEL/ITEM/@attr </Path>
<Path> /LEVEL/ITEM/@attr[@attr2="value"] </Path>
```

The following examples illustrate the attribute forms. For more information about the examples, see “Specifying a Location Path on the PATH Element” on page 44.

- The following location path tells the XML engine to scan the XML markup until it finds the TEAM element. The XML engine retrieves the value from the abbrev= attribute.

```
<Path syntax="XPath"> /NHL/CONFERENCE/DIVISION/TEAM/@abbrev </Path>
```

- The following location path tells the XML engine to scan the XML markup until it finds the TEAM element. The XML engine retrieves the value from the founded= attribute where the value of the abbrev= attribute is ATL. The two attributes must be for the same element.

```
<Path> /NHL/CONFERENCE/DIVISION/TEAM[@founded="ATL"] </Path>
```

### Requirements

Whether the PATH element is required or allowed is determined by the class="ORDINAL" attribute for the COLUMN element. If the class="ORDINAL" attribute is not specified, which is the default, PATH is required and INCREMENT-PATH, DECREMENT-PATH, and RESET-PATH are not allowed. If the class="ORDINAL" attribute is specified, PATH is not allowed, INCREMENT-PATH or DECREMENT-PATH is required, and RESET-PATH is optional.

If an XML namespace is defined with the NAMESPACES element, you must include the identification number in the location path.
The XPath construction is a formal specification that puts a path description similar to UNIX on each element of the XML structure. XPath syntax is case sensitive. For example, if an element tag name is uppercase, it must be uppercase in the location path. If it is lowercase, it must be lowercase in the location path. All location paths must begin with the root-enclosing element (denoted by a slash '/'), or with the "any parent" variant (denoted by double slashes '//'). Other W3C documented forms are not currently supported.

**Example**

“Specifying a Location Path on the PATH Element” on page 44

**INCREMENT-PATH syntax="type" beginend="BEGIN | END"** XMLV2 and XML specifies a location path for a counter variable, which is established by specifying the COLUMN element attribute class="ORDINAL". The location path tells the XML engine where in the input data to increment the accumulated value for the counter variable by 1.

**syntax="type"**

is an optional attribute that specifies the type of syntax in the location path. The syntax is valid XPath construction in compliance with the W3C specifications. The XPath form supported by the XML engine allows elements and attributes to be individually included in the generated SAS data set. Here is an example:

```
syntax="XPath"
```

**Default** XPath

**Requirements**

The value must be XPath or XPathENR.

If an XML namespace is defined with the NAMESPACES element, you must specify the type of syntax as XPathENR (XPath with Embedded Namespace Reference). This is because the syntax is different from the XPath specification. Here is an example:

```
syntax="XPathENR"
```

**beginend="BEGIN | END"**

is an optional attribute that specifies to stop processing when either the element start tag is encountered or the element end tag is encountered.

**Default** BEGIN

**Requirements**

If an XML namespace is defined with the NAMESPACES element, you must include the identification number in the location path preceding the element that is being defined. The identification number is enclosed in braces. Here is an example:

```
INCREMENT-PATH syntax="XPathENR">/Table/Hurricane/{1}Month</INCREMENT-PATH>
```

The XPath construction is a formal specification that puts a path description similar to UNIX on each element of the XML structure. Note that XPath syntax is case sensitive. For example, if an element tag name is uppercase, it must be uppercase in the location path. If it
is lowercase, it must be lowercase. All location paths must begin with the root-enclosing element (denoted by a slash '/') or with the "any parent" variant (denoted by double slashes '/*'). Other W3C documented forms are not currently supported.

If the variable is not a counter variable, PATH is required and INCREMENT-PATH and RESET-PATH are not allowed. If the variable is a counter variable, PATH is not allowed and either INCREMENT-PATH or DECREMENT-PATH is required.

Example

"Including a Key Field with Generated Numeric Keys" on page 33

RESET-PATH syntax="type" beginend="BEGIN | END" XMLV2 and XML

specifies a location path for a counter variable, which is established by specifying the COLUMN element attribute class="ORDINAL". The location path tells the XML engine where in the XML document to reset the accumulated value for the counter variable to zero.

syntax="type"

is an optional attribute that specifies the type of syntax in the location path. The syntax is valid XPath construction in compliance with the W3C specifications. The XPath form supported by the XML engine allows elements and attributes to be individually included in the generated SAS data set. Here is an example:

syntax="XPATH".

Default XPath

Requirements The value must be XPath or XPathENR.

If an XML namespace is defined with the NAMESPACES element, you must specify the type of syntax as XPathENR (XPath with Embedded Namespace Reference). This is because the syntax is different from the XPath specification. Here is an example: syntax="XPathENR".

beginend="BEGIN | END"

is an optional attribute that specifies to stop processing when either the element start tag is encountered or the element end tag is encountered.

Default BEGIN

Requirements If the variable is not a counter variable, RESET-PATH is not allowed. If the variable is a counter variable, RESET-PATH is optional.

If an XML namespace is defined with the NAMESPACES element, you must include the identification number in the location path preceding the element that is being defined. The identification number is enclosed in braces. Here is an example: <RESET-PATH syntax="XPathENR">/Table/Hurricane/{1}Month</RESET-PATH>.

The XPath construction is a formal specification that puts a path description similar to UNIX on each element of the XML structure. Note that XPath syntax is case sensitive. For example, if an element tag name is uppercase, it must be uppercase in the location path. If it is lowercase, it must be lowercase. All location paths must begin
DECREMENT-PATH syntax="type" beginend="BEGIN | END" XMLV2 and XML specifies a location path for a counter variable, which is established by specifying the COLUMN element attribute class="ORDINAL". The location path tells the XML engine where in the input data to decrement the accumulated value for the counter variable by 1.

**syntax="type"**

is an optional attribute that specifies the type of syntax in the location path. The syntax is valid XPath construction in compliance with the W3C specifications. The XPath form supported by the XML engine allows elements and attributes to be individually included in the generated SAS data set. Here is an example:

```xml
<DECREMENT-PATH syntax="XPath">/Table/Hurricane/{1}Month</DECREMENT-PATH>
```

The XPath construction is a formal specification that puts a path description similar to UNIX on each element of the XML structure. XPath syntax is case sensitive. For example, if an element tag name is uppercase, it must be uppercase in the location path. If it is lowercase, it must be lowercase in the location path. All location paths must begin with the root-enclosing element (denoted by a slash '/'), or with the "any parent" variant (denoted by double slashes /**/). Other W3C documented forms are not currently supported.
Chapter 11
Using SAS XML Mapper to Generate and Update an XMLMap

What Is SAS XML Mapper? ................................................................. 113
Using the Windows ................................................................. 114
Using the Menu Bar ................................................................. 114
Using the Toolbar ................................................................. 114
How Do I Get SAS XML Mapper? ............................................. 115
To Start SAS XML Mapper ..................................................... 115

What Is SAS XML Mapper?

SAS XML Mapper is an XMLMap support tool for the XMLV2 and XML engines. Based on Java, SAS XML Mapper is a stand-alone application that removes the tedium of creating and modifying an XMLMap.

SAS XML Mapper provides a graphical interface that you can use to generate the appropriate XML elements. SAS XML Mapper analyzes the structure of an XML document or an XML schema and generates basic XML syntax for the XMLMap.

The interface consists of windows, a menu bar, and a toolbar. Using SAS XML Mapper, you can display an XML document or an XML schema, create and modify an XMLMap, and generate example SAS programs.
Using the Windows

The XML window and the XMLMap window are the two primary windows. The XML window, which is on the left, displays an XML document in a tree structure. The XMLMap window, which is on the right, displays an XMLMap in a tree structure. The map tree displays three layers: the top level is the map itself, the second tier includes tables, and the leaf nodes are columns. The detail area at the top displays information about the currently selected item, such as attributes for the table or column. The information is subdivided into tabs.

There are several source windows on the bottom of the interface, such as the XML source window, the XMLMap source window, the SAS Code Example window, and so on.

Using the Menu Bar

The menu bar provides menus to request functionality. For example, select the File menu, and then Open XML to display a browser so that you can select an XML document to open.

Using the Toolbar

The toolbar contains icons for shortcuts to several items on the menu bar. For example, the first icon from the left is the Open an XML file icon. Select it to display a browser so that you can select an XML document to open.
How Do I Get SAS XML Mapper?

SAS XML Mapper can be installed from the installation media for Windows and UNIX platforms, or it can be downloaded from the SAS website at http://support.sas.com/demosdownloads/setupcat.jsp?cat=Base+SAS+Software. For SAS Viya releases, SAS XML Mapper is not included in the installation media. You can download it from the SAS website to install it.

The latest version of SAS XML Mapper, which is SAS 9.4, can be downloaded and used with SAS 9.4 or with versions of SAS prior to SAS 9.4. There are some features that can be used only with SAS 9.3 and SAS 9.4 SAS XML Mapper, such as the 2.1 XMLMap version.

SAS XML Mapper has online Help attached, which includes usage examples. From the menu bar, select Help, and then Help Topics.

For a quick tutorial of SAS XML Mapper, see the video How to Automatically Generate XMLMap Files (video) on the Base SAS Focus Area at http://support.sas.com/rnd/base/xmlengine. Look for the heading XML Mapper and click on the link to the video.

To Start SAS XML Mapper

To start SAS XML Mapper:

• In a Windows environment, launch SAS XML Mapper from your desktop. Typically, you can launch SAS XML Mapper by selecting Start ⇒ All Programs ⇒ SAS ⇒ SAS XML Mapper 9.4.

• In a UNIX environment, start SAS XML Mapper from the UNIX command prompt.
Part 4

Appendixes

Appendix 1

Usage Not Supported for the XMLV2 Engine ........................................ 119

Appendix 2

Example CDISC ODM Document .......................................................... 137
Appendix 1
Usage Not Supported for the XMLV2 Engine

Exporting an XML Document for Use by Oracle ........................................ 119
Exporting Numeric Values ........................................................................ 121
Exporting an XML Document in CDISC ODM Markup .......................... 126
Importing an XML Document with Numeric Values ................................. 127
Importing an XML Document Created by Microsoft Access ..................... 128
Importing Concatenated XML Documents .............................................. 132
Importing a CDISC ODM Document ....................................................... 134

Exporting an XML Document for Use by Oracle

This example exports an XML document from a SAS data set for use by Oracle. By specifying the ORACLE markup type, the XML engine generates tags that are specific to Oracle standards. XMLTYPE=ORACLE is not supported by the XMLV2 engine.

The following output shows the SAS data set MyFiles.Class to be exported to Oracle.
Output A1.1  SAS Data Set MyFiles.Class

The following SAS program exports an XML document from the SAS data set MyFiles.Class:

```
libname myfiles 'SAS-library'; /*1*/

libname trans xml 'XML-document' xmltype=oracle; /*2*/

data trans.class; /*3*/
  set myfiles.class;
run;
```

1 The first LIBNAME statement assigns the libref MyFiles to the physical location of the SAS library that stores the SAS data set Class. The V9 engine is the default.

2 The second LIBNAME statement assigns the libref Trans to the physical location of the file (complete pathname, filename, and file extension) that will store the exported XML document and specifies the XML engine. The engine option XMLTYPE=ORACLE produces tags that are equivalent to the Oracle 8i XML implementation.
The DATA step reads the SAS data set MyFiles.Class and writes its content in Oracle XML markup to the specified XML document.

Here is the resulting XML document.

**Output A1.2 XML Document Exported from MyFiles.Class to Be Used by Oracle**

```xml
<ROWSET>
  <ROW>
    <Name> Alfred </Name>
    <Gender> M </Gender>
    <Age> 14 </Age>
    <Height> 69 </Height>
    <Weight> 112.5 </Weight>
  </ROW>
  <ROW>
    <Name> Alice </Name>
    <Gender> F </Gender>
    <Age> 13 </Age>
    <Height> 56.5 </Height>
    <Weight> 84 </Weight>
  </ROW>
  ...
  <ROW>
    <Name> William </Name>
    <Gender> M </Gender>
    <Age> 15 </Age>
    <Height> 66.5 </Height>
    <Weight> 112 </Weight>
  </ROW>
</ROWSET>
```

---

**Exporting Numeric Values**

This example uses a small SAS data set, with a numeric variable that contains values with a high precision. The following SAS program creates the data set with an assigned user-defined format, and then exports two XML documents to show the difference in output.

This example is appropriate for the XML engine only. The XMLV2 engine ignores any assigned format and displays the value using BEST16.

```sas
libname format xml 'C:\Output\format.xml'; /*3*/
libname prec xml 'C:\Output\precision.xml' xmldouble=internal; /*2*/
data npi; /*3*/
do n=1 to 10;
  n_pi = n*3.141592653589793;
  output;
end;
format n_pi f14.2;
```
run;

data format.dbltest; /*4*/
    set npi;
run;

data prec.rawtest; /*5*/
    set npi;
run;

title 'Drops the Precision'; /*6*/
proc print data=format.dbltest;
    format n_pi f14.10;
run;

title 'Keeps the Precision'; /*7*/
proc print data=prec.rawtest;
    format n_pi f14.10;
run;

1 The first LIBNAME statement assigns the libref Format to the file that will store the generated XML document Format.XML. The default behavior for the engine is that an assigned SAS format controls numeric values.

2 The second LIBNAME statement assigns the libref Prec to the file that will store the generated XML document Precision.XML. The XMLDOUBLE= option specifies INTERNAL, which causes the engine to retrieve the stored raw values.

3 The DATA step creates the temporary data set NPI. The data set has a numeric variable that contains values with a high precision. The variable has an assigned user-defined format that specifies two decimal points.

4 The DATA step creates the data set Format.DblTest from Work.NPI.

5 The DATA step creates the data set Prec.RawTest from Work.NPI.

6 From the data set Format.DblTest, the PRINT procedure generates the XML document Format.XML, which contains numeric values controlled by the SAS format. See Output A1.3 on page 123.

7 For the PRINT procedure output, a format was specified to show the precision loss. In the output, the decimals after the second digit are zeros. See Output A1.4 on page 124.

8 From the data set Prec.RawTest, the PRINT procedure generates the XML document Precision.XML, which contains the stored numeric values. See Output A1.5 on page 125.

9 For the PRINT procedure output, a format was specified to show the retained precision. See Output A1.6 on page 126.
Output A1.3   XML Document Format.XML

<?xml version="1.0" encoding="iso-8859-1" ?>
<TABLE>
   <DBLTEST>
      <n>1</n>
      <n_pi>3.14</n_pi>
   </DBLTEST>
   <DBLTEST>
      <n>2</n>
      <n_pi>6.28</n_pi>
   </DBLTEST>
   <DBLTEST>
      <n>3</n>
      <n_pi>9.42</n_pi>
   </DBLTEST>
   <DBLTEST>
      <n>4</n>
      <n_pi>12.57</n_pi>
   </DBLTEST>
   <DBLTEST>
      <n>5</n>
      <n_pi>15.71</n_pi>
   </DBLTEST>
   <DBLTEST>
      <n>6</n>
      <n_pi>18.85</n_pi>
   </DBLTEST>
   <DBLTEST>
      <n>7</n>
      <n_pi>21.99</n_pi>
   </DBLTEST>
   <DBLTEST>
      <n>8</n>
      <n_pi>25.13</n_pi>
   </DBLTEST>
   <DBLTEST>
      <n>9</n>
      <n_pi>28.27</n_pi>
   </DBLTEST>
   <DBLTEST>
      <n>10</n>
      <n_pi>31.42</n_pi>
   </DBLTEST>
</TABLE>
### Output A1.4  PRINT Procedure Output for Format.DblTest

<table>
<thead>
<tr>
<th>Obs</th>
<th>N_PI</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.1400000000</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>6.2800000000</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>9.4200000000</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>12.5700000000</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>15.7100000000</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>18.8500000000</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>21.9900000000</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>25.1300000000</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>28.2700000000</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>31.4200000000</td>
<td>10</td>
</tr>
</tbody>
</table>
Output A1.5  XML Document Precision.XML

```xml
<?xml version="1.0" encoding="iso-8859-1" ?>
<TABLE>
  <RAWTEST>
    <n rawvalue="QRAAAAAAAAA=">1</n>
    <n_pi rawvalue="QTJD9qiIWjA=">3.14</n_pi>
  </RAWTEST>
  <RAWTEST>
    <n rawvalue="QSAAAAAAAAA=">2</n>
    <n_pi rawvalue="QWSH7VEQcGA=">6.28</n_pi>
  </RAWTEST>
  <RAWTEST>
    <n rawvalue="QTAAAAAAAAA=">3</n>
    <n_pi rawvalue="QZbL4/mZDpA=">9.42</n_pi>
  </RAWTEST>
  <RAWTEST>
    <n rawvalue="QUAAAAAAAAA=">4</n>
    <n_pi rawvalue="QckP2qIhaMA=">12.57</n_pi>
  </RAWTEST>
  <RAWTEST>
    <n rawvalue="QVAAAAAAAAA=">5</n>
    <n_pi rawvalue="QftT0UqpwvA=">15.71</n_pi>
  </RAWTEST>
  <RAWTEST>
    <n rawvalue="QWAAAAAAAAA=">6</n>
    <n_pi rawvalue="QhLZfH8zIdI=">18.85</n_pi>
  </RAWTEST>
  <RAWTEST>
    <n rawvalue="QXAAAAAAAAA=">7</n>
    <n_pi rawvalue="QhX9u+m7p3U=">21.99</n_pi>
  </RAWTEST>
  <RAWTEST>
    <n rawvalue="QYAAAAAAAAA=">8</n>
    <n_pi rawvalue="QhkY1RLRg=">25.13</n_pi>
  </RAWTEST>
  <RAWTEST>
    <n rawvalue="QZAAAAAAAAA=">9</n>
    <n_pi rawvalue="QhxGOr7Msrs=">28.27</n_pi>
  </RAWTEST>
  <RAWTEST>
    <n rawvalue="QaAAAAAAAAA=">10</n>
    <n_pi rawvalue="Qh9qeilV0F4=">31.42</n_pi>
  </RAWTEST>
</TABLE>

Exporting Numeric Values
Exporting an XML Document in CDISC ODM Markup

This example exports the SAS data set that is imported in “Importing a CDISC ODM Document” on page 134 back to an XML document that is in CDISC ODM markup. Because the CDISCODM markup type is specified, the XML engine generates tags that are specific to the CDISC Operational Data Model. XMLTYPE=CDISCODM is not supported by the XMLV2 engine.

The following SAS program exports an XML document from the SAS data set Odm.AE:

```
filename output 'C:\myoutput.xml'; /*1*/
libname output xml xmltype=CDISCODM formatactive=yes; /*2*/
data output.AE2;
    set odm.AE;
run;
```

1 The FILENAME statement assigns the fileref Output to the physical location of the external file (complete pathname, filename, and file extension) to which the exported information will be written.

2 The LIBNAME statement uses the fileref Output as the output location and specifies the XML engine. It includes the following engine options:
   - XMLTYPE=CDISCODM supports the markup standards for CDISC ODM 1.2.
   - FORMATACTIVE=YES specifies to convert SAS formats to the corresponding CDISC ODM CodeList elements.
The output is the same as the XML document that is shown in Appendix 2, “Example CDISC ODM Document,” on page 137.

---

**Importing an XML Document with Numeric Values**

This example imports the XML document Precision.XML, which was exported in “Exporting Numeric Values” on page 121. This example illustrates how you can change the behavior for importing numeric values.

The first SAS program imports the XML document using the default behavior, which retrieves parsed character data (PCDATA) from the element:

```
libname default xml 'C:\Output\precision.xml';

title 'Default Method';
proc print data=default.rawtest;
   format n_pi f14.10;
run;
```

The result of the import is the SAS data set Default.RawTest.

**Output A1.7** PRINT Procedure Output for Default.RawTest

![Default Method]

The second SAS program imports the XML document using the XMLDOUBLE= option to change the behavior, which retrieves the value from the rawdata= attribute in the element. (This example uses the XML engine. When importing, XMLDOUBLE=INTERNAL is not supported by the XMLV2 engine.)

```
libname new xml 'C:\Output\precision.xml' xmldouble=internal;

title 'Precision Method';
proc print data=new.rawtest;
   format n_pi f14.10;
run;
```
The result of the import is SAS data set New.RawTest

**Output A1.8** PRINT Procedure Output for New.RawTest

---

**Importing an XML Document Created by Microsoft Access**

This example imports the following XML document, which was created from a Microsoft Access database. Because the XML document contains an embedded XML schema, you must specify the MSACCESS markup type rather than the default GENERIC type. MSACCESS obtains a variable's attributes from the embedded schema. XMLTYPE=MSACCESS is not supported by the XMLV2 engine.

```xml
<xml version="1.0" encoding="windows-1252"/>
<root xmlns:xs="http://www.w3.org/2001/XMLSchema"
     xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
     xmlns:od="urn:schemas-microsoft-com:officedata">
  <xs:schema>
    <xs:element name="dataroot">
      <xs:complexType>
        <xs:sequence>
          <xs:element ref="SUPPLIERS" minOccurs="0" maxOccurs="unbounded" />
        </xs:sequence>
      </xs:complexType>
    </xs:element>
    <xs:element name="SUPPLIERS">
      <xs:complexType>
        <xs:sequence>
          <xs:element name="HOMEPAGE" minOccurs="0" od:jetType="text" od:sqlSType="nvarchar">
            <xs:simpleType>
              <!-- Content of the simple type -->
            </xs:simpleType>
          </xs:element>
        </xs:sequence>
      </xs:complexType>
    </xs:element>
  </xs:schema>
</root>
```
Importing an XML Document Created by Microsoft Access

<x:simpleType>
  <xs:restriction base="xs:string">
    <xs:maxLength value="94" />
  </xs:restriction>
</xs:simpleType>

<x:element name="FAX" minOccurs="0"
  od:jetType="text" od:sqlSType="nvarchar">
  <xs:simpleType>
    <xs:restriction base="xs:string">
      <xs:maxLength value="15" />
    </xs:restriction>
  </xs:simpleType>
</x:element>

<x:element name="PHONE" minOccurs="0"
  od:jetType="text" od:sqlSType="nvarchar">
  <xs:simpleType>
    <xs:restriction base="xs:string">
      <xs:maxLength value="15" />
    </xs:restriction>
  </xs:simpleType>
</x:element>

<x:element name="COUNTRY" minOccurs="0"
  od:jetType="text" od:sqlSType="nvarchar">
  <xs:simpleType>
    <xs:restriction base="xs:string">
      <xs:maxLength value="11" />
    </xs:restriction>
  </xs:simpleType>
</x:element>

<x:element name="POSTALCODE" minOccurs="0"
  od:jetType="text" od:sqlSType="nvarchar">
  <xs:simpleType>
    <xs:restriction base="xs:string">
      <xs:maxLength value="8" />
    </xs:restriction>
  </xs:simpleType>
</x:element>

<x:element name="REGION" minOccurs="0"
  od:jetType="text" od:sqlSType="nvarchar">
  <xs:simpleType>
    <xs:restriction base="xs:string">
      <xs:maxLength value="8" />
    </xs:restriction>
  </xs:simpleType>
</x:element>

<x:element name="CITY" minOccurs="0"
  od:jetType="text" od:sqlSType="nvarchar">
  <xs:simpleType>
    <xs:restriction base="xs:string">
      <xs:maxLength value="13" />
    </xs:restriction>
  </xs:simpleType>
</x:element>

<x:element name="ADDRESS" minOccurs="0"
  od:jetType="text" od:sqlSType="nvarchar">
  <xs:simpleType>
    <xs:restriction base="xs:string">
      <xs:maxLength value="94" />
    </xs:restriction>
  </xs:simpleType>
</x:element>
<xs:restriction base="xs:string">
  <xs:maxLength value="45" />
</xs:restriction>
</xs:simpleType>
</xs:element>
<xs:element name="CONTACTTITLE" minOccurs="0"
  od:jetType="text" od:sqlSType="nvarchar">
  <xs:simpleType>
    <xs:restriction base="xs:string">
      <xs:maxLength value="28" />
    </xs:restriction>
  </xs:simpleType>
</xs:element>
<xs:element name="CONTACTNAME" minOccurs="0"
  od:jetType="text" od:sqlSType="nvarchar">
  <xs:simpleType>
    <xs:restriction base="xs:string">
      <xs:maxLength value="26" />
    </xs:restriction>
  </xs:simpleType>
</xs:element>
<xs:element name="COMPANYNAME" minOccurs="0"
  od:jetType="text" od:sqlSType="nvarchar">
  <xs:simpleType>
    <xs:restriction base="xs:string">
      <xs:maxLength value="38" />
    </xs:restriction>
  </xs:simpleType>
</xs:element>
<xs:element name="SUPPLIERID" minOccurs="0"
  type="xs:double" />
</xs:sequence>
</xs:complexType>
</xs:element>
</xsd:schema>
<dataroot>
  <SUPPLIERS>
    <HOMEPAGE/>
    <PAX/>
    <PHONE>(272) 444-2222</PHONE>
    <COUNTRY>UK</COUNTRY>
    <POSTALCODE>EC1 4SD</POSTALCODE>
    <REGION/>
    <CITY>London</CITY>
    <ADDRESS>49 Franklin St.</ADDRESS>
    <CONTACTTITLE>Purchasing Manager</CONTACTTITLE>
    <CONTACTNAME>Charlotte Smith</CONTACTNAME>
    <COMPANYNAME>Exotic Flowers</COMPANYNAME>
    <SUPPLIERID>1</SUPPLIERID>
  </SUPPLIERS>
  <SUPPLIERS>
    <HOMEPAGE>#MYCAJUN.HTM#</HOMEPAGE>
    <PAX/>
    <PHONE>(512) 284-3677</PHONE>
    <COUNTRY>USA</COUNTRY>
The following SAS program interprets the XML document as a SAS data set:

```
libname access xml '/u/myid/XML/suppliers.xml'
   xmltype=msaccess xmlmeta=schemadata; /*1*/
proc print data=access.suppliers (obs=2); /*2*/
   var contactname companyname;
run;
```

1 The LIBNAME statement assigns the libref Access to the physical location of the XML document (complete pathname, filename, and file extension) and specifies the XML engine. By default, the XML engine expects GENERIC markup, so you must include the XMLTYPE= option to read the XML document in MSACCESS markup and to obtain a variable's attributes from the embedded schema. The option XMLMETA=SCHEMADATA specifies to import both data and metadata-related information from the input XML document.

2 The PRINT procedure produces the output. The procedure uses the OBS= data set option to print only the first two observations, and the VAR statement to print only specific variables (columns).

Output A1.9  PRINT Procedure Output for Access.Suppliers

<table>
<thead>
<tr>
<th>Obs</th>
<th>CONTACTNAME</th>
<th>COMPANYNAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Charlotte Smith</td>
<td>Exotic Flowers</td>
</tr>
<tr>
<td>2</td>
<td>Shelley Martin</td>
<td>New Orleans Cajun Foods</td>
</tr>
</tbody>
</table>

Using the CONTENTS procedure, the output displays the file's attributes, as well as the attributes of each interpreted column (variable), such as the variable's type and length, which are obtained from the embedded XML schema. Without the embedded XML schema, the results for the attributes would be default values.

```
proc contents data=access.suppliers;
run;
```
Importing Concatenated XML Documents

For a file that is a concatenation of multiple XML documents, you can use the XMLV2 engine without the XMLCONCATENATE= option, as long as the elements pass XML markup parsing rules such as case sensitivity.

To use the older XML engine to import the file, simply specify the LIBNAME statement option XMLCONCATENATE=YES.

**Note:** Use XMLCONCATENATE=YES cautiously. If an XML document consists of concatenated XML documents, the content is not standard XML construction. The option is provided for convenience, not to encourage invalid XML markup.

This example imports the following file named ConcatStudents.XML, which consists of two XML documents:
First, using the default XML engine behavior, which does not support concatenated XML documents (XMLCONCATENATE=NO), the following SAS program imports the first XML document, which consists of three observations, and produces an error for the second XML document:

```sas
libname concat xml 'C:\Example\ConcatStudents.xml';
proc datasets library=concat;
```
Log A1.1  SAS Log Output

NOTE: Libref CONCAT was successfully assigned as follows:
    Engine:        XML
    Physical Name: C:\Example\ConcatStudents.xml
20   proc datasets library=concat;
ERROR: "xml" is illegal as a processing-instruction target name.
ERROR: encountered during XMLMap parsing
     occurred at or near line 23, column 7

Specifying the LIBNAME statement option XMLCONCATENATE=YES enables the XML engine to import the concatenated XML documents as one SAS data set:

   libname concat xml 'C:\Example\ConcatStudents.xml' xmlconcatenate=yes;

   proc print data=concat.students;
   run;

Output A1.11  PRINT Procedure Output for Concat.Students

The SAS System

<table>
<thead>
<tr>
<th>Obs</th>
<th>CITY</th>
<th>SCHOOL</th>
<th>NAME</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Houston</td>
<td>Bellaire</td>
<td>Linda Kay</td>
<td>1345</td>
</tr>
<tr>
<td>2</td>
<td>Houston</td>
<td>Sam Houston</td>
<td>Chas Wofford</td>
<td>2456</td>
</tr>
<tr>
<td>3</td>
<td>Houston</td>
<td>Sharpstown</td>
<td>Jerry Kolar</td>
<td>3567</td>
</tr>
<tr>
<td>4</td>
<td>Austin</td>
<td>Reagan</td>
<td>Brad Martin</td>
<td>1234</td>
</tr>
<tr>
<td>5</td>
<td>Austin</td>
<td>Westwood</td>
<td>Zac Hanwell</td>
<td>2345</td>
</tr>
<tr>
<td>6</td>
<td>Austin</td>
<td>Bowie</td>
<td>Walter Smith</td>
<td>3456</td>
</tr>
</tbody>
</table>

Importing a CDISC ODM Document

This example imports the XML document that is shown in Appendix 2, “Example CDISC ODM Document,” on page 137. The document conforms to Version 1.2 of the CDISC Operational Data Model (ODM). To import a CDISC ODM document, you specify CDISCODM as the XML markup type, and you can specify values for the FORMATACTIVE= option, FORMATLIBRARY= option, and FORMATNOREPLACE= option. Note that XMLTYPE=CDISCODM is not supported by the XMLV2 engine.

The following SAS program imports the XML document as a SAS data set:

   filename odm 'C:\Example\AE.XML'; /*
   libname odm xml xmltype=CDISCODM /*
       FormActive=YES /*
       FormatNoReplace=NO /*
       FormatLibrary="Work"; /*
   */
The FILENAME statement assigns the fileref Odm to the physical location of the XML document (complete pathname, filename, and file extension).

The LIBNAME statement uses the fileref Odm to reference the XML document and specifies the XML engine. If the fileref matches the libref, you do not need to specify the physical location of the XML document in the LIBNAME statement. By default, the XML engine expects GENERIC markup, so you must include the XMLTYPE= option to read the XML document in CDISCODM markup.

FORMATACTIVE=YES specifies to convert CDISC ODM CodeList elements in the document to SAS formats.

FORMATNOREPLACE=NO specifies to replace any existing SAS formats in the format catalog that have the same name as the converted formats.

FORMATLIBRARY="Work" specifies to create the format catalog in the temporary Work library. The Work library is also the default if you omit the FORMATLIBRARY= option.

The output from the CONTENTS procedure displays the file's attributes as well as the attributes of each interpreted column (variable), such as the variable's type and length. The attributes are obtained from the embedded ODM metadata content. The VARNUM option causes the variables to be printed first in alphabetical order and then in the order of their creation.
### The SAS System

#### The CONTENTS Procedure

<table>
<thead>
<tr>
<th>Data Set Name</th>
<th>CONTENTS</th>
<th>Observations</th>
<th>Variables</th>
<th>23</th>
</tr>
</thead>
<tbody>
<tr>
<td>Member Type</td>
<td>DATA</td>
<td>Engine</td>
<td>JSL</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Imported</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Observation</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variables</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Projection</td>
<td>Compressed</td>
<td>NO</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Data Set</td>
<td>Sorted</td>
<td>NO</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Label</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Data</td>
<td>Representation</td>
<td>Default</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Encoding</td>
<td>Default</td>
<td></td>
</tr>
</tbody>
</table>

#### Variables in Creation Order

<table>
<thead>
<tr>
<th>#</th>
<th>Variable</th>
<th>Type</th>
<th>Len</th>
<th>Format</th>
<th>Internal</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>__STUDYID</td>
<td>Char</td>
<td>100</td>
<td></td>
<td></td>
<td>__STUDYID</td>
</tr>
<tr>
<td>2</td>
<td>__REDACTAVERSIONID</td>
<td>Char</td>
<td>100</td>
<td></td>
<td></td>
<td>__REDACTAVERSIONID</td>
</tr>
<tr>
<td>3</td>
<td>__SUBJECTKEY</td>
<td>Char</td>
<td>100</td>
<td></td>
<td></td>
<td>__SUBJECTKEY</td>
</tr>
<tr>
<td>4</td>
<td>__STUDYEVENTID</td>
<td>Char</td>
<td>100</td>
<td></td>
<td></td>
<td>__STUDYEVENTID</td>
</tr>
<tr>
<td>5</td>
<td>__STUDYEVENTREPEATKEY</td>
<td>Char</td>
<td>100</td>
<td></td>
<td></td>
<td>__STUDYEVENTREPEATKEY</td>
</tr>
<tr>
<td>6</td>
<td>__FORIAD</td>
<td>Char</td>
<td>100</td>
<td></td>
<td></td>
<td>__FORIAD</td>
</tr>
<tr>
<td>7</td>
<td>__FORIADREPEATKEY</td>
<td>Char</td>
<td>100</td>
<td></td>
<td></td>
<td>__FORIADREPEATKEY</td>
</tr>
<tr>
<td>8</td>
<td>__TEHONDFOOD</td>
<td>Char</td>
<td>100</td>
<td></td>
<td></td>
<td>__TEHONDFOOD</td>
</tr>
<tr>
<td>9</td>
<td>__TEHONDFOODREPEATKEY</td>
<td>Char</td>
<td>100</td>
<td></td>
<td></td>
<td>__TEHONDFOODREPEATKEY</td>
</tr>
<tr>
<td>10</td>
<td>TEAM</td>
<td>Char</td>
<td>5</td>
<td></td>
<td></td>
<td>TEAM</td>
</tr>
<tr>
<td>11</td>
<td>PNO</td>
<td>Char</td>
<td>15</td>
<td>Protocol Number</td>
<td></td>
<td>Protocol Number</td>
</tr>
<tr>
<td>12</td>
<td>STCTY</td>
<td>Char</td>
<td>4</td>
<td></td>
<td></td>
<td>STCTY</td>
</tr>
<tr>
<td>13</td>
<td>F_STATUS</td>
<td>Char</td>
<td>1</td>
<td></td>
<td></td>
<td>F_STATUS</td>
</tr>
<tr>
<td>14</td>
<td>LINE_ID</td>
<td>Num</td>
<td>2</td>
<td>Line Number</td>
<td></td>
<td>Line Number</td>
</tr>
<tr>
<td>15</td>
<td>ABTERM</td>
<td>Char</td>
<td>100</td>
<td></td>
<td></td>
<td>ABTERM</td>
</tr>
<tr>
<td>16</td>
<td>ABSEND</td>
<td>Num</td>
<td>2</td>
<td>Start Month - Enter Two Digits 01-12</td>
<td></td>
<td>Start Month - Enter Two Digits 01-12</td>
</tr>
<tr>
<td>17</td>
<td>ABSEND</td>
<td>Num</td>
<td>2</td>
<td>Start Day - Enter Two Digits 01-31</td>
<td></td>
<td>Start Day - Enter Two Digits 01-31</td>
</tr>
<tr>
<td>18</td>
<td>ABSEND</td>
<td>Num</td>
<td>2</td>
<td>Start Year - Enter Four Digits Year</td>
<td></td>
<td>Start Year - Enter Four Digits Year</td>
</tr>
<tr>
<td>19</td>
<td>ABSEND</td>
<td>Num</td>
<td>2</td>
<td>Stop Month - Enter Two Digits 01-12</td>
<td></td>
<td>Stop Month - Enter Two Digits 01-12</td>
</tr>
<tr>
<td>20</td>
<td>ABSEND</td>
<td>Num</td>
<td>2</td>
<td>Stop Day - Enter Two Digits 01-31</td>
<td></td>
<td>Stop Day - Enter Two Digits 01-31</td>
</tr>
<tr>
<td>21</td>
<td>ABSEND</td>
<td>Num</td>
<td>2</td>
<td>Stop Year - Enter Four Digits Year</td>
<td></td>
<td>Stop Year - Enter Four Digits Year</td>
</tr>
<tr>
<td>22</td>
<td>ABSEND</td>
<td>Num</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>ABSEND</td>
<td>Num</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>ABSEND</td>
<td>Num</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>ABSEND</td>
<td>Num</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>ABSEND</td>
<td>Num</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>ABSEND</td>
<td>Num</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>ABSEND</td>
<td>Num</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>ABSEND</td>
<td>Num</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>ABSEND</td>
<td>Num</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Legend: Char = Character, Num = Numeric, F = Fixed, Var = Variable*
Appendix 2

Example CDISC ODM Document

Here is an example of an XML document that is in the CDISC ODM format. This document is used in “Importing a CDISC ODM Document” on page 134 and in “Exporting an XML Document in CDISC ODM Markup” on page 126.
Clinical Data Interchange Standards Consortium (CDISC)
Operational Data Model (ODM) for clinical data interchange

You can learn more about CDISC standards efforts at
http://www.cdisc.org/standards/index.html

- <ODM xmlns="http://www.cdisc.org/ns/odm/v1.2"
  xmlns:ds="http://www.w3.org/2000/09/xmldsig#"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.cdisc.org/ns/odm/v1.2 ODM1-2-0.xsd"
  ODMVersion="1.2"
  FileOID="000-00-0000"
  FileType="Snapshot"
  Description="Adverse events from the CTChicago file"
  AsOfDateTime="2009-03-31T14:01:41"
  CreationDateTime="2009-03-31T14:01:41">
  - <Study OID="STUDY.StudyOID">
    - <GlobalVariables>
      <StudyName>CDISC Connect-A-Thon Test Study III</StudyName>
      <StudyDescription>This file contains test data from a previous CDISC Connect-A-Thon.</StudyDescription>
    </GlobalVariables>
    - <BasicDefinitions />
    - <MetaDataVersion OID="v1.1.0" Name="Version 1.1.0">
      - <Protocol>
        <StudyEventRef StudyEventOID="SE.VISIT1" OrderNumber="1" Mandatory="Yes" />
      </Protocol>
      <StudyEventDef OID="SE.VISIT1" Name="Study Event Definition" Repeating="Yes" Type="Common">
        <FormRef FormOID="FORM.AE" OrderNumber="1" Mandatory="No" />
      </StudyEventDef>
      <FormDef OID="FORM.AE" Name="Form Definition" Repeating="Yes">
        <ItemGroupRef ItemGroupOID="IG.AE" Mandatory="No" />
      </FormDef>
    </MetaDataVersion />
  - <GlobalVariables>
    <StudyName>CDISC Connect-A-Thon Test Study III</StudyName>
    <StudyDescription>This file contains test data from a previous CDISC Connect-A-Thon.</StudyDescription>
  </GlobalVariables>
  <BasicDefinitions />
  - <MetaDataVersion OID="v1.1.0" Name="Version 1.1.0">
    - <Protocol>
      <StudyEventRef StudyEventOID="SE.VISIT1" OrderNumber="1" Mandatory="Yes" />
    </Protocol>
    - <StudyEventDef OID="SE.VISIT1" Name="Study Event Definition" Repeating="Yes" Type="Common">
      <FormRef FormOID="FORM.AE" OrderNumber="1" Mandatory="No" />
    </StudyEventDef>
    - <FormDef OID="FORM.AE" Name="Form Definition" Repeating="Yes">
      <ItemGroupRef ItemGroupOID="IG.AE" Mandatory="No" />
    </FormDef>
    - <ItemGroupDef OID="IG.AE" Repeating="Yes" SASDatasetName="AE" Name="Adverse Events" Domain="AE"
      Comment="Some adverse events from this trial">
      <ItemRef ItemOID="ID.TAREA" OrderNumber="1" Mandatory="No" />  
      <ItemRef ItemOID="ID.PNO" OrderNumber="2" Mandatory="No" />  
      <ItemRef ItemOID="ID.SCTRY" OrderNumber="3" Mandatory="No" />  
      <ItemRef ItemOID="ID.F_STATUS" OrderNumber="4" Mandatory="No" />  
      <ItemRef ItemOID="ID.LINE_NO" OrderNumber="5" Mandatory="No" />  
      <ItemRef ItemOID="ID.ATERM" OrderNumber="6" Mandatory="No" />  
      <ItemRef ItemOID="ID.AESTMON" OrderNumber="7" Mandatory="No" />  
      <ItemRef ItemOID="ID.AESTDAY" OrderNumber="8" Mandatory="No" />  
      <ItemRef ItemOID="ID.AESTYR" OrderNumber="9" Mandatory="No" />  
      <ItemRef ItemOID="ID.AESTDT" OrderNumber="10" Mandatory="No" />  
      <ItemRef ItemOID="ID.AEENMON" OrderNumber="11" Mandatory="No" />  
      <ItemRef ItemOID="ID.AEENDAY" OrderNumber="12" Mandatory="No" />  
      <ItemRef ItemOID="ID.AEENYR" OrderNumber="13" Mandatory="No" />  
      <ItemRef ItemOID="ID.AEENDT" OrderNumber="14" Mandatory="No" />  
      <ItemRef ItemOID="ID.AEREVENT" OrderNumber="15" Mandatory="No" />  
      <ItemRef ItemOID="ID.AEREL" OrderNumber="16" Mandatory="No" />  
      <ItemRef ItemOID="ID.AEOUT" OrderNumber="17" Mandatory="No" />  
      <ItemRef ItemOID="ID.AEOUTDT" OrderNumber="18" Mandatory="No" />  
      <ItemRef ItemOID="ID.AECONTRT" OrderNumber="19" Mandatory="No" />
    </ItemGroupDef>
Example CDISC ODM Document

- <!-- Column attributes as defined in the table -->
  
  - <ItemDef OID="ID.TAREA" SASFieldName="TAREA" Name="Therapeutic Area" DataType="text" Length="4">
      <CodeListRef CodeListOID="CL.$TAREAF" />
    </ItemDef>
  
  - <ItemDef OID="ID.PNO" SASFieldName="PNO" Name="Protocol Number" DataType="text" Length="15" />
  
  - <ItemDef OID="ID.SCTRY" SASFieldName="SCTRY" Name="Country" DataType="text" Length="4">
      <CodeListRef CodeListOID="CL.$SCTRYF" />
    </ItemDef>
  
  - <ItemDef OID="ID.F_STATUS" SASFieldName="F_STATUS" Name="Record status, 5 levels, internal use" DataType="text" Length="1">
      <CodeListRef CodeListOID="CL.$F_STATU" />
    </ItemDef>
  
  - <ItemDef OID="ID.LINE_NO" SASFieldName="LINE_NO" Name="Line Number" DataType="integer" Length="2" />
  
  - <ItemDef OID="ID.AETERM" SASFieldName="AETERM" Name="Conmed Indication" DataType="text" Length="100" />
  
  - <ItemDef OID="ID.AESTMON" SASFieldName="AESTMON" Name="Start Month - Enter Two Digits 01-12" DataType="integer" Length="2" />
  
  - <ItemDef OID="ID.AESTDAY" SASFieldName="AESTDAY" Name="Start Day - Enter Two Digits 01-31" DataType="integer" Length="2" />
  
  - <ItemDef OID="ID.AESTYR" SASFieldName="AESTYR" Name="Start Year - Enter Four Digit Year" DataType="integer" Length="4" />
  
  - <ItemDef OID="ID.AESTDT" SASFieldName="AESTDT" Name="Derived Start Date" DataType="date" />
  
  - <ItemDef OID="ID.AEENMON" SASFieldName="AEENMON" Name="Stop Month - Enter Two Digits 01-12" DataType="integer" Length="2" />
  
  - <ItemDef OID="ID.AEENDAY" SASFieldName="AEENDAY" Name="Stop Day - Enter Two Digits 01-31" DataType="integer" Length="2" />
  
  - <ItemDef OID="ID.AEENYR" SASFieldName="AEENYR" Name="Stop Year - Enter Four Digit Year" DataType="integer" Length="4" />
  
  - <ItemDef OID="ID.AEENDT" SASFieldName="AEENDT" Name="Derived Stop Date" DataType="date" />
  
  - <ItemDef OID="ID.AESEV" SASFieldName="AESEV" Name="Severity" DataType="text" Length="1" />
  
  - <ItemDef OID="ID.AEREL" SASFieldName="AEREL" Name="Relationship to study drug" DataType="text" Length="1" />
  
  - <ItemDef OID="ID.AEOUT" SASFieldName="AEOUT" Name="Outcome" DataType="text" Length="1" />
  
  - <ItemDef OID="ID.AEACTTRT" SASFieldName="AEACTTRT" Name="Actions taken re study drug" DataType="text" Length="1" />
  
  - <ItemDef OID="ID.AECONTRT" SASFieldName="AECONTRT" Name="Actions taken, other" DataType="text" Length="1" />
  
  <!-- Translation to ODM markup for any PROC FORMAT style user defined or SAS internal formatting specifications applied to columns in the table -->
  
  - <CodeList OID="CL.$TAREAF" SASFormatName="$TAREAF" Name="$TAREAF" DataType="text">
    <CodeListItem CodedValue="ONC" />
  </CodeList>
  
  <!-- END -->
- <Decode>
  <TranslatedText xml:lang="en">Oncology</TranslatedText>
  </Decode>
- <CodeList OID="CL.$SCTRYF" SASFormatName="$SCTRYF" Name="$SCTRYF" DataType="text">
  - <CodeListItem CodedValue="USA">
    <TranslatedText xml:lang="en">United States</TranslatedText>
  </CodeListItem>
  </CodeList>
- <CodeList OID="CL.$F_STATU" SASFormatName="$F_STATU" Name="$F_STATU" DataType="text">
  - <CodeListItem CodedValue="S">
    <TranslatedText xml:lang="en">Source verified, not queried</TranslatedText>
  </CodeListItem>
  - <CodeListItem CodedValue="V">
    <TranslatedText xml:lang="en">Source verified, queried</TranslatedText>
  </CodeListItem>
  </CodeList>
- <CodeList OID="CL.$AESEV" SASFormatName="$AESEV" Name="$AESEV" DataType="text">
  - <CodeListItem CodedValue="1">
    <TranslatedText xml:lang="en">Mild</TranslatedText>
  </CodeListItem>
  - <CodeListItem CodedValue="2">
    <TranslatedText xml:lang="en">Moderate</TranslatedText>
  </CodeListItem>
  - <CodeListItem CodedValue="3">
    <TranslatedText xml:lang="en">Severe</TranslatedText>
  </CodeListItem>
  - <CodeListItem CodedValue="4">
    <TranslatedText xml:lang="en">Life Threatening</TranslatedText>
  </CodeListItem>
  </CodeList>
- <CodeList OID="CL.$AEREL" SASFormatName="$AEREL" Name="$AEREL" DataType="text">
  - <CodeListItem CodedValue="0">
    <TranslatedText xml:lang="en">None</TranslatedText>
  </CodeListItem>
  - <CodeListItem CodedValue="1">
    <TranslatedText xml:lang="en">Unlikely</TranslatedText>
  </CodeListItem>
  - <CodeListItem CodedValue="2">
    <TranslatedText xml:lang="en">Possible</TranslatedText>
  </CodeListItem>
  - <CodeListItem CodedValue="3">
    <TranslatedText xml:lang="en">Probable</TranslatedText>
  </CodeListItem>
  </CodeList>
- <TranslateItem xml:lang="en">Oncology</TranslateItem>
- <List>
  - <ListItem>United States</ListItem>
  - <ListItem>Source verified, not queried</ListItem>
  - <ListItem>Mild</ListItem>
  - <ListItem>Moderate</ListItem>
  - <ListItem>Severe</ListItem>
  - <ListItem>Life Threatening</ListItem>
  - <ListItem>None</ListItem>
  - <ListItem>Unlikely</ListItem>
  - <ListItem>Possible</ListItem>
  - <ListItem>Probable</ListItem>
</List>
Resolved, no residual effects
Continuing
Resolved, residual effects
Death
Discontinued permanently
Reduced
Interrupted
Medication required
Hospitalization required or prolonged
Other
- <!-- Administrative metadata
-->
<AdminData />
- <!-- Clinical Data : AE

Adverse Events

Some adverse events from this trial
-->

- <ClinicalData StudyOID="STUDY.StudyOID" MetaDataVersionOID="v1.1.0">
  - <SubjectData SubjectKey="001">
    - <StudyEventData StudyEventOID="SE.VISIT1" StudyEventRepeatKey="1">
      - <FormData FormOID="FORM.AE" FormRepeatKey="1">
        - <ItemGroupData ItemGroupOID="IG.AE" ItemGroupRepeatKey="1">
          <ItemData ItemOID="ID.TAREA" Value="ONC" />
          <ItemData ItemOID="ID.PNO" Value="143-02" />
          <ItemData ItemOID="ID.SCTRY" Value="USA" />
          <ItemData ItemOID="ID.F_STATUS" Value="V" />
          <ItemData ItemOID="ID.LINE_NO" Value="1" />
          <ItemData ItemOID="ID.AETERM" Value="HEADACHE" />
          <ItemData ItemOID="ID.AESTMON" Value="06" />
          <ItemData ItemOID="ID.AESTDAY" Value="10" />
          <ItemData ItemOID="ID.AESTYR" Value="1999" />
          <ItemData ItemOID="ID.AESTDT" Value="1999-06-10" />
          <ItemData ItemOID="ID.AEENMON" Value="" />
          <ItemData ItemOID="ID.AEENDAY" Value="" />
          <ItemData ItemOID="ID.AEENYR" Value="" />
          <ItemData ItemOID="ID.AEENDT" Value="" />
          <ItemData ItemOID="ID.AESEV" Value="1" />
          <ItemData ItemOID="ID.AEREL" Value="0" />
          <ItemData ItemOID="ID.AEOUT" Value="1" />
          <ItemData ItemOID="ID.AEACTTRT" Value="0" />
          <ItemData ItemOID="ID.AECONTRT" Value="1" />
        </ItemGroupData>
        - <ItemGroupData ItemGroupOID="IG.AE" ItemGroupRepeatKey="2">
          <ItemData ItemOID="ID.TAREA" Value="ONC" />
          <ItemData ItemOID="ID.PNO" Value="143-02" />
          <ItemData ItemOID="ID.SCTRY" Value="USA" />
          <ItemData ItemOID="ID.F_STATUS" Value="V" />
          <ItemData ItemOID="ID.LINE_NO" Value="2" />
          <ItemData ItemOID="ID.AETERM" Value="CONGESTION" />
          <ItemData ItemOID="ID.AESTMON" Value="06" />
          <ItemData ItemOID="ID.AESTDAY" Value="11" />
          <ItemData ItemOID="ID.AESTYR" Value="1999" />
          <ItemData ItemOID="ID.AESTDT" Value="1999-06-11" />
          <ItemData ItemOID="ID.AEENMON" Value="" />
          <ItemData ItemOID="ID.AEENDAY" Value="" />
          <ItemData ItemOID="ID.AEENYR" Value="" />
          <ItemData ItemOID="ID.AEENDT" Value="" />
          <ItemData ItemOID="ID.AESEV" Value="1" />
          <ItemData ItemOID="ID.AEREL" Value="0" />
          <ItemData ItemOID="ID.AEOUT" Value="2" />
          <ItemData ItemOID="ID.AEACTTRT" Value="0" />
          <ItemData ItemOID="ID.AECONTRT" Value="1" />
        </ItemGroupData>
      </FormData>
    </StudyEventData>
  </SubjectData>
</ClinicalData>
</ODM>
Recommended Reading

Here is the recommended reading list for this title:

- *The Little SAS Book: A Primer*
- *SAS Language Reference: Concepts*
- *SAS Global Statements: Reference*
- *SAS Data Set Options: Reference*
- SAS Companion that is specific to your operating environment
- Base SAS focus area at support.sas.com/base
- For information about XML (Extensible Markup Language), see the website www.w3.org/XML

For a complete list of SAS publications, go to [sas.com/store/books](http://sas.com/store/books). If you have questions about which titles you need, please contact a SAS Representative:

SAS Books
SAS Campus Drive
Cary, NC 27513-2414
Phone: 1-800-727-0025
Fax: 1-919-677-4444
Email: sasbook@sas.com
Web address: [sas.com/store/books](http://sas.com/store/books)
American Standard Code for Information Interchange
See ASCII.

ASCII (American Standard Code for Information Interchange)
a 7-bit encoding standard that provides a basic set of 128 characters, supporting a variety of computer systems. ASCII encodes the uppercase and lowercase letters of the English alphabet, punctuation marks, the digits 0-9, and control characters. This set of 128 characters is also included in most other encodings. See also EBCDIC, encoding.

CDISC
See Clinical Data Interchange Standards Consortium.

CDISC ODM
See CDISC Operational Data Model.

CDISC Operational Data Model (CDISC ODM)
a vendor-neutral, platform-independent format that supports the electronic acquisition, exchange, and archival of clinical trials data for the medical and biopharmaceutical industries.

Clinical Data Interchange Standards Consortium (CDISC)
an organization that develops pharmaceutical industry standards. The standards support the electronic acquisition, exchange, and archival of clinical trials data and metadata for medical and biopharmaceutical product development. CDISC defines several data models for the interoperability of clinical data exchange.

Coordinated Universal Time (UTC)
by international agreement, the time standard by which clocks and time are regulated (for example, for computers and online services). Specifically, it is the local time at the prime meridian, which passes through Greenwich, England. UTC is the term that now supersedes Greenwich Mean Time.

data set
See SAS data set.

data view
See SAS data view.
Document Object Model (DOM)
a cross-platform and language-independent convention for representing and interacting with objects in documents. Traditionally, the DOM applies to HTML, XHTML, and XML documents.

document type definition (DTD)
a set of markup declarations that define a document type for a markup language (SGML, XML, or HTML). The DTD specifies how the tags should be interpreted by an application that displays or prints the documents.

DOM
See Document Object Model.

DTD
See document type definition.

EBCDIC (Extended Binary Coded Decimal Interchange Code)
a family of single-byte and multi-byte encodings for the representation of data on IBM mainframe and mid-range computers. See also ASCII, encoding.

encoding
a mapping of a coded character set to code values.

Extended Binary Coded Decimal Interchange Code
See EBCDIC.

Extensible Markup Language
See XML.

file reference
See fileref.

File Transfer Protocol (FTP)
a telecommunications protocol that is used for transferring files from one computer to another over a network.

fileref (file reference)
a name that is temporarily assigned to an external file or to an aggregate storage location such as a directory or a folder. The fileref identifies the file or the storage location to SAS. See also libref.

format
See SAS format.

FTP
See File Transfer Protocol.

graphics template
See ODS template.

hierarchical database
a database that is organized as a tree structure of segments. A DL/I database has a hierarchical data structure.

informat
See SAS informat.
key field
   See sequence field.

library reference
   See libref.

libref (library reference)
a SAS name that is associated with the location of a SAS library. For example, in the name MYLIB.MYFILE, MYLIB is the libref, and MYFILE is a file in the SAS library. See also SAS library.

markup language
   a set of codes that are embedded in text in order to define layout and certain content.

metadata
descriptive data about data that is stored and managed in a database, in order to facilitate access to captured and archived data for further use.

observation
   a row in a SAS data set. All of the data values in an observation are associated with a single entity such as a customer or a state. Each observation contains either one data value or a missing-value indicator for each variable.

ODS template (graphics template)
a description of how output should appear when it is formatted. ODS templates are stored as compiled entries in a template store (item store). Common template types include STATGRAPH, STYLE, CROSSTABS, TAGSET, and TABLE.

parsed character data (PCDATA)
a data definition that is used in XML to specify mixed content—that is, an element containing character data and child elements in arbitrary order and number of occurrences.

PCDATA
   See parsed character data.

SAS data file
   a type of SAS data set that contains data values as well as descriptor information that is associated with the data. The descriptor information includes information such as the data types and lengths of the variables, as well as the name of the engine that was used to create the data. See also SAS data set, SAS data view.

SAS data set (data set)
a file whose contents are in one of the native SAS file formats. There are two types of SAS data sets: SAS data files and SAS data views. See also SAS data file, SAS data view.

SAS data view (data view)
a type of SAS data set that retrieves data values from other files. A SAS data view contains only descriptor information such as the data types and lengths of the variables (columns) plus other information that is required for retrieving data values from other SAS data sets or from files that are stored in other software vendors' file formats.
SAS format (format)
a type of SAS language element that is used to write or display data values according to the data type: numeric, character, date, time, or timestamp.

SAS informat (informat)
a type of SAS language element that is used to read data values according to the data's type: numeric, character, date, time, or timestamp.

SAS library
one or more files that are defined, recognized, and accessible by SAS, and that are referenced and stored as a unit. Each file is a member of the library.

SAS variable (variable)
an identifier that is output as a column in a SAS data set or in a SAS data view. The data values for each variable describe a single characteristic for all observations (rows).

SAS XML Mapper
a graphical interface that can be used to create and modify XMLMaps for use by the SAS XMLV2 LIBNAME engine. The SAS XML Mapper analyzes the structure of an XML document and generates basic XML markup for the XMLMap.

SAX
See Simple API for XML.

sequence field (key field)
a field that identifies and provides access to segments in a database. It contains the record's key, which is located in the same position in each record of a key-sequenced data set.

Simple API for XML (SAX)
an event-based, sequential access parser API that was developed for XML documents by members of the XML-DEV group. SAX provides a mechanism for reading data from an XML document that is an alternative to that provided by the Document Object Model (DOM).

tagset
a template that defines how to create a type of markup language output from a SAS format. Tagsets produce markup output such as Hypertext Markup Language (HTML), Extensible Markup Language (XML), and LaTeX. See also markup language.

Uniform Resource Identifier
See URI.

Uniform Resource Locator
See URL.

URI (Uniform Resource Identifier)
a string that identifies resources such as files, images, and services on the World Wide Web. A URL is a type of URI.

URL (Uniform Resource Locator)
a character string that is used by a web browser or other software application to access or identify a resource on the internet or on an intranet. The resource could be a web page, an electronic image file, an audio file, a JavaServer page, or any other
type of electronic object. The full form of a URL specifies which communications
protocol to use for accessing the resource, as well as the directory path and file name
of the resource.

**UTC**

See *Coordinated Universal Time*.

**variable**

See *SAS variable*.

**W3C**

See *World Wide Web Consortium*.

**World Wide Web Consortium (W3C)**

an international community that develops open standards to ensure the long-term

**XML (Extensible Markup Language)**

a markup language that structures information by tagging it for content, meaning, or
use. Structured information contains both content (for example, words or numbers)
and an indication of what role the content plays. For example, content in a section
heading has a different meaning from content in a database table.

**XML namespace**

a W3C specification that provides a simple method for qualifying element and
attribute names used in Extensible Markup Language (XML) documents by
associating them with namespaces identified by URI references. XML namespaces
avoid element name conflicts. For example, if an XML document contains a
CUSTOMER element and a PRODUCT element, and both elements contain a nested
ID element, XML namespaces make each nested ID element unique.

**XML Path Language (XPath)**

a query language that describes how to locate and process items in an XML
document. An XPath specification assigns a path description, similar to what is used
in UNIX, to each element of the XML structure.

**XML schema**

defines the structure, content, and semantics of an XML document. An XML schema
is typically expressed in terms of constraints on the structure and content of the
document type.

**XMLMap file**

a file that contains XML tags that tell the SAS XMLV2 LIBNAME engine or the
SAS XML LIBNAME engine how to interpret an XML document.

**XPath**

See *XML Path Language*.

**XPath with Embedded Namespace Reference (XPathENR)**

an XMLMap element that is used by the XMLV2 engine for XML namespace
support.

**XPathENR**

See *XPath with Embedded Namespace Reference*. 
Index

A
ampersand
  importing XML documents with 11, 85
apostrophe (’)
  importing XML documents with 11, 85
ATTRIBUTE element 98
AUTOMAP= option
  LIBNAME statement 76

B
beginend= attribute
  DECREMENT-PATH element 111
INCREMENT-PATH element 109
RESET-PATH element 110
TABLE-END-PATH element 102

C
CDISC ODM markup 86
  CodeList elements 77
example document 137
exporting XML documents 126
importing XML documents 134
character data
  non-escaped 11, 85
character sets
  specifying 79
CHARMULTIPLIER= option
  LIBNAME statement 77
class= attribute
  COLUMN element 103
COLUMNS element 103
col values
  syntax 102
concatenated data
  avoiding 36
concatenated XML documents 81
  importing 132
count= attribute
  NAMESPACES element 97
create processing 5

D
data investigation 23
data sets, exporting as XML documents
  See exporting XML documents
data sets, importing XML documents as
  See importing XML documents
  See importing XML documents with XMLMap
data sets, importing XML documents as
  See importing XML documents
  See importing XML documents with XMLMap
DATASETS procedure 4
DATATYPE element 104
date values
  exporting XML documents containing 15
  ISO 8601 informats and formats for importing 39
datetime values
  exporting XML documents containing 15
  DECREMENT-PATH element 111
DEFAULT element 105
DESCRIPTION element 106
description= attribute
  SXLEMAP element 96
TABLE element 99
DOM application 6
double quotation marks
  importing XML documents with 11, 85

E
elements for columns 102
elements for exporting 98
elements for tables 99
elements for XML namespaces 96
enclosed attribute format 81
encoding 6, 82
engine names 75
ENUM element 105
errors
  when importing XML documents 7
exporting elements
  syntax 98
exporting XML documents 5, 15
CDISC ODM markup 126
customized tagset for 60
date, time, and datetime values 15
for Oracle 119
metadata information in separate file 16, 84
numeric values 121
exporting XML documents with
XMLMap 55
hierarchical structure 55
external files
for metadata-related information 85

F
filerefs 83
   URL access method for referencing 43
format catalog
   libref for 78
   replacing format entries 78
FORMAT element 106
FORMATACTIVE= option
   LIBNAME statement 77
FORMATLIBRARY= option
   LIBNAME statement 78
FORMATNOREPLACE= option
   LIBNAME statement 78
formats
   importing dates 39

G
generated numeric keys
   including key field with 33
generating XMLMap 49
GENERIC markup 70, 86
exporting XML documents containing
date, time, and datetime values 15
importing XML documents 9
physical structure for importing XML
documents 20

H
HEADING element 98
hierarchical data
   importing as related data sets 29
hierarchical structure
   exporting XML documents with 55

I
id= attribute
   NS element 97
importing XML documents 4, 9

K
key field
   including with generated numeric keys 33

L
LENGTH element 106
LIBNAME statement 69
assigning a libref 4
engine names 7, 69, 75
options 70
required arguments 75
syntax 73
librefs 75
assigning 4
for format catalog 78
location path, specifying on PATH element 44

M
map files
See XMLMap
markup languages
See tagsets
document bar
SAS XML Mapper 114
metadata
exporting XML documents with separate metadata 16, 84
external file for 85
Microsoft Access documents
importing 128
MSACCESS markup 86
importing XML documents 128

N
name= attribute
ATTRIBUTE element 98
COLUMN element 103
SXLEMAP element 96
TABLE element 99
TABLEREF element 98
namespace elements
importing XML documents with XMLMap 47
NAMESPACEs element 96
ndec= attribute
FORMAT element 106
INFORMAT element 106
non-escaped character data 11, 85
NS element 97
numeric keys
including key field with 33
numeric values 81
exporting 121
importing XML documents with 127

O
observation boundary
avoiding concatenated data 36
ODS MARKUP destination
XMLV2 engine versus 7
ODSCHARET= option
LIBNAME statement 79
ODSRECSEP= option
LIBNAME statement 79
ODSTRANTAB= option
LIBNAME statement 80
one-to-many relationship 29
open element format 81
Oracle
exporting XML documents for 119
ORACLE markup 87
OUTPUT element 98
output files
translation tables for 80
output processing 5
overriding tagsets 80

P
PATH element 107
specifying location path on 44
physical structure
importing XML documents with GENERIC markup 20
prefix= attribute
NS element 97
PREFIXATTRIBUTES= option
LIBNAME statement 80

R
read processing 5
record separators
XML documents 79
RESET-PATH element 110
retain= attribute
COLUMN element 103

S
SAS processing
supported by XMLV2 engine 5
SAS XML Mapper 113
generating and updating XMLMap 113
installing 115
menu bar 114
starting 115
toolbar 114
windows 114
SAX application 6
sequential access engine 6
Simple API for XML (SAX) 6
single quotation mark (')
importing XML documents with 11, 85
special characters
importing XML documents with 11, 85
starting SAS XML Mapper 115
SXLEMAP element
syntax 95
syntax= attribute
  DECREMENT-PATH element 111
  INCREMENT-PATH element 109
  PATH element 107
  RESET-PATH element 110
  TABLE-END-PATH element 101
  TABLE-PATH element 100

T
  TABLE element 99
  table elements
    syntax 99
  TABLE-DESCRIPTION element 102
  TABLE-END-PATH element 100
  TABLE-PATH element 99
  TABLEREF element 98
  TAGSET= option
    LIBNAME statement 80
tagsets 59
customized 59
customized, defining with TEMPLATE procedure 60
customized, exporting XML documents 60
  overriding 80
  TEMPLATE procedure 59
defining customized tagsets 60
time values
  exporting XML documents containing 15
  importing with time zone 40
time zones
  importing time values with 40
tool bar
    SAS XML Mapper 114
  transferring XML documents across environments 6
translation tables
  for output files 80
  TYPE element 104

U
  update processing 5
  updating XMLMap 113
  URL access method
    referencing filerefs 43

V
  validating XML documents 6
  value= attribute
    ATTRIBUTE element 98
  version= attribute
  SXLEMAP element 95

W
  W3C specifications 11, 70
  width= attribute
    FORMAT element 106
    INFORMAT element 106
  windows
    SAS XML Mapper 114

X
  XML documents 3
    CDISC ODM format 137
    concatenated 81
    not in required physical structure 23
    record separators 79
    transferring across environments 6
    validating 6
  XML documents, exporting
    See exporting XML documents
  XML documents, importing
    See importing XML documents
    See importing XML documents with XMLMap
  XML Mapper
    See SAS XML Mapper
  XML markup 83, 86
  XML namespace elements
    importing XML documents with XMLMap 47
    syntax 96
  XML window 114
  XMLCONCATENATE= option
    LIBNAME statement 81, 132
  XMLDATAFORM= option
    LIBNAME statement 81
  XMLDOUBLE= option
    LIBNAME statement 81, 127
  XMLENCODING= option
    LIBNAME statement 81
  XMLFILEREF= option
    LIBNAME statement 83
  XMLMap
    See also exporting XML documents with XMLMap
    See also importing XML documents with XMLMap
column elements 102
  comparing functionality 91
  exporting elements 98
  generating with AUTOMAP= option 76
generating with SAS XML Mapper 113
SXLEMAP element 95
syntax 83
table elements 99
updating with SAS XML Mapper 113
versions 70
XML namespace elements 96
XMLMAP markup 87
XMLMap window 114
XMLMAP= option
   LIBNAME statement 83
XMLMETA= option
   LIBNAME statement 16, 84
XMLPROCESS= option
   LIBNAME statement 11, 85
XMLSCHEMA= option
   LIBNAME statement 16, 85
XMLTYPE= option
   LIBNAME statement 86
XMLV2 and XML engine comparison 7, 69
XPath attribute
   DECREMENT-PATH element 111
   PATH element 107
   RESET-PATH element 110
   TABLE-END-PATH element 101
   TABLE-PATH element 100
XPathENR attribute
   DECREMENT-PATH element 111
   PATH element 107
   RESET-PATH element 110
   TABLE-END-PATH element 101
   TABLE-PATH element 100
Gain Greater Insight into Your SAS® Software with SAS Books.

Discover all that you need on your journey to knowledge and empowerment.