Reading Microsoft Sharepoint List Data Using SAS

Harry Droogendyk
Stratia Consulting Inc
harry@stratia.ca
What is SharePoint?

- Microsoft “web application platform”
- Enterprise collaboration software
  - Content management, portal
  - Forms, data collection
  - Document management (versioning)
- Easily integrated with MS Office
What is Sharepoint?

- server-based technology
  - browser access, thin client
  - search
  - communities
- security model
  - owners, contributors, consumers
  - distributed administration
What is Sharepoint?
What is a Sharepoint List?
What is a Sharepoint List?
Adding Sharepoint List Data
Adding Sharepoint List Data
## Editing Sharepoint List Data

### H&A and Travel Absenteeism and Vacation

<table>
<thead>
<tr>
<th>Item</th>
<th>Start Date</th>
<th>End Date</th>
<th>Emploee</th>
<th>Staffing Code</th>
<th>Bank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modify Record</td>
<td>6/4/2012</td>
<td>6/6/2012</td>
<td></td>
<td>Vacation</td>
<td></td>
</tr>
<tr>
<td>Modify Record</td>
<td>5/26/2012</td>
<td>6/15/2012</td>
<td></td>
<td>Vacation</td>
<td>1.0</td>
</tr>
<tr>
<td>Modify Record</td>
<td>6/4/2012</td>
<td>6/5/2012</td>
<td></td>
<td>Sick</td>
<td>1.0</td>
</tr>
<tr>
<td>Modify Record</td>
<td>6/4/2012</td>
<td>6/5/2012</td>
<td></td>
<td>Sick</td>
<td>1.0</td>
</tr>
<tr>
<td>Modify Record</td>
<td>5/31/2012</td>
<td>6/1/2012</td>
<td></td>
<td>Sick</td>
<td>1.0</td>
</tr>
<tr>
<td>Modify Record</td>
<td>6/5/2012</td>
<td>6/11/2012</td>
<td></td>
<td>Sick</td>
<td>1.0</td>
</tr>
<tr>
<td>Modify Record</td>
<td>6/7/2012</td>
<td>6/18/2012</td>
<td></td>
<td>Sick</td>
<td>1.0</td>
</tr>
<tr>
<td>Modify Record</td>
<td>6/12/2012</td>
<td>6/12/2012</td>
<td></td>
<td>Sick</td>
<td>Blank</td>
</tr>
<tr>
<td>Modify Record</td>
<td>6/11/2012</td>
<td>6/11/2012</td>
<td></td>
<td>Sick</td>
<td>1.0</td>
</tr>
<tr>
<td>Modify Record</td>
<td>6/18/2012</td>
<td>6/22/2012</td>
<td></td>
<td>Vacation</td>
<td>1.0</td>
</tr>
<tr>
<td>Modify Record</td>
<td>6/12/2012</td>
<td>6/12/2012</td>
<td></td>
<td>Vacation</td>
<td>1.0</td>
</tr>
<tr>
<td>Modify Record</td>
<td>6/11/2012</td>
<td>6/11/2012</td>
<td></td>
<td>Family Responsibility</td>
<td>1.0</td>
</tr>
</tbody>
</table>
What is a “Web Service”?

- WWW machine to machine communication
- REST
  - Representational State Transfer
- SOAP
  - Simple Object Access Protocol
- blah blah blah
What is a “Web Service”?

- API
  - Application Programming Interface
- Communication protocol
  - invoked with specific syntax, parms
  - returns data / functionality
- HTTP request messages
  - URL and parms
  - XML returned
Sharepoint “Web Service”

- `owssvr.dll`
  - Microsoft web service
  - appropriate parms
- `http://server.../_vti_bin/owssvr.dll?cmd=Display&list=\text{list~GUID}\&view=\text{view~GUID}\&XMLDATA=TRUE&FilterField1=\text{some~field}\&FilterValue1=\text{some~value}`
Upfront Work

- define a view
  - return all rows
  - optional filter / sort criteria
- list GUID (Globally Unique IDentifier)
- view GUID
Define a Sharepoint List View

### Actions
- **Start Date**
  - 5/22/2012
- **5/28/2012**
- **List Settings**
  - Manage settings such as permissions, columns, views, and policy.

### Settings
- **Create Column**
  - Add a column to store additional information about each item.
- **Create View**
  - Create a view to select columns, filters, and other display settings.

### View Name:
- HA Travel Absent/Vac All rows

### Web address of this view:
- http://.../H08_WFMProductionSupportBusiness/MOR/Lists/HA and Travel Absenteeism and Vacation/HA Travel AbsentVac A.aspx

- **Make this the default view**
  - (Applies to public views only)

### Item Limit
- Use an item limit to limit the amount of data that is returned to users of this view. You can either make this an absolute limit, or allow users to view all the items in the list in batches of the specified size.

- **Number of items to display:** 214748
  - Display items in batches of the specified size.
  - Limit the total number of items returned to the specified amount.
List GUID

- grab the URL when viewing list in browser

http://server/_layouts/listedit.aspx?List=%7BF6A1B978%2D9B0F%2D4046%2DB9E5%2D8CA44086163A%7D
View GUID

- open the view in the browser
- select “Edit in Datasheet”
grab the URL when viewing datasheet in browser

http://server/Lists/HA%20and%20Travel%20Absenteeism%20and%20Vacation/HA%20Travel%20AbsentVac%20All%20rows.aspx?ShowInGrid=True&View=%7B5CB7A2C4%2DA123%2D4EC4%2DBB9C%2D515F563DE94E%7D
%let url = http://server/_vti_bin/owssvr.dll;
%let ListID = {F6A1B978-9B0F-4046-B9E5-8CA44086163A};
%let ViewID = {5CB7A2C4-A123-4EC4-BB9C-515F563DE94E};

%let request = %nrstr(Cmd=Display&List=)&ListID
               %nrstr(&View=)&ViewID
               %nrstr(&XMLDATA=TRUE);

* filters only test equality ;

%let filter = %nrstr(&FilterField1=Modified
               &FilterValue1)=mm/dd/yyyy;
_count file references for HTTP request and XML output location

%let dir = &root\&env\xml;
%let xmlfile = ha_data.xml;

filename listout "&dir\&xmlfile";
filename request temp;
SAS Reading Sharepoint

write the request data

```sas
data _null_;  
  file request; 
  
  length request $256; 
  request = "&request&filter"; 
  put request; 
  putlog 'Request: ' request; 
  
  stop; 
  run;
```
issue the HTTP request to the web service

```sas
proc http
    in = request
    out = listout
    url = "&url"
    method = "get"
;
run;
```

filerefs and &url established earlier
SAS Reading Sharepoint

- XML schema info

```xml
<xml xmlns:s="uuid:BDC6E3F0-6DA3-11d1-A2A3-00AA00C14882"
      xmlns:dt="uuid:C2F41010-65B3-11d1-A29F-00AA00C14882"
      xmlns:rs="urn:schemas-microsoft-com:rowset"
      xmlns:z="#RowsetSchema">
  <s:Schema id="RowsetSchema">
    <s:ElementType name="row" content="eltOnly" rs:CommandTimeout="30">
      <s:AttributeType name="ows_LinkTitle" rs:name="Edit Item" rs:number="1">
        <s:datatype dt:type="string" dt:maxLength="512" />
      </s:AttributeType>
      <s:AttributeType name="ows_Start_x0020_Date" rs:name="Start Date" rs:number="2">
        <s:datatype dt:type="datetime" dt:maxLength="8" />
      </s:AttributeType>
      <s:AttributeType name="ows_End_x0020_Date" rs:name="End Date" rs:number="3">
        <s:datatype dt:type="datetime" dt:maxLength="8" />
      </s:AttributeType>
    </s:ElementType>
  </s:Schema>
</xml>
```
SAS Reading Sharepoint

- Sharepoint List data in XML format

```xml
<rs:data>
  <z:row ows_LinkTitle="Modify Record" ows_Start_x0020_Date="2012-09-14 00:00:00" ows_End_x0020_Date="2012-09-14 00:00:00"
    ows_Employee="355;#Employee1" ows_Staffing_x0020_Code="Vacation"
    ows_Hours="7.50000000000000" ows_Shift_x0020_Length="7.5"
    ows_Created="2012-07-10 12:45:31" ows_Author="175;#Userid12"
    ows_Modified="2012-10-22 08:32:57" ows_Editor="176;#Userid13"
    ows_Note="<div></div>" />
  <z:row ows_LinkTitle="Modify Record" ows_Start_x0020_Date="2012-04-09 00:00:00" ows_End_x0020_Date="2012-04-13 00:00:00"
    ows_Employee="392;#Employee2" ows_Staffing_x0020_Code="Purchase Vacation"
    ows_Hours="37.50000000000000" ows_Shift_x0020_Length="7.5"
    ows_Created="2012-07-10 15:05:28" ows_Author="176;#Userid13"
    ows_Modified="2012-10-22 16:24:18" ows_Editor="101;#Userid12"
    ows_Note="<div></div>" />
</rs:data>
```
SAS and XML

- SAS cannot read Sharepoint XML directly
- must create an XML map
open the XML file just created
SAS XML Mapper

XML source:

```xml
<xml xmlns:s='uuid:BD06E3F0-5DA3-11d1-A2A3-00A00C14882'
     xmlns:r='urn:schema:microsoft-com:rowset'
     xmlns:m='Rowsetschema'>
  <s:Schema id='RowSetSchema'>
    <s:ElementType name='row' content='eltOnly'/>
  </s:Schema>
</xml>
```

XML file loaded: ha_absenteeism_data.xml
SAS XML Mapper

XML elements:
- xml
  - s:Schema
    - rs:data
SAS XML Mapper
XML map is only more XML

```xml
<COLUMN name="ows_Start_x0020_Date">
  <PATH syntax="XPath">/xml/rs:data/z:row/@ows_Start_x0020_Date</PATH>
  <TYPE>numeric</TYPE>
  <DATATYPE>datetime</DATATYPE>
  <FORMAT width="20">DATETIME</FORMAT>
  <INFORMAT width="19">IS8601DT</INFORMAT>
</COLUMN>
```
Using the XML Map

```plaintext
filename HAabsent "&dir\&xmlfile";
filename sxlemap "&root\&env\HA.map";
libname HAabsent xml xmlmap=sxlemap
access=readonly;

data mylib.sp_list_data;
  set HAabsent.z_row ;
run;
```
Post Processing

data sp_list_data_massaged ( drop = ows_: );
  if 0 then set mylib.sp_list_data_final;
  set mylib.sp_list_data
    ( drop = rs_data_ordinal z_row_ordinal ows_LinkTitle );

  start_dt = datepart(ows_start_x0020_date);
  end_dt = datepart(ows_end_x0020_date);

  employee_nm = scan(scan(ows_employee,-1,'#'),1,'-');
  employee_no = scan(scan(ows_employee,-1,'#'),-1,'-');

  absent_type = ows_staffing_x0020_code;
  absent_hours= ows_hours;
  shift_length= input(ows_shift_x0020_length,?? 32.);
  sick_days = input(scan(ows_sick_x0020_days,-1,'#'),32.);
  occur_type = ows_occurrence_x0020_type;

run;
Final Thoughts

- only runs in Windows
  - secret, double Windows handshake
  - Unix access possible in 9.4
- Sharepoint lists are slow
  - not for large data
- HTTP request construct is picky
- error messages aren’t helpful
Questions?

Harry Droogendyk
Stratia Consulting Inc.

harry@stratia.ca