

## Paper 37-28

## Using a Dynamic SAS/IntrNet® Application to Create Statistical Comparison Reports and Download as SAS® Data Sets.

John Owusu-Djamboe, US Census Bureau, Washington DC

### ABSTRACT

As part of an internal review process by analysts working on the American Community Survey at the US Census Bureau, current year sample estimate values are compared to previous year estimate values to test the difference in proportions between two independent samples and their statistical significance in a distribution. Detailed and summary comparison statistical reports are created and further analysis done. Since these reports are calculated using various complex statistical tests, analysts want to be able to download the reports as SAS datasets and perform further analysis by running SAS programs against these reports. Dynamic SAS/IntrNet® Application provides this capability. It also allows the user to select multiple parameters and multiple values within a parameter concurrently. This paper describes a method for downloading HTML reports generated within a SAS/IntrNet application directly into SAS data sets, eliminating the intermediate step of downloading to a spreadsheet before reading it back into SAS. This paper will also describe the necessary SAS code and the application utilities used to accomplish this task.

### INTRODUCTION

The Continuous Measurements Office, a branch of the Demographic Survey Division at the Census Bureau is responsible for the American Community Survey (ACS) that provides accurate and up-to-date profiles of American communities on a yearly basis, instead of every ten years as the decennial census currently provides. The Tabulations group creates tables and profiles of demographic, social, economic, and housing estimates that are published on the Census web site for use by local communities, as well as by state and federal agencies. This group also produces comparison reports of profile tables of previous and current year estimates in the data review process. These reports include the results of statistical tests that reveal significant difference in an estimate or in a distribution from one year to another. Differences may be due to a problem in collecting and processing the data or they may simply reflect the change in responses to the survey.

### SAS/IntrNet Overview

The SAS system is accessed from the web through SAS/IntrNet Application Methods: data services, which allow users to update and query data from a Web browser using SAS SQL protocol and compute services, which allow users to execute SAS programs from a Web browser. SAS/IntrNet Application Dispatcher is a compute service that allows users to pass parameter selections from a web page to the appropriate SAS program that executes on a server, sending HTML results back to the user's Web browser. This paper is limited to applications intended to run with the use of the Application Dispatcher over an intranet or internet. Other SAS/IntrNet products will not be discussed here.

### PROBLEM

The analysts who review the data want to be able to download the reports directly as SAS data sets without going through the intermediate process of downloading to a spreadsheet before writing code to create SAS data sets. While SAS has provided a proprietary code on SAS web site and also in the SAS Institutes Instructor Based Training Course Notes manuals for downloading HTML reports as CSV, PDF or RTF files, there is no such package for downloading as SAS data sets. The SAS/IntrNet

Application source data sets fsauth.sas7bdat and fshead.sas7bdat, that determine what kind of files that can be downloaded, do not provide for download as a SAS data set.

### SOLUTION

There are various techniques that can be used to download an HTML report as a SAS data set. This paper focuses on using the FILESRV Dispatch utility to solve this problem. This technique involves making changes to source files and scripts used by the SAS/IntrNet Application software to determine file extensions that can be downloaded. The development methodology is outlined below.

### DEVELOPMENT METHODOLOGY

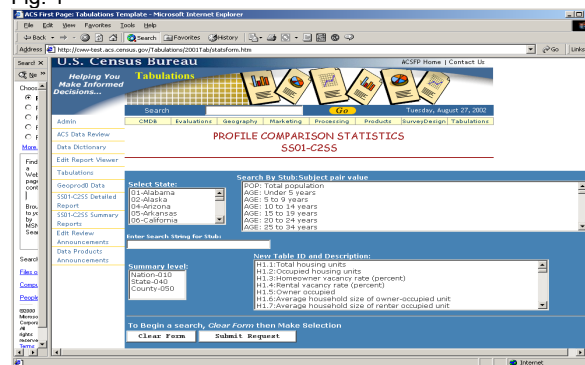
There are three development tasks in this application. Each one depends on the other for the success of the application.

- Design/layout of the web page.
- Update of two datasets: (1). The Authorization dataset, fsauth.sas7bdat, used to determine which files can be served and also used to set patterns (or masks) for the files that can be served, and (2). The HTTP Headers dataset, fshead.sas7bdat, used to determine which HTTP and MIME headers are served with the file and the mode in which the file is open. The FILESRV Autocall Macro is then invoked.
- Modification of the exec line in the start.pl script with an "insert SASHELP 'path to your dataset' " code, compilation and restart of the Application Server.

### Web Page layout/design

The figure below shows the web page as the user sees it.

Fig. 1



The screen has four select boxes and a text box to enter a search string for the "stub" (descriptor of a data cell). The user can request output by selecting single or multiple values from a select box or request multiple combinations from all select boxes. When the user enters text in the "Enter search string" box, no value is selected from the "Search by stub" box because you want to retrieve a stub document by pattern matching. Even if you select a value from the "Search by stub" box, the SAS program ignores it. When the Submit button is clicked the values selected will be passed to the SAS program as macro variables.

Part of the code that produced the Web form is shown below.

```

<form method=get
action=".../sasscripts/broker.exe?"><table><tr><td colspan="2"
rowspan="1">
<b><font color="white">Select State:</font>
</b><select name=code multiple size=5>
<option value=01>01-Alabama
<option value=01>02-Alaska
<option value=01>03-Arizona
</select></td>
<td colspan=2><b><font color=white>Search By Stub:Subject
pair value</font></b>
<select name=stibt multiple size=7>
<option value="POP:Total population">POP: Total population
<option value="AGE:Under 5 years">AGE: Under 5 years
<option value="AGE:5 to 9 years">AGE: 5 to 9 years
</select></td></tr></Table><Table><tr>
<td><b><font color="white">Summary
level:</font></b><br><select name=lev multiple size=4><option
value=010>Nation-010
value=040>State-040
</td><td colspan=2><b><fontcolor="white">
New Table ID and Description:</font></b><br>
<select name=descr multiple size=7>
<option value="H1.1:Total housing units">H1.1:Total housing
units
</select></td></tr><tr><input class="LargeButtons" type=reset
value="Clear Form">
<input class="LargeButtons" type=submit value="Submit
Request">
<input type=hidden name=_debug value=0>
<input type=hidden name=goback value=1>
<input type=hidden name=_service value=cms05><input
type=hidden name=_program value=geotmp.geostats.sas>
</form>

```

In the HTML code above is a hidden form variable named `_program`. This variable tells our broker CGI which program to run. In this case the program is `geostats.sas`. The `"value=geotmp.geostats.sas"` string contains `"geotmp"` which is a file reference pointing to the directory containing the SAS program. This file reference is allocated by the APPSTART.SAS program, which is not discussed here.

## SAS PROGRAM

The SAS program contains some macro variables that I want to mention. Since the stubs contain special characters that require special treatment to resolve, I used SAS functions to ensure that these variables resolve as expected.

The macro variables that are passed from the Web form to the program include `stibt`, which contain values shown below. These special characters can cause serious problems and may not resolve. Examples of values of `stibt` are shown below:

```

<select name=stibt multiple size=7>
<option value="COM2:Car,truck,or van-drove alone">COM2:Car,
truck, or van -- drove alone
<option value="COM2:Car, truck, or van -- carpoled">COM2:
Car, truck, or van -- carpoled
<option value="HHINC:$10,000 to $14,999">HHINC:$10,000 to
$14999

```

The Application Server is set up (by default) to hide certain special characters such as quotation marks and semicolons when passed into parameters. Also in that list of characters are `&` and `%`. To handle those characters properly, one can restore those characters with the following `%LET` statement after the `%GLOBAL` statement:

```

%let stibt=%qsysfunc(appsrv_unsafe(stibt));
%let descr=%qsysfunc(appsrv_unsafe(descr));

```

## MULTI-SELECT

The user has the option of selecting multiple values from multiple parameters. If we select two state values (Alaska, Utah) from the

state box and set `_debug=131`, we will see that the original macro variable for `state`, `&code`, equal to the first selection, "Alaska", but now there are three other macro variables: `&code0`, `&code1`, `&code2`. `&code0` contains the number of state variables passed while `&code1` and `&code2` contain the value of the first and second state. The open code below shows how the SAS program uses these macro variables to retrieve the data.

```

%if %superq(stibt) ne %then %do;
%let stibtxt=; %end;
%if %superq(descry) ne %then %do;
%let stibtxt=;
%end;
%if (&code ne and %superq(stibt) ne and &lev= and
%superq(stibtxt) = and %superq(descr) = ) %then %do;
proc sql;
create table stgeo as
select
geoname,st,stibt,sumlevel,variable,tbldescr,p_work,z_work,
newtbl,cell,numcells
from stt.stats_srtid
where st in
( %if &code and &code0= %then %do; "&code" %end; %else
%if &code0 %then %do; %do i=1 %to &code0;
"&&code&i" %if &i ne &code0 %then ;; %end; %end;) and
variable in (%if &stibt0= %then
%do; "%scan(%nrbrquote(&stibt),1,;)" %end;%else %if
&stibt0 ne %then %do; %do i=1 %to &stibt0;
"%scan(%nrbrquote(&&stibt&i),1,;)" %if &i ne &stibt0 %then
;,%end;%end;); and stibt in (%if %superq(stibt) ne and &stibt0=
%then %do; "%scan(%nrbrquote(&stibt),2,;)" %end;%else %if
&stibt0 ne %then %do; %do i=1 %to &stibt0;
"%scan(%nrbrquote(&&stibt&i),2,)" %if &i ne &stibt0 %then ;;
%end;%end;); quit;

```

The report that is returned to the Web by the program is shown in Fig. 2. You have the option to download to a spreadsheet (CSV file) or as SAS data set. This paper is focused on downloading such a report as a SAS data set.

Fig. 2

| geoname | State | Stubs          | sumlevel | Subject | Table Description | p-value   | z-score   | New Table ID | New Lineno | numcells |
|---------|-------|----------------|----------|---------|-------------------|-----------|-----------|--------------|------------|----------|
| Alaska  | 02    | 5 to 9 years   | 040      | AGE     | Age               | 0.8531223 | -0.185136 | P1.3         | 2          | 13       |
| Arizona | 04    | 5 to 9 years   | 040      | AGE     | Age               | 0.1173004 | 1.5662061 | P1.3         | 2          | 13       |
| Alaska  | 02    | 10 to 14 years | 040      | AGE     | Age               | 0.5716771 | -0.565583 | P1.3         | 3          | 13       |
| Arizona | 04    | 10 to 14 years | 040      | AGE     | Age               | 0.0029836 | 2.9694184 | P1.3         | 3          | 13       |

The code behind the "Download as SAS Data" button is outlined below:

```

data _null_;
file _webout;

```

```

put '<form action="" "&_url" "">';
put '<input type=hidden name=_service '
    " value=&_service>";
put '<input type=hidden name=_program '
    "value=&_pgmlib..download2sas.sas>";
put '<input type=hidden name=_sessionid '
    " value=&_sessionid>";
put '<input type=submit '
    ' value="Download as SAS Data">';
put '</form>';
run;

```

The proprietary code behind the "Download as Spreadsheet" button is outlined below:

```

Data _null_;
File _webout;
put '<form action="" "&&_url" "">';
put '<input type=hidden name=_service' "value=&_service>";
put '<input type=hidden name=_program' "value =
&_pgmlib..csvroutes.sas>";
put '<input type=hidden name =_sessionid' value="
&_sessionid>";
put '<input type=hidden name= csvfile value=' "SS01-C2SS>";
put '<input type=submit' 'value="Download as Spreadsheet">';

```

## DOWNLOAD AS SAS DATA SET

Downloading into SAS data set can be challenging because there is no proprietary code to use like the one for spreadsheet shown below:

```

/* SAS Code for CSV Download */
/*csvroutes.sas */
%global csvfile;
/* send CSV output */
%ds2csv(data=save.csvdata,conttype=y,contdisp=y,
savefile=&csvfile..csv,csvref=_webout,runmode=s,
openmode=replace);

```

To download as SAS data set through SAS/IntrNet when using the SAVED library, the download program must use the same Dispatcher session as the program that creates the data set. A session is created with a call to the APPSERV\_SESSION function as shown below.

```
%let rc =%sysfunc(appsrv_session(create));
```

The code below will create a data set in the SAVED library that will be used in the download routine.

```
/* create dataset saved with the session and download as SAS
Dataset */
```

```

data save.sasdat;
set stgeo; /* Data set that created the web report */
run;

```

## THE FILESRV UTILITY

The Filesrv Dispatcher utility is used to download the HTML report as SAS data set. In addition, there are issues related to the HTTP headers like Content-type to make the browser understand what it is receiving. And, in particular, if the file is downloaded using Internet Explorer, the (default) filename will contain characters (left and right square brackets – [ and ]) that are not valid within SAS. So this utility must be used carefully.

The FILESRV program is documented very well on the SAS Web site: <http://www.sas.com/rnd/web/intrnet/filesrv/filesrv.html>. You can run the Filesrv utility from an address line on a browser window as shown in the example below:

<http://pp224.acs.gov/cgi-bin/broker.exe? service=default& program=sashelp.webprog.file>

[srv.scl& filetype=e& file=/tab3/2001/pdata/reports.sas7bdat& deb ug=0](http://www.sas.com/rnd/web/intrnet/filesrv/examds.html)

## OR

invoke as a macro from a SAS program as shown below:

```

%let extfile=/tab3/2000/pdata/reports.sas7bdat
%filesrv(file=&extfile, filetype=e);

```

where reports.sas7bdat is the dataset that you want to download. Along with the FILESRV program you must also update two files, fsauth.sas7bdat and fshead.sas7bdat, as mentioned earlier.

## AUTHORIZATION DATA SET.

The authorization data set is FSAUTH.sas7bdat. It is located in the SASHELP library. It is provided with the Filesrv program. It determines which files can be downloaded. An example of the authorization data set is shown in the table below. It is documented on SAS web site at:

<http://www.sas.com/rnd/web/intrnet/filesrv/examds.html>

As you can see in the table a user has authorization to request a file from the */\*/\*public\_html/\** tree. If a user requests a file named people.bat from this tree, the file is not served because the HTTP headers data set does not list a bat extension. To customize your application the value of variable PATH has to be modified to reference the location where the data in the SAVED directory will be stored for users to download to their local machine.

| FILETYPE | PATH                     |
|----------|--------------------------|
| E        | ~/":"/":~/public_html/": |
| C        | ~"sashelp.":             |
| C        | ~"permdata.":            |

## THE HTTP HEADER DATA SET.

The HTTP header data set is FSHEAD.sas7bdat. It is also provided with the Filesrv program. It determines which HTTP and MIME headers will be served with the file and the mode in which the file is open. An example of the HTTP header data set is shown in the table below. It is documented on SAS web site at: <http://www.sas.com/rnd/intrnet/filesrv/examds.html>

| EXT | DATATYPE | FILETYPE | HEADER                                |
|-----|----------|----------|---------------------------------------|
| Bmp | b        | e        | Content-type:image/bmpc               |
| csv | t        | e        | Content-type:application/vdn.ms-excel |
| doc | b        | e        | Content-type:application/msword       |

In the above table, external files with the extensions csv are listed as text files and are served with headers that invoke the appropriate helper applications, if they are available on the client. Similarly external files with extensions bmp and doc are listed as binary files. When you open FSHEAD.sas7bdat (SASHELP.FSHEAD), you will find that sas7bdat entry can not be served because the HTTP headers data set does not have an entry for this extension type. This means that a SAS8 dataset with extension ".sas7bdat" cannot be served with headers that invoke the appropriate helper applications, if they are available on the client. You can not download as SAS data set when you reference SASHELP.FSAUTH and SASHELP.FSHEAD in your application.

## MODIFICATION

The two data sets FSAUTH.sas7bdat and FSHEAD.sas7bdat have to be modified to include entries for sas7bdat extensions with a binary data type. If your Application Server is running on a Unix platform, as in my case, these datasets are located at the

SASHELP directory, ie, `/usr/local/sas8/sashelp/`. It is not recommended to change any data set in this library. I copied both of these data sets from SASHELP to another location and modified them there.

You can use a data step or PROC SQL to do the update.

The following PROC SQL code inserts values in your copy of FSAUTH.sas7bdat.

```
Libname test '/tab3/2001/papp';
Proc sql;
Insert into test.fsauth
Set Filetype='e',
Path=""'/tab3/2001/pdata'";
quit;
```

Insert the following values into your copy of FSHEAD.sas7bdat.

```
Proc sql;
Insert into test.fshead
Set
filetype='e',
datatype='b',
ext='sas7bdat',
header='Content-type: application/x-sas-data'
```

```
Set
filetype='e',
datatype='b',
ext='sas7bdat',
header='Content-disposition:attachment;
filename=temp.sas7bdat';
```

The filename, temp.sas7bdat is the default name for the dataset created after download. You can change it to a name of your choice.

We are not done yet. The Application Server does not know where these updated data sets are since we will not be using the version in the /sashelp directory but our modified versions. To point the Application Server to these updated files we have to modify the script start.pl. Your Unix administrator can do this. Make sure you stop the Dispatcher before updating the script.

## START.PL script

This script has to be modified to run your copy of FSAUTH.sas7bdat and FSHEAD.sas7bdat.

```
#!/usr/local/sas8/install/perl/bin/perl5
#This code will start each of the servers in the service as
#daemon processes. The double fork will cause the servers to
#reparent to process id 1 and therefore they will not be killed if
#the shell running this script exists. There is no need to run this
#script in the background or with nohup.
#Make list of server ports for this service.
@ports=qw(5001);
#Now loop over each server port.
for $port (@ports){
  unless (fork) {
    #this is the first child
    unless (fork) {
      #this is the second child
      sleep 1 until getppid == 1;
      #now suppress the display, change to the tmp directory
      and exec the server
      undef($ENV{'DISPLAY'});
      chdir "/usr/local/sas8/internet/default/tmp";
      exec "/usr/local/sas8/sas ../appstart_$port.sas -insert SASHELP
'/tab3/2001/pdata' -noterminal -rsasuser -log ../appstart_$port.log
> ../appstart_$port.out 2>&1";
      exit 0; } exit 0; }
  wait; #wait for the first child to be reaped }
```

The part that you insert into the script is: “`-insert SASHELP '/tab3/2001/pdata'`”

The directory “`/tab3/2001/pdata`” will contain the data that the users will download.

Stop the Application Server. Then restart the application server to allow the newly updated files FSAUTH.sas7bdat and FSHEAD.sas7bad to take effect as well as the start.pl script.

## Download2sas.sas

The last task is to write the program Download2sas.sas that invokes the FILESRV Autocall Macro. Code is show below.

```
/* download2sas.sas */

%macro load;
libname dload v8 '/tab3/2001/pdata';
data dload.sasdat;
set save.sasdat; /* dataset was saved in the save directory */
run;

/* Invoke the FILESRV Utility in a Macro */

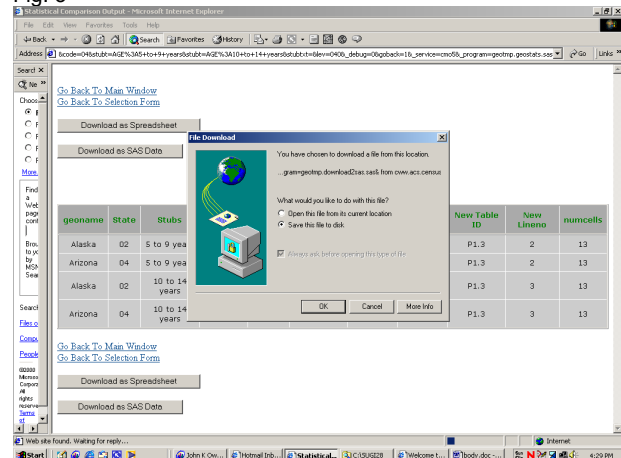
%let extfile= /tab3/2001/pdata/sasdat.sas7bdat;
%filesrv(file=&extfile,filetyp=e);

proc datasets library=dload;
delete sasdat; /* delete dataset after download */
run;
%mend;
%load;
```

## DOWNLOAD PROCEDURE

When you click on the ‘Download as SAS data’ button on the report window in Fig. 2, the File download window pops up as show in Fig. 3 below.

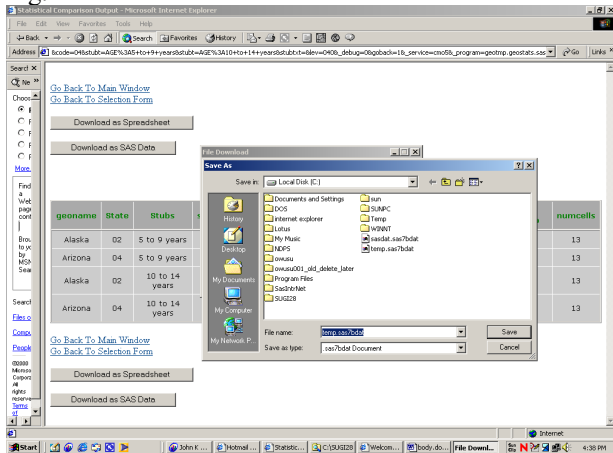
Fig. 3



You have the option to open the file in its current location or Save file to disk. Opening the dataset in the current location can be challenging if you don't have SAS software installed locally on your machine. If SAS is on the network and you are accessing the executable you may not have all the parts you need to open successfully. In my case, SAS software is not installed on my Ultra5 Sun WorkStation so I saved the file to disk. When you click on the OK button in Fig. 3, the Save As window pops up as

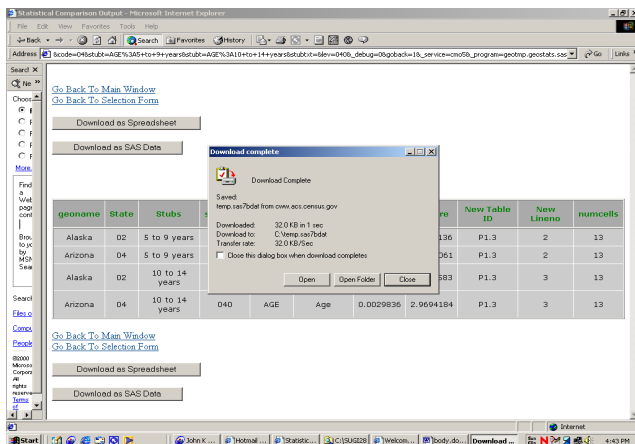
shown in Fig. 4. Select a directory of your choice and click on the Save button as shown below in Fig. 4

Fig. 4



When you click the Save button, the “Download Complete” window pops up as shown in Fig 5.

Fig 5.



Again if SAS software is not installed locally on your computer then click the close button. You can later open up PC SAS or UNIX SAS and assign a library reference to that directory then open up the data set . When you click on the Open Folder button, it opens up the directory containing the data set that you saved.

**CONCLUSION**

This paper has discussed some of the technical issues that must be addressed before downloading an HTML report as SAS data set in Web Application.. The focus was on the FILESRV Dispatcher utility as well as the Authorization and HTTP header files that needed to be updated in order to download an HTML report as SAS data set. Even though files have to be updated and a script modified before a successful download will occur, it is an efficient way to provide the user with data sets on the fly without first creating CSV files.

**REFERENCES**

1. SAS Institute’s Web site  
<http://www.sas.com/rnd/web/internet/filesrv/filesrv.html>  
<http://www.sas.com/rnd/web/internet/filesrv/auth.html>  
<http://www.sas.com/rnd/web/internet/filesrv/httphead.html>  
<http://www.sas.com/rnd/web/internet/filesrv/examds.html>
2. SUGI 27 Paper 39-27: A look at the Development Process for SAS/IntrNet Application – Andrew Rosenbaum
3. SUGI 27 Paper 182-27: Connecting the SAS system to the Web: A hand-on Introduction to SAS/IntrNet Application Dispatcher: Vincent Timbers

**ACKNOWLEDGEMENT**

- My thanks to the following people:
1. Douglas Hillmer, leader of the Tabulations Team for reviewing and editing this paper.
  2. Charles Jones of the Tabulations Team for his contribution in creating the statistical reports and also reviewing this paper.
  3. The Technical Support Staff at SAS Institute, especially Tim Besselman, who helped me with the Filesrv utility.

**CONTACT INFORMATION**

Your comments and questions are valued and encouraged. Contact the author at:  
 John Owusu-Djamboe  
 US Census Bureau  
 Demographic Survey Division  
 Continuous Measurement Office, Room 1678-3  
 Washington DC 20233  
 Work Phone: (301) 763-6211  
 Email: [john.k.owusu.djamboe@census.gov](mailto:john.k.owusu.djamboe@census.gov)