

## **Spring 2010 OASUS Questions and Answers**

The following answers are provided to the benefit of the OASUS Users Group and are not meant to replace SAS Technical Support.

**Question 1 : Can I use DDE on SAS Enterprise Guide? I tried and it did not work.**

The answer is No! You cannot use DDE with Enterprise Guide. This is clearly explained in the SAS note 20767:

*Submitting DDE code to a remote server is not supported under SAS Enterprise Guide or SAS/CONNECT software.*

*The remote SAS sessions accessed with SAS Enterprise Guide and SAS/CONNECT software run in a 'displayless' environment. Microsoft does not support running their Office products in these environments (which they refer to as "server-side").*

*Refer to the following Microsoft Knowledgebase article for further details:*

*<http://support.microsoft.com/default.aspx?scid=KB;EN-US;Q257757&>*

**Question 2 : Has anyone come across problems in reading SAS files generated in a 32-bit system on a 64-bit system? How was the problem solved?**

When a file is created in a 32-bit environment and you wish to read it in a 64-bit environment, a conversion must take place. Essentially, before SAS can process data on given host computer (Windows 32-bit, Linux 64-bit, AIX 64 bit, etc) it has to match the representation native to the host. In other words, if the data has been created such that the representation does not match the native representation of the host (i.e. using a foreign representation) where the data is being processed, a conversion to that latter native format must take place. When going from a 32-bit to a 64-bit (even if it is Windows in both cases), the native formats will not match.

SAS provides a number of facilities to deal with various host-based representations:

**OUTREP= Option**

This option is applicable to a SAS file (as a dataset option) or to an entire library (as a LIBNAME option). It allows you to create SAS files using a foreign representation. Without the OUTREP option, SAS will always create files using the native representation. The OUTREP option accepts one value that specifies the data representation to be used. These values are defined in the SAS documentation. As an example, the value for 32-bit Windows representation is WINDOWS\_32 and for 64-bit Windows is WINDOWS\_64.

```
/* Create a 64-bit Windows SAS file using OUTREP= option */
data CLASS_W64 (outrep= WINDOWS_64);

    set sashelp.class;

run;
```

**Reading a SAS File using Cross Environment Data Access (CEDA)**

Although the OUTREP option does wonders when creating file using a specific representation, it is of no help when reading a file. So how can you deal with a SAS file that does not match the native representation? By simply reading it as if it was a native file! When you do so, a SAS facility called Cross Environment Data Access (CEDA) converts the file for you. Normally, a SAS note will be written out in the log explaining that CEDA was invoked and that performance may be reduced. Although CEDA is somewhat magical, it does require some computing resources. For very large files, this could be an important issue so a conversion strategy should be contemplated.

### *SAS Notes when CEDA is invoked*

```
NOTE: Data file WORK.CLASS_W64.DATA is in a format that is native to another host, or the file
      encoding does not match the session encoding. Cross Environment Data Access will be used,
      which might require additional CPU resources and might reduce performance.
NOTE: There were 19 observations read from the data set SASHELP.CLASS.
NOTE: The data set WORK.CLASS_W64 has 19 observations and 5 variables.
NOTE: DATA statement used (Total process time):
      real time          0.01 seconds
      cpu time           0.01 seconds
```

### **PROC CPORT and CIMPORT**

CEDA is limited to SAS data files and cannot be used for any other types of files such as catalogs and indexes. If you need to convert these, an option is to use transport files. A transport file is a host-neutral container to store various SAS files in order to move them from one host to the other.

You create a transport file using PROC CPORT. The transport file must then be copied over the target host where you want the various SAS files contained in the transport file be re-created (if you use FTP, make sure that a binary transfer is performed). On the target host, you run PROC CIMPORT to re-create the SAS files.

More information and examples on transport files can be found in the SAS online documentation under the CPORT and CIMPORT procedures.

### **References:**

Gilsen, Bruce. CEDA: The Invisible Hand, SAS Users Group International. Paper 225-31.

SAS Institute Inc. 2008. Moving and Accessing SAS® 9.2 Files. Cary, NC: SAS Institute Inc.

**Question 3 : Is there a function in SAS that allows us to create a subset of data based on a variable, selecting only the top 30%? For example, there is a data set with numbers of passengers processed in an hour at a screening line. The values vary by date and hours. Purpose: select top 30% of records and calculate an average of those values**

This can be done easily with PROC RANK. PROC RANK is a little SAS gem that is good to know about.

PROC RANK computes ranks from one or more numeric variables across multiple observations. Assuming that there are no ties in the data, PROC RANK will assign a rank from 1 to n (n equal the number of observations) to each observation for a given variable. If there is a tie, then PROC RANK will assign the average rank for the ties. For example:

```
proc rank data=sashelp.class out=work.r_class;
  var height;
  ranks r_height;
run;
```

Produces:

	Name	Sex	Age	Height	r_height	Weight
1	Alfred	M	14	69	18	112.5
2	Alice	F	13	56.5	3	84
3	Barbara	F	13	65.3	14	98
4	Carol	F	14	62.8	10	102.5
5	Henry	M	14	63.5	11	102.5
6	James	M	12	57.3	4	83
7	Jane	F	12	59.8	7	84.5
8	Janet	F	15	62.5	8.5	112.5
9	Jeffrey	M	13	62.5	8.5	84
10	John	M	12	59	6	99.5
11	Joyce	F	11	51.3	1	50.5
12	Judy	F	14	64.3	12	90
13	Louise	F	12	56.3	2	77
14	Mary	F	15	66.5	15.5	112
15	Philip	M	16	72	19	150
16	Robert	M	12	64.8	13	128
17	Ronald	M	15	67	17	133
18	Thomas	M	11	57.5	5	85
19	William	M	15	66.5	15.5	112

It is possible to handle ties differently by using the TIES= option but for the problem at hand the default of TIES=MEANS is fine.

By default the smallest value receives a rank of 1. It is possible to reverse the direction of the ranks by using the DESCENDING option (in which case the largest value gets assign a rank of 1).

If you are interested in groups of values rather than the individual rank of each observations, you can do so by using the GROUPS= option. The GROUPS option specifies the number of groups. Each group value is calculated using the individual rank as followed:

$$FLOOR(rank * k/(n+1))$$

where:

FLOOR is the FLOOR function

rank is the rank of a given value

k is the value of GROUPS=

n is the number of observation having nonmissing values

The following example creates groups 0 to 9 (10 groups) in descending order of height:

```
proc rank data=sashelp.class out=work.r_class descending
groups=10;
var height;
ranks r_height;
run;
```

Produces

	Name	Sex	Age	Height	r_height	Weight
1	Alfred	M	14	69	1	112.5
2	Alice	F	13	56.5	8	84
3	Barbara	F	13	65.3	3	98
4	Carol	F	14	62.8	5	102.5
5	Henry	M	14	63.5	4	102.5
6	James	M	12	57.3	8	83
7	Jane	F	12	59.8	6	84.5
8	Janet	F	15	62.5	5	112.5
9	Jeffrey	M	13	62.5	5	84
10	John	M	12	59	7	99.5
11	Joyce	F	11	51.3	9	50.5
12	Judy	F	14	64.3	4	90
13	Louise	F	12	56.3	9	77
14	Mary	F	15	66.5	2	112
15	Philip	M	16	72	0	150
16	Robert	M	12	64.8	3	128
17	Ronald	M	15	67	1	133
18	Thomas	M	11	57.5	7	85
19	William	M	15	66.5	2	112

To answer the question explicitly, rather than creating groups per se we can ask PROC RANK to create a percentage value based on the rank value (assuming that TIES=MEAN) as followed:

$$(rank/n)/100$$

where:

rank is the rank of a given value

n is the number of observation having nonmissing values

Therefore to get the top 30%, one only needs to request the percentage (via the PERCENT option) along with the DESCENDING option:

```
proc rank data=sashelp.class
    out=work.r_class(where=(r_height <= 30))
    descending percent;
var height;
ranks r_height;
run;
```

Produces the following dataset for the top 30% based on height. Deriving the means from that dataset becomes a trivial exercise.

	Name	Sex	Age	Height	r_height	Weight
1	Alfred	M	14	69	10.526315789	112.5
2	Mary	F	15	66.5	26.315789474	112
3	Philip	M	16	72	5.2631578947	150
4	Ronald	M	15	67	15.789473684	133
5	William	M	15	66.5	26.315789474	112

#### References:

Bilenas, Jonas V. (2009). Using PROC RANK and PROC UNIVARIATE to Rank or Deciles variables, North Eastern SAS Users Group. Paper 225-31.

***Question 4 : Does SAS work with Open Office the way it works with MS Office?***

The SAS Addin for MS-Office is not currently supported with Open Office. The most recent release of the Addin (version 4.3) supports the following MS-Office suites:

- Microsoft Office 2010 (32 - bit version only)
- Microsoft Office 2007
- Microsoft Office 2003

***Question 5 : Can you let us know of any tools that we can use to confirm what objects (tables, jobs, stored processes) we have changed or modified in a SAS DI environment besides the manual means we are currently employing.***

The Impact Analysis function of DI may help in this regard. This question should however be directed to the OPUS group since it is related to the SAS platform.

***Question 6 : Can we install more than one metadata server on the same Unix server which has more than one SAS base, e.g. SAS 9.1 and SAS 9.2? (One metadata server is for production and the other one is for development.)***

Yes it is possible. This question should be directed to the OPUS group since it is related to the SAS platform.