PROC CONTENTS

Proc contents gives descriptive information for all variables found in SAS dataset such as type, length, position, format, informat and label.

Using this procedure we can see description information of dataset.

It is useful for document purpose. (Documenting SAS datasets stored in libraries)

Syntax:-

```
PROC CONTENTS <OPTIONS>;
RUN;
```

Examples:-

```
PROC CONTENTS;
RUN;
```

It prints content information for latest dataset.

Options:-

Data: - Specify the input data set for which dataset we are finding descriptor information.

```
PROC CONTENTS DATA=SASHELP.CLASS;
RUN;
```

Out: - Specify the output data set

It creates description information portion in output dataset.

```
PROC CONTENTS DATA=SASHELP.CLASS OUT=WORK.DATASET;
RUN;
```

Noprint: - Suppress the printing of the output

```
PROC CONTENTS DATA= DATA=SASHELP.CLASS OUT=DATASET NOPRINT;
RUN;
```

Position: - Default variables in the dataset are printed out an alphabetic order, when this position option specified, a second list of variables is output according to their position in the dataset.

```
PROC CONTENTS DATA=SASHELP.CLASS POSITION;
RUN;
```

But don't want in Alphabetic order, Need only list of variables with their attributes in dataset order use

Varnum: - Print a list of the variables by their logical position in the data set

```
PROC CONTENTS DATA=SASHELP.CLASS VARNUM;
RUN;
```

Short: - Prints out only names of variables (without attributes) in SAS dataset.

```
PROC CONTENTS DATA=SASHELP.CLASS SHORT;
RUN;
```

Default it prints list of variables in Alphabetic order, but need to print in both alphabetic and dataset order

```
PROC CONTENTS DATA=SASHELP.CLASS SHORT POSITION;
RUN;
```

Default it prints list of variables in Alphabetic order, but need to print in only dataset order

```
PROC CONTENTS DATA=SASHELP.CLASS SHORT   /*POSITION*/ VARNUM;
RUN;
```

Centiles: - Print centiles information for indexed variables

```
/*SIMPLE INDEX */
PROC SQL;
CREATE INDEX NAME ON SASHELP.CLASS (NAME);
QUIT;
/* COMPOSITE INDEX*/
PROC SQL;
CREATE INDEX NAME ON SASHELP.CLASS (NAME,SEX,AGE,HEIGHT,WEIGHT);
QUIT;
PROC CONTENTS DATA= SASHELP.CLASS CENTILES;
RUN;
```

Out2:- Specify the name of an output data set to contain information about indexes and integrity constraints

```
PROC CONTENTS DATA= SASHELP.CLASS OUT =DATASET1 OUT2=DATASET2;
RUN;
```

Data=_all_:- Prints descriptor information for all the datasets in library.

```
PROC CONTENTS DATA=WORK._ALL_;
RUN;
```

Nods:-Suppress the printing of individual files

NODS can use with only DATA=_ALL_.

```
PROC CONTENTS DATA=WORK._ALL_ NODS;
RUN;
```

Fmtlen: - Print the length of a variable's informat or format

```
PROC CONTENTS DATA= SASHELP.CLASS FMTLEN;
RUN;
```

Directory: - Print a list of the SAS files in the SAS library

```
PROC CONTENTS DATA= SASHELP.CLASS DIRECTORY;
RUN;
```

Order: - COLLATE | CASECOLLATE | IGNORECASE | VARNUM

Collate - Prints list of variables in alphabetical order beginning with uppercase and then lowercase names.

Casecollate - Prints list of variables in alphabetic order even if they include mixed-case names and numerics.

Ignorecase - Prints list of variables in alphabetical order ignoring the case of letters.

Varnum - It is same like Varnum option (means prints variables in dataset order).

```
PROC CONTENTS DATA= SASHELP.CLASS ORDER=VARNUM;
RUN;
```

Memtype: -

The contents statement produces output only for member types like **DATA**, **VIEW**, and **ALL**, which includes DATA and VIEW.

```
PROC DATASETS MEMTYPE=DATA;
CONTENTS DATA=SASHELP._ALL_;
RUN;
```

```
PROC DATASETS MEMTYPE=DATA;
CONTENTS DATA=SASHELP._ALL_;
RUN;
```

To get contents off all catalogs

```
PROC CONTENTS MEMTYPE=LIB.CATALOG;
RUN;
```

To get contents off all Views

```
PROC CONTENTS MEMTYPE=LIB.VIEW;
RUN;
```

How can you Calculate Data Set Size

To estimate the amount of disk space that is needed for a SAS data set:

Create a dummy SAS data set that contains information from sashelp.class (which contains 19 observations and 5 variables) that you need run the CONTENTS procedure using the dummy data set

Determine the data set size by performing simple math using information from the CONTENTS procedure output.

```
DATA DS;
SET SASHELP.CLASS;
RUN;
```

```
PROC CONTENTS DATA=DS;
TITLE 'EXAMPLE FOR CALCULATING DATA SET SIZE';
RUN;
```

These statements generate the output shown in Example for Calculating Data Set Size with PROC CONTENTS.

EXAMPLE FOR CALCULATING DATA SET SIZE
The CONTENTS Procedure

Data Set Name	WORK.DS	Observations	19
Member Type	DATA	Variables	5
Engine	V9	Indexes	0
Created	Monday, October 27, 2008 08:59:50 PM	Observation Length	40
Last Modified	Monday, October 27, 2008 08:59:50 PM	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	WINDOWS_32		
Encoding	wlatin1 Western (Windows)		

Engine/Host Dependent Information	
Data Set Page Size	4096
Number of Data Set Pages	1
First Data Page	1
Max Obs per Page	101
Obs in First Data Page	19
Number of Data Set Repairs	0
Filename	C:\DOCUME~1\ADMINI~1\LOCALS~1\Temp\SAS Temporary Files_TD2300\ds.sas7bdat
Release Created	9.0201M0
Host Created	XP_PRO

Alphabetic List of Variables and Attributes			
#	Variable	Type	Len
3	Age	Num	8
4	Height	Num	8
1	Name	Char	8
2	Sex	Char	1
5	Weight	Num	8

The size of the resulting data set depends on the data set page size and the number of observations. You can use your PROC CONTENTS output and the following formula to estimate the data set size:
Number of data pages = 1 + (floor (Observations / Max Obs per Page))
Size = 256 + (Data Set Page Size * number of data pages)
(Floor represents a function that rounds the value down to the nearest integer.)

Taking the information that is shown in Example for Calculating Data Set Size with PROC CONTENTS, you can calculate the size of the example data set:
Number of data pages = 1 + (floor(1/101))
Size = 256 + (4096 * 1) = 4352

Thus, the example data set uses 4,352 bytes of storage space.

Interview Questions

1) What is proc content? Syntax?

2) Why should we use Proc content?

3) Use proc contents for below dataset WORK.DS and observe output clearly

```
DATA WORK.DS;
SET SASHELP.CLASS;
RUN;
```

4) Use proc contents with below options and define all options with examples?
   A) Data=Datasetname
   B) Out=Datasetname
   C) Out2=Datasetname
   D) Data=_All_
   E) Position
   F) Varnum
   G) Short
   H) Nods
   I) Directory
   J) Centiles

5) What is the output if I use SHORT & POSITION options in Proc contents?

6) What is the output if I use SHORT & VARNUM options in Proc contents?

7) How can you find out below ds dataset size?

```
DATA WORK.DS;
SET SASHELP.CLASS;
RUN;
```

8) What is the significance of below program?

```
Proc contents data=work._all_;
Run;
```

9) What is Member type? And what are the Member types do you know?

10) How can you get contents of all Catalogs in a library?

10) How can you get contents of all datasets in a library?

11) How can you get contents of all views in a library?

12) How can you get contents of all Datasets, Catalogs, Views in a library?

13) What is Index ? How can you create it ?

14) What is View? How can you create it ?

15) What is difference between Dataset & View ?

16) How can you calculate a dataset size?

17) How can you list the column names from SAS Dataset?

18) If you get a huge dataset for manipulation? How can you understand that data ?