greatlearning
Learning for Life

SAS Procedures

greatlearning
Learning for Life

Module Content

- PROC PRINT
- PROC FREQ
- PROC CONTENT
- PROC DATASET
- PROC FORMAT
- PROC CHART
- PROC REPORT
- PROC TABULATE
- ODS

Procedure Print

greatlearning
Learning for Life

- ✓ PROC PRINT?
  - ✓ Prints the observations in a SAS dataset using all or some of the variables
  - ✓ A variety of reports can be created which includes summarization across various groups.
  - ✓The output is in numeric format
- ✓ Print Procedure -- Syntax
  - ✓ PROC PRINT <other options>;
    BY <DESCENDING> variable-1 …… <DESCENDING> variable-n;
    VAR variable(s);
    SUM variable(s);
    PAGEBY BY-variable;
    SUMBY BY-variable;
    LABEL variable-name = ‘’;

Example of Proc Print

EXAMPLE OF PROC PRINT

Obs	customer_ id	credit_lmt
1	A1	$6,100
2	A2	$4,600
3	A3	$1,000
4	A4	$3,000
5	A5	$1,800
6	A6	$3,100
7	A7	$5,300
8	A8	$4,000
9	A9	$2,600
10	A10	$2,700

Procedure Print -Details

greatlearning
Learning for Life

- ✓ Print Procedure – Syntax…
  - ✓ PROC PRINT <other options>;
    - ✓ DATA =
      - ✓ Specify the input dataset
      - ✓ If this option is not used, the procedure prints the latest dataset
    - ✓ OBS =
      - ✓ Specifies number of observations
    - ✓ DOUBLE
      - ✓ Puts a blank line between values Alias : D
    - ✓ NOOBS
      - ✓ Suppresses the observation number in the output
    - ✓ ROUND
      - ✓ rounds unformatted numeric values to 2 decimal places. ALIAS : R
    - ✓ N
      - ✓ Prints the no. of observations in the dataset at the end of the printed output
    - ✓ LABEL
      - ✓ Uses variables' labels as column headings
      - ✓ If the LABEL is omitted, proc print uses default option as the variable's name ALIAS : L

Procedure Print -Details

greatlearning
Learning for Life

- ✓ Print Procedure -- Syntax
  - ✓ PROC PRINT <other options>;
    BY <**DESCENDING**> variable-1 …… <DESCENDING> variable-n;
    VAR variable(s);
    SUM variable(s);
    PAGEBY BY-variable;
    SUMBY BY-variable;
    LABEL variable-name = ‘’;

BY indicates to create separate section for each group

NOTSORTED : Specifies that observations are not necessarily in alphabetic/numeric order.

If the VAR statement is omitted, then the procedure prints all the variables in the dataset

Procedure FREQ

greatlearning
Learning for Life

- ✓ What does the FREQ Procedure do ?
  - ✓ Produces frequency and cross tabulation
  - ✓ tests and measures of association for 2 –way tables
  - ✓ stratified analysis, providing statistics within and across strata
  - ✓ output can be saved as SAS datasets
- ✓ Freq Procedure -- Syntax
  - ✓ PROC FREQ *<option(s)>;*
    TABLES requests *</option(s)>;*
    OUTPUT <OUT = SAS-dataset> *options;*

Example

```
data new;
 set perf;
 if rscore > 800 then risk_level='L';
 else if rscore > 600 then risk_level='M';
 else risk_level='H';
 zone1=substr(zipcode,1,1);
run;

proc freq data=new;
 tables segment risk_level*zone1/list
   missing ;
run;
```

Procedure FREQ-Details

greatlearning
Learning for Life

- ✓ Freq Procedure – Syntax…
  - ✓ PROC FREQ <*option(s)*>;
    - ✓ DATA =
      - ✓ Specify the input SAS dataset
      - ✓ If this option is omitted, then PROC FREQ uses the most recently created one
    - ✓ PAGE
      - ✓ outputs one table in each page
      - ✓ If not provided multiple tables per page are displayed
    - ✓ NOPRINT
      - ✓ Suppresses the display of all output
    - ✓ ORDER = ‘
      - ✓DATA | FORMATTED | FREQ | INTERNAL
      - ✓ indicates order in which frequency variables are to be displayed

orders according to their original order in input dataset

orders by formatted values

orders by descending frequency values

DEFAULT option: orders by their unformatted values

Procedure FREQ-Details

greatlearning
Learning for Life

- ✓ Freq Procedure – Syntax…
  - ✓ TABLES requests *</option(s)>*;
    - ✓ requests
      - ✓ frequency and crosstabulation tables that needs to be produced can be specified
      - ✓ can be one variable name or multiple variable names separated using asterisks
    - ✓ options
      - ✓ LIST : 2-way to n-way tables as list
      - ✓ NOCOL : No column percentage
      - ✓ NOROW : No row percentage
      - ✓ NOFREQ : No frequency count
      - ✓ NOPERCENT : No %ages, row % and column % in cross tabulation tables or percentages & cumulative % in 1-way frequency tables and list format
      - ✓ NOCUM : No cumulative frequencies and cumulative %
      - ✓ OUT = : specifies variables to output dataset
    - ✓ OUTPUT <OUT = SAS-dataset> *options*;
      - ✓ Creates a SAS dataset containing statistics computed by PROC FREQ

Proc Freq – SAS Output

greatlearning
Learning for Life

Frequency Distribution of Segment

The FREQ Procedure

segment	Frequency	Percent	Cumulative Frequency	Cumulative Percent
S1	11	22.00	11	22.00
S2	18	36.00	29	58.00
S3	21	42.00	50	100.00

Cross Frequency of Risk Level Vs Zone

With list option

risk_level	zone1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
H	A	4	8.00	4	8.00
H	B	5	10.00	9	18.00
L	A	12	24.00	21	42.00
L	B	6	12.00	27	54.00
M	A	8	16.00	35	70.00
M	B	15	30.00	50	100.00

Without list option

Table of risk_level by zone1

risk_level / zone1 Frequency Percent Row Pct Col Pct	A	B	Total
H	4 8.00 44.44 16.67	5 10.00 55.56 19.23	9 18.00
L	12 24.00 66.67 50.00	6 12.00 33.33 23.08	18 36.00
M	8 16.00 34.78 33.33	15 30.00 65.22 57.69	23 46.00
Total	24 48.00	26 52.00	50 100.00

Proc Contents

- Provides information on
  - Attributes of the dataset- Name, # of records, Date created etc.
  - Attributes of variable- Name, Label, Type etc
  - If the data set is sorted
  - Other information- like engine etc.
- Syntax

```
PROC CONTENTS DATA=<data set>;
RUN
```

Example

```
PROC CONTENTS DATA =perf;
Run;
```

Proc Contents – Output

greatlearning
Learning for Life

```
                                The SAS System          00:00 Thursday, July 1, 1993   22

                              The CONTENTS Procedure

Data Set Name        WORK.PERF                              Observations          50
Member Type          DATA                                   Variables             10
Engine               V9                                     Indexes               0
Created              1:52 Wednesday, December 15, 2010      Observation Length    72
Last Modified        1:52 Wednesday, December 15, 2010      Deleted Observations  0
Protection                                                  Compressed            NO
Data Set Type                                               Sorted                NO
Label
Data Representation  WINDOWS
Encoding             wlatin1  Western (Windows)


                          Engine/Host Dependent Information

Data Set Page Size          8192
Number of Data Set Pages    1
First Data Page             1
Max Obs per Page            113
Obs in First Data Page      50
Number of Data Set Repairs  0
File Name                   C:\Users\Owner\AppData\Local\Temp\SAS Temporary Files\_TD520\perf.sas7bdat
Release Created             9.0000M0
Host Created                WIN_PRO


                       Alphabetic List of Variables and Attributes

#    Variable     Type    Len    Format        Informat
5    credit_lmt   Num       8    BEST12.       BEST32.
9    curr_mth     Num       8    BEST12.       BEST32.
1    customer_id  Char      4    $4.           $4.
10   delq_flag    Num       8    BEST12.       BEST32.
2    dob          Num       8    MMDDYY10.     MMDDYY10.
4    doj          Num       8    MMDDYY10.     MMDDYY10.
8    rscore       Num       8    BEST12.       BEST32.
3    segment      Char      2    $2.           $2.
6    spend        Num       8    BEST12.       BEST32.
7    zipcode      Char      5    $5.           $5.
```

Variable Information

PROC Datasets

greatlearning
Learning for Life

- What does the DATASETS Procedure do ?
  - With this procedure, one can
    - copy SAS files from one library to another
    - append, rename, delete and list SAS files in a SAS library
    - modify attributes of SAS datasets and variables within datasets
- Datasets Procedure -- Syntax
  - PROC DATASETS *<option(s)>*;
    APPEND BASE = <*libref.*>SAS-dataset DATA = <*libref.*>SAS-dataset
    <FORCE>;
    CHANGE previous-name1 = new name1 <*…previous-name-n = new-name-n* >;
    CONTENTS < *option(s)*>;
    COPY OUT = libref-1;
    SELECT *SAS-file(s);*
    DELETE *SAS-file(s);*
    MODIFY *SAS-file <(option(s))>;*
    RENAME previous-name1 = new-name1 <*…previous-name-n = new-name-n* >;
    SAVE *SAS-file(s) </ MEMTYPE = mtype>;*

PROC Datasets – Details

greatlearning
Learning for Life

- ✓ Datasets Procedure – Syntax…
  - ✓ PROC DATASETS *<option(s)>*;
    - ✓ LIBRARY =
      - ✓ Specify the procedure input library
    - ✓ KILL
      - ✓ Deletes all SAS files in the library that are available for processing
    - ✓ MEMTYPE =
      - ✓ Limits processing to a particular type of SAS file
      - ✓ If MEMTYPE = DATA, the PROC DATASETS limits processing to SAS datasets in the library specified or the default one
    - ✓CHANGE previous-name1 = new-name1 *<…previous-name-n = new-name-n >*;
    - ✓ Renames one or more SAS files in the same SAS data library
  - ✓ CONTENTS *<option(s)>*;
    - ✓ DATA = : Specify the input dataset
    - ✓ OUT = : Specify the name for an output dataset
    - ✓SHORT : displays only variable names
  - ✓ COPY OUT = library;
    - ✓ Copies SAS files to specified SAS library
  - ✓ SELECT *SAS-file(s);*
    - ✓ Select the SAS files for copying

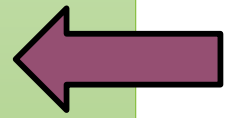
Proc Format

- Proc format is used to define user defined formats.
- The formats can be numeric or character
- Useful for grouping of variable

```
proc format <options>;
value <$> name <format options>;
run;
```

http://www.ats.ucla.edu/stat/sas/library/nesug00/bt3001.pdf

```
proc print data=new (obs=10);
   var rlevel rscore new_seg segment;
run;
Proc freq data=new;
 table rlevel new_seg*segment /list missing;
Run;
```



Example

```
Proc format;
Value level
    LOW-600="H" 600-800="M"
    800-HIGH="L"
    Other="X" ;
Value $SEG
    "S1"="GroupI"
    "S2","S3"="GroupII"
    Other="X" ;
    Run;
Data new;
    Set perf;
    rlevel =put (rscore, level.);
    new_seg =put (segment, $seg.);
    Run;
```

Proc Format

greatlearning
Learning for Life

Print of 10 Observations

Obs	rlevel	rscore	new_seg	segment
1	T	945	GroupII	S3
2	M	790	GroupI	S1
3	H	530	GroupII	S3
4	M	606	GroupII	S2
5	M	621	GroupI	S1
6	M	743	GroupII	S3
7	M	747	GroupII	S3
8	H	573	GroupI	S1
9	M	610	GroupII	S3
10	M	668	GroupII	S2

Frequency

The FREQ Procedure

rlevel	Frequency	Percent	Cumulative Frequency	Cumulative Percent
H	9	18.00	9	18.00
M	23	46.00	32	64.00
T	18	36.00	50	100.00

new_seg	segment	Frequency	Percent	Cumulative Frequency	Cumulative Percent
GroupI	S1	11	22.00	11	22.00
GroupII	S2	18	36.00	29	58.00
GroupII	S3	21	42.00	50	100.00

Proc Export

- Export SAS data to flat files (CSV, Text Delimited Files)

```
PROC EXPORT DATA=<dataset>
<options>
OUTFILE="filename“
<DBMS=identifier> <REPLACE>;
RUN;
```

If DBMS=dlm, then also specify delimiter!

Example

- Export summ data created in previous example to csv and pipe delimited text file.

```
proc export data=summ outfile =
  ‘<filepath>/output.csv' dbms =csv replace;
run;

proc export data=summ outfile =
  ‘<filepath>/output.txt' dbms =dlm replace;
delimiter='|';
run;
```

Proc Chart

greatlearning
Learning for Life

Proc chart is a procedure that produces text-based bar charts as well as pie and star charts.

```
PROC CHART data = test;
        vbar var1;
RUN;

PROC CHART data = test;
        pie var1;
RUN;
```

Proc Report

greatlearning
Learning for Life

Proc report is a procedure that produces output as reports

```
proc report data=sashelp.class nowindows headline headskip;
Columns name age sex height weight;
Define name / display "Name_of_Staff" width = 7;
Define sex /group "Staff_Gender" width = 9;
Define age / analysis mean "Age_of_Staff" width = 5;
Define Height / analysis mean "Height_of_Staff" format= 7.2;
Define weight / analysis mean "Weight_of_Staff" format =
7.2;
Define BMI /computed format = 7.2;
Compute BMI ;
BMI = height.mean/weight.mean;
Endcompute;
Break after sex/skip summarize dol dul;
Run;
```

Name_of_Staff	Age_of_Staff	Staff_Gender	Height_of_Staff	Weight_of_Staff
Alice	13	F	56.50	84.00
Barbara	13		65.30	98.00
Carol	14		62.80	102.50
Jane	12		59.80	84.50
Janet	15		62.50	112.50
Joyce	11		51.30	50.50
Judy	14		64.30	90.00
Louise	12		56.30	77.00
Mary	15		66.50	112.00
	13.222222	F	60.59	90.11
Alfred	14	M	69.00	112.50
Henry	14		63.50	102.50
James	12		57.30	83.00
Jeffrey	13		62.50	84.00
John	12		59.00	99.50
Philip	16		72.00	150.00
Robert	12		64.80	128.00
Ronald	15		67.00	133.00
Thomas	11		57.50	85.00
William	15		66.50	112.00
	13.4	M	63.91	108.95

Proc Tabulate

greatlearning
Learning for Life

Proc Tabulate generates a professional looking table.

```
PROCTABULATE data=one;
        CLASS gender fulltime;
        VAR income;
        TABLE gender , Income * fulltime * ( n mean) ;
RUN;
```

Descriptive Statistics	Quantile Statistics
COLPCTN	MEDIAN | P50
PCTSUM	P1
COLPCTSUM	Q3 | P75
MAX	P90
ROWPCTN	P95
MEAN	P5
ROWPCTSUM	P10
MIN	P99
STDDEV / STD	Q1 | P25
N	QRANGE
STDERR	
NMISS	
SUM	Hypothesis Testing
PAGEPCTSUM	
PCTN	ProbT
VAR	T

Output Delivery System - ODS

greatlearning
Learning for Life

ODS opens a whole new world of choices in generating high-quality, detailed presentation output from SAS. With ODS, you can create various file types including HTML, Rich Text Format (RTF), PostScript (PS), Portable Document Format (PDF), and SAS data sets.

ODS <File Type> File = ' ';

SAS Procedure(s);

ODS <File Type> CLOSE;

Where File Type = HTML, RTF, Listing, PDF, ..

Example1:

```
ODS HTML FILE = 'c:\test\output.xls'
                STYLE = minimal;

PROC PRINT DATA=Data1;
 VAR1 VAR2;
RUN;
ODS HTML CLOSE;
```

Example2:

```
ODS HTML FILE = 'c:\test\output_style.html'
                STYLE = barrettsblue;
PROC PRINT DATA=Data1;
 VAR1 VAR2;
RUN;
ODS HTML CLOSE;
```