|  |  |  |  |
| --- | --- | --- | --- |
| **Steps for Creating a SAS Datasets** | | | **Scope** |
| 1 | **LIBNAME** libref ‘<Path>’; | Reference a SAS data library | Global |
| 2 | **FILENAME** fileref ‘<Path>’; | Reference (Temp) an external file | Global |
| 3 | **DATA** ‘SASDataSetName’; | Name a SAS data set |  |
| 4 | **INFILE** ‘filename/fileref’  ***OBS***=10; | Identify an external file using INFILE statement  **OBS** mention the ***range till which data needs to be read.*** Can be used in data and proc print. Used to ***verify Data*** reading without affecting RAM space much. |  |
| 5 | **INPUT** <informats>; | Describe data |  |
| 6 | **RUN;** | Execute the DATA step |  |
| 7 | PROC PRINT Data=<DS>; | List the data |  |
| 8 | RUN; | Execute the final program step |  |

**Column Style:** **[Standard Data + *Well Ordered in Column]***

1--------10---------20---------30--------40---------50---------60---------70---------80--------90

124 61 Mod Male Pradeep United States

123 76 Ded Female Sruthi India

142 89 Reg Male Sathyamurthy United Kingdom

|  |  |  |
| --- | --- | --- |
| **PROC PRINT DATA=DATASETNAME**  **NOOBS \*used to avoid printing observation column while printing;**  **DOUBLE \*print double spacing in SAS Output and not in SAS Report;**  **(OBS=3) \* Print only the first 3 observation of the dataset in print;** | | **Scope** |
| **Sum** <Col Name>; | Calculate the sum of the column | Local |
| **VAR** <Col Name>; | Mention the variable and its ***order*** of printing | Local |
| **Label** <Col Name>=’’; | Define label name for a column  Can mention up to ***256 char***  ***Can be defined in single or multiple lines*** | **Local** |
| **Where** <column condi>  **CONTAINS** ‘str’;  **?** ‘str’;  **IN**(‘str1’,’str2’); | Defines the column condition  ***=, ^=, >, <, >=, <=***  ***CONTAINS*** *is string comparison*  ***AND, OR*** *operator used along with col name each time*  ***IN*** *operator is used as SQL style in comparison.* | Local |
| **ID** <Col Names>; | Act as a primary key, ***replace OBS*** column without explicitly mention of NOOBS.  ID used along with ***Var*** will ***display*** a ***column twice***. | Local |
| **SUM** <Col Name>; | Will provide the total of the column specified. | Local |
| **BY** <Col Name>; | Col Name should be ***same as one that is sorted before*** using this. ***Subset results***. | Local |
| **BY** <Col Name1>;  **ID** <Col Name1>; | When ID used along with BY it will:   1. ***Supress OBS*** column 2. ID/BY variable name is printed in left col 3. Each ***ID/BY value is printed only once*** at the start of each by group and on the line, that has group sub-total. | Local |
| **By** <Col Name1>;  **PAGEBY**<Col Name1>; | Mostly used along with sum-by-id.  ***Column used in PAGEBY should be same as one used in BY***.  Used to ***print each sub-total on a separate page***. | Local |
| **FORMAT** <Col Name>; | When defined ***inside PROC it scopes within it***. To make it ***permanent FORMAT or Labels*** need ***to be defined in DATA step*** | Local/Global |
| **TITLE** ‘str1’; | Generally, ***need to be defined outside a PROC step***.  However, it can be used inside PROC too  ***TITLE is global.*** Once defined will stay forever until title statement is modified, cancelled or end SAS session.  ***Cancel of title*** is done by ***title;*** | Global |
| **FOOTNOTE** ‘str2’; | Used to print note below a table/graph  It is same as TITLE function, up to ***10 footnotes can be defined*** in SAS.  ***Cancel of footnote*** is done by:  ***Footnote;*** | Global |

|  |  |  |
| --- | --- | --- |
| **PROC SORT** **DATA=DATASETNAME**  **OUT=DATASETNAME \*o/p SAS dataset** | |  |
| **by** <Col Name>;  **by descending** <col1> | Sorted by the column mentioned, sort takes place from right to left columns mentioned.  If used with descending it will apply to ***column which is immediately after it***, rest of the other columns will be sorted in ascending order. | Local |
| **NOTSORTED**; | To explicitly mention not to sort if the ***values are equal*** based on by condition. | Local |
|  |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
| **PROC FORMAT LIB=library**  **LIBRARY/LIB \*Defines the SAS library that needs to be referred;**  **FMTLIB \*print all the user defined format present in the Library mentioned;** | | | **Scope** |
| 1 | **LIBNAME** ***library*** ‘<Path>’; | Reference a SAS data library | Permanent |
| 2 | **PROC FORMAT LIB**=library  **FMTLIB**; | Library can be the SAS library referred above or it can be a catalog like ***library.catalog.***  ***FMTLIB*** will list all the user defined format present in the library. ***formats.sas7bcat*** file is created in the path mentioned in library. | Permanent |
| 3 | **Value** <format-name> | Format name must begin with ***$ for Char*** var  Cannot be > 8 char in length  Cannot be the name of existing SAS format  Cannot end with a number  Does not end with a period when defined | Permanent |
|  | Range1=’label1’ | Range1= Actual Column Data  Label1= Description of Range1  Numeric => 102=’Manager’  Character => ‘A’=’Good Performance’  Range => low-<12=’Not Teen Age’ | Permanent |
|  | Range2=’label2’; | Always the last Range must be ended with; which implies SAS that PROC FORMAT statement ends. | Permanent |
| 4 | **PROC FORMAT;** | This format will be created in the ***work directory*** which ***means temporary***. | Temporary |
|  | Value <format-name> | Scope within that SAS session only | Temporary |
|  | Range1=’label1’ | Scope within that SAS session only | Temporary |
|  | Range2=’label2’; | Scope within that SAS session only | Temporary |
| 5 | **PROC CATALOG;** | You can delete the user defined format | Permanent |

|  |  |  |  |
| --- | --- | --- | --- |
| **PROC REPORT DATA=<DATASETNAME>**  **WD/NOWD \*Decides should the o/p be printed in a dedicated report window;**  **DOUBLE \*print double spacing in SAS Output and not in SAS Report;**  **SPLIT=’<symbol>’ \* Symbol can be \*, # $ etc., Used to define the label split in reporting;** | | | **Scope** |
| 1 | **COLUMN** <Col Names> | Used ***to subset the column*** that is needed to be displayed in the report. | Local |
| 2 | **WHERE** <Col Condi/Name>  **In** (‘value1’,’value2’) | Used to ***filter out the data*** required  In ***used along with where to filter the data*** based on values provided, SQL style usage. | Local |
| 3 | **DEFINE** <Col1>/<usage>  DEFINE <Col2>/<attribute>  DEFINE <Col3>/<options>  DEFINE <Col3>/<Justification>  DEFINE <Col3>/<Col Heading> | Used to ***build column definitions*** in report like column space and width, etc.,  Let to ***define more than one column attribute*** at a time.  Column can be defined ***in any order*** and list ***options within it in any order as well***.  ***Usage*** specifies ***how to use the variables***:   1. Across 2. Analysis 3. Computed 4. Display 5. Group 6. Order   ***Attributes*** specifies the ***look*** of each column:  Width and spacing has its ***effect only in o/p window*** and doesn’t affect HTML window.   1. **Format** – define SAS/user format, default is ***its variable type*** 2. **Width** – width of col, default is ***Max*** 3. **Spacing** – No of blank char, default is ***2***   ***Options*** specifies the ***further formatting*** option:   1. DESCENDING 2. NOPRINT 3. NOZERO 4. PAGE   ***Justification*** specifies ***arrangements*** of column:   1. Center 2. Left 3. Right   ***Column Heading*** is the ***label definition***. ***Split*** in report definition is used to ***split the column label*** as needed. (e.g. ***SPLIT=’\*’;***) define col/c\*t; | Local |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |