

1
10:34 Thursday, January 21, 2021

The SAS System

```
1          ;*';*";*//quit;run;
2          OPTIONS PAGENO=MIN;
3          %LET _CLIENTTASKLABEL='01.15.2020_(HPC)_Basic_SAS_Test';
4          %LET _CLIENTPROCESSFLOWNAME='Process Flow';
5          %LET _CLIENTPROJECTPATH='';
6          %LET _CLIENTPROJECTPATHHOST='';
7          %LET _CLIENTPROJECTNAME='';
8          %LET _SASPROGRAMFILE='Q:\DC_Analyst\Charu\Testing
codes\01.15.2020_(HPC)_Basic_SAS_Test.sas';
9          %LET _SASPROGRAMFILEHOST='HS-CTXAPP01';
10
11          ODS _ALL_ CLOSE;
12          OPTIONS DEV=PNG;
13          GOPTIONS XPIXELS=0 YPIXELS=0;
14          FILENAME EGSR TEMP;
15          ODS tagsets.sasreport13(ID=EGSR) FILE=EGSR
16              STYLE=HTMLBlue
17
18          STYLESHEET=(URL="file:///C:/Program%20Files%20(x86)/SASHome/x86/SASEnterp
riseGuide/7.1/Styles/HTMLBlue.css")
19          NOGTITLE
20          NOGFOOTNOTE
21          GPATH=&sasworklocation
22          ENCODING=UTF8
23          options(rolap="on")
24          ;
NOTE: Writing TAGSETS.SASREPORT13(EGSR) Body file: EGSR
25
26          GOPTIONS ACCESSIBLE;
27
28          /*****
29          *****/
30          Project: High Performance Computing- Testing
31          Program: Running some basic functions
32          Run twice
33          Dataset: nhanes3 (smaller), nhanes_bigdata(big data)
34          Analyst: Charu Verma
35          Date: 01.15.2021*/
36          libname lib "Q:\DC_Analyst\Charu\Dataset";
NOTE: Libref LIB was successfully assigned as follows:
      Engine:          V9
      Physical Name: Q:\DC_Analyst\Charu\Dataset
37          /* Start timer*/
38          %let _timer_start= %sysfunc(datetime());
39
40          data test;
41          set lib.nhanes_bigdata;/*Change dataset name here*/
42          run;
```

NOTE: There were 80204010 observations read from the data set
LIB.NHANES_BIGDATA.

NOTE: The data set WORK.TEST has 80204010 observations and 43 variables.
NOTE: DATA statement used (Total process time):
 real time 2:36.25
 cpu time 44.51 seconds

```
40      proc contents data=test;
41      run;
```

NOTE: PROCEDURE CONTENTS used (Total process time):
 real time 0.04 seconds
 cpu time 0.03 seconds

2
10:34 Thursday, January 21, 2021 The SAS System

```
42      options number pageno=1 ls=79 ps=59 ;
43      options formchar = " |----|+|----+=|-\<>*" ;
44      proc printto new
45      !
print='Q:\DC_Analyst\Charu\Dataset\01212020_(HPC)_BasicSAS_test_HSE(
44      ! 2bprog).txt';
45      run;
```

NOTE: PROCEDURE PRINTTO used (Total process time):
 real time 0.01 seconds
 cpu time 0.01 seconds

```
46      /*Descriptive statistics*/
47      title 'Descriptive statistics';
48      proc means data = test;
49      var _numeric_;
50      run;
```

NOTE: There were 80204010 observations read from the data set WORK.TEST.
NOTE: The PROCEDURE MEANS printed pages 1-2.
NOTE: PROCEDURE MEANS used (Total process time):
 real time 3:15.89
 cpu time 7:11.84

```
51      proc anova data=test;
52      class sex;
53      model age education sbp dbp = sex;
54      run;
```

NOTE: The PROCEDURE ANOVA printed pages 3-7.
NOTE: PROCEDURE ANOVA used (Total process time):
 real time 4:09.39
 cpu time 52.45 seconds

```

55      proc freq data=test;
56      tables sex  dmaethnr dmaracer dmarethn dmpmetro death
dentalhealth
56      ! hipfract marital
57      cancer cholesterol colon cvd diabetes hypertension health
stroke;
58      run;

```

NOTE: There were 80204010 observations read from the data set WORK.TEST.
NOTE: The PROCEDURE FREQ printed pages 8-11.
NOTE: PROCEDURE FREQ used (Total process time):
real time 1:07.14
cpu time 44.64 seconds

```

59      title 'Correlation-Age and Diastolic BP';
60      proc corr data=test nomiss spearman;
61      var age dbp;
62      run;

```

NOTE: The PROCEDURE CORR printed page 12.
NOTE: PROCEDURE CORR used (Total process time):
3 The SAS System
10:34 Thursday, January
21, 2021

```

real time 4:56.15
cpu time 2:00.67

```

```

63      proc summary data=test;
64      var hipfract cancer colon cvd stroke;
65      class seqn;
66      output out=work.summary1 (drop=_type_freq_)
67      sum= hipfracttot cancertot colontot cvdtot stroketot;
68      run;

```

WARNING: The variable _type_freq_ in the DROP, KEEP, or RENAME list has never

been referenced.

NOTE: There were 80204010 observations read from the data set WORK.TEST.
NOTE: The data set WORK.SUMMARY1 has 4011 observations and 8 variables.
NOTE: PROCEDURE SUMMARY used (Total process time):
real time 3:15.70
cpu time 1:12.64

```

69      proc sort data=work.summary1 out=total;by seqn;run;

```

NOTE: There were 4011 observations read from the data set WORK.SUMMARY1.
NOTE: The data set WORK.TOTAL has 4011 observations and 8 variables.
NOTE: PROCEDURE SORT used (Total process time):
real time 0.00 seconds
cpu time 0.00 seconds

```
70      proc sort data=test; by seqn;run;
```

NOTE: There were 80204010 observations read from the data set WORK.TEST.

NOTE: The data set WORK.TEST has 80204010 observations and 43 variables.

NOTE: PROCEDURE SORT used (Total process time):

```
      real time          4:07.45
```

```
      cpu time           2:41.81
```

```
71      proc sql;
```

```
72      create table tot_summary as
```

```
73      select *
```

```
74      from summary1  a inner join test b
```

```
75      on a.seqn =b.seqn;
```

WARNING: Variable SEQN already exists on file WORK.TOT_SUMMARY.

NOTE: Table WORK.TOT_SUMMARY created, with 80204010 rows and 50 columns.

```
76      quit;
```

NOTE: PROCEDURE SQL used (Total process time):

```
      real time          6:40.88
```

```
      cpu time           1:09.12
```

```
77      proc tabulate data= tot_summary;
```

```
78      title 'Table';
```

```
79      class sex;
```

```
80      var hipfracttot  cancertot  colontot  cvdtot  stroketot;
```

```
4                                     The SAS System
                                     10:34 Thursday, January
```

```
21, 2021
```

```
81      table sex * (N Mean Max) ,
```

```
82      hipfracttot  cancertot  colontot  cvdtot  stroketot;
```

```
83      run;
```

NOTE: There were 80204010 observations read from the data set
WORK.TOT_SUMMARY.

NOTE: The PROCEDURE TABULATE printed pages 13-14.

NOTE: PROCEDURE TABULATE used (Total process time):

```
      real time          4:49.75
```

```
      cpu time           34.04 seconds
```

```
84      proc format;
```

```
85      value race
```

```
86      1=      "Non-Hispanic white"
```

```
87      2=      "Non-Hispanic black"
```

```
88      3=      "Mexican-American"
```

```
89      4=      "Other"
```

```
90      ;
```

NOTE: Format RACE has been output.

```
91      value gender
```

```

92          1=      "Male"
93          2=      "Female"
94          ;
NOTE: Format GENDER has been output.
95          value health
96          1=      "Excellent"
97          2=      "Very good"
98          3=      "Good"
99          4=      "Fair"
100         5=      "Poor"
101         8=      "Blank but applicable"
102         9=      "Dont know"
103         ;
NOTE: Format HEALTH has been output.
104
105         value smoke
106         1=      "Current"
107         2=      "Former"
108         3=      "Never"
109         ;
NOTE: Format SMOKE has been output.
110         run;

NOTE: PROCEDURE FORMAT used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


111         data summary_rep;
112         set tot_summary;
113         format sex gender. smoking smoke. health health. dmarethn
race.;
114         run;

NOTE: There were 80204010 observations read from the data set
WORK.TOT_SUMMARY.
NOTE: The data set WORK.SUMMARY_REP has 80204010 observations and 50
variables.
NOTE: DATA statement used (Total process time):
5                                     The SAS System
                                     10:34 Thursday, January
21, 2021

      real time           7:55.58
      cpu time            55.10 seconds


115
116         proc report data= summary_rep;
117         title 'Data Report';
118         column sex health smoking hipfracttot cancertot colontot
cvdttot
119         ! stroketot;
119         define sex/group;

```

```

120      define health/group;
121      define smoking/group;
122      define hipfracttot/analysis sum width=5;
123      define cancertot/analysis sum width=5;
124      define colontot/analysis sum width=5;
125      define stroketot/analysis sum width=5;
126
127      run;

```

NOTE: There were 80204010 observations read from the data set
WORK.SUMMARY_REP.

NOTE: The PROCEDURE REPORT printed page 15.

NOTE: PROCEDURE REPORT used (Total process time):

```

      real time          4:20.15
      cpu time           47.85 seconds

```

```

128      /* Stop timer */
129      data _null_;
130      dur = datetime() - &_timer_start;
131      put 30*'-'/' ' Total Duration:' dur time13.2 / 30*'-'';
132      run;

```

```

-----
Total Duration:    0:47:14.55
-----

```

NOTE: DATA statement used (Total process time):

```

      real time          0.00 seconds
      cpu time           0.00 seconds

```

```

133
134
135      GOPTIONS NOACCESSIBLE;
136      %LET _CLIENTTASKLABEL=;
137      %LET _CLIENTPROCESSFLOWNAME=;
138      %LET _CLIENTPROJECTPATH=;
139      %LET _CLIENTPROJECTPATHHOST=;
140      %LET _CLIENTPROJECTNAME=;
141      %LET _SASPROGRAMFILE=;
142      %LET _SASPROGRAMFILEHOST=;
143
144      ;*';*";*/;quit;run;
145      ODS _ALL_ CLOSE;
146
147
148      QUIT; RUN;

```

```

6
The SAS System
10:34 Thursday, January
21, 2021

```

```

149

```