```
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
STYLESHEET=(URL="file:///C:/Program%20Files%20(x86)/SASHome/x86/SASEnterp
riseGuide/7.1/Styles/HTMLBlue.css")
18              NOGTITLE
19              NOGFOOTNOTE
20              GPATH=&sasworklocation
21              ENCODING=UTF8
22              options(rolap="on")
23          ;
NOTE: Writing TAGSETS.SASREPORT13(EGSR) Body file: EGSR
24
25          GOPTIONS ACCESSIBLE;
26
/*********************************************************************
*******************
27          Project: High Performance Computing- Testing
28          Program: Running some basic functions
29          Run twice
30          Dataset: nhanes3 (smaller), nhanes_bigdata(big data)
31          Analyst: Charu Verma
32          Date: 01.15.2021*/
33          libname lib "Q:\DC_Analyst\Charu\Dataset";
NOTE: Libref LIB was successfully assigned as follows:
      Engine:        V9
      Physical Name: Q:\DC_Analyst\Charu\Dataset
34          /* Start timer*/
35          %let _timer_start= %sysfunc(datetime());
36
37          data test;
38          set lib.nhanes_bigdata;/*Change dataset name here*/
39          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             2:23.98
      cpu time              42.93 seconds


40         proc contents data=test;
41         run;

NOTE: PROCEDURE CONTENTS used (Total process time):
      real time             0.06 seconds
      cpu time              0.01 seconds
```

```
42         /*options number pageno=1 ls=79 ps=59 ;
43         options formchar = "|----|+|---+=|-/\<>*";
44         proc printto new
print='Q:\DC_Analyst\Charu\Dataset\01212020_(HPC)_BasicSAS_test_HSE(3prog
-diff).txt';
45         run;*/
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: PROCEDURE MEANS used (Total process time):
      real time             4:22.14
      cpu time              7:13.78


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

NOTE: PROCEDURE ANOVA used (Total process time):
      real time             4:36.81
      cpu time              51.57 seconds


55         proc freq data=test;
56          tables sex  dmaethnr dmaracer dmarethn dmpmetro death
dentalhealth 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: PROCEDURE FREQ used (Total process time):
```

```
          real time              44.07 seconds
          cpu time               43.95 seconds


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

NOTE: PROCEDURE CORR used (Total process time):
          real time              1:54.04
          cpu time               1:53.95


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.
3                                                        The SAS System
11:54 Thursday, January 21, 2021

NOTE: PROCEDURE SUMMARY used (Total process time):
          real time              25.45 seconds
          cpu time               1:07.35


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              0.01 seconds
          cpu time               0.01 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              2:37.87
          cpu time               2:31.45


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           5:03.14
      cpu time            1:03.21


77      proc tabulate data= tot_summary;
78      title 'Table';
79          class sex;
80       var hipfracttot cancertot colontot cvdtot stroketot;
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: PROCEDURE TABULATE used (Total process time):
      real time           3:23.62
      cpu time            34.00 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):
      real time             5:10.34
      cpu time              48.01 seconds


115
116          proc report data= summary_rep;
117           title 'Data Report';
118           column sex health smoking hipfracttot cancertot colontot
cvdtot 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: PROCEDURE REPORT used (Total process time):
```

```
      real time             3:42.93
      cpu time              46.03 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:    0:34:24.61
         ------------------------------
         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;
149
```