```
1   /*************************************************************************
2    Program Name: jblubau1_hw14_script
3    Date Created: 11/27/2016
4    Author: Joseph Blubaugh
5    Purpose: Homework Assignment 14
6   ************************************************************************* */
7
8   libname datadb 'C:\Users\Joseph\Projects\learning\Statistics\STAT_604\Materials'
8  ! access=readonly;
NOTE: Libref DATADB was successfully assigned as follows:
      Engine:        V9
      Physical Name: C:\Users\Joseph\Projects\learning\Statistics\STAT_604\Materials
9   libname output 'C:\Users\Joseph\Projects\learning\Statistics\STAT_604\Data';
NOTE: Libref OUTPUT was successfully assigned as follows:
      Engine:        V9
      Physical Name: C:\Users\Joseph\Projects\learning\Statistics\STAT_604\Data
10
11   filename outpdf
11 ! 'C:\Users\Joseph\Projects\learning\Statistics\STAT_604\Homework\jblubau1_hw14_output.pdf';
12
13   * 1) Reference the file with a fileref;
14   filename school
14 ! 'C:\Users\Joseph\Projects\learning\Statistics\STAT_604\Materials\OKSchools.csv';
15
16   * 2) Set up page in landscape layout. Date is to only be displayed on the final section of the
16 !  output;
17   options nodate number dtreset orientation=landscape pageno=2;
18   ods pdf file=outpdf contents=no startpage=yes;
NOTE: Writing ODS PDF output to DISK destination "OUTPDF", printer "PDF".
19
20   *3) create the format function for the division of each school;
21   proc format;
22       value division
23           1251 - high = '6A'
24           721 - 1250 = '5A'
25           375 - 720 = '4A'
26           181 - 374 = '3A'
27           107 - 180 = '2A'
28           70 - 106 = 'A'
29           0 - 69 = 'B'
30           . = 'Non-HS';
NOTE: Format DIVISION has been output.
31   * 4) create the format function for the size of class for each school;
32       value clsize
33           . = 'Unknown'
34           0 -< 10 = 'Very Small'
35           10 -< 14 = 'Small'
36           14 -< 18 = 'Medium'
37           18 -< 22 = 'Large'
38           22 - high = 'Very Large';
NOTE: Format CLSIZE has been output.
39   run;

NOTE: PROCEDURE FORMAT used (Total process time):
      real time           0.07 seconds
      cpu time            0.01 seconds
```

```
40
41   * 5) Import School;
42   data schools;
43       infile school dlm=',' firstobs=2 DSD;
44       input School: $50. LocCity: $50. MailCity: $50. County: $50.
45       Teachers Grade7 Grade8 Grade9 Grade10 Grade11 Grade12
46       Ungraded PreTotal ElemTotal HSTotal STRatio;
47   run;

NOTE: The infile SCHOOL is:
      Filename=C:\Users\Joseph\Projects\learning\Statistics\STAT_604\Materials\OKSchools.csv,
      RECFM=V,LRECL=32767,File Size (bytes)=134141,
      Last Modified=27Nov2016:14:30:53,
      Create Time=27Nov2016:14:30:53

NOTE: 1785 records were read from the infile SCHOOL.
      The minimum record length was 55.
      The maximum record length was 110.
NOTE: The data set WORK.SCHOOLS has 1785 observations and 16 variables.
NOTE: DATA statement used (Total process time):
      real time           0.03 seconds
      cpu time            0.03 seconds


48
49   * 6) Print the first 30 obs;
50   proc print data=schools (obs=30) noobs;
NOTE: Writing HTML Body file: sashtml.htm
51       title Oklahoma School Analysis;
52       title2 Partial Listing;
53       footnote Based on NCES Data;
54   run;

NOTE: There were 30 observations read from the data set WORK.SCHOOLS.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.32 seconds
      cpu time            0.23 seconds


55
56   * 7) Create frequency table using temporary labels created in step 4;
57   proc freq data=schools;
58       tables STRatio / nocum missing;
59       format STRatio clsize.;
60       label STRatio="Class Size";
61       title2 Distribution of Class Sizes Based on Student/Teacher Ratio;
62   run;

NOTE: There were 1785 observations read from the data set WORK.SCHOOLS.
NOTE: PROCEDURE FREQ used (Total process time):
      real time           0.08 seconds
      cpu time            0.01 seconds


63
64
```

```
65   * 8) Use summary procedure to create an average STRatio by division;
66   proc summary data=schools missing;
67       var STRatio;
68       class HSTotal;
69       format HSTotal division.;
70       output out=sum1 n=Schools mean=Ratio;
71   run;

NOTE: There were 1785 observations read from the data set WORK.SCHOOLS.
NOTE: The data set WORK.SUM1 has 9 observations and 5 variables.
NOTE: PROCEDURE SUMMARY used (Total process time):
      real time           0.02 seconds
      cpu time            0.01 seconds


72
73   * 9) Print the summary table;
74   options date number;
75   proc print data=sum1 noobs;
76       where _TYPE_ = 1;
77       var HSTotal Schools Ratio;
78       label HSTotal = 'Division';
79       format Ratio 3.1;
80       title2;
81       title3 Average Student-Teacher Ratio by School Division;
82   run;

NOTE: There were 8 observations read from the data set WORK.SUM1.
      WHERE _TYPE_=1;
NOTE: At least one W.D format was too small for the number to be printed. The decimal may be shifted by the "BEST" format.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.14 seconds
      cpu time            0.03 seconds


83
84   * 10) Housekeeping;
85   title;
86   title2;
87   title3;
88   footnote;
89
90   ods pdf close;
NOTE: ODS PDF printed 4 pages to C:\Users\Joseph\Projects\learning\Statistics\STAT_604\Homework\jblubau1_hw14_output.pdf.
```