```
1   /*************************************************************************
2    Program Name: jblubau1_hw10_script
3    Date Created: 10/24/2016
4    Author: Joseph Blubaugh
5    Purpose: Homework Assignment 10
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_hw10_output.pdf';
12
13   * 2) Read in tabled1x2016 ;
14   data monthly_jobs;
15      set datadb.tabled1x2016;
16      * b) Fix spelling;
17      if industry = 'TRADE, TRANSPORTATION, AND UITLITIES' then do
18         industry = 'TRADE, TRANSPORTATION, AND UTILITIES';
19      end;
20      * c) Change industry to propper case;
21      industry = propcase(industry);
22      where state ne '';
23      if Aug__2015 ne . then;
24         year = '2015';
25         month = 'August    ';
26         Jobs = Aug__2015;
27         output;
28      if Sept__2015 ne . then;
29         year = '2015';
30         month = 'September ';
31         Jobs = Sept__2015;
32         output;
33      if Oct__2015 ne . then;
34         year = '2015';
35         month = 'October   ';
36         Jobs = Oct__2015;
37         output;
38      if Nov__2015 ne . then;
39         year = '2015';
40         month = 'November  ';
41         Jobs = Nov__2015;
42         output;
43      if Dec__2015 ne . then;
44         year = '2015';
45         month = 'December  ';
46         Jobs = Dec__2015;
47         output;
48      if Jan__2016 ne . then;
49         year = '2016';
```

```
50          month = 'January   ';
51          Jobs = Jan__2016;
52          output;
53       if Feb__2016 ne . then;
54          year = '2016';
55          month = 'February  ';
56          Jobs = Feb__2016;
57          output;
58       if Mar__2016 ne . then;
59          year = '2016';
60          month = 'March     ';
61          Jobs = Mar__2016;
62          output;
63       if Apr__2016 ne . then;
64          year = '2016';
65          month = 'April     ';
66          Jobs = Apr__2016;
67          output;
68       if May_2016 ne . then;
69          year = '2016';
70          month = 'May       ';
71          Jobs = May_2016;
72          output;
73       if June_2016 ne . then;
74          year = '2016';
75          month = 'June      ';
76          Jobs = June_2016;
77          output;
78       if July_2016 ne . then;
79          year = '2016';
80          month = 'July      ';
81          Jobs = July_2016;
82          output;
83       if Aug__2016 ne . then;
84          year = '2016';
85          month = 'August    ';
86          Jobs = Aug__2016;
87          output;
88       keep industry state month year jobs;
89    run;

NOTE: There were 424 observations read from the data set DATADB.TABLED1X2016.
      WHERE state not = ' ';
NOTE: The data set WORK.MONTHLY_JOBS has 5512 observations and 5 variables.
NOTE: DATA statement used (Total process time):
      real time           0.04 seconds
      cpu time            0.04 seconds


90
91    * 3) Create 6 data sets from the bls data;
92    data
93       large (keep=industry state average_jobs)
94       medium (keep=industry state average_jobs)
95       small (keep=industry state average_jobs)
96       government (keep=state average_jobs market_size)
97       goods (keep=industry state average_jobs market_size)
```

```
98      services (keep=industry state average_jobs market_size);
99      set datadb.bls_jobs1516;
NOTE: Data file DATADB.BLS_JOBS1516.DATA is in a format that is native to another host, or the
      file encoding does not match the session encoding. Cross Environment Data Access will be
      used, which might require additional CPU resources and might reduce performance.
100     * b) fix name;
101     if industry = 'TRADE, TRANSPORTATION, AND UITLITIES' then do ;
102        industry = 'TRADE, TRANSPORTATION, AND UTILITIES';
103     end;
104     * c) compute average;
105     average_jobs = sum(of Aug__2015--Aug__2016)/13;
106     format average_jobs 8.1;
107     label average_jobs = 'Average Jobs' market_size = 'Market Size';
108     * d) do not process missing values;
109     if missing(average_jobs) then delete;
110     * e) Separate the 3 datasets based on market size;
111     if average_jobs > 1000 then do;
112        market_size = 'Large';
113        output large;
114     end;
115     else if 100 <= average_jobs <= 1000 then do;
116        market_size = 'Med.';
117        output medium;
118     end;
119     else do;
120        market_size = 'Small';
121        output small;
122     end;
123     * f) Use select statement to create 3 more data sets;
124     select (industry);
125        when ('GOVERNMENT') do;
126           output government ;
127        end;
128        when ('CONSTRUCTION', 'MANUFACTURING') do;
129           output goods;
130        end;
131        otherwise do;
132           output services;
133        end;
134     end;
135  run;

NOTE: Missing values were generated as a result of performing an operation on missing values.
      Each place is given by: (Number of times) at (Line):(Column).
      7 at 105:20   7 at 105:48
NOTE: There were 424 observations read from the data set DATADB.BLS_JOBS1516.
NOTE: The data set WORK.LARGE has 24 observations and 3 variables.
NOTE: The data set WORK.MEDIUM has 259 observations and 3 variables.
NOTE: The data set WORK.SMALL has 134 observations and 3 variables.
NOTE: The data set WORK.GOVERNMENT has 53 observations and 3 variables.
NOTE: The data set WORK.GOODS has 102 observations and 4 variables.
NOTE: The data set WORK.SERVICES has 262 observations and 4 variables.
NOTE: DATA statement used (Total process time):
      real time          0.09 seconds
      cpu time           0.03 seconds
```

```
136
137  * 4) Setup pdf;
138  ods pdf file=outpdf bookmarkgen=yes bookmarklist=hide;
NOTE: Writing ODS PDF output to DISK destination "OUTPDF", printer "PDF".
139
140  * 5) Print first 50 and last 50 from step 2;
141
142  title '5a - First 50 Observations from Monthly Jobs Data Set';
143
144  proc print data=monthly_jobs (obs=50) noobs;
NOTE: Writing HTML Body file: sashtml.htm
145      var industry state month year jobs;
146  run;

NOTE: There were 50 observations read from the data set WORK.MONTHLY_JOBS.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.37 seconds
      cpu time            0.21 seconds


147
148  title '5b - Last 50 Observations from Monthly Jobs Data Set';
149
150  proc print data=monthly_jobs (firstobs=5463 obs=5512) noobs;
151      var industry state month year jobs;
152  run;

NOTE: There were 50 observations read from the data set WORK.MONTHLY_JOBS.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.07 seconds
      cpu time            0.04 seconds


153
154  * 6) Print observations from 3) data sets;
155
156  * 6a) Print 30 obs from small;
157  title '6a - First 30 Observations of Small Markets';
158  proc print data=small (obs=30) label;
159  run;

NOTE: There were 30 observations read from the data set WORK.SMALL.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.04 seconds
      cpu time            0.01 seconds


160
161  * 6b) Print 30 obs from medium;
162  title '6b - First 30 Observations of Medium Markets';
163  proc print data=medium (obs=30) label;
164  run;

NOTE: There were 30 observations read from the data set WORK.MEDIUM.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.04 seconds
      cpu time            0.01 seconds
```

```
165
166  * 6c) Print all obs from large;
167  title '6c - Large Markets';
168  proc print data=large label;
169  run;

NOTE: There were 24 observations read from the data set WORK.LARGE.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.06 seconds
      cpu time            0.01 seconds


170
171  * 6d) Print 30 obs beginning at ob 75 from goods data set, no obs numbers;
172  title '6d - Selected Observations from Goods Industry';
173  proc print data=goods (firstobs=75 obs=105) label noobs;
174  run;

NOTE: There were 28 observations read from the data set WORK.GOODS.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.06 seconds
      cpu time            0.01 seconds


175
176  * 6e) Print 30 obs from small market in services data set;
177  title '6e - Small Markets in the Services Industry';
178  proc print data=services (obs=30) label;
179      where market_size = 'Small';
180  run;

NOTE: There were 30 observations read from the data set WORK.SERVICES.
      WHERE market_size='Small';
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.07 seconds
      cpu time            0.04 seconds


181
182  * 6f) Print all obsservations from the government data set;
183  title '6f - Government Industry';
184  proc print data=government label;
185  run;

NOTE: There were 53 observations read from the data set WORK.GOVERNMENT.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.07 seconds
      cpu time            0.03 seconds


186
187  * 7) Datasets in work library;
188  title '7 - Data Sets in WORK Library';
189  proc print data=sashelp.vtable label noobs;
190      where libname = 'WORK';
```

```
191      var libname memname crdate nobs nvar;
192  run;

NOTE: Data file DATADB.BLS_JOBS1516.DATA is in a format that is native to another host, or the
      file encoding does not match the session encoding. Cross Environment Data Access will be
      used, which might require additional CPU resources and might reduce performance.
NOTE: The map data sets in library MAPSGFK are based on the digital maps from GfK GeoMarketing and
      are covered by their Copyright. For additional information, see
      http://support.sas.com/mapsonline/gfklicense.
NOTE: There were 7 observations read from the data set SASHELP.VTABLE.
      WHERE libname='WORK';
NOTE: PROCEDURE PRINT used (Total process time):
      real time          0.56 seconds
      cpu time           0.28 seconds


193
194  ods pdf close;
NOTE: ODS PDF printed 12 pages to
      C:\Users\Joseph\Projects\learning\Statistics\STAT_604\Homework\jblubau1_hw10_output.pdf.
```