```
1          OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
NOTE: ODS statements in the SAS Studio environment may disable some output features.
69
70         /*********************************************************************************/
71         /* Program Name: STAT 604 HW#11   */
72         /* Date Created: 11/1/2021   */
73         /* Author: Jack Rodoni   */
74         /* Purpose: STAT 604 HW#11   */
75         /* Date Modified: 11/04/2021    */
76         /* Location: /home/u59649056/Homeworks/JRodoni_Homework11.sas   */
77         /*********************************************************************************/
78
79
80         /* 1. ) The first Covid19 case in Texas was reported on February 12, 2020. Below the program header, */
81         /*      include a macro assignment statement to create a macro variable that contains this date in a */
82         /*      manner that can be used throughout the program in data step statements and in titles. */
83
84         %let firstcasedate=12Feb2020;
85
86         /* 2.) Include housekeeping statements to clear titles and footnotes and suppress the printing of */
87         /*     procedure titles. */
88
89         TITLE;
90         FOOTNOTE;
91         ods noproctitle;
92
93         /* 3.) Assign a libref to the mylib folder containing your permanent data sets. Downloaded homework */
94         /*     files must be in a separate folder from the mylib folder. Assign a libref to the homework data */
95         /*     folder and add access=readonly to the end of the statement to prevent accidental corruption of */
96         /*     the original data. Create a fileref to the pdf file for output. */
97
98         libname mylib "/home/u59649056/Homeworks/mylib";
NOTE: Libref MYLIB refers to the same physical library as _TEMP2.
NOTE: Libref MYLIB was successfully assigned as follows:
      Engine:        V9
      Physical Name: /home/u59649056/Homeworks/mylib
99         libname HWDATA "/home/u59649056/Homeworks/Homework Data" access = readonly;
NOTE: Libref HWDATA refers to the same physical library as _TEMP3.
NOTE: Libref HWDATA was successfully assigned as follows:
      Engine:        V9
      Physical Name: /home/u59649056/Homeworks/Homework Data
100        filename HW11pdf "/home/u59649056/Homeworks/mylib/JRodoni_HW11_Output.pdf";
101
102
103        /* 4.) Write a single SAS step that will use the “All Texas” permanent data set as input and create a */
104        /*     permanent “Jobs” data set in mylib with the following modifications: */
105        data mylib.Jobs;
106        set HWDATA.alltx(rename=(COUNTY_FIPS_NUMBER = TempColumn));
107
108        /* (a) Change the way the following variables are displayed without changing the underlying */
109        /*        data: Percent Fatal Cases (DEATH_PERCENT for me) as a percentage with 3 decimal places, Report_Date like */
110        /*        10/29/21, death_count and positive_cases_count with comma separators and no */
111        /*         decimal places. */
112
113        format PCT_FATAL_CASES PERCENT8.3;
114        format REPORT_DATE MMDDYY8.;
115        format DEATH_COUNT POSITIVE_CASES_COUNT COMMA.;
116
117
118
119        /* (b) Convert the County_FIPS_Number variable to character. It must have the same name in */
120        /*         the output data set and use no more spaces than necessary. There is to be no note in */
121        /*         the log about numeric to character conversion. */
122
123        COUNTY_FIPS_NUMBER = put(TempColumn, 5.);
124        drop TempColumn;
125
126
127        /* (c) Create a new variable that contains the full weekday name of the Report_Date. This can */
128        /*         be done with a slight modification to one of the conversion expressions demonstrated in */
129        /*         the lecture slides */
130
131        Weekday_Name = put(Report_Date, DOWNAME9.);
132
133
134        /* (d) Create a new “Covid Week” variable that contains the week number of the Report_Date */
135        /*         relative to the date of the first Covid case. In other words, all dates reported in the */
136        /*         same week as Feb. 12, 2020, will be week 0. Those in the prior week will be -1, etc. Use */
137        /*         the macro variable in this expression so we can change the reference point if we want. */
138        start_date = "&firstcasedate"d;
```

```
139        Covid_Week = intck('week',start_date, Report_Date, 'd');
140        drop start_date;
141        RUN;

NOTE: There were 153255 observations read from the data set HWDATA.ALLTX.
NOTE: The data set MYLIB.JOBS has 153255 observations and 11 variables.
NOTE: DATA statement used (Total process time):
      real time           0.22 seconds
      user cpu time       0.12 seconds
      system cpu time     0.04 seconds
      memory              3607.25k
      OS Memory           28840.00k
      Timestamp           11/04/2021 07:49:50 PM
      Step Count                        24  Switch Count  6
      Page Faults                       0
      Page Reclaims                     823
      Page Swaps                        0
      Voluntary Context Switches        392
      Involuntary Context Switches      5
      Block Input Operations            23840
      Block Output Operations           28936


141     !
142
143
144        /* 5.) Write a single step that will use the tabled1x data set as input and create a permanent data set */
145        /*     in mylib with the following modifications: */
146
147        data mylib.statesJobs;
148        set HWDATA.tabled1x(rename=(STATE = STATENAME));
NOTE: Data file HWDATA.TABLED1X.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.
149
150        /*     a. For efficiency, do not read into the PDV any observations that have a missing state */
151        /*        value */
152
153
154        where StateName is NOT missing;
155
156        /*     b. Some of the state names have a footnote number appended to them in the form of a */
157        /*       number enclosed in a set of parentheses. We want the value in the variable named */
158        /*       State to contain only the actual name of the state. But we want to preserve the original */
159        /*   value. Use a data set option to change the name of the original state variable. When */
160        /*   the original state value ends with the number in parentheses, assign the portion of the */
161        /*   value prior to the parenthesis to the State variable. Otherwise, assign the original value */
162        /*   to the State variable. */
163
164        if substr(StateName, length(StateName),1) = ")" then State = substr(StateName,1,length(StateName)-3);
165        else State = StateName;
166        drop StateName;
167
168        /*    (c) Use a variable list in the mean function to create a new variable that is the average of */
169        /*    the values in Aug_2017 and Aug_2018. Make sure the name will not cause a “circular” */
170        /*    reference should variable lists be used on the new data set. */
171
172        Avg = mean(of Aug_2017 Aug_2018);
173
174        /*    (d) Include a statement that will delete the row and return to the top of the data step when */
175        /*    the new average value is missing. */
176
177        if Avg = . then delete;
178
179        /*    (e) Use a variable list in the sum function to create a new variable with the total of jobs */
180        /*    from all of the 2017 months. */
181
182        Total2017 = sum(of Aug_2017--Dec_2017);
183
184        /*    (f) Use a variable list in the sum function to create a new variable with the total of jobs */
185        /*    from all of the 2018 months */
186
187        Total2018 = sum(of Jan_2018--Aug_2018);
188        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).
      4 at 172:8
NOTE: There were 424 observations read from the data set HWDATA.TABLED1X.
      WHERE StateName is not null;
NOTE: The data set MYLIB.STATESJOBS has 420 observations and 18 variables.
NOTE: DATA statement used (Total process time):
```

```
      real time           0.02 seconds
      user cpu time       0.01 seconds
      system cpu time     0.00 seconds
      memory              1205.87k
      OS Memory           26532.00k
      Timestamp           11/04/2021 07:49:50 PM
      Step Count                        25  Switch Count  5
      Page Faults                       0
      Page Reclaims                     234
      Page Swaps                        0
      Voluntary Context Switches        59
      Involuntary Context Switches      0
      Block Input Operations            160
      Block Output Operations           272



189
190
191         /* 6.) Close all output destinations. Open a PDF destination to receive your output. Suppress the */
192         /*     creation of bookmarks in the PDF file. */
193
194         ods _ALL_ CLOSE;
195         ods pdf file=HW11pdf
196         bookmarklist=OFF;
NOTE: Writing ODS PDF output to DISK destination "HW11PDF", printer "PDF".
197
198         /* 7.) Write a PROC step that will report the descriptor portion of the first permanent data set created */
199         /* above in step 4. Use “Texas Covid History” as the first title and “Descriptor Portion” as the */
200         /* second title.  */
201
202
203         proc contents data=mylib.Jobs;
204         TITLE1 "Texas Covid History";
205         TITLE2 "Descriptor Portion";
206         run;

NOTE: PROCEDURE CONTENTS used (Total process time):
      real time           0.03 seconds
      user cpu time       0.03 seconds
      system cpu time     0.01 seconds
      memory              3308.46k
      OS Memory           26804.00k
      Timestamp           11/04/2021 07:49:50 PM
      Step Count                        26  Switch Count  2
      Page Faults                       0
      Page Reclaims                     727
      Page Swaps                        0
      Voluntary Context Switches        27
      Involuntary Context Switches      0
      Block Input Operations            288
      Block Output Operations           24



207
208         /* 8.) Produce a report from this permanent data set where the county_fips_number is 48029 and the */
209         /* covid week value is between -1 and 1. This fips number is from Bexar County where the first */
210         /*     Covid case in Texas was reported. Change only the second title to be “Bexar County Data */
211         /*     around 12Feb2020”. Use the macro variable instead of the literal date to construct the title. */
212
213         proc print data = mylib.Jobs;
214         where county_fips_number = "48029" and -1<=Covid_Week<=1;
215         Title2 "Bexar County Data around &firstcasedate";
216         run;

NOTE: There were 21 observations read from the data set MYLIB.JOBS.
      WHERE (county_fips_number='48029') and (Covid_Week>=-1 and Covid_Week<=1);
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.08 seconds
      user cpu time       0.04 seconds
      system cpu time     0.02 seconds
      memory              2655.78k
      OS Memory           27316.00k
      Timestamp           11/04/2021 07:49:50 PM
      Step Count                        27  Switch Count  2
      Page Faults                       0
      Page Reclaims                     463
      Page Swaps                        0
      Voluntary Context Switches        249
      Involuntary Context Switches      5
      Block Input Operations            28672
      Block Output Operations           16
```

```
217
218         /* 9.) Print the descriptor portion of the permanent data set created in step 5. The printout must list */
219         /*    the variables in creation order. Use “2017-2018 Jobs Data” as the first title and “Descriptor */
220         /*    Portion” as the second title. */
221
222         proc contents data = mylib.statesjobs varnum;
223         TITLE1 "2017-2018 Jobs Data";
224         TITLE2 "Descriptor Portion";
225         run;

NOTE: PROCEDURE CONTENTS used (Total process time):
      real time           0.01 seconds
      user cpu time       0.02 seconds
      system cpu time     0.00 seconds
      memory              927.03k
      OS Memory           26804.00k
      Timestamp           11/04/2021 07:49:50 PM
      Step Count                        28  Switch Count  2
      Page Faults                       0
      Page Reclaims                     101
      Page Swaps                        0
      Voluntary Context Switches        23
      Involuntary Context Switches      0
      Block Input Operations            288
      Block Output Operations           16


226
227         /* 10.) */
228         proc print data=mylib.statesjobs;
229         Title2 "Data Portion";
230         run;

NOTE: There were 420 observations read from the data set MYLIB.STATESJOBS.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           1.13 seconds
      user cpu time       1.14 seconds
      system cpu time     0.00 seconds
      memory              2357.06k
      OS Memory           28336.00k
      Timestamp           11/04/2021 07:49:51 PM
      Step Count                        29  Switch Count  1
      Page Faults                       0
      Page Reclaims                     465
      Page Swaps                        0
      Voluntary Context Switches        13
      Involuntary Context Switches      7
      Block Input Operations            0
      Block Output Operations           384


231
232         /* 11.) */
233         ods pdf close;
NOTE: ODS PDF printed 31 pages to /home/u59649056/Homeworks/mylib/JRodoni_HW11_Output.pdf.
234
235         /* 12.) Use the information you discovered about the downloaded data, the log and the report */
236         /* information contained in your PDF output document to find the answers to the questions below */
237         /* and include the answers in a comment section at the bottom of your program file: */
238
239         /* a. On what day of the week was the first case reported in Bexar County? */
240
241         /* Wednesday                                                                  */
242
243         /*b. What was the Positive_Cases_Count on Saturday of Covid Week 1 in Bexar County? */
244
245         /*    4 */
246
247         /* c. How many observations are in the tabled1x data set, how many were read in by the */
248         /*    data step and how many were written out? */
249
250         /*        519, 424, 420 */
251
252         /* d. How does the average number of August Government jobs in the District of Columbia */
253         /*    compare with Texas? (I know it’s hard to follow the split table. Use the Obs value to */
254         /*    link the two sections together.) */
255
256         /*         on average texas has about 8 times as many government jobs in August than in DC */
257
258         /* 13.) Save the final version of the program and convert it to a PDF file with a name like */
259         /*     FKincheloe_HW11_prog.pdf. Convert the log to PDF. */
```

```
260
261         /* 14.) Upload and submit the three documents to the assignment on Canvas. */
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284         OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
294
```