```
/*********************************************************************************/
/* Program Name: STAT 604 HW#11                                                  */
/* Date Created: 11/1/2021                                                       */
/* Author: Jack Rodoni                                                           */
/* Purpose: STAT 604 HW#11                                                       */
/* Date Modified: 11/04/2021                                                     */
/* Location: /home/u59649056/Homeworks/JRodoni_Homework11.sas                    */
/*********************************************************************************/


/* 1. ) The first Covid19 case in Texas was reported on February 12, 2020. Below the program header, */
/*      include a macro assignment statement to create a macro variable that contains this date in a */
/*      manner that can be used throughout the program in data step statements and in titles. */

%let firstcasedate=12Feb2020;

/* 2.) Include housekeeping statements to clear titles and footnotes and suppress the printing of */
/*     procedure titles. */

TITLE;
FOOTNOTE;
ods noproctitle;

/* 3.) Assign a libref to the mylib folder containing your permanent data sets. Downloaded homework */
/*     files must be in a separate folder from the mylib folder. Assign a libref to the homework data */
/*     folder and add access=readonly to the end of the statement to prevent accidental corruption of */
/*     the original data. Create a fileref to the pdf file for output. */

libname mylib "/home/u59649056/Homeworks/mylib";
libname HWDATA "/home/u59649056/Homeworks/Homework Data" access = readonly;
filename HW11pdf "/home/u59649056/Homeworks/mylib/JRodoni_HW11_Output.pdf";


/* 4.) Write a single SAS step that will use the “All Texas” permanent data set as input and create a */
/*     permanent “Jobs” data set in mylib with the following modifications: */
data mylib.Jobs;
    set HWDATA.alltx(rename=(COUNTY_FIPS_NUMBER = TempColumn));

/*  (a) Change the way the following variables are displayed without changing the underlying */
/*          data: Percent Fatal Cases (DEATH_PERCENT for me) as a percentage with 3 decimal places, Report_Date like */
/*          10/29/21, death_count and positive_cases_count with comma separators and no */
/*         decimal places. */

    format PCT_FATAL_CASES PERCENT8.3;
    format REPORT_DATE MMDDYY8.;
    format DEATH_COUNT POSITIVE_CASES_COUNT COMMA.;



/*  (b) Convert the County_FIPS_Number variable to character. It must have the same name in */
/*         the output data set and use no more spaces than necessary. There is to be no note in */
/*         the log about numeric to character conversion. */

    COUNTY_FIPS_NUMBER = put(TempColumn, 5.);
    drop TempColumn;


/*  (c) Create a new variable that contains the full weekday name of the Report_Date. This can */
/*         be done with a slight modification to one of the conversion expressions demonstrated in */
/*         the lecture slides */

    Weekday_Name = put(Report_Date, DOWNAME9.);


/*  (d) Create a new “Covid Week” variable that contains the week number of the Report_Date */
/*         relative to the date of the first Covid case. In other words, all dates reported in the */
/*         same week as Feb. 12, 2020, will be week 0. Those in the prior week will be -1, etc. Use */
/*         the macro variable in this expression so we can change the reference point if we want. */
    start_date = "&firstcasedate"d;
    Covid_Week = intck('week',start_date, Report_Date, 'd');
    drop start_date;
RUN;
```

```
/* 5.) Write a single step that will use the tabled1x data set as input and create a permanent data set */
/*     in mylib with the following modifications: */

data mylib.statesJobs;
    set HWDATA.tabled1x(rename=(STATE = STATENAME));

/*     a. For efficiency, do not read into the PDV any observations that have a missing state */
/*        value */


    where StateName is NOT missing;

/*     b. Some of the state names have a footnote number appended to them in the form of a */
/*        number enclosed in a set of parentheses. We want the value in the variable named */
/*        State to contain only the actual name of the state. But we want to preserve the original */
/*        value. Use a data set option to change the name of the original state variable. When */
/*        the original state value ends with the number in parentheses, assign the portion of the */
/*        value prior to the parenthesis to the State variable. Otherwise, assign the original value */
/*        to the State variable. */

    if substr(StateName, length(StateName),1) = ")" then State = substr(StateName,1,length(StateName)-3);
    else State = StateName;
    drop StateName;

/*     (c) Use a variable list in the mean function to create a new variable that is the average of */
/*         the values in Aug_2017 and Aug_2018. Make sure the name will not cause a “circular” */
/*         reference should variable lists be used on the new data set. */

    Avg = mean(of Aug_2017 Aug_2018);

/*     (d) Include a statement that will delete the row and return to the top of the data step when */
/*         the new average value is missing. */

    if Avg = . then delete;

/*     (e) Use a variable list in the sum function to create a new variable with the total of jobs */
/*         from all of the 2017 months. */

    Total2017 = sum(of Aug_2017--Dec_2017);

/*     (f) Use a variable list in the sum function to create a new variable with the total of jobs */
/*         from all of the 2018 months */

    Total2018 = sum(of Jan_2018--Aug_2018);
run;


/* 6.) Close all output destinations. Open a PDF destination to receive your output. Suppress the */
/*     creation of bookmarks in the PDF file. */

ods _ALL_ CLOSE;
ods pdf file=HW11pdf
    bookmarklist=OFF;

/* 7.) Write a PROC step that will report the descriptor portion of the first permanent data set created */
/*  above in step 4. Use “Texas Covid History” as the first title and “Descriptor Portion” as the */
/*  second title.  */


proc contents data=mylib.Jobs;
TITLE1 "Texas Covid History";
TITLE2 "Descriptor Portion";
run;

/* 8.) Produce a report from this permanent data set where the county_fips_number is 48029 and the */
/*  covid week value is between -1 and 1. This fips number is from Bexar County where the first */
/*     Covid case in Texas was reported. Change only the second title to be “Bexar County Data */
/*     around 12Feb2020”. Use the macro variable instead of the literal date to construct the title. */

proc print data = mylib.Jobs;
where county_fips_number = "48029" and -1<=Covid_Week<=1;
```

```
Title2 "Bexar County Data around &firstcasedate";
run;

/* 9.) Print the descriptor portion of the permanent data set created in step 5. The printout must list */
/*     the variables in creation order. Use “2017-2018 Jobs Data” as the first title and “Descriptor */
/*     Portion” as the second title. */

proc contents data = mylib.statesjobs varnum;
TITLE1 "2017-2018 Jobs Data";
TITLE2 "Descriptor Portion";
run;

/* 10.) */
proc print data=mylib.statesjobs;
Title2 "Data Portion";
run;

/* 11.) */
ods pdf close;

/* 12.) Use the information you discovered about the downloaded data, the log and the report */
/*      information contained in your PDF output document to find the answers to the questions below */
/*      and include the answers in a comment section at the bottom of your program file: */

/*      a. On what day of the week was the first case reported in Bexar County? */

/*      Wednesday                                                                  */

/*      b. What was the Positive_Cases_Count on Saturday of Covid Week 1 in Bexar County? */

/*         4 */

/*      c. How many observations are in the tabled1x data set, how many were read in by the */
/*         data step and how many were written out? */

/*         519, 424, 420 */

/*      d. How does the average number of August Government jobs in the District of Columbia */
/*         compare with Texas? (I know it’s hard to follow the split table. Use the Obs value to */
/*         link the two sections together.) */

/*         on average texas has about 8 times as many government jobs in August than in DC */

/* 13.) Save the final version of the program and convert it to a PDF file with a name like */
/*      FKincheloe_HW11_prog.pdf. Convert the log to PDF. */

/* 14.) Upload and submit the three documents to the assignment on Canvas. */
```