

Stat 604

Assignment 12 - SAS

OBJECTIVES: In this assignment you will practice using functions, using the PRINT procedure options and creating user defined formats.

You should have all the information you need to complete this assignment by viewing the first 11 SAS lectures.

This assignment will use the permanent, modified jobs data set that was created from the tabled1x data set in the previous assignment. If you had difficulty creating this data set, the professor's version, **jobs2018**, is available on Canvas and in the Fall2021 folder on SoDA.

1. Add a header comment section to the beginning of a new program in your SAS session. Be sure to include a comment line above each section of the program that identifies the associated assignment step and a brief description of what the section is doing. Include housekeeping statements to clear titles and footnotes and suppress the printing of procedure titles.
2. Assign a libref to the **mylib** folder containing your permanent data sets. Create a fileref to the pdf file for output.
3. Open the PDF destination to receive your output.
4. Create a temporary custom format that can be applied to any of the columns containing the number of jobs. All values less than 10 will be displayed as 'Very Low'. Values between 10 and 100 inclusive display 'Low'. Values above 100 through 200 are 'Medium'. Values above 200 through 500 are 'Medium High'. Values above 500 through 1000 are 'High' and all values above 1000 are 'Very High'.
5. Write a single SAS step that will use the "Jobs" data set from the previous assignment as input and create a temporary data set with the following modifications:
 - a. Without creating a new variable, recalculate the value in the variable containing the 2017 Jobs total by subtracting the Aug_2017 value from the existing value. Use one of the numeric functions to ensure that a missing August value does not cause the total to be missing. (Hint: You can cause "subtraction" to occur by making one of your arguments to the function negative.)
 - b. Without creating a new variable, change the Industry values to proper case.
 - c. Since words like "And" should not be capitalized, replace them with "and" in the Industry variable.
6. Without creating a new data set, reorder the rows in the new temporary data set by state and industry to accommodate the reports in the next two steps.
7. Produce a report listing the categories based on the "August Average" column by applying your custom format to that column. The first few rows of the report are shown below. Your output must match this sample. Use statements, proc options, temporary labels and temporary formats as needed to accomplish this purpose. Note: The state column is used in place of the default OBS column to identify records in the report.

Jobs Analysis by August Categories

State	Industry	August Average Jobs
Alabama	Construction	Low
Alabama	Education and Health Services	Medium High
Alabama	Financial Activities	Low
Alabama	Government	Medium High
Alabama	Leisure and Hospitality	Medium High
Alabama	Manufacturing	Medium High
Alabama	Professional and Business Services	Medium High
Alabama	Trade, Transportation, and Utilities	Medium High

8. Produce a report of jobs summaries from the midwestern states of 'Texas', 'Oklahoma', 'Kansas', 'Nebraska', 'South Dakota', and 'North Dakota' based on the values in the Total 2017 and Total 2018 columns. The first table of the report are shown below. Your output, including titles, must match this sample. Use statements, proc options, temporary labels and temporary formats as needed to accomplish this purpose. Note: The var statement can be used without any variables to exclude unwanted columns from appearing in the report. State and Industry are used to replace the OBS column for identifying rows.

Midwest States Job Summary Thousands of Jobs

State	Industry	Sep. - Dec. 2017	Jan. - Aug. 2018
Kansas	Construction	238	480
	Education and Health Services	792	1,585
	Financial Activities	312	631
	Government	1,025	2,066
	Leisure and Hospitality	519	1,049
	Manufacturing	650	1,317
	Professional and Business Services	716	1,460
	Trade, Transportation, and Utilities	1,076	2,167
Kansas		5,327	10,754

9. In a single PROC step, create a copy of the temporary data set created earlier in the assignment, reordered by the descending values of Aug_2018. The copy will also be a temporary data set.
10. Write a single PROC step that will list and report the descriptor portion of all data sets in the WORK library. Supply an appropriate title.
11. Use the last data set created to print a "Top 10" list of the industries and states with the highest number of jobs in August 2018. Suppress the printing of observation numbers. Include only the Aug_2018, Industry, and State columns in that order. Supply an appropriate first title and use

“Thousands of Jobs” as the second title line. Give Aug_2018 a label of “August 2018 Employment” and show the values with a comma separator.

12. Close the PDF destination.
13. Use the log and 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. Which categories of jobs are represented by the August average in Texas and how many of each category?
 - b. Which of the Midwest states had the fewest total jobs in the Sep. – Dec. 2017 time period? What was that total?
 - c. What other objects besides data sets are listed in the output from step 10?
 - d. Which state and industry have the most jobs in 2018?
 - e. How many different states are in the top 10 list?
 - f. Which state appears the most in the top 10 list?
14. Save the final version of the program and convert it to a PDF file with a name like FKincheloe_HW12_prog.pdf. Convert the log to PDF.
15. Upload and submit the three documents to the assignment on Canvas.