Stat 604
Assignment 6 - SAS

This assignment reinforces the concepts covered in lectures SAS01 through SAS11. Specifically, it focuses on the use of arrays and variable lists to process data. You will use these along with other methods to rotate data. It will also allow you to practice using the match merge process to combine data.

Perform each of the exercises listed below. To the extent you have been taught to control it, your output should match that in the PDF file posted on eCampus. In many cases you will be given the specific steps necessary to accomplish the task. However, as the assignment progresses, you may not be told all of the preliminary or detailed tasks required to accomplish an objective. In those situations, it is up to you to figure out how to prepare for and accomplish the objective. Download the files **chopnjoe19.sas7bdat** and **charities.sas7bdat** from the Assignment Files section on eCampus to your SAS homework library folder for use in this assignment.

Chop and Joe Remodeling Co. encourages its employees to support charitable organizations. It allows employees to use payroll deductions to automatically make contributions to over 250 different charities. Currently, each employee may contribute to up to 10 different charities. The **chopnjoe19** data set is already sorted by employee_id and contains information about each employee along with last year's contribution amounts and the charitable organizations to which they have donated. Even though the available charities represent a broad range of categories from aging to religious organizations, Chop and Joe is specifically interested in helping those in poverty. The **charities** data set contains the list of available charities and the category to which each organization belongs. You need to consolidate this information and restructure it to facilitate analysis of giving to the desired categories. All of your code should be written so that Chop and Joe can increase the number of charities to which a person contributes without you needing to upgrade your code. (In other words, the number 10 never appears in your program code.)

1. Begin your program with the required header, filename, and libname statements. As always, your program must include comments in the appropriate places. In addition to the comment boxes you have been using above data and proc steps, use single comment lines to identify the sub-steps that are being accomplished within the data step.
2. From the **chopnjoe19** data set, create a narrow data set containing the employee id and organization name. Use a data step and an array to accomplish this transformation. For efficiency, only read the variables you need from the input data set. The resulting data set should not contain any missing values for the organization name. You may find it handy later on to have the index value from your array so keep it in the output data set. The resulting data set should have 2243 observations and 3 variables. The first few values are shown below:

Obs	Employee_ID	i	Organization
1	120101	1	Appalachian Trail Conservancy
2	120102	1	Living Water International
3	120102	2	American Refugee Committee
4	120102	3	Give Kids The World
5	120103	1	Ronald McDonald House Charities

3. Sort the new data set in place so that it can be merged with the data contained in the **charities** data set.
4. In your work library create a sorted copy of the **charities** data set for merging with the new data set created in the previous steps.
5. Use match merge to create a new data set that combines the charity category with the employee/organization data. Make sure no charities are included in the data when there have been no contributions made to that organization. The resulting data set should have 2243 observations and 4 variables. The first few values are shown below:

Obs	Employee_ID	i	Organization	Category
1	120196	4	AARP Foundation	Aging
2	120271	10	AARP Foundation	Aging
3	120714	1	AARP Foundation	Aging
4	120741	6	AARP Foundation	Aging
5	121027	1	AARP Foundation	Aging

6. Use the transpose procedure to transform this data set back into a wide data set of charity categories that can be merged with the original contributions data set. The resulting data set should have 424 observations and 11 variables. The first few values are shown below:

	Employee_ID	Donee_Type1	Donee_Type2	Donee_Type3	Donee_Type4	Donee_Type5	Donee_Type6	Donee_Type7
1	120101	Environment						
2	120102	Hunger	Relief	Children				
3	120103	Children						
4	120104	Environment	Disease	Children	Religious	Religious	Hunger	Disease
5	120105	Disease	Children	Environment	Disease	Children	Religious	Animals
6	120106	Aging	Religious	Children	Disease	Disease	Children	
7	120107	Children	Religious					

(NOTE: Not all columns are visible due to space limitations.)

7. Use a single data step to perform a match merge that combines the organization types with the original **chopnjoe19** data set and creates some summary variables. Apply the appropriate labels to your new variables. The successful solution will most likely use some arrays. Create a variable that contains the amount the employee contributed to Relief charities based on the organization type value loaded from the **charities** data set. Likewise, create a variable for the total the employee donated to charities in the Hunger category. Create a third variable that contains the total amount the employee donated to all charities. Create a fourth variable that shows the percentage of the employee's salary that was contributed. Label and format the percentage as shown in the output data.
8. Print the descriptor portion and then the data portion of the final data set. When you print the data portion, suppress the printing of observation numbers and print only the variables shown in the output posted on eCampus.
9. Convert the program and log to PDF files and submit them to eCampus along with your SAS output.