SAS Data set for Q1-Q4: WORK.Creditacc_files

Q1: Create a sub data set based on the condition age between 45 and 55 or income is more than $40000.

Q2: Create a sub data set based on the condition age between 45 and 55 or income is not missing and less than $40000.

Q3: Create a sub data set containing the top 30 records. If values of income are missing, then impute (replace) them as zero.

Q4: Obtain a sub data set based on the condition utilization>40 or credit_limit is missing or 0. Additionally,

- Create a field 'utilization' using 'revolve_bal'/'credit limit' if 'credit_limit'>0
- Impute (replace) all missing values of age and income by zero
- Only keep the columns id, age, income, revolve_bal, credit_limit and utilization in the resulting table

SAS Data set for Q5-Q8: WORK.Fin_account

Browsing the data set 'WORK.Fin_account' which contains the credit card account information (snapshot) at certain month. Please solve the following questions (for each question you will create a new SAS data set) using data step and conditional statements.

Q5. Create a column 'utilization' based on the calculation 100*Tot_New_Bal1012/ Credit_Lim1012, if Credit_Lim1012 is missing or 0 then assign value 999 to the new variable 'utilization'.

Q6. Create a new table 'normal_fin_acc' only containing the records with the restriction condition '0=<utilization<=100'.

Q7. Following Q6, apply 'SELECT...when' statement on the variable 'utilization' of table 'normal_fin_acc' to generate new variable 'ut_level', whose value is determined by the following rules:

70 to 100→'HIGH'

35 to 69→ 'MEDIUM'

Less than 35→'LOW'

Finally only keep the following columns in the resulting data set: 'ACCT_KEY', 'Tot_New_Bal1012', 'Credit_Lim1012', 'utilization', 'ut_level' and 'Purchases'.

Q8. Following Q7, create a new table 'acc_ind' using the condition 'Purchases>0' to restrict the observations in the resulting data set. In addition,

- Create the following missing dummy variables based on 'Tot_New_Bal1012' and 'Credit_Lim1012', i.e. the value of variable 'bal_miss' is 1 if the value of 'Tot_New_Bal1012' is missing otherwise is 0, and the value of 'limit_miss' is 1 if the value of 'Credit_Lim1012' is missing otherwise is 0
- Only keep the columns 'ACCT_KEY', 'ut_level', 'bal_miss' and 'limit_miss'