Pre Data Quality Macro

PREDQ Macro:

Description	This macro is used to produce a summary report containing number of non-missing records in each variable for two datasets and highlight the following: 1) Decline in number of non-missing values in a variable. 2) Any format changes for a variable in two datasets. 3) New or dropped variables. This produces output in excel format or in "Results – SAS Report" tab (in SAS EG). >>Orange implies 0% to 10% drop in number of records >>Red implies 10% to 100% drop in number of records >>Yellow on "Old vs New Data Type" column cell implies data type change >>Light Green on "Old vs New Data Type" column cell implies data type is similar but format is different "Old vs New Data Type" will not appear if NEW_DTSET is blank. Script Location : G:\dqmacro\predq.sas
Syntax	%PREDQ(DOMAIN= ,DB= ,NEW_DTSET= ,OUTPUT_DIR=);
Parameters	%PREDQ(DOMAIN = Name of the library like health, social etc.. ,DB = Database name to compare to. It will automatically pull last member in the cluster unless providing memnum in both DB & NEW_DTSET parameters. ,NEW_DTSET=New Database/cluster member. Keep this field blank if both the databases to compare are in same cluster. Provide memnum/dataset name to compare specific member. ,OUTPUT_DIR=Location of output.
Example	1) When both comparing datasets are in same cluster: %PREDQ(DOMAIN=HEALTH, DB=MHMHLTH_MHID_2010APR, NEW_DTSET=, OUTPUT_DIR=G:\dmusers\username); 2) When both comparing datasets are in different cluster: %PREDQ(DOMAIN=common, DB=conc_2006, NEW_DTSET=conc_2011, OUTPUT_DIR=G:\dmusers\username); 3) When want to compare specific members in same/different cluster: %PREDQ(DOMAIN=REGISTRY, DB=VSA_MMDF_2000JAN(MEMNUM=2), NEW_DTSET=VSA_MMDF_2000JAN(MEMNUM=3), OUTPUT_DIR=G:\dmusers\username);
Sample Output	*Orange implies 0% to 10% drop in number of records* *Red implies 10% to 100% drop in number of records* *Yellow on "Old vs New Data Type" column implies data type change* *Light Green on "Old vs New Data Type" column implies data type is similar but format is different* ***Latest Cluster Member: conc_2011(MEMNUM=1)*** ***Previous Cluster member: conc_2006(MEMNUM=1)***

Variables In Cluster	Previous Cluster Member Count	Latest Cluster Member Count	Percent Change	Previous vs Latest Data Type
DAUID	.	2179	Variable not in Previous Cluster	Char(8)
CC_F_POP	54529	2176	-96.01%	Num(8)(BEST12) VS Num(8)
DB_IR2011	.	2179	Variable not in Previous Cluster	Char(1)
CTNAME	54626	2179	-96.01%	Char(9)($9) VS Char(7)
CTCODE	.	2179	Variable not in Previous Cluster	Char(4)
DA_UID	54626	.	Variable not in Latest Cluster	Char(10)($10)