## 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=MHMHLT_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 | ![](https://via.placeholder.com/150) |