Paper 075-2009

Renaming in Batches

Vincent Weng, Educational Testing Service, Princeton, NJ
Ying Feng, CTB/McGraw-Hill, Monterey, CA

ABSTRACT

Renaming SAS® variables may sound easy, but what if you have hundreds of variables that must be renamed? In this instance it will become an annoying and error-prone process if you simply use Rename statements to rename each variable individually. This paper addresses how you can perform the renaming of many variables quickly and error-free.

A group of "renaming" macros will be created, which perform different renaming jobs, i.e. renaming all the variables in a SAS data set, renaming partial variables, adding prefixes, adding suffixes, and/or replacing the prefix or the suffix. It's a very handy tool for tailoring variable names in batches quickly to fit different needs.

INTRODUCTION

Renaming SAS® variables happens a lot in our real life, especially in the Data Analysis area. Sometimes when we are handling the real data, for either security issues or per clients' requests, variables are needed to be renamed from common names such as "var 1", "var 2", …, "var n", to new names with prefix or suffix of subject or types. But with hundreds of variables, it becomes very annoying and time consuming to rename them individually. Different macros of renaming variables will be discussed here, which add a prefix, or replace prefixes and suffixes. SAS DICTIONARY tables are read-only SAS data views that contain lists of things that are related to the current SAS session. We can retrieve a number of variables and their names from the DICTIONARY tables. PROC SQL's DICTIONARY.TABLES and COLUMNS are utilized to rename all the variables in a SAS data set. The same technique can also be used to rename only selected variables. The macro function %SYSFUNC allows access by the macro processor to most data step functions and several SCL functions, which allows you to access dataset observations. The data set functions, OPEN, CLOSE and VARNAME of %SYSFUNC will be used to replace the prefix and suffix.

DATA:

```
/* Creating a dataset */
DATA A;
input id $4. before_var1_after before_var2_after before_var3_after;
datalines;
i001 1 2 3
i002 3 4 5
i003 6 7 8
i004 9 10 12
;
run;
```

MACRO 1: ADD PREFIX ON ALL VARIABLES

Extract number of variables from PROC SQL's DICTIONARY.TABLES and the names of the variables from DICTIONARY.COLUMNS, and then attach a prefix to each variable name.

```
/* Adding Prefix on all variables */

%macro rename(lib,dsn,newname);
proc contents data=&lib..&dsn;
title 'before renaming';
run;

proc sql noprint;
```

```
 select nvar into :num_vars
 from dictionary.tables
 where libname="&LIB" and memname="&DSN";

 select distinct(name) into :var1-:var%trim(%left(&num_vars))
 from dictionary.columns
 where libname="&LIB" and memname="&DSN";
 quit;
run;

proc datasets library = &LIB;
modify &DSN;
rename
%do i = 1 %to &num_vars.;
&&var&i = &newname._&&var&i.
%end;
;
quit;
run;

proc contents data=&lib..&dsn.;
title 'after renaming';
run;
%mend rename;

DATA B;
set A;
run;

%rename(WORK,B,Try1);
```

Partial OUTPUT of MACRO 1:

```
                    before renaming

            Alphabetic List of Variables and Attributes

                #    Variable             Type     Len

                1    id                   Char       4
                2    before_var1_after    Num        8
                3    before_var2_after    Num        8
                4    before_var3_after    Num        8

                       after renaming

            Alphabetic List of Variables and Attributes

                #    Variable                   Type     Len

                1    Try1_id                    Char       4
                2    Try1_before_var1_after     Num        8
                3    Try1_before_var2_after     Num        8
                4    Try1_before_var3_after     Num        8
```

MACRO 2: ADD PREFIX ON SELECTED VARIABLES

Same approach of renaming all variables but using start and end positions to rename only the selected variables. Variable list need to be sorted before running this macro.

```
/* Adding Prefix on Selected Variables */

%macro addprefix(lib,dsn,start,end,newname);
proc contents data=&lib..&dsn;
title 'before renaming';
run;

proc sql noprint;
 select nvar into :num_vars
 from dictionary.tables
 where libname="&LIB" and memname="&DSN";

 select distinct(name) into :var1-:var%trim(%left(&num_vars))
 from dictionary.columns
 where libname="&LIB" and memname="&DSN";
quit;
run;

proc datasets library = &LIB;
modify &DSN;
rename
%do i = &start. %to &end.;
&&var&i = &newname_&&var&i.
%end;
;
quit;
run;

proc contents data=&lib..&dsn;
title 'Adding Prefix on Selected variables';
run;
%mend addprefix;

DATA C;
set A;
run;

%addprefix(WORK,C,2,4,Try2);
```

Partial OUTPUT of MACRO2:

```
                    Adding Prefix on Selected variables

                 Alphabetic List of Variables and Attributes

                  #    Variable                  Type     Len

                  1    id                        Char       4
                  2    Try2_before_var1_after    Num        8
                  3    Try2_before_var2_after    Num        8
                  4    Try2_before_var3_after    Num        8
```

MACRO 3: REPLACE PREFIX ON SELECTED VARIABLES

The %SYSFUNC macro has allowed access to the SAS component language inside of traditional DATA step programming. We can easily retrieve variable information by using the data set functions, OPEN, CLOSE and VARNAME and replace the existing prefix with the new name.

```
/* Replacing Prefix on Selected Variables */

%macro replaceprefix(lib,dsn,start,end,oldprefix,newprefix);
proc contents data=&lib..&dsn.;
title 'before renaming';
run;

data temp;
set &lib..&dsn.;
run;

%LET ds=%SYSFUNC(OPEN(temp,i));
%let ol=%length(&oldprefix.);
%do i=&start %to &end;
 %let dsvn&i=%SYSFUNC(VARNAME(&ds,&i));
 %let l=%length(&&dsvn&i);
 %let vn&i=&newprefix.%SUBSTR(&&dsvn&i,&ol+1,%EVAL(&l-&ol));
%end;
data &lib..&dsn.;
set temp;

%do i=&start %to &end;
&&vn&i=&&dsvn&i;
drop &&dsvn&i;
%end;
%let rc=%SYSFUNC(CLOSE(&ds));
proc contents data=&lib..&dsn.;
title 'Replacing Prefix on Selected variables ';
run;
%mend replaceprefix;

DATA D;
set A;
run;

%replaceprefix(WORK,D,2,4,before,Try3);
```

Partial OUTPUT of MACRO 3:

Replacing Prefix on Selected variables

Alphabetic List of Variables and Attributes

#	Variable	Type	Len
1	id	Char	4
2	Try3_var1_after	Num	8
3	Try3_var2_after	Num	8
4	Try3_var3_after	Num	8

Reproduce MACRO 2 result by replacing “before” with BLANK input in MACRO 3.

```
/* Adding Prefix = Replacing BLANK Prefix on Selected Variables */

DATA D;
set A;
run;

%replaceprefix(WORK,D,2,4, ,Try3_);
```

Partial OUTPUT of MACRO 3:

Replacing Prefix on Selected variables

Alphabetic List of Variables and Attributes

#	Variable	Type	Len
1	id	Char	4
2	Try3_before_var1_after	Num	8
3	Try3_before_var2_after	Num	8
4	Try3_before_var3_after	Num	8

MACRO 4: REPLACE SUFFIX ON SELECTED VARIABLES

We can use same approach of Macro 3, Using %SYSFUNC to retrieve the dataset and variable information and replace the existing suffix with new name.

```
/* Replacing Suffix on Selected Variables */

%macro replacesuffix(lib,dsn,start,end,oldsuffix,newsuffix);
proc contents data=&lib..&dsn.;
title 'before renaming';
run;

data temp;
set &lib..&dsn.;
run;

%LET ds=%SYSFUNC(OPEN(temp,i));
%let ol=%length(&oldsuffix.);
%do i=&start %to &end;
 %let dsvn&i=%SYSFUNC(VARNAME(&ds,&i));
 %let l=%length(&&dsvn&i);
 %let vn&i=%SUBSTR(&&dsvn&i,1,%EVAL(&l-&ol))&newsuffix.;
%end;
data &lib..&dsn.;
set temp;

%do i=&start %to &end;
&&vn&i=&&dsvn&i;
drop &&dsvn&i;
%end;
%let rc=%SYSFUNC(CLOSE(&ds));
proc contents data=&lib..&dsn.;
title ' Replacing Suffix on Selected variables ';
run;
%mend replacesuffix;

DATA E;
set A;
run;

%replacesuffix(WORK,E,2,4,after,Try4);
```

Partial OUTPUT of MACRO 4:

Replacing Suffix on Selected variables

Alphabetic List of Variables and Attributes

```
                          #    Variable              Type      Len

                          1    id                    Char       4
                          2    before_var1_Try4      Num        8
                          3    before_var2_Try4      Num        8
                          4    before_var3_Try4      Num        8
```

Using MACRO 4 to Add Suffix on Selected Variables.

```
/* Adding Suffix = Replacing BLANK Suffix on Selected Variables */

DATA F;
set A;
run;

%replacesuffix(WORK,E,2,4, ,_Try4);
```

Partial OUTPUT of MACRO 4:

```
                        Replacing Suffix on Selected variables

                    Alphabetic List of Variables and Attributes

                     #    Variable                    Type      Len

                     1    id                          Char       4
                     2    before_var1_after_Try4      Num        8
                     3    before_var2_after_Try4      Num        8
                     4    before_var3_after_Try4      Num        8
```

CAVEATS

There are some caveats of using PROC SQL's DICTIONARY.TABLES and COLUMNS. The library and dataset names must in Uppercase. Using the "upcase" function is highly recommended here. Besides that, only the prefix can be concatenated, we need to find a way to concatenate the suffix.

CONCLUSION

As demonstrated above, with the help of renaming macros we can perform different renaming jobs with hundreds of variable names in batches. Using the macro function %SYSFUNC is the winner, because it can perform all the renaming jobs we need for our routine work. There are some limitations of using PROC SQL's DICTIONARY here but it is still a very useful tool, which provides us an alternative way of renaming the variables.

REFERENCES

P. Ravi(2003). "Renaming All Variables in a SAS Data Set Using the Information from PROC SQL's Dictionary Tables." Proceedings of the Twenty-Eighth Annual SAS® Users Group International Conference, Seattle, 2003

D. Morgan(2003). "%Fun&With%SYSFUNC" Proceedings of the Twenty-Eighth Annual SAS® Users Group International Conference, Seattle, 2003

ACKNOWLEDGMENTS

Many thanks to my Co-Author Ying Feng and the following people: Cathy Trapani, David Williamson, Bruce Kaplan and Ted Blew for reviewing and providing constructive feedback.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Vincent Weng
Enterprise: Educational Testing Services
Address: Rosedale Road
City, State ZIP: Princeton, NJ 08541
Work Phone: (609)7345635
Fax: (609)7341090
E-mail: vweng@ets.org