

Paper 056-29

Reading Multiple Complex Report Files

Tugluke Abdurazak, PricewaterhouseCoopers LLP

ABSTRACT

In auditing, accounting, financial analysis and inventory analysis, we frequently receive complex report files from clients that have to be converted to SAS® data sets. For example, a financial statement from a company may contain assets, income, expenses, taxes, gross profit and other summarized information, all in one report. In order to analyze the financial statement, the information in the report may have to be read into different data sets. If the content of the reports is simple, or the quantity of reports is small, this work can easily be done by using Microsoft Excel, Ultra Edit or Monarch for example. However, when there are a large number of complex reports that need to be converted, the SAS data step and SAS macro programming are applicable and effective methods.

This paper describes the construction of a SAS program to read a complex financial statement. A basic knowledge of the SAS data step and the macro language is required.

DIFFERENT WAYS TO READ MULTIPLE FILES

As a preliminary for the subsequent material, I would like to introduce some general methods to read multiple files. Suppose one received nine files (file1.dat, file2.dat, ..., file9.dat) with exactly the same layout. The following SAS code can be used to read the list of files and output a combined data set name FILE_OUT.

```
filename flist ( "c:\temp\file1.dat"
                "c:\temp\file2.dat"
                "c:\temp\file3.dat"
                "c:\temp\file4.dat"
                "c:\temp\file5.dat"
                "c:\temp\file6.dat"
                "c:\temp\file7.dat"
                "c:\temp\file8.dat"
                "c:\temp\file9.dat" );

data file_out;
  infile flist;
  input a b c d e;
run;
```

Alternatively, the following SAS macro code can be used to read the list of files. Note that the PROC APPEND procedure is used to combine separate data sets. The macro tracks the number of files being read during the execution.

```
*--- List of files to be read;
%let flist = file1 file2 file3 file4 file5 file6 file7 file8 file9;

*--- Macro readit loop through the list of file names and read them in;
%macro readit( varlist =, fn= );
  %local i ;
  %let i = 1;
  %do %until ( %scan(&varlist,&i,%str( )) = %str() );
    %let var&i = %scan(&varlist,&i,%str( ));
    data &&var&i;
      length source $10. ;
      infile "c:\temp\&&var&i...dat" pad;
      *---Input statement: May need to change;
      input a b c d e;
      source = "&&var&i" ;
    run;

  *---Combine all data sets;
  %if &i = 1 %then
  %do;
    data &fn ;
```

```

        set &&var&i;
        run;
    %end;
    %else
    %do;
        proc append base=&fn data=&&var&i;
            run;
        %end;
    %let i = %eval(&i+1);

%end;

%let varcnt = %eval(&i-1);
%put The number of data sets being read are &varcnt ;

%mend loopit;

*--- Execute macro;
options obs=max;
%loopit( varlist =&flist, fn=testds );

```

BRIEF DESCRIPTION

Appendix A is an example of a financial statement report which includes ASSETS, LIABILITIES, INCOME, EXPENSES, SERVICE and GROSS PROFIT. Due to the complexity of the report layout, it is difficult to isolate the different financial components. However, after carefully examining the report, one may notice that each type of financial summary (category) has certain patterns. For example, the liability information always starts at column 66 and ends at column 117. Also, other than one exceptional case (monthly summary), it is clear that all liabilities follow a pattern. Therefore, I used the column input statement to read the data lines.

For illustration purposes, the SAS program reads the liability information from the financial statements and outputs a SAS data set named 'liability' (provided in Appendix B). The actual program used for the project is much more complex in nature because the program reads all information and parses it into separate analytic data sets for the different financial categories.

THE SAS PROGRAM

The complete program consists of the following three sections:

Section I: Defines libraries containing SAS data sets ; assigns folder locations and defines a macro variable to contain a list of file names to be converted.

Section II: SAS Macros to be used repeatedly during the process to read liability information from each files, combine all data sets to form a single data set.

Section III: Converts combined data set and outputs it as an Excel file for further analysis.

The complete program is shown below with a detailed description of every section and each step by using macro comments '%*'. The program has proven to be an excellent device for reading multiple financial statements. The program outputs an analytical file as soon as we receive financial statements from the client. It eliminated the tedious and repetitive work of manual conversion.

```

/*****
/* Program: read_liability.sas
/* Purpose: Read Financial Statement and output liability information to a SAS data sets.*/
/* Programmer: Tuqluke Abdurazak
/* PricewaterHouseCoopers LLP.
/* Date Created: 09/13/2003
/*****

/* Section I */

options ls=150 ps=60 pageno=1 mprint;

*---Folder to store converted SAS data sets;
libname out "c:\sugi29\sasdata";

*---Folder to store rawdata and converted data in Excel;
%let fpath = c:\sugi29\financial_statements ;

title "Reading Financial Statements";

footnotel "Sample Program For SUGI 29";

```

```

/* Section II */

*---File names to be read to SAS;
%let flist = FS1 FS2 F3 FS4 FS5 FS6 FS7 FS8 FS9 FS10 FS11 FS12 FS13 FS14 FS15;

*---Macro read_ld read liability data files;
%macro read_ld( flist = );

%*---Define local macro variables;
%local i ;

%let i = 1;
%*---Scan the list of filenames (flist) and sequentially process each files;
%do %until (%scan(&flist,&i,%str( )) = %str());
  %let var&i = %scan(&flist,&i,%str( ));

%*---Parse Liability Data;
data liability_&var&i ;
  length filename $10. id $8. month $25. year $8. dealer $5. sales_area $5. region $5.
    district $5. type $15. liability_type $50. ;
  retain month year dealer sales_area region district type ;
  infile "&fpath\&&var&i...txt" truncover;
  input @40 dum1 $45. @2 dum2 $50. @65 c65 $1. @68 c68 $1. @66 c66 $2. @;
  if trim(left(dum2)) in ( '*****' ) then delete;
  else if trim(left(dum1)) = 'FINANCIAL STATEMENT AUDIT REPORT' then
    do;
      input @91 month $25. @109 year $6. @124 dealer $6. /
        @127 sales_area $4. /
        @125 region $10. /
        @125 district $10. ////
        @2 type $15. ;
    end;
  if (c65 = '*' and c68 = '*') and c66 ne '**' then
    do;
      input @69 c69 $3. @112 c112 $1. @113 c113 $1. @118 c118 $1. @;
      if c69 in ('JAN','FEB','MAR','APR','MAY','JUN','JUL','AUG','SEP','OCT','NOV',
        'DEC','TOT') then
        input @66 id $2. @69 mo $3. +1 new_nis_t $ new_oth_t $ use_ret_t $
          use_whis_t $ profit_loss_t $ @119 liability_amt_t $119-133;
        else if ( c112 = '*' and c118 = '*' ) or ( c113 = '*' and c118 = '*' ) then
          input @66 id $2. @69 liability_type $69-112 @113 liability_acct $5.
            @119 liability_amt_t $119-133;
        else
          do;
            input @69 c69_90 $69-90 @;
            if trim(left(translate(c69_90,' ','-'))) in ('NET CASH POSITION',
              'NET WORKING CAPITAL' 'EFFECTIVE NET WORTH') then
              input @66 id $2. @69 liability_type $69-90 @91 liability_amt_t $91-108;
              else input @66 id $2. @69 liability_type $69-117;
            end;
          end;
    end;

  filename = "&&var&i" ;

  if trim(left(id)) in ('','**','LN') then delete;

  array cn(*) new_nis_t new_oth_t use_ret_t use_whis_t profit_loss_t liability_amt_t;
  array nn(*) new_nis new_oth use_ret use_whis profit_loss liability_amt;
  do i=1 to dim(cn);
    nn(i) = input( translate(cn(i),' ','*.'),comma12. );
  end;
  liability_acct = translate(liability_acct,' ','*.');

  if trim(left(liability_acct)) = 'OSS' then delete;

  drop type dum1 dum2 c65 c66 c68 c69 c112 c113 c118 c69_90 i
    new_nis_t new_oth_t use_ret_t use_whis_t profit_loss_t liability_amt_t;
run;

%*---Combine all data sets using PROC APPEND;
%if &i = 1 %then
%do;
  data liability ;
    set liability_&&var&i;

```

```

        run;
    %end;
    %else
    %do;
        proc append base=liability data=liability_&&var&i;
            run;
        %end;

    %let i = %eval(&i+1);

%end;

%let varcnt = %eval(&i-1);
%put The number of data sets being read are &varcnt ;

%mend ;

%*---Execute Macro;
%read_ng( flist = &flist );

/* Section III */

%*---Output to Excel using PROC DBLOAD;
proc dbload dbms=excel data=Liability;
    path="&fpath\liability.xls";
    putnames yes;
    limit=0 ;
    load;
run;

/*****The End of Program*****/

```

SUMMARY

Combining SAS data step and macro language to read complex report files has proven to be a very effective solution, especially when a large number of complex report files need to be analyzed. When a report file appears to be very complex and information appears to be scattered everywhere, a programmer should carefully look through the report, identify the pattern, and use powerful SAS data step programming techniques to extract information, rather than relying on application software that may not solve the current business issue.

REFERENCES

1. Professional SAS Programming Secrets by Rick Aster & Rhena Seidman. ISBN 0-07-913095-X (pbk.)
2. SAS Guide to macro Processing, Version 6, Second Edition (SAS Institute Inc.). ISBN 1-55544-382-6
3. SAS macro Facility Tip & Techniques, Version 6, First Edition (SAS Institute Inc.). ISBN 1-55544-605-1

ACKNOWLEDGEMENT

The author would like to express his appreciation to Dr. Jessica Pollner for her assistance in preparing this paper.

CONTACT INFORMATION

Tugluke Abdurazak
Economics & Statistics Practice
PricewaterhouseCoopers LLP
1900 K Street, NW
Washington D.C. 20006
Tel: 202-822-5592, Fax: 202-822-5800
tugluke.abdurazak@us.pwcglobal.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.

Appendix A: Sample of Financial Statements Need to be Analyzed

(Note: The sample data in the following tables has been completely altered. No real meaning should be derived from any value given in these tables.)

ASSETS		*ACCT*	AMOUNT	*LN*	LIABILITIES	*ACCT*	AMOUNT	
DATE: 05/11/1888 TIME: 04:31.44 DATE PROCESSED NDP: 11/11/88 HERITAGE HHHH/BBBBBBBB # 9999 REPORT NO. XXXX00-R1 PAGE 1 A FINANCIAL STATEMENT AUDIT REPORT - MONTH OF DECEMBER 1988 DEALER # 3333 SALES AREA: E REGION 99 DISTRICT 88								

CURRENT ASSETS				CURRENT LIABILITIES				
CASH				1*	ACCOUNTS PAYABLE			
ON HAND		*201*	888*	2*	TRADE CREDITORS	*300*	33,333	
IN BANK		*202*	8,888,888*	3*	PRODUCT LIEN PAYOFF	*303*		
CONTRACTS IN TRANSIT		*203*	888,888*	4*	CUSTOMER DEPOSITS	*304*	88,888	
* TOTAL CASH AND CONTRACTS		*	8,888,888*	5*	LICENSE FEES	*303*	88,888	
RECEIVABLES				6*	WARRANTY CLAIMS ADVANCE	*307*		
PRODUCT ACCOUNTS		*210*	88,888*	7*	* TOTAL ACCOUNTS PAYABLE	*	888,888	
SERVICE, PARTS & BDY. SHP. ACCTS		*220*	88,888*	8*	NOTES PAYABLE			
* TOTAL CUSTOMER RECEIVABLES		*	888,888*	9*	NEW PRODUCTS & DEMONSTRATORS	*310*	8,888,888	
LESS ALLOWANCE FOR DOUBTFUL ACCTS		*330*		10*	USED PRODUCTS	*312*		
* NET CUSTOMER RECEIVABLES		*	999,999*	11*	DRIVER EDUCATION CARS	*314*		
FACTORY RECEIVABLES		*221*	99,999*	12*	LEASE PRODUCTS	*317*		
WARRANTY & SERV. CONT. RECEIVABLES-NISSAN		*222*	9,999*	13*	RENTAL PRODUCTS	*317*		
WARRANTY & SERV. CONT. RECEIVABLES-OTHER		*223*	99,999*	14*	CURRENT AMT - LONG TERM DEBT	*314*		
* TOTAL RECEIVABLES		*	999,999*	15*	OTHER	*311*	9,999	
INVENTORIES * SHOW NUMBER OF UNITS				16*	* TOTAL NOTES PAYABLE	*	9,999,999	
DEMOS. NISSAN		*230*		17*	ACCRUED LIABILITIES			
NEW CAR NISSAN 200SX/ / SENTRA/ 24/*231*			999,999*	18*	INTEREST	*320*	99,999	
ALT/ 10/MX/ 7/QST/ 2/ PTH/ 7/*				19*	PAYROLL	*321*	99,999	
ZX/ / SX/ /XTR/ 14/ / /*				20*	INSURANCE	*322*	-999	
NEW TRUCK NISSAN 2WD/ 1/ 4WD/ 3/*232*			999,999*	21*	TAXES - OTHER THAN PAYROLL, SALES, INCOME	*323*	9,999	
2WDCC/ 2/ 4WDCC/ 2/*				22*	TAXES - SALES	*324*	9,999	
DEMOS OTHER		*233*	99,999*	23*	TAXES - PAYROLL	*323*		
NEW PRODUCT OTHER		*234*	9,999,999*	24*	INCOME TAXES - PREVIOUS YEAR	*327*		
* TOTAL NEW CAR & TRUCK			9,999,999*	25*	INCOME TAXES - CURRENT YEAR	*327*		
NEW VEH. OTHER THAN CAR & TRUCK		*237*		26*	BONUSES - EMPLOYEES	*324*		
USED VEH. NISSAN/ / OTHER/ /*				27*	BONUSES - OWNERS	*321*		
USED VEH. TOTAL/ /OVER 30/ /*237*				28*	PENSION FUND	*330*		
USED VEH. OTHER THAN CAR & TRUCK		*231*		29*	RESERVE FOR REPOSSESSION LOSSES	*331*		
DRIVER EDUCATION CARS		*240*		30*	OTHER	*332*		
LESS LIFO RESERVE - NEW VEH.		*241*	99,999*	31*	* TOTAL ACCRUED LIABILITIES	*	99,999	
LESS LIFO RESERVE - USED VEH.		*241U*		32*	* TOTAL CURRENT LIABILITIES	*	9,999,999	
PARTS & ACCESSORIES		*242*	999,999*	33*	LONG TERM DEBT *****			
LESS LIFO RESERVE - P & A		*243*		34*	MORTGAGES PAYABLE	*340*		
GAS, OIL & GREASE		*244*	999*	35*	OFFICERS/OWNERS	*341*		
BODY SHOP MATERIALS		*243*		37*	COMPANY CARS & SERVICE PRODUCTS	*342*		
SUBLET REPAIRS		*247*	9,999*	38*	OTHER	*343*		
WORK IN PROCESS-LABOR		*247*	9,999*	39*	* TOTAL LONG TERM DEBT	*		
NO - AUTOMOTIVE MERCHANDISE		*244*		40*	TOTAL LIABILITIES	*	9,999,999	

PAGE 1 B				DEALER # 3712 FOR 12-00 KBMD300-R1 PAGE 1 B				

ASSETS		*ACCT*	AMOUNT	*LN*	LIABILITIES	*ACCT*	AMOUNT	

* TOTAL OTHER INVENTORIES		*	8,888*	40*	NET CASH POSITION -		1,037,124	
* TOTAL ALL INVENTORIES		*	8,888,888*	41*	CALC: PG 1, ASSETS/L3, 23, 27/ - LIAB/L3, 4, 3, & 1/*			
OTHER CURRENT ASSETS				42*	- DFF IF L2 + L7 EXCEEDS CURR MD EXP PG. 2 L33*			
SECURITIES		*270*		43*	NET WORKING CAPITAL -		1,374,027	
FINANCE INCOME & INS. COMM REC.		*271*	88,888*	44*	CALC: ASSET/L32/ + LIFO ACCTS/L31, L32, L34/			
PREPAID EXPENSES		*272*	88,888*	43*	- LIAB/32/			
* TOTAL OTHER CURRENT ASSETS		*	88,888*	47*	EFFECTIVE NET WORTH -		1,442,732	
LEASE AND RENTAL PRODUCTS				47*	CALC: N.W./L73/ + 70% LIFO RESERVES/L31, L32, L34/*			
LEASE PRODUCTS		*277*		44*				
RENTAL PRODUCTS		*274*		41*	NET WORTH			
LESS ACCUM DEPR. - LEASE & RENTAL		*377*		30*	CORPORATION			
* TOTAL LEASE & RENTAL PRODUCTS		/		31*	CAPITAL STOCK	*310*	9,999	
* TOTAL CURRENT ASSETS		*	8,888,888*	32*	PAID IN CAPITAL	*313*	9,999,999	
FIXED ASSETS - AUTO BUSINESS ONLY				33*	LESS TREASURY STOCK	*317*		

* ACCT * COST *LESS ACC DPR*				34*	RETAINED EARNINGS	*311*	-99,999	

LAND	*240*	*		33*	LESS DIVIDENDS	*312*		
BULDS/IMPRVM	*241*341*	*		37*	PROPRIETORSHIP OR PARTNERSHIP	*		
MACH & SHIP	*242*342*	99,999*	99,999*	37*	INVESTMENT	*313*		
PARTS & ACC.	*243*343*	99,999*	99,999*	34*	LESS DRAWING	*314*		

FURN & FLX.	*244*344*	99,999*	99,999*	31*	NEW NIS*NEW OTH*USE RET*USE WILS*PROFIT/LOSS	*		

SVC UNITS	*243*343*	*		70*	JAN*		-9,999*	
LEASE HLDS	*247*347*	9,999*	999*	71*	FEB*		9,999*	
SIGNS	*247*347*	*		72*	MAR*		99,999*	
* TOTAL FIXED ASSETS	* 999,999*	99,999*	888,888*	73*	APR*		99,999*	

OTHER ASSETS				74*	MAY*		99,999*	
FIN. -INS. -SERV. CONT. INC. REC. - DEFERRED	*241*			73*	JUN*		99,999*	
DEPOSITS ON CONTRACTS	*210*			77*	JUL*	43*	11*	9,999*
LIFE INSURANCE - CASH VALUE	*211*			77*	AUG*	23*	31*	-99,999*
NOTES & ACCT REC - OFFICERS/OWNERS	*213*			74*	SEP*	24*	37*	-99,999*

• ADVANCES TO EMPLOYEES	*214 *	888*71*OCT*	13*	41*	*	*	-99,999*
• OTHER NOTES & ACCT REC	*213 *	*70*NOV*	14*	32*	*	*	-99,999*
• OTHER NON-AUTOMOTIVE ASSETS	*217 *	*71*DEC*	17*	22*	*	*	-99,999*
• * TOTAL OTHER ASSETS	* *	888*72*TOT*	133*	270*	*	*	*311 * -99,999
• * TOTAL NON-CURRENT ASSETS	* *	888,888*73*	* TOTAL NET WORTH		*	*	9,999,999
* TOTAL ASSETS	* *	8,888,888*74*	TOTAL LIABILITIES & NET WORTH		*	*	9,999,999

Appendix B: Converted Financial Statements

(Note: The sample data in the following tables has been completely altered. No real meaning should be derived from any value given in these tables.)

Table b.1: Example of Financial Statement by ID

filename	id	month	year	dealer	Sales_area	region	district	liability_type	libility_acct	libility_amt
FS	1	DECEMBER	1988	DL	EAST	88	99	TRADE CREDITORS	1	99999.00
FS	2	DECEMBER	1988	DL	EAST	88	99	PAYOFF	2	99999.00
FS	3	DECEMBER	1988	DL	EAST	88	99	CUSTOMER DEPOSITS	3	99999.00
FS	4	DECEMBER	1988	DL	EAST	88	99	FEES	4	99999.00
FS	5	DECEMBER	1988	DL	EAST	88	99	WARRANTY CLAIMS ADVANCE	5	99999.00
FS	6	DECEMBER	1988	DL	EAST	88	99	INCOME TAXES - PREVIOUS YEAR	6	999999.00
FS	7	DECEMBER	1988	DL	EAST	88	99	INCOME TAXES - CURRENT YEAR	7	99999.00
FS	8	DECEMBER	1988	DL	EAST	88	99	EDUCATION	8	
FS	9	DECEMBER	1988	DL	EAST	88	99	LEASE	9	
FS	10	DECEMBER	1988	DL	EAST	88	99	RENTAL	10	
FS	11	DECEMBER	1988	DL	EAST	88	99	CURRENT AMT - LONG TERM DEBT	11	99999.00
FS	12	DECEMBER	1988	DL	EAST	88	99	OTHER	12	
FS	13	DECEMBER	1988	DL	EAST	88	99	INTEREST	13	99999.00
FS	14	DECEMBER	1988	DL	EAST	88	99	PAYROLL	14	99999.00
FS	15	DECEMBER	1988	DL	EAST	88	99	INSURANCE	15	999.00
FS	16	DECEMBER	1988	DL	EAST	88	99	TAXES - OTHER THAN PAYROLL,SALES,INCOME	16	99999.00
FS	17	DECEMBER	1988	DL	EAST	88	99	TAXES - SALES	17	999999.00
FS	18	DECEMBER	1988	DL	EAST	88	99	TAXES - PAYROLL	18	
FS	19	DECEMBER	1988	DL	EAST	88	99	BONUSES - EMPLOYEES	19	
FS	20	DECEMBER	1988	DL	EAST	88	99	BONUSES	20	

Table b.2: Example of Financial Statement by Month

filename	id	year	dealer	sales_area	region	district	mo	new_nis	new_oth	use_ret	use_whis	profit_loss
FS	1	1988	DL	EAST	88	99	JAN					99999.00
FS	2	1988	DL	EAST	88	99	FEB	99.00	88.00	99.00	99.00	99999.00
FS	3	1988	DL	EAST	88	99	MAR	99.00	88.00	99.00	99.00	99999.00
FS	4	1988	DL	EAST	88	99	APR	99.00	88.00	99.00	99.00	99999.00
FS	5	1988	DL	EAST	88	99	MAY	99.00	88.00	99.00	99.00	99999.00
FS	6	1988	DL	EAST	88	99	JUN	99.00	88.00	99.00	99.00	99999.00
FS	7	1988	DL	EAST	88	99	JUL	99.00	88.00	99.00	99.00	99999.00
FS	8	1988	DL	EAST	88	99	AUG	99.00	88.00	99.00	99.00	99999.00
FS	9	1988	DL	EAST	88	99	SEP	99.00	88.00	99.00	99.00	99999.00
FS	10	1988	DL	EAST	88	99	OCT	99.00	88.00	99.00	99.00	99999.00
FS	11	1988	DL	EAST	88	99	NOV	99.00	88.00	99.00	99.00	99999.00
FS	12	1988	DL	EAST	88	99	DEC	99.00	88.00	99.00	99.00	99999.00
FS		1988	DL	EAST	88	99	TOT	999.00	888.00	999.00	999.00	999999.00