

Automated double data entry comparisons (AD-047)

Melinda MacDougall is a Senior Data Programmer at Cincinnati Children's Hospital in the Division of General and Community Pediatrics. She has been using SAS[®] software for 15+ years as both a biostatistician and data programmer. She has worked on both NIH-funded clinical research studies as well as contracts for the Centers for Medicare & Medicaid Services.



Automated double data entry comparisons (AD-047)

Melinda MacDougall, MS
Cincinnati Children's Hospital Medical Center
Senior Data Programmer



Double Data Entry – Example Data

	ID	Char	Num
1	1	Text text Text	13
2	2	Text w/ blank	18.12
3	3	n/a	12.1
4	4	Match	10
5	5	Only in 1	8.9

	ID	Char	Num
1	1	TEXT TEXT TEXT	13.01
2	2	Text w/ blank	18.1
3	3	N/A	12.1
4	4	Match	10
5	6	2 only	14

	ID	Char	Num	Char_v2	Num_v2
1	1	Text text Text	13	TEXT TEXT TEXT	13.01
2	2	Text w/ blank	18.12	Text w/ blank	18.1
3	3	n/a	12.1	N/A	12.1
4	4	Match	10	Match	10
5	5	Only in 1	8.9		.
6	6		.	2 only	14



Pull variable list

```
proc contents data=test noprint out=vars; run;
```

```
proc sql number;
```

```
select name, type  
into :varlist separated by ' ',  
:typelist separated by ' ';
```

Pull variable and type
(numeric vs character)

```
from vars
```

```
where lowercase(name) not in('ID')
```

Exclude ID field(s)

```
AND scan(lowercase(name), -1, '_') ^= 'v2'
```

```
order by varnum
```

Only pull base variable names

```
;
```

```
quit;
```

Compare data

```
%macro flag;  
  %do i=1 %to %num_tokens(&varlist);  
    %let var1 = %scan(&varlist,&i);  
    %let var2 = %scan(&varlist,&i)_v2;  
    %let type = %scan(&typelist,&i);  
  
    %if &type=1 %then %do; *numeric;  
      Diff_&var1 = (&var1 ^= &var2);  
    %end;  
    %else %if &type=2 %then %do; *character;  
      var1 = upcase(compress(&var1));  
      var2 = upcase(compress(&var2));  
      Diff_&var1 = (var1 ^= var2);  
    %end;  
  %end;  
%mend flag;
```



Transpose data

```
%macro wide2long;
  %do i=1 %to %num_tokens(&varlist);
    %let var1 = %scan(&varlist,&i); %let var2 = %scan(&varlist,&i)_v2;
    %let type = %scan(&typelist,&i);
    if Diff_&var1=1 then do;
      Field = "&var1";
      %if &type=1 %then %do; *numeric;
        if not missing(&var1) then Entry_1 = left(put(&var1,8.));
        else call missing(Entry_1);
        if not missing(&var2) then Entry_2 = left(put(&var2,8.));
        else call missing(Entry_2);
      %end;
      %else %if &type=2 %then %do; *character;
        Entry_1 = &var1; Entry_2 = &var2;
      %end;
    end;
  %end;
%mend wide2long;
```



Output Data

```
ods excel file="&filepath.\Differences.xlsx"  
options(Sheet_Name="Diff" Frozen_Headers='1'  
        flow='tables' Absolute_Column_Width='8,8,12,12');  
proc print data=diff noobs;  
    var ID Field Entry_1 Entry_2  
    / style(data)=[just=left VERTICALALIGN=top]  
      style(header)=[just=left];  
run;  
ods excel close;
```

	A	B	C	D
1	ID	Field	Entry_1	Entry_2
2	1	Num	13.00	13.01
3	2	Num	18.12	18.10
4	3	Char	n/a	Not applicable
5	5	Char	Only in 1	
6	5	Num	8.90	
7	6	Char		2 only
8	6	Num		14.00



References

- Text Utility Macros for Manipulating Lists of Variable Names [Robert J. Morris]
 - <https://support.sas.com/resources/papers/proceedings/proceedings/sugi30/029-30.pdf>



Thank You!

- Melinda MacDougall, MS
- Cincinnati Children's Medical Center
- Senior Data Programmer
- Melinda.MacDougall@cchmc.org



Trademark Info

The output for this paper was generated using SAS software. Copyright © 2016 SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA.

