SAS Learning Module – Proc Transpose
(http://www.ats.ucla.edu/stat/sas/modules/ltow_transpose.htm)

How to reshape data long to wide using proc transpose

1. Transposing one variable

Sometimes you need to reshape your data which is in a long format (shown below)

```
famid year faminc
1    96   40000
1    97   40500
1    98   41000
2    96   45000
2    97   45400
2    98   45800
3    96   75000
3    97   76000
3    98   77000
```

into a wide format (shown below).

```
famid faminc96 faminc97 faminc98
1     40000    40500    41000
2     45000    45400    45800
3     75000    76000    77000
```

Below is an example of using SAS **proc transpose** to reshape the data from a long to a wide format.

```
data long1 ;
input famid year faminc ;
cards ;
1 96 40000
1 97 40500
1 98 41000
2 96 45000
2 97 45400
2 98 45800
3 96 75000
3 97 76000
3 98 77000  ;
run;

proc transpose data=long1 out=wide1 prefix=faminc;
by famid ;
id year;
var faminc;
run;
proc print data = wide1;
run;
```

Obs	famid	_NAME_	faminc96	faminc97	faminc98
1	1	faminc	40000	40500	41000
2	2	faminc	45000	45400	45800
3	3	faminc	75000	76000	77000

Notice that the option **prefix= faminc** specifies a prefix to use in constructing names for transposed variables in the output data set. SAS automatic variable **_NAME_** contains the name of the variable being transposed.

2. Transposing two variables

With only a few modifications, the above example can be used to reshape two (or more) variables.
The approach here is to use **proc transpose** multiple times as needed.
The multiple transposed data files then are merged back.

```
data long2;
input famid year faminc spend;
cards;
1 96 40000 38000
1 97 40500 39000
1 98 41000 40000
2 96 45000 42000
2 97 45400 43000
2 98 45800 44000
3 96 75000 70000
3 97 76000 71000
3 98 77000 72000;
run ;

proc transpose data=long2 out=widef prefix=faminc;
by famid;
id year;
var faminc;
run;

proc transpose data=long2 out=wides prefix=spend;
by famid;
id year;
var spend;
run;

data wide2;
merge  widef(drop=_name_) wides(drop=_name_);
by famid;
run;

proc print data=wide2;
run;
```

Obs	famid	faminc96	faminc97	faminc98	spend96	spend97	spend98
1	1	40000	40500	41000	38000	39000	40000
2	2	45000	45400	45800	42000	43000	44000
3	3	75000	76000	77000	70000	71000	72000

3. Reshaping data with two variables that identify the wide record

Sometimes, there is no variable in the data set that uniquely identifies each observation.
Rather, two or more variables are necessary to uniquely identify each observation.
In this situation, we have to specify these variables in the **by** statement.

```
data long3;
INPUT famid birth age ht;
cards;
1 1 1 2.8
1 1 2 3.4
1 2 1 2.9
1 2 2 3.8
1 3 1 2.2
1 3 2 2.9
2 1 1 2.0
2 1 2 3.2
2 2 1 1.8
2 2 2 2.8
2 3 1 1.9
2 3 2 2.4
3 1 1 2.2
3 1 2 3.3
3 2 1 2.3
3 2 2 3.4
3 3 1 2.1
3 3 2 2.9
;
run;

proc transpose data=long3 out=wide3 prefix=ht;
by famid birth;
id age;
var ht;
run;

proc print data=wide3;
run;
```

Obs	famid	birth	_NAME_	ht1	ht2
1	1	1	ht	2.8	3.4
2	1	2	ht	2.9	3.8
3	1	3	ht	2.2	2.9
4	2	1	ht	2.0	3.2
5	2	2	ht	1.8	2.8
6	2	3	ht	1.9	2.4
7	3	1	ht	2.2	3.3
8	3	2	ht	2.3	3.4
9	3	3	ht	2.1	2.9

4. A more realistic example

The following example is a more realistic example that uses a data file having 300 records in long format (50 wide records and six time points).

```
data long4;
input id year inc;
cards;
1 90 66483
1 91 69146
1 92 74643
1 93 79783
1 94 81710
1 95 86143
2 90 17510
2 91 17947
2 92 19484
2 93 20979
2 94 21268
2 95 22998
3 90 57947
3 91 62964
3 92 68717
3 93 70957
3 94 75198
3 95 75722
4 90 64831
4 91 71060
4 92 71918
4 93 72514
4 94 73100
4 95 74379
5 90 18904
5 91 19949
5 92 21335
5 93 22237
5 94 23829
5 95 23913
6 90 32057
6 91 34770
6 92 35834
6 93 37387
6 94 40899
6 95 42372
```

7 90 60551
7 91 64869
7 92 67983
7 93 70498
7 94 71253
7 95 75177
8 90 16553
8 91 18189
8 92 18349
8 93 19815
8 94 21739
8 95 22980
9 90 32611
9 91 33465
9 92 35961
9 93 36416
9 94 37183
9 95 40627
10 90 61379
10 91 66002
10 92 67936
10 93 70513
10 94 74405
10 95 76009
11 90 24065
11 91 24229
11 92 25709
11 93 26121
11 94 26617
11 95 28142
12 90 32975
12 91 36185
12 92 37601
12 93 41336
12 94 43399
12 95 43670
13 90 69548
13 91 71341
13 92 72455
13 93 76552
13 94 80538
13 95 85330
14 90 50274
14 91 53349

14 92 55900
14 93 59375
14 94 61216
14 95 63911
15 90 72011
15 91 73334
15 92 76248
15 93 77724
15 94 78638
15 95 80582
16 90 18911
16 91 20046
16 92 21343
16 93 21630
16 94 22330
16 95 23081
17 90 68841
17 91 75410
17 92 80806
17 93 81327
17 94 81571
17 95 86499
18 90 28099
18 91 30716
18 92 32986
18 93 36097
18 94 39124
18 95 39866
19 90 17302
19 91 18778
19 92 18872
19 93 19884
19 94 20665
19 95 21855
20 90 16291
20 91 16674
20 92 16770
20 93 17182
20 94 17979
20 95 18917
21 90 43244
21 91 46545
21 92 47633
21 93 50744

21 94 54734
21 95 59075
22 90 56393
22 91 59120
22 92 60801
22 93 61404
22 94 63111
22 95 69278
23 90 47347
23 91 49571
23 92 50101
23 93 51345
23 94 56463
23 95 56927
24 90 16076
24 91 17217
24 92 17296
24 93 17900
24 94 18171
24 95 18366
25 90 65906
25 91 69679
25 92 76131
25 93 77676
25 94 81980
25 95 85426
26 90 58586
26 91 61188
26 92 66542
26 93 69267
26 94 71063
26 95 74549
27 90 61674
27 91 66584
27 92 69185
27 93 75193
27 94 78647
27 95 81898
28 90 31673
28 91 31883
28 92 32774
28 93 34485
28 94 36929
28 95 39751

29 90 63412
29 91 67593
29 92 69911
29 93 73092
29 94 80105
29 95 81840
30 90 27684
30 91 28439
30 92 30861
30 93 31406
30 94 32960
30 95 35530
31 90 71873
31 91 76449
31 92 80848
31 93 88691
31 94 94149
31 95 97431
32 90 62177
32 91 63812
32 92 64235
32 93 65703
32 94 69985
32 95 71136
33 90 37684
33 91 38258
33 92 39208
33 93 39489
33 94 39745
33 95 41236
34 90 64013
34 91 66398
34 92 71877
34 93 75610
34 94 76395
34 95 79644
35 90 16011
35 91 16847
35 92 17746
35 93 19123
35 94 19183
35 95 19996
36 90 49215
36 91 52195

36 92 52343
36 93 56365
36 94 58752
36 95 59354
37 90 15774
37 91 16643
37 92 17605
37 93 18781
37 94 18996
37 95 19685
38 90 29106
38 91 31693
38 92 31852
38 93 34505
38 94 35806
38 95 36179
39 90 25147
39 91 26923
39 92 28785
39 93 30987
39 94 34036
39 95 34106
40 90 71978
40 91 79144
40 92 80453
40 93 86580
40 94 95164
40 95 96155
41 90 46166
41 91 47579
41 92 49455
41 93 53849
41 94 56630
41 95 57473
42 90 55810
42 91 59443
42 92 65291
42 93 66065
42 94 69009
42 95 74365
43 90 49642
43 91 50603
43 92 53917
43 93 54858

43 94 58470
43 95 59767
44 90 21348
44 91 22361
44 92 23412
44 93 24038
44 94 24774
44 95 25828
45 90 44361
45 91 48720
45 92 51356
45 93 54927
45 94 56670
45 95 58800
46 90 56509
46 91 60517
46 92 61532
46 93 65077
46 94 69594
46 95 73089
47 90 39097
47 91 40293
47 92 43237
47 93 44809
47 94 48782
47 95 53091
48 90 18685
48 91 19405
48 92 20165
48 93 20316
48 94 22197
48 95 23557
49 90 73103
49 91 76243
49 92 76778
49 93 82734
49 94 86279
49 95 86784
50 90 48129
50 91 49267
50 92 53799
50 93 58768
50 94 63011
50 95 66461

```
;
run;
proc transpose data=long4 out=wide4 prefix=inc;
by id;
id year;
var inc;
run;

proc print data=wide4 (obs=10);
run;
```

Obs	id	_NAME_	inc90	inc91	inc92	inc93	inc94	inc95
1	1	inc	66483	69146	74643	79783	81710	86143
2	2	inc	17510	17947	19484	20979	21268	22998
3	3	inc	57947	62964	68717	70957	75198	75722
4	4	inc	64831	71060	71918	72514	73100	74379
5	5	inc	18904	19949	21335	22237	23829	23913
6	6	inc	32057	34770	35834	37387	40899	42372
7	7	inc	60551	64869	67983	70498	71253	75177
8	8	inc	16553	18189	18349	19815	21739	22980
9	9	inc	32611	33465	35961	36416	37183	40627
10	10	inc	61379	66002	67936	70513	74405	76009

5. Reshaping data with numeric and character variables

The following example shows how to reshape multiple variables, some of which are numeric and other that are character (i.e., string) variables. The approach here is the same as in Example 2 that **proc transpose** is used multiple times and the data files are then merged together.

```
data long5;
length debt $ 3;
input famid year faminc spend debt $ ;
cards;
1 96 40000 38000 yes
1 97 40500 39000 yes
1 98 41000 40000 no
2 96 45000 42000 yes
2 97 45400 43000 no
2 98 45800 44000 no
3 96 75000 70000 no
3 97 76000 71000 no
3 98 77000 72000 no;
run;

proc transpose data=long5 out=widef prefix=faminc;
by famid;
id year;
var faminc;
run;

proc transpose data=long5 out=wides prefix=spend;
by famid;
```

```
id year;
var spend;
run;

proc transpose data=long5 out=wided prefix=debt;
by famid;
id year;
var debt;
run;

data wide5;
merge widef (drop=_name_) wides (drop =_name_) wided (drop=_name_);
by famid;
run;

proc print data=wide5;
run;
```

Ob	famid	faminc96	faminc97	faminc98	spend96	spend97	spend98	debt96	debt97	debt98
1	1	40000	40500	41000	38000	39000	40000	yes	yes	no
2	2	45000	45400	45800	42000	43000	44000	yes	no	no
3	3	75000	76000	77000	70000	71000	72000	no	no	no