Chapter 17
Restructure Data Wizard for Data Classified the Wrong Way (20 Patients)
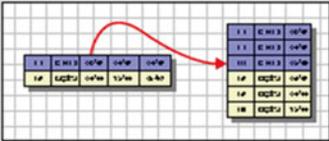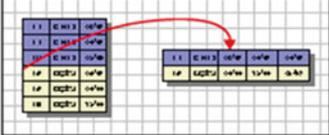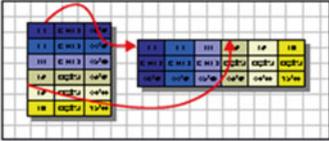
General Purpose

Underneath the opening page of the Restructure Data Wizard in SPSS is given. In the current chapter this tool will be applied for restructuring multiple variables in a single case to multiple cases with a single variables.

Suppose in a study the treatment outcome has been measured several times instead of once. In current clinical research repeated measures in a single subject are common. The problem with repeated measures is, that they are more close to one another than unrepeated measures. If this is not taken into account, then data analysis will lose power. The underneath table gives an example of a 2 group parallel-group study comparing two treatments for cholesterol reduction of 5 weeks. The example is taken from Chap. 6, Mixed linear models, pp 65–77, in: Machine learning in medicine part one, Springer Heidelberg Germany, 2013, from the same authors.

It shows that 5 different variables present the 5 subsequent outcome measurements in each patient. In order to analyze these data in appropriately the table has to be restructured with each week given a separate row. This is a pretty laborious exercise, and it will get really annoying if you have 100 or more patients instead of 20. The restructure data wizard, however, should do the job within seconds.

patient no	week 1	week 2	week 3	week 4	week 5	treatment modality
1	1,66	1,62	1,57	1,52	1,50	0,00
2	1,69	1,71	1,60	1,55	1,56	0,00
3	1,92	1,94	1,83	1,78	1,79	0,00
4	1,95	1,97	1,86	1,81	1,82	0,00
5	1,98	2,00	1,89	1,84	1,85	0,00
6	2,01	2,03	1,92	1,87	1,88	0,00
7	2,04	2,06	1,95	1,90	1,91	0,00
8	2,07	2,09	1,98	1,93	1,94	0,00
9	2,30	2,32	2,21	2,16	2,17	0,00
10	2,36	2,35	2,26	2,23	2,20	0,00
11	1,57	1,82	1,83	1,83	1,82	1,00
12	1,60	1,85	1,89	1,89	1,85	1,00
13	1,83	2,08	2,12	2,12	2,08	1,00
14	1,86	2,11	2,16	2,15	2,11	1,00
15	2,80	2,14	2,19	2,18	2,14	1,00
16	1,92	2,17	2,22	2,21	2,17	1,00
17	1,95	2,20	2,25	2,24	2,20	1,00
18	1,98	2,23	2,28	2,27	2,24	1,00
19	2,21	2,46	2,57	2,51	2,48	1,00
20	2,34	2,51	2,55	2,55	2,52	1,00

week 1=hdl-cholesterol level after 1 week of trial
treatment modality=treatment modality (0=treatment 0, 1=treatment 1)

Example   103

Primary Scientific Question

Can the restructure data wizard provide a table suitable for testing treatment efficacies adjusted for the repeated nature of the outcome data.

Example

The above data file is entitled "restructure.sav", and is in extras.springer.com. Start by opening the data file in SPSS statistical software.

Command:

click Data....click Restructure....mark Restructure selected variables into cases.... click Next....mark One (for example, w1, w2, and w3)....click Next....Name: id (the patient id variable is already provided)....Target Variable: enter "firstweek, second-week...... fifthweek"....Fixed Variable(s): enter treatment....click Next.... How many index variables do you want to create?....mark One....click Next....click Next again....click Next again....click Finish....Sets from the original data will still be in use…click OK.

Return to the main screen and observe that there are now 100 rows instead of 20 in the data file. The first 10 rows are given underneath.

id	treatment	Index1	Trans1
1	0,00	1	1,66
1	0,00	2	1,62
1	0,00	3	1,57
1	0,00	4	1,52
1	0,00	5	1,50
2	0,00	1	1,69
2	0,00	2	1,71
2	0,00	3	1,60
2	0,00	4	1,55
2	0,00	5	1,56

id = patient id
treatment = treatment modality
Index1 = week of treatment (1–5)
Trans1 = outcome values

We will now perform a mixed linear analysis of the data.

Command:

Analyze….mixed models….linear….specify subjects and repeated….subject: enter id ….continue….linear mixed model….dependent: Trans1….factors: Index1, treatment….fixed….build nested term….treatment ….add….Index1….add…. Index1 build term by* treatment….Index1 *treatment….add….continue….OK (*=sign of multiplication).

The underneath table shows the main results from the above analysis. After adjustment for the repeated nature of the outcome data the treatment modality 0 performs much better than the treatment modality 1. The results from alternative analyses for these data were not only less appropriate but also less sensitive. The discussion of this is beyond the scope of the current chapter, but it can found in the Chap. 6, Mixed linear models, pp 65–77, in: Machine learning in medicine part one, Springer Heidelberg Germany, 2013, from the same authors.

Type III tests of fixed effects[a]				
Source	Numerator df	Denominator df	F	Sig.
Intercept	1	76,570	6988,626	,000
Week	4	31,149	,384	,818
Treatment	1	76,570	20,030	,000
Week*treatment	4	31,149	1,337	,278

[a]Dependent variable: outcome

Conclusion

The restructure data wizard provides a table suitable for testing treatment efficacies adjusted for the repeated nature of the outcome data. It is particularly pleasant if your data file is big, and has many (repeated) observations.

Note

More background, theoretical and mathematical information of restructuring data files is in the Chap. 6, Mixed linear models, pp 65–77, in: Machine learning in medicine part one, Springer Heidelberg Germany, 2013, from the same authors.