

第4章 文件级别的数据管理

- 数据管理:

- 对数据文件的结构进行重新调整或转换，以便适合于相应的统计方法。

- 数据管理包括的内容:

- 计算新变量、变量值的分组合并、变量的重新编码、变量的编秩，加权个案、文件合并与拆分、分类汇总、数据文件重组、等等。

- 数据管理有两个级别：

- 变量级别（Transform菜单即“转换”菜单）

- 计算变量：对变量进行四则运算生成新变量。
 - 变量转换：Recode（重新编码）、Visual Bander（可视离散化）、Count（计数）、Rank Cases（排秩个案）、Automatic Recode（自动重新编码）等。

- 文件级别（Data菜单即“数据”菜单）

- 简单命令：插入变量，插入个案，复制数据集，等。
 - 常用过程：个案排序、筛选和加权，拆分文件，等。
 - 数据重组：长宽型数据相互转换，数据转置。
 - 文件合并：将几个数据文件合并为一个数据文件。

学习目标

- 能利用 “Sort Cases” (排序个案) 进行排序
- 能利用 “Split File” (拆分文件) 进行记录拆分
- 能利用 “Select Cases” (选择个案) 进行记录筛选
- 能利用 “Weight Cases” (加权个案) 进行记录加权
- 能利用 “Aggregate” (分类汇总) 进行分类汇总
- 能利用 “Merge Files” (合并文件) 对两个数据文件进行合并
- 能利用 “Restructure” (重组) 对数据文件重新排列

主要内容

- 4.1 几个常用过程
- 4.2 多个数据文件的合并
- 4.3 数据文件的重组与转置

4.1 几个常用过程

排序个案 (Sort Cases)

- 应用场景

- 将记录按一个或多个变量值的升序或降序排列，有利于进行查找、修改等操作，也可以方便用户了解数据，进行数据分析。
- 缺失值在排序中会排在最小值的前面，因此可以通过排序的方法很快发现哪些记录为缺失。

- 进入实现过程

- 在数据视图界面通过鼠标右击变量名快速实现；
- “数据” → “排序个案”。

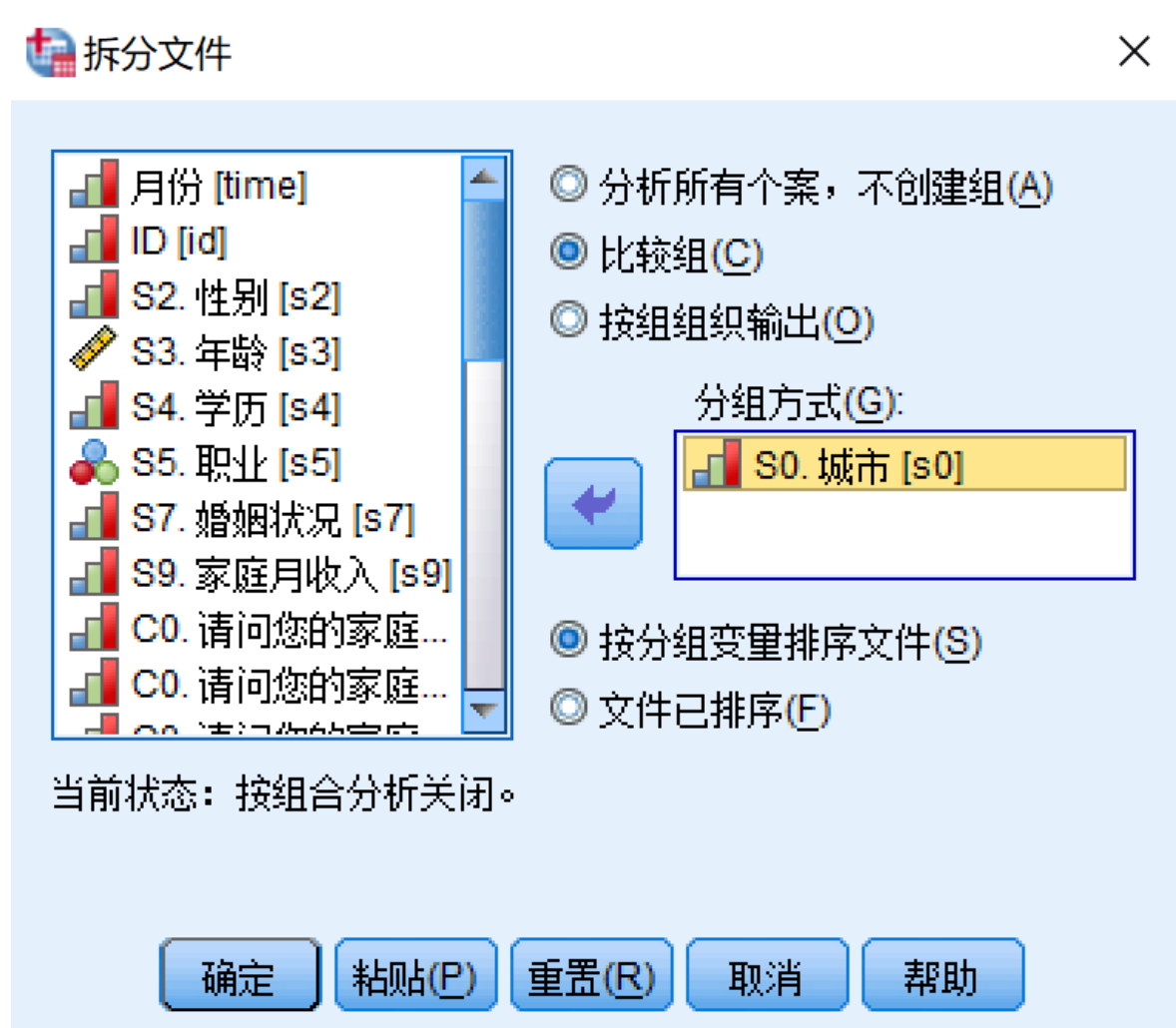
排序个案 (Sort Cases)



拆分文件 (Split File)

- **应用场景：**将文件中的记录按某个**分类变量**进行分组，从而对数据进行分组处理分析。
- **几种分割方式**
 - Analyze all cases: 不拆分文件。
 - Compare groups: 按所选分类变量分割文件，不同组的同一类的分析结果紧挨在一起，便于相互比较。
 - Organize output by groups: 按所选分类变量分割文件，同一组的所有分析结果放置在一起。
- **进入实现过程**
 - 快捷工具栏的“拆分文件”按钮。
 - “数据” → “拆分文件”。

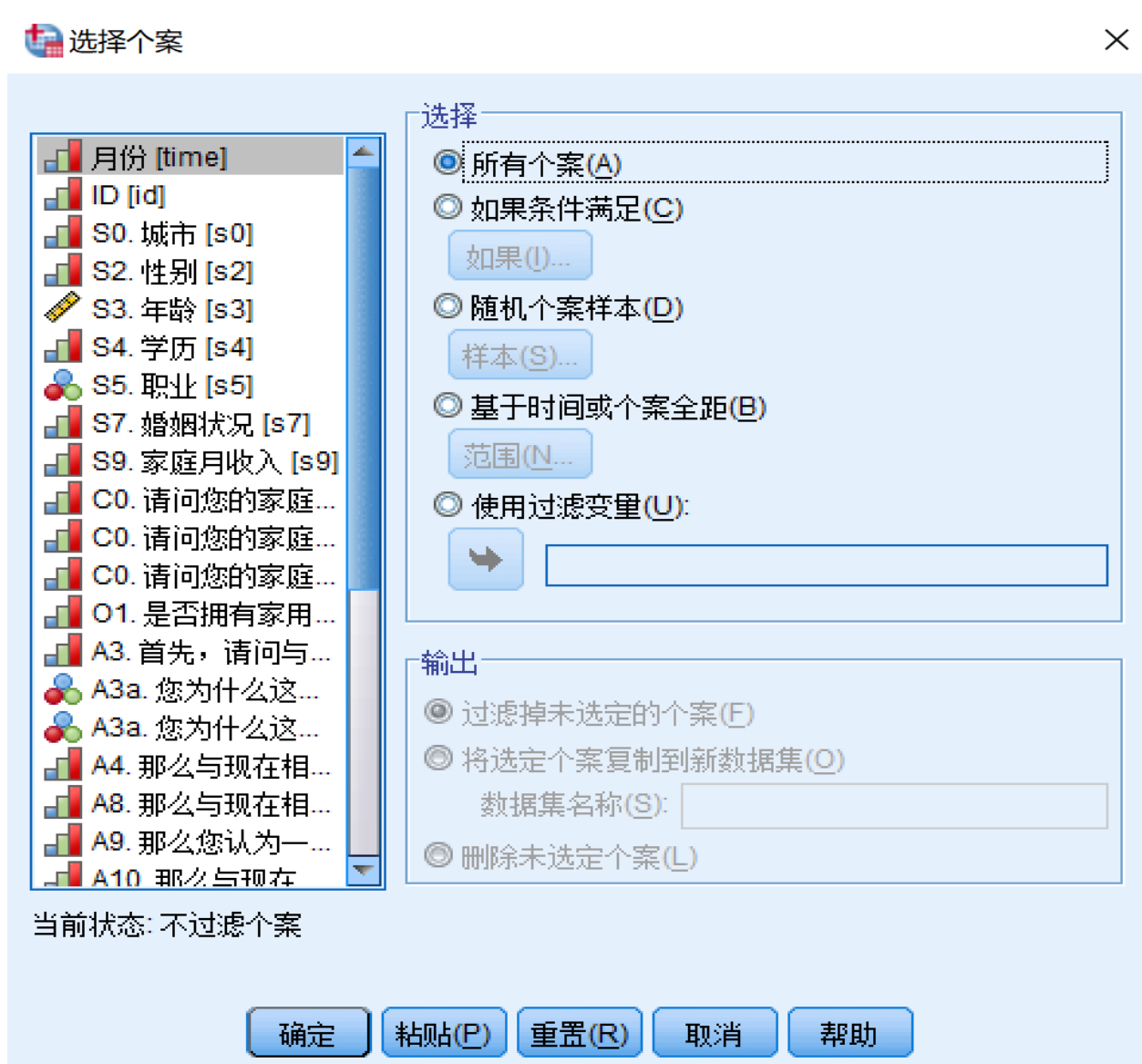
拆分文件 (Split File)



选择个案 (Select Cases)

- 应用场景：按要求筛选出数据集的一部分参与数据分析。
- 几种筛选方式
 - 条件筛选：可以使用组合条件进行筛选。
 - 随机筛选：可以直接在数据集中按百分比随机抽样，或从前 n 个记录中随机抽取 m 个记录。
 - 基于个案范围筛选：基于记录号来选择记录。
 - 通过变量筛选：可以通过某个变量来筛选记录，该变量取值为非0的记录将被选中。
- 进入实现过程
 - 快捷工具栏的“选择个案”按钮。
 - “数据” → “选择个案”。

选择个案 (Select Cases)



加权个案 (Weigh Cases)

- 数据录入格式

- 枚举格式

	学号	姓名	成绩
1	141070001	阿依古丽木	95.00
2	141070002	白雪	96.00
3	141070003	白央	96.00
4	141070004	陈辰	96.00
5	141070005	陈驰	97.00

- 频数格式

	性别	人数	平均成绩
1	男	41	82.00
2	女	50	85.00

- 应用场景：频数格式数据，或枚举格式中每条记录的权重不同。

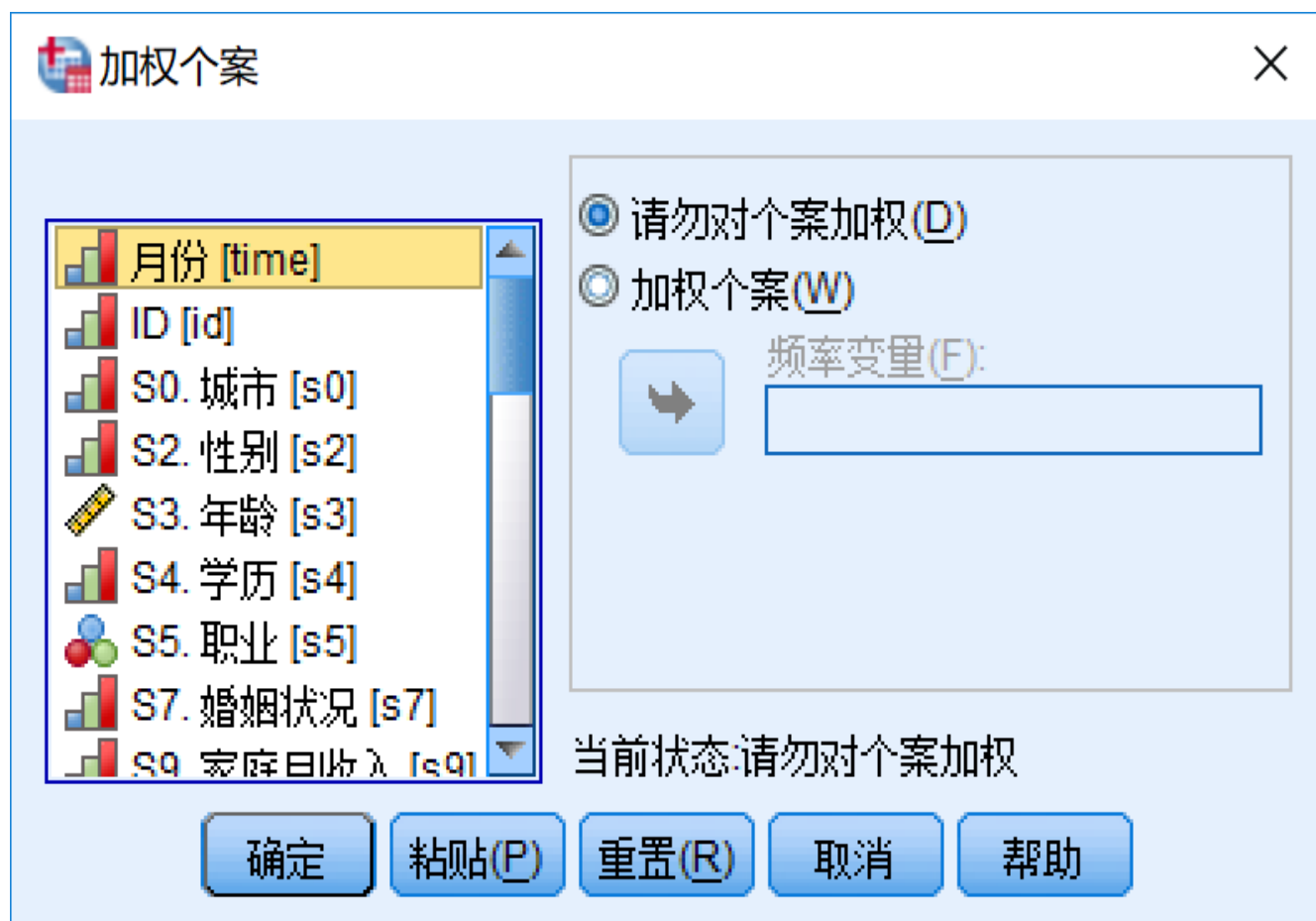
加权个案 (Weigh Cases)

- 进入实现过程

- 快捷工具栏的“加权个案”按钮。
- “数据” → “加权个案”。

- 只有先进行加权设定，统计软件才能正确识别和处理频数数据，将频数设置为权重。


加权个案 (Weigh Cases)





分类汇总 (Aggregate)


- 所谓分类汇总是指按指定的**分类变量**对个案进行分组，对每组记录的各变量求指定的**描述统计量**（和、平均值等）。
- 实现过程
 - “数据” → “汇总”
 - 指定分组变量和汇总变量
 - 设置汇总变量的描述统计量
 - 输出汇总结果


分类汇总 (Aggregate)


 汇总数据 ×


 月份 [time]


 ID [id]


 S2. 性别 [s2]


 S3. 年龄 [s3]


 S4. 学历 [s4]


 S5. 职业 [s5]


 S7. 婚姻状况 [s7]


 S9. 家庭月收入 [s9]


 C0. 请问您的家庭目...


 C0. 请问您的家庭目...


 C0. 请问您的家庭目...

 O1. 是否拥有家用轿...


 A3. 首先, 请问与一...

 A3a. 您为什么这么...

 A3a. 您为什么这么...

 A4. 那么与现在相比...

分组变量(B):

 S0. 城市 [s0]

汇总变量

变量摘要(S):

s3_mean = MEAN(s3)

函数(F)...

名称与标签(N)...

☐ 个案数(C)

名称(M): N_BREAK

保存

☒ 将汇总变量添加到活动数据集(D)

☐ 创建只包含汇总变量的新数据集(E)

数据集名称(D):

☐ 写入只包含汇总变量的新数据文件(W)

文件(L)...

E:\我的教学\管理统计学教学\17-18年第二学期本科\...\laggr.sav

分类汇总 (Aggregate)

- 案例

- 数据文件: CCSS_Sample.sav。
- 要求: 按time月份和s0城市对变量index1进行均数汇总, 并将结果输出到Sum_index1中。
- 实现过程:
 - “数据” → “汇总”。
 - 指定分类变量为time (第1分类变量) 和s0 (第2分类变量)。
 - 将变量index1选入汇总变量框。
 - “函数” 按钮点击进去后选择 “平均值”。
 - “保存” 框组中, 勾选第2项, 数据集名称设为 Sum_index1。

A8. 那么与现任相比...

A9. 那么您认为一年...

A10. 那么与现在相...

A16. 对于大宗耐用...

家庭收入2级 [Ts9]

Qs9

Qa3

Qa4

Qa8

Qa9

Qa10

Qa16

总指数 [index1]

现状指数 [index1a]

预期指数 [index1b]

time <= 200704 (FL...

分组变量(B):

月份 [time]

S0. 城市 [s0]

汇总变量

变量摘要(S):

index1_mean = MEAN(index1)

函数(F)...

名称与标签(N)...

☒ 个案数(C)
 名称(M): N_BREAK

保存

☐ 将汇总变量添加到活动数据集(D)
 ☒ 创建只包含汇总变量的新数据集(E)
 数据集名称(D): Sum_index1

☐ 写入只包含汇总变量的新数据文件(W)
 文件(L)... E:\我的教学\管理统计学教学\18-19年第二学期本科...\Sum.sav

适用于大型数据集的选项

☐ 文件已经按分组变量排序(A)
 ☐ 在汇总之前排序文件(G)

确定

粘贴(P)

重置(R)

取消

帮助

1

A

ABC

s7

s9

c0_1

c0_2

c1

未婚

2000-2999元

2 无

2 无

汇总数据: 汇总函数

汇总统计

特定值

个案数

☒ 平均值(M)
 ☐ 第一个(F)
 ☐ 加权(E)

☐ 中位数(N)
 ☐ 最后一个(L)
 ☐ 加权缺失(G)

☐ 合计(S)
 ☐ 最小值(U)
 ☐ 未加权

☐ 标准差(R)
 ☐ 最大值(X)
 ☐ 未加权缺失(H)

百分比、分数、计数

☐ 百分比(P)
 ☐ 分数(T)
 ☐ 计数(C)

☐ 以上(A)
 值:

☐ 以下(B)
 低:
 高:

☐ 内部(D)

☐ 外部(O)

继续

取消

帮助

已婚	5000-5999元	2 无	2 无
已婚	15000-199...	2 无	2 无
已婚	4000-4999元	2 无	2 无
已婚	4000-4999元	2 无	2 无



	time	s0	index1_mean	N_BREAK	变量	变量
1	200704	100北京	100.05	100		
2	200704	200上海	97.79	100		
3	200704	300广州	97.16	100		
4	200712	100北京	97.13	101		
5	200712	200上海	92.10	101		
6	200712	300广州	93.19	102		
7	200812	100北京	91.97	102		
8	200812	200上海	87.68	102		
9	200812	300广州	91.70	100		
10	200912	100北京	102.58	75		
11	200912	200上海	102.56	84		
12	200912	300广州	100.86	80		
13						

4.2 多个数据文件的合并

纵向拼接和横向合并

- 纵向拼接

- 从外部数据文件中**增加记录**到当前数据文件
- 相对而言比较简单，操作上没有需要特别注意之处

- 横向合并

- 从外部数据文件**增加变量**到当前数据文件

a.sav [数据集3] - IBM SPSS Statistics 数据编辑器

文件(F) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A)



	id	sex	age	height	变量
1	2	1	16	158	
2	4	1	34	164	
3	6	2	56	170	
4	8	1	68	172	
5	10	2	25	178	
6					
7					

c.sav [数据集5] - IBM SPSS Statistics 数据编辑器

文件(F) 编辑(E) 视图(V) 数据(D)



	id	weight	变量
1	2	46	
2	6	92	
3	8	51	
4	12	70	
5			

b.sav [数据集4] - IBM SPSS Statistics 数据编辑器

文件(F) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A) 直接输入



	id	sex	age	h	w	变量
1	1	2	19	165	53	
2	3	1	30	175	70	
3	5	2	28	162	48	
4	7	1	44	169	68	
5						

纵向拼接

- **数据文件：** a. sav（当前数据文件，教材自带）。
b. sav（外部数据文件，教材自带）。
- **要求：** 将b. sav的数据添加至a. sav中。
- **实现过程**
 - “数据” → “合并文件” → “**添加个案**” → 选择外部数据文件。
 - 将非同名但含义相同的变量**强行配对**进入当前数据文件：
height(*)和h(+)
 - 外部文件有而当前文件没有的变量可以进入也可以不进入当前文件:w(+)
 - 为个案来源设置标识变量：勾选“将个案源表示为变量”。

纵向拼接

 添加个案，来源 数据集4 ×

非成对变量(U):
height(*)
h(+)
w(+)


配对(A)

新的活动数据集中的变量(V):
id
sex
age

重命名(E)...

☒ 将个案源表示为变量(I):
源01

(*)=活动数据集
(+)=数据集4

确定

粘贴(P)

重置(R)

取消

帮助

纵向拼接

添加个案，来源 数据集4

非成对变量(U):

height(*)
h(+)
w(+)

新的活动数据集中的变量(V):

id
sex
age

配对(A)

☒ 将个案源表示为变量(I):


源01

重命名(E)...

(*)=活动数据集
(+)=数据集4

确定 粘贴(P) 重置(R) 取消 帮助

纵向拼接

 添加个案，来源 数据集4

非成对变量(U):

w(+)

重命名(E)...

新的活动数据集中的变量(V):

id
sex
age
height & h

☒ 将个案源表示为变量(I):
源01

配对(A)

(*)=活动数据集
(+)=数据集4

确定

粘贴(P)

重置(R)

取消

帮助

纵向拼接

*a.sav [数据集3] - IBM SPSS Statistics 数据编辑器









文件(F) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A) 直销(M)

11:

	id	sex	age	height	源01	变量
1	2	1	16	158	0	
2	4	1	34	164	0	
3	6	2	56	170	0	
4	8	1	68	172	0	
5	10	2	25	178	0	
6	1	2	19	165	1	
7	3	1	30	175	1	
8	5	2	28	162	1	
9	7	1	44	169	1	
10						

纵向拼接

*a.sav [数据集3] - IBM SPSS Statistics 数据编辑器

文件(F)	编辑(E)	视图(V)	数据(D)	转换(T)	分析(A)	直销(M)	图形(G)
							
12 : w							
	id	sex	age	height	w	源01	变量
1	2	1	16	158	.	0	
2	4	1	34	164	.	0	
3	6	2	56	170	.	0	
4	8	1	68	172	.	0	
5	10	2	25	178	.	0	
6	1	2	19	165	53	1	
7	3	1	30	175	70	1	
8	5	2	28	162	48	1	
9	7	1	44	169	68	1	
10							

横向合并

- **数据文件：** a. sav（当前数据文件，教材自带）。
c. sav（外部数据文件，教材自带）。
- **要求：** 将c. sav的数据添加至a. sav中。
- **实现过程**
 - “数据” → “合并文件” → “**添加变量**” → 选择外部数据文件。
 - 勾选“匹配关键变量的个案” → 勾选“**两个数据集中的个案都是按关键变量的顺序进行排序**” → 勾选“两个文件都提供个案”。
 - 将id(+)选入“关键变量”。
 - 为个案来源设置标识变量：勾选“将个案源表示为变量”。

横向合并

- 注意：三个选项

- “两个文件都提供个案”：两个文件中的记录都将进入合并后的文件。
- “非活动数据集为基于关键字的表”：保留当前文件所有数据，但丢弃只在外部文件出现的数据。
- “当前数据集为基于关键字的表”：保留外部文件所有数据，但丢弃只在当前文件出现的数据。

横向合并

添加变量从 数据集5



已排除的变量(E):

id(+)

重命名(A)...

☒ 匹配关键变量的个案(M)

☒ 两个数据集中的个案都是按关键变量的顺序进行排序(C)

☐ 非活动数据集为基于关键字的表(K)

☐ 活动数据集为基于关键字的表(A)

☒ 两个文件都提供个案(B)

☒ 将个案源表示为变量(I): 源01

(*)=活动数据集

(+)=数据集5

新的活动数据集(N):

sex(*)

age(*)

height(*)

weight(+)

id(*)

关键变量(V):

确定


粘贴(P)

重置(R)

取消

帮助

横向合并

 添加变量从 数据集5 ×

已排除的变量(E):

重命名(A)...

☒ 匹配关键变量的个案(M)
☒ 两个数据集中的个案都是按关键变量的顺序进行排序(C)

☐ 非活动数据集为基于关键字的表(K)
☐ 活动数据集为基于关键字的表(A)
☒ 两个文件都提供个案(B)

☒ 将个案源表示为变量(I): 源01

(*)=活动数据集
(+)=数据集5

确定

粘贴(P)

重置(R)

取消

帮助

新的活动数据集(N):

sex(*)
age(*)
height(*)
weight(+)

关键变量(V):

id

横向合并

*a.sav [数据集3] - IBM SPSS Statistics 数据编辑器

文件(F) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A) 直销(M) 图形(G)

11 : weight

	id	sex	age	height	weight	源01	变量
1	2	1	16	158	46	1	
2	4	1	34	164	.	0	
3	6	2	56	170	92	1	
4	8	1	68	172	51	1	
5	10	2	25	178	.	0	
6	12	.	.	.	70	1	
7							

横向合并

保留当前文件所有数据，
但丢弃只在外部文件出
现的数据

a: id	c: id
2	2
4	6
6	8
8	12
10	

添加变量从 数据集5

已排除的变量(E):

新的活动数据集(N):

- sex(*)
- age(*)
- height(*)
- weight(+)

匹配关键变量的个案(M)

- ☒ 两个数据集中的个案都是按关键变量的顺序进行排序(C)
- ☒ 非活动数据集为基于关键字的表(K)
- ☐ 活动数据集为基于关键字的表(A)
- ☐ 两个文件都提供个案(B)

将个案源表示为变量(I): 源01

(*)=活动数据集
(+)=数据集5

关键变量(V):

- id

确定 粘贴(P) 重置(R) 取消 帮助

横向合并

*a.sav [数据集3] - IBM SPSS Statistics 数据编辑器

文件(F) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A) 直销(M)

10 :

	id	sex	age	height	weight	变量	
1	2	1	16	158	46		
2	4	1	34	164	.		
3	6	2	56	170	92		
4	8	1	68	172	51		
5	10	2	25	178	.		
6							
7							

横向合并

保留外部文件所有数据，
但丢弃只在当前文件出
现的数据。

a: id	c: id
2	2
4	6
6	8
8	12
10	

添加变量从 数据集5

已排除的变量(E):

新的活动数据集(N):

sex(*)
age(*)
height(*)
weight(+)

重命名(A)...

☒ 匹配关键变量的个案(M)

☒ 两个数据集中的个案都是按关键变量的顺序进行排序(C)

☐ 非活动数据集为基于关键字的表(K)

☒ 活动数据集为基于关键字的表(A)

☐ 两个文件都提供个案(B)

将个案源表示为变量(I): 源01

(*)=活动数据集
(+)=数据集5

确定 粘贴(P) 重置(R) 取消 帮助

关键变量(V):

id

横向合并

*a.sav [数据集3] - IBM SPSS Statistics 数据编辑器

文件(F) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A) 直销(M)

12:

	id	sex	age	height	weight	变量
1	2	1	16	158	46	
2	6	2	56	170	92	
3	8	1	68	172	51	
4	12	.	.	.	70	
5						
6						



南京大學
NANJING UNIVERSITY

4.3 数据文件的重组与转置

长型数据格式和宽型数据格式

- **长型：**在重复测量个体数据过程中，同一个体的每一次测量被单独记录为一个case。 Anxiety.sav
- **宽型：**在重复测量个体数据过程中，每一个个体被记录为一个case，他的所有测量被记录在不同的变量中。
Anxiety2.sav
- 上述两个文件记录的都是12名精神病患者在接受治疗后的4个时间点的精神状态评分。变量有“Subject(病人编号)”、“Score(评分)”、“Trial(跟踪测量的时间点编号)”、“Anxiety(治疗前是否焦虑)”、“Tension(治疗前是否紧张)”。

长型格式转换为宽型格式

- 数据文件：Anxiety.sav。
- 要求：将长型数据重组为宽型数据。
- 实现过程
 - “数据” → “重组”。
 - 选择格式变换类型：“将选定个案重组为变量”。
 - 选择识别个案组的变量：Subject。
 - 选择索引变量：Trial。

长型格式转换为宽型格式

重组数据向导

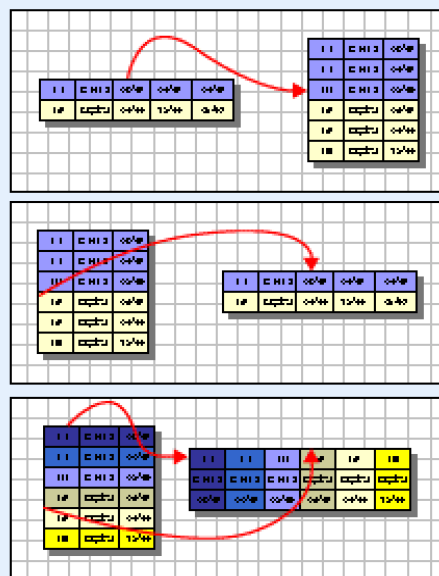


欢迎使用“重组数据向导”！

此向导帮您将数据从单个个案中的多个变量（列）重组到相关个案（行）组合，反之亦然，或者您可以选择变换数据。



向导使用重组数据替换当前数据集。注意不能取消数据重组。



您希望做什么？

☐ 将选定变量重组为个案(C)

当您的当前数据中每个个案有一些您想重新安排到新数据集中相关个案组合的变量时，使用此功能。

☒ 将选定个案重组为变量(V)

当您想重新安排相关个案组合，从而来自每个组合的数据表示为新数据集中的单个个案时，使用此功能。

☐ 变换所有数据(T)

所有个案将成为变量，所选变量将成为新数据集中的个案。（选择此选项将结束向导，变换对话框将出现。）

< 上一步(B)

下一步 >(N)

完成

取消

帮助

长型格式转换为宽型格式

重组数据向导 - 第 2 步 (共 5 步)



个案到变量: 选择变量

将当前文件中的个案组数据重组为新文件中的单个个案。

通过将这些变量移动到标识变量列表中，选择识别个案组合的变量。或者您还可选择索引变量。



保持在当前文件变量列表中的变量包含在个案组合内变化的数据或不会变化的数据。

拥有变化数据的变量将成为重组文件中的新变量组合。拥有不会变化数据的变量将复制到新文件中。

当前文件中的变量(C):

Anxiety [anxiety]
 Tension [tension]
 Score [score]



标识变量(D):

Subject [subject]

索引变量(X):

Trial [trial]



< 上一步(B)

下一步 >(N)

完成

取消

帮助

宽型格式转换为长型格式

- 数据文件: Anxiety2.sav
- 要求: 将宽型数据重组为长型数据
- 实现过程
 - “数据” → “重组”
 - 选择格式变换类型: “将选定变量重组为个案”
 - 确定需重组的变量组个数: 一个
 - 确定重组后数据个案组的标识变量: Subject
 - 选择需被转换的变量组: Trial 1—Trial 4
 - 设置重组后文件中变量组对应到的一个目标变量的名称: score
 - 设置固定变量: Anxiety 和 Tension
 - 设置索引变量的数目: 一个

宽型格式转换为长型格式

重组数据向导

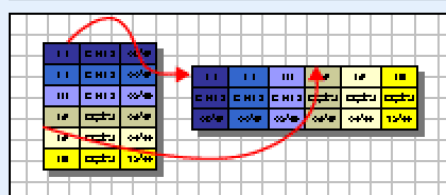
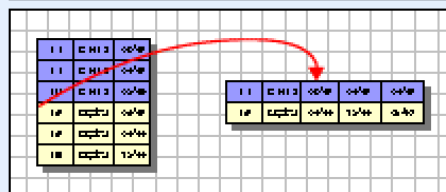
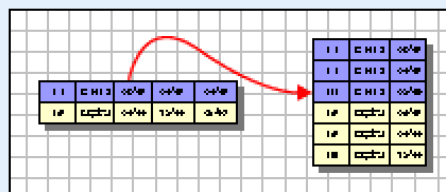


欢迎使用“重组数据向导”！

此向导帮您将数据从单个个案中的多个变量（列）重组到相关个案（行）组合，反之亦然，或者您可以选择变换数据。



向导使用重组数据替换当前数据集。注意不能取消数据重组。



您希望做什么？

☒ 将选定变量重组为个案(C)

当您的当前数据中每个个案有一些您想重新安排到新数据集中相关个案组合的变量时，使用此功能。

☐ 将选定个案重组为变量(V)

当您想重新安排相关个案组合，从而来自每个组合的数据表示为新数据集中的单个个案时，使用此功能。

☐ 变换所有数据(T)

所有个案将成为变量，所选变量将成为新数据集中的个案。（选择此选项将结束向导，变换对话框将出现。）

< 上一步(B)

下一步 >(N)

完成

取消

帮助

宽型格式转换为长型格式

重构数据向导 - 第 2 步 (共 7 步)



变量到个案: 变量组数量

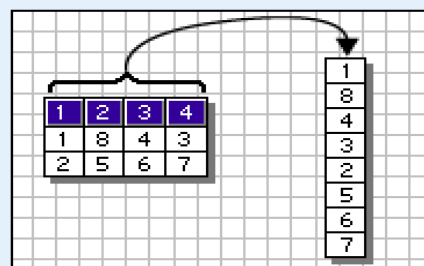
您已经选择将选定变量重组为新文件中的相关个案组。

一组相关变量 (称为变量组) 表示一个变量的多个测量结果。



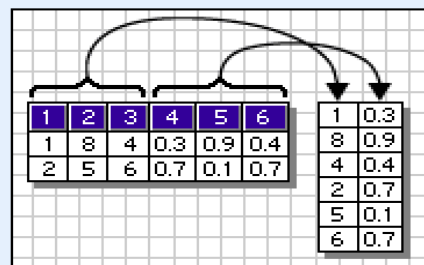
例如, 变量可以是宽度。如果用三种不同测量记录, 每个代表时间上的不同点—w1、w2 和 w3, 然后在变量组合中重新安排数据。

如果在文件中有超过一个变量, 同时通常记录在变量组合中, 例如高度, 记录为 h1、h2 和 h3。



您希望重组多少个变量组?

☒ 一个 (例如, w1、w2 和 w3) (O)



☐ 多个 (例如, w1、w2 和 w3 以及 h1、h2 和 h3 等等) (M)
数量(H)?

< 上一步(B)

下一步 >(N)

完成

取消

帮助

宽型格式转换为长型格式

重构数据向导 - 第 3 步 (共 7 步)



变量到个案: 选择变量

对于当前数据中的每个变量组, 已重组的文件中都会有一个目标变量。

在此步骤中, 选择如何识别重组数据中的个案组合, 并选择哪些变量属于每个目标变量。

或者, 您还可以选择将变量作为“固定变量”复制到新文件中。

当前文件中的变量(C):

Subject [subject]
 Anxiety [anxiety]
 Tension [tension]
 Trial 1 [trial1]
 Trial 2 [trial2]
 Trial 3 [trial3]
 Trial 4 [trial4]

个案组标识(G):

使用选定变量



变量(R): Subject [subject]

要变换的变量

目标变量(T):

score



Trial 1 [trial1]
 Trial 2 [trial2]
 Trial 3 [trial3]
 Trial 4 [trial4]

固定变量(F):



Anxiety [anxiety]
 Tension [tension]

< 上一步(B)

下一步 >(N)

完成

取消

帮助

宽型格式转换为长型格式

重构数据向导 - 第 4 步 (共 7 步)

×

变量到个案: 创建索引变量

在当前数据中, 变量组合的值以多个变量的形式出现在单个个案中。例如, 单个个案包含 w1、w2 和 w3 的值。

在新数据中, 变量组合的值将以单个变量的形式出现在多个个案中。例如, 有三个个案, 每个有 w1、w2 和 w3。

索引是识别从初始个案创建的新个案组合的新变量。例如, 名为 "w" 的索引将有值 1、2 和 3。

1	1	1	0.07
1	1	2	0.11
1	1	3	0.05
2	1	1	0.08
2	1	2	0.04
2	1	3	0.06

您希望创建多少索引变量?

☒ 一个(O)

当变量组合记录单个因子、处理方法或条件的效果时, 使用此功能。

☐ 多个(M)

多少(H)?

2

当变量组合记录多个因子、处理方法或条件的效果时, 使用此功能。

☐ 无(E)

如果要将索引信息保存在要变换的变量集中, 使用此功能。

1	1	1	1	0.07
1	1	1	2	0.11
1	1	1	3	0.05
1	1	2	1	0.08
1	1	2	2	0.04
1	1	2	3	0.06

1	1	0.08	2	0.07
2	1	0.11	2	0.11
3	1	0.07	2	0.05
4	1	0.06	2	0.08
5	1	0.09	2	0.04
6	1	0.02	2	0.06

< 上一步(B)


下一步 >(N)

完成

取消

帮助

宽型格式转换为长型格式

 重构数据向导 - 第 5 步 (共 7 步)

×

变量到个案: 创建一个索引变量

您已经选择创建一个索引变量。变量的值可以是连续数字或组合中的变量名称。
在表中，您可以指定索引变量的名称和标签。

索引值是什么类型?

☒ 有序数(S)
索引值(D): 1, 2, 3, 4

☐ 变量名(A)
索引值(D): trial1, trial2, trial3, trial4

编辑索引变量名称和标签(X):

	名称	标签	级别	索引值
1	trial		4	1, 2, 3, 4

< 上一步(B)

下一步 >(N)

完成

取消

帮助

数据转置

- **数据转置：**对数据进行**行列互换**，即将记录转为变量，将变量转为记录后，重新显示在编辑窗口中。
- **实现过程**
 - “数据” → “变换”
 - 选择需要转置的变量
 - 选择转置后的变量名

数据转置

文档关键词.sav [数据集2] - IBM SPSS Statistics 数据编辑器

文件(F) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A) 直销(M) 图形(G) 实用程序(U) 窗口(W) 帮助(H)



13 : keyword6

	文档编号	keyword1	keyword2	keyword3	keyword4	keyword5	keyword6	变量
1	d1	1	1	0	0	1	0	
2	d2	0	1	1	1	0	0	
3	d3	1	1	0	1	0	1	
4	d4	0	0	1	1	1	0	
5								
6								
7								
8								
9								
10								
11								
12								

数据转置



数据转置

*未标题2 [数据集3] - IBM SPSS Statistics 数据编辑器

文件(F) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A) 直销(M) 图形(G) 实用程序(U) 窗口(W) 帮助(H)



	CASE_LBL	d1	d2	d3	d4	变量
1	keyword1	1.00	.00	1.00	.00	
2	keyword2	1.00	1.00	1.00	.00	
3	keyword3	.00	1.00	.00	1.00	
4	keyword4	.00	1.00	1.00	1.00	
5	keyword5	1.00	.00	.00	1.00	
6	keyword6	.00	.00	1.00	.00	
7						
8						
9						
10						
11						
12						

THE END