

赋值错误更正

本模块可用于更正数据错误，给缺失值赋值。

数据错误定义方式：

1. 通过条件语句定义，如 “> 100”
2. 异常值定义可采用：
 - 大于均数+3 倍标准差，或小于均数减 3 倍标准差。
 - 大于某（如 99%）百分位数，或小于某（如 1%）百分位数。

输出新变量，以确保原数据没有改动。

例：将 HEIGHT>2.1 的改为缺失值，输出新变量 HEIGHT.NEW：

原变量名	新变量名
HEIGHT	HEIGHT.NEW
更正条件(如 >100, ==200, >"A")	正确赋值 (如 100, NA, "A")
<input type="text" value=">2.1"/> <input type="checkbox"/> 缺失	<input type="text" value="NA"/>

例：将 FEV1 3 倍标准差之外的改为缺失，输出新变量 FEV1.NEW

原变量名	新变量名
FEV1	FEV1.NEW
更正条件(如 >100, ==200, >"A")	正确赋值 (如 100, NA, "A")
<input type="text"/> <input type="checkbox"/> 缺失	<input type="text"/>
<input type="text"/> <input type="checkbox"/> 缺失	<input type="text"/>
<input type="text"/> <input type="checkbox"/> 缺失	<input type="text"/>
<input checked="" type="checkbox"/> 定义极端值如下，替换极端值为缺失	
<input checked="" type="checkbox"/> < 均数 - <input type="text" value="3"/> SD (均数减n倍的标准差)	<input type="checkbox"/> < <input type="text" value="1"/> % 百分位数
<input checked="" type="checkbox"/> > 均数 + <input type="text" value="3"/> SD (均数加n倍的标准差)	<input type="checkbox"/> > <input type="text" value="99"/> % 百分位数
<input type="button" value="保存"/>	<input type="button" value="取消"/>

例：将 DBP 最低 1%与最高 1%改为缺失，输出新变量 DBP.NEW1：

原变量名		新变量名	
DBP		DBP.NEW1	
更正条件(如 >100, ==200, >"A")		正确赋值(如 100, NA, "A")	
<input type="text"/>	<input type="checkbox"/> 缺失	<input type="text"/>	
<input type="text"/>	<input type="checkbox"/> 缺失	<input type="text"/>	
<input type="text"/>	<input type="checkbox"/> 缺失	<input type="text"/>	
<input type="checkbox"/> 定义极端值如下, 替换极端值为缺失			
<input type="checkbox"/> < 均数 - <input type="text" value="3"/> SD (均数减n倍的标准差)		<input checked="" type="checkbox"/> < <input type="text" value="1"/> %百分位数	
<input type="checkbox"/> > 均数 + <input type="text" value="3"/> SD (均数加n倍的标准差)		<input checked="" type="checkbox"/> > <input type="text" value="99"/> %百分位数	

例：将 AGE 缺失值赋值为 35，输出新变量 AGE.NEW

原变量名		新变量名	
AGE		AGE.NEW	
更正条件(如 >100, ==200, >"A")		正确赋值(如 100, NA, "A")	
<input type="text"/>	<input checked="" type="checkbox"/> 缺失	<input type="text" value="35"/>	

与其它数据合并（增加变量）

把几个数据文件按记录编号（研究对象编号）侧面相合并，这样结果数据文件变量数就是所有数据文件变量数（除研究对象编号变量外）的总和。

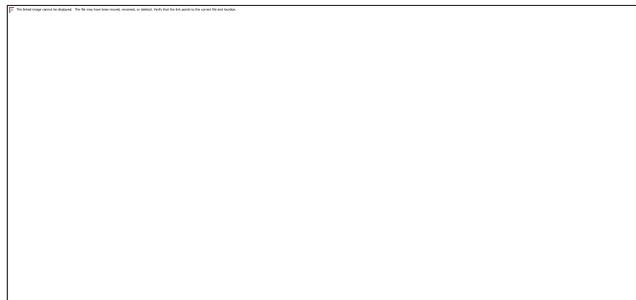
需要给出研究对象编号变量，系统会自动检查每个数据文件的每个研究对象的编号是否有不同，根据选择可以输出所有研究对象或共同的研究对象或部分研究对象。

假设数据文件 D1、D2、D3，研究对象编号如下表所示：

	D1	D2	D3
.....			
A01	√	√	√
A02		√	√
A03	√		√
A04	√	√	
.....			

A01 出现在所有数据文件中，A02 不在 D1 中，A03 不在 D2 中，A04 不在 D3 中。

点击“增加”选择所要合并的数据文件名，选择数据文件名，右击“输出选择”列，系统提示输出选择：



选择“1”表示输出文件中的所有研究对象都应存在于该数据文件中
选择“0”表示输出文件中的所有研究对象都不应存在于该数据文件中
选择“.”表示输出文件中的所有研究对象可在可不在该数据文件中

上表中，如输出选择为：“...”，即数据文件 D1 的输出选择为“.”，D2 为“.”，D3 为“.”，则 A01-A04 都满足条件；如输出选择为：“.10”，即数据文件 D1 的输出选择为“.”，D2 为“1”，D3 为“0”，则 A04 满足，A01-A03 都不满足条件；如输出选择为：“111”，即数据文件 D1 的输出选择为“1”，D2 为“1”，D3 为“1”，则 A01 满足，A02-A04 都不满足条件；如输出选择为：“.1.”，即数据文件 D1 的输出选择为“.”，D2 为“1”，D3 为“.”，则 A01、A02、A04 满足，A03 不满足条件；如输出选择为：“11.”，即数据文件 D1 的输出选择为“1”，D2 为“1”，D3 为“.”，则 A01、A04 满足，

A02、A03 不满足条件。

