

## 在每条记录内计算统计量

在每条记录内计算几个变量的平均数、中位数、百分位数、极差、总和、等等，可以通过“创建新变量”窗口，如下图所示：在 DEMO 数据中要计算每个对象四个呼吸症状有几个（求和）

创建新变量

新变量名: NSYPTOMS 连续性 (点击切换) ?

\* 变量名应是大写, 点击此处查看函数、操作符、表达式输入法  自动转换表达式

由表达式产生 (如 wt./ht./ht):  
NSYPTOMS =

保存

由两个日期变量计算年龄 (日期2-日期1)

日期2 变量名 日期2 格式如: 日期1 变量名 日期1 格式如:  
 1/21/1999 -  1/21/1999 /365.25

取消

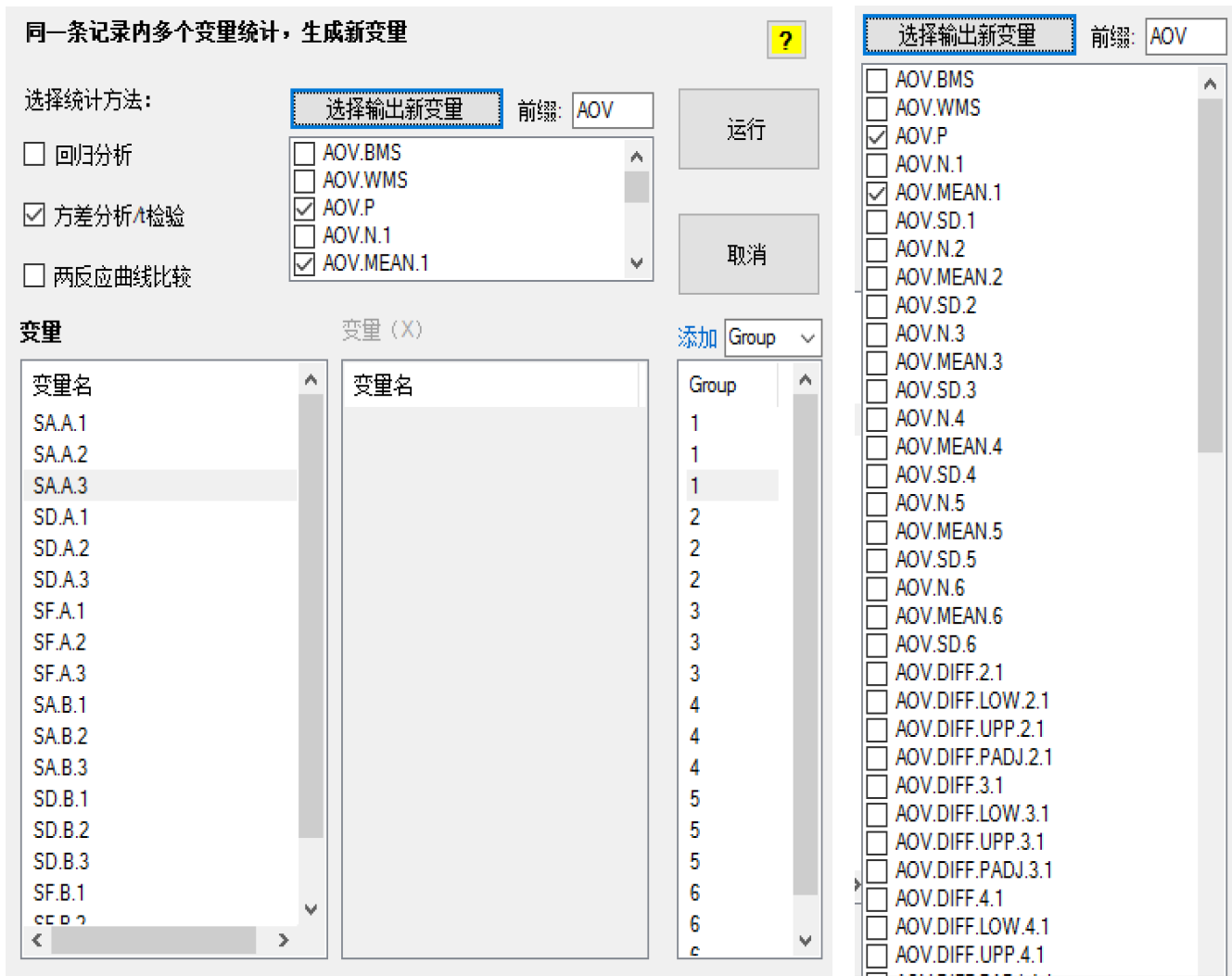
由几个变量计算统计量 (如均数, 总和, 等) 原变量清单 (变量1, 变量2, ...)

总和 of COUGH, PHLEGM, WHEEZE, SOB

重复测量资料，如果每个研究对象一条记录，在同一条记录内有多次测量结果，则可能需要对多个测量值进行更复杂的统计，如计算变化斜率、组间比较等。

另一种常见情况如基因芯片数据，一般是每个基因一条记录，每个样品一列。根据课题设计，可能有 2 个或多组样本或是时间系列样本（不同时间点收集的样品），这时要比较样本组，或分析基因表达与时间的关系，就可以使用该模块。

例，一基因芯片研究比较两组样本基因表达有无差异：有 3 例患者（"SA"、"SD"和"SF"），每个病人提供甲醛固定石蜡包埋（FFPE）样本和冷冻样本，看不同样本处理对结果有多大影响。每个样品重复了三次，因此每个基因有 18 个检测值，分 6 组。下图示，如何对这 6 组数据进行方差分析并做两两比较。



上图中的组别 1、2、3、4、5、6 是手工输入的，右击输入框，输入 1:6\*3 即表示从 1 到 6，每个重复 3 次。点击选择输出变量显示右边的输出变量窗口。

例，观察一组儿童身高增长，每个儿童测量了 15 次，每次记录年龄、身高。现在要计算每个儿童身高年增长率，即在每条记录内对这 15 次测量的身高值与年龄进行回归分析，求得每个人的回归系数。输入界面如下。

**同一条记录内多个变量统计，生成新变量** ?

选择统计方法：

回归分析

方差分析/t检验

两反应曲线比较

选择输出新变量 前缀: REG

REG.SLOPE

REG.SE

REG.P

运行

取消

应变量 (Y)

自变量 (X)

添加 Time

变量名	变量名	Time
HT1	AGE1	
HT2	AGE2	
HT3	AGE3	
HT4	AGE4	
HT5	AGE5	
HT6	AGE6	
HT7	AGE7	
HT8	AGE8	
HT9	AGE9	
HT10	AGE10	
HT11	AGE11	
HT12	AGE12	
HT13	AGE13	
HT14	AGE14	
HT15	AGE15	

如果每个人每次测量的时间是固定的，即每条记录的 AGE1、AGE2、.....、AGE15 都是相同的，这时数据中 AGE1-AGE15 这些变量就可以省去。没有这些 X 变量，要进行回归，就需要手工输入每次测量的时间值。点击“添加”时间，输入时间值即可。