

IBM SPSS 20.0 统计软件应用

南方医科大学生物统计学系

(http://www.echobelt.org/)

刘薇

danmin64@163.com



Outline

- ▶ 1. 绪论、数据文件的建立及整理 (第一-三章)
- > 2. 数据转换 (第四章)
- > 3. 数据描述 (第六章1-3节)
- ▶ 4. 均数比较 (第八章)
- > 5. 一般线性模型 (第九章)
- ▶ 6. 习题课(前半程教学内容)
- 7. 统计图形绘制与编辑(第十五章)
- 8. 卡方检验(第六章第四节)
- > 9. 非参数检验 (第十三章)
- ▶ 10.相关与回归(第十章1-6节)
- ▶ 11.习题课(后半程教学内容)
- ▶ 12.总复习(全部授课章节)
- ▶ 13.考试



主要内容

- > 概述
- > 数据文件的建立、导入和导出
- > 数据文件的整理
- > SPSS的应用注意事项



第一章 概述



概述

- SPSS (Statistical Package for Social Sciences) 于 1968年由美国斯坦福大学的三位研究生研发而成,并以此名称创立SPSS公司。
- 经过多年发展,SPSS软件已成为国际上最权威的统计软件之一。其他主流的统计分析软件包括 SAS软件、Stata软件、R软件等等。
- 2009年7月28日宣布被IBM公司收购,改名IBM SPSS. 当前最高版本为IBM SPSS 26.0。



概述

Time	Version	Time	Version
<u>1968</u>	SPSS 1	2006	SPSS 15.0
1984	SPSS/PC+	2007	SPSS 16.0
1992	SPSS 5.0	2009	SPSS 17.0
1996	SPSS 6.0/7.0	2010	IBM SPSS 18/19
1997	SPSS 8.0	2011	IBM SPSS 20
1998	SPSS 9.0	2012	IBM SPSS 21
1999	SPSS 10.0	2013	IBM SPSS 22
2002	SPSS 11.0	2014	IBM SPSS 23
2003	SPSS 12.0	2015	IBM SPSS 24
2004	SPSS 13.0	2016	IBM SPSS 25
2005	SPSS 14.0	2017	IBM SPSS 26



概述

■SPSS软件的优点:

- 权威性(被美国FDA等权威机构认可)
- ■操作简单,界面友好(WINDOWS操作理念)
- ■输出规范并图形化(可编辑与共享),输出格式多样:



IBM SPSS 20主要窗口及其功能

数据编辑窗 数据文件的建立,整理

结果输出窗

分析结果的展示

■ 程序编程窗

一般供专业人员使用

部分高级功能需编程实现



IBM SPSS20主要窗口及其功能

■数据编辑窗 .sav

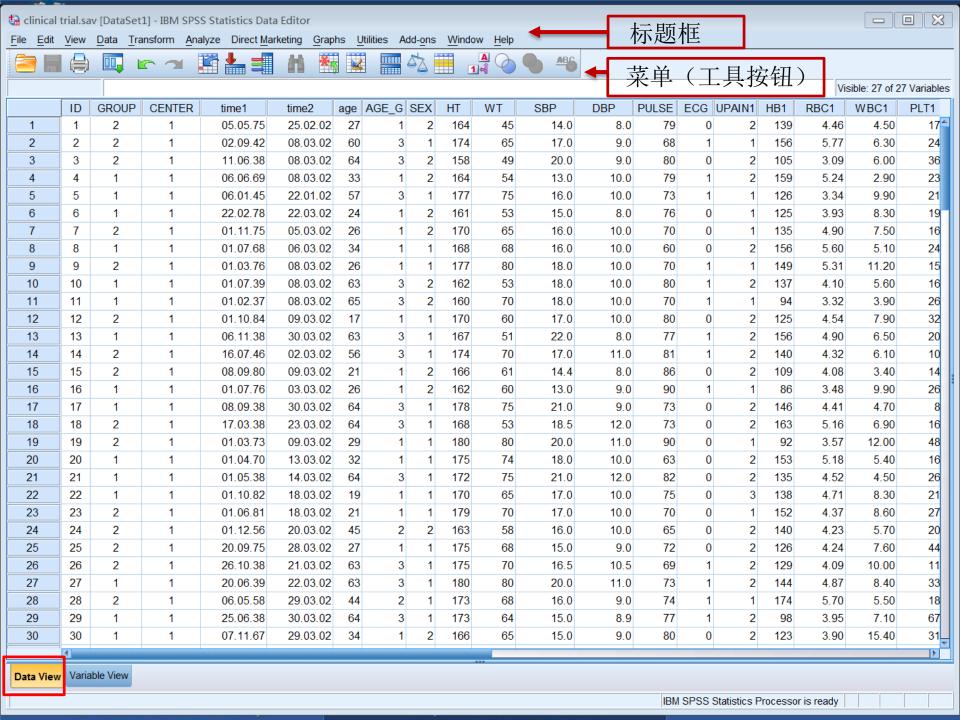


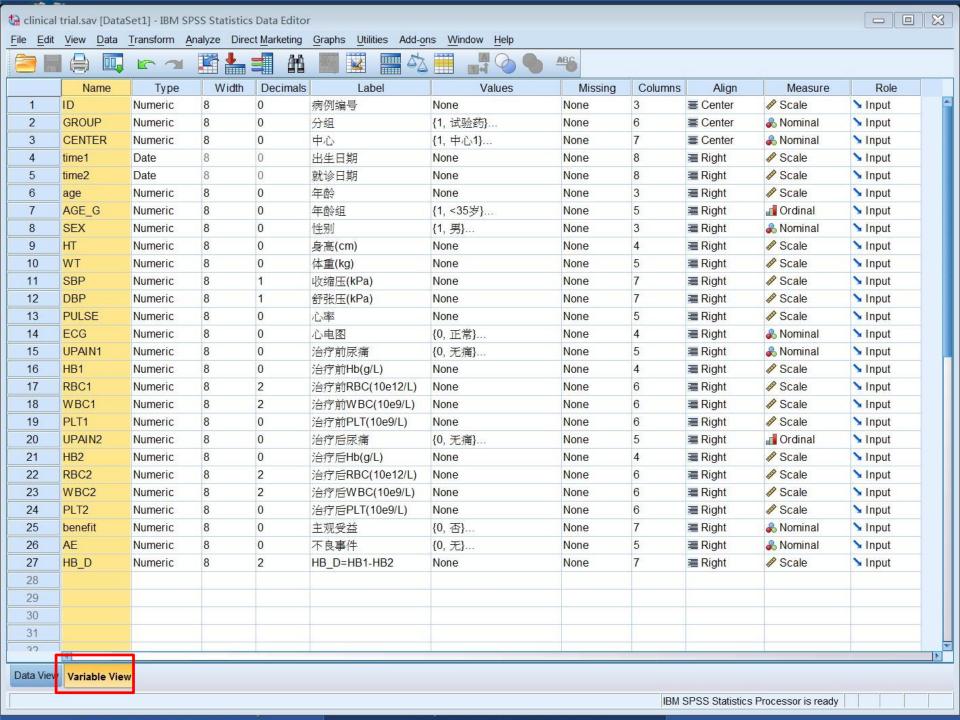
ANCOVA

文件类型:

SPSS Statistics Data Document (.sav)

- ■功能:建立新的数据文件,编辑和显示已有的数据文件
- ■内容
 - ■数据窗口(Data View):显示和编辑数据
 - 变量窗口(Variable View): 定义、显示和 编辑变量特征







IBM SPSS20主要窗口及其功能

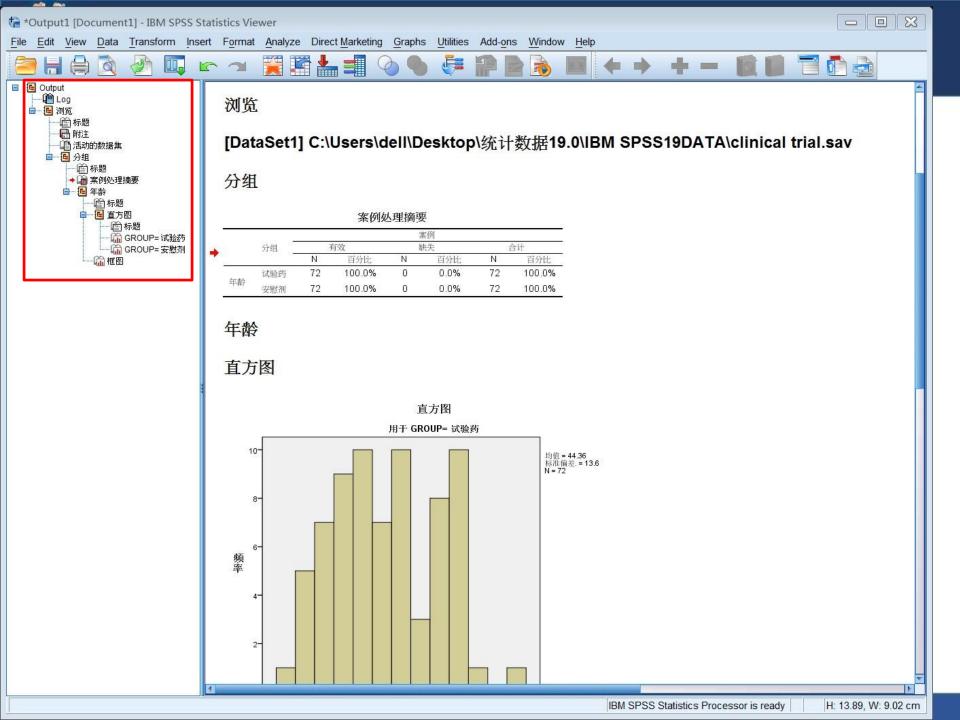
■ 结果输出窗口 .spv



Output1

文件类型: SPSS Statistics Output Document (.spv)

■ 统计分析结果:内容有文本、图形、表格





IBM SPSS20主要窗口及其功能

■程序编辑窗 .sps



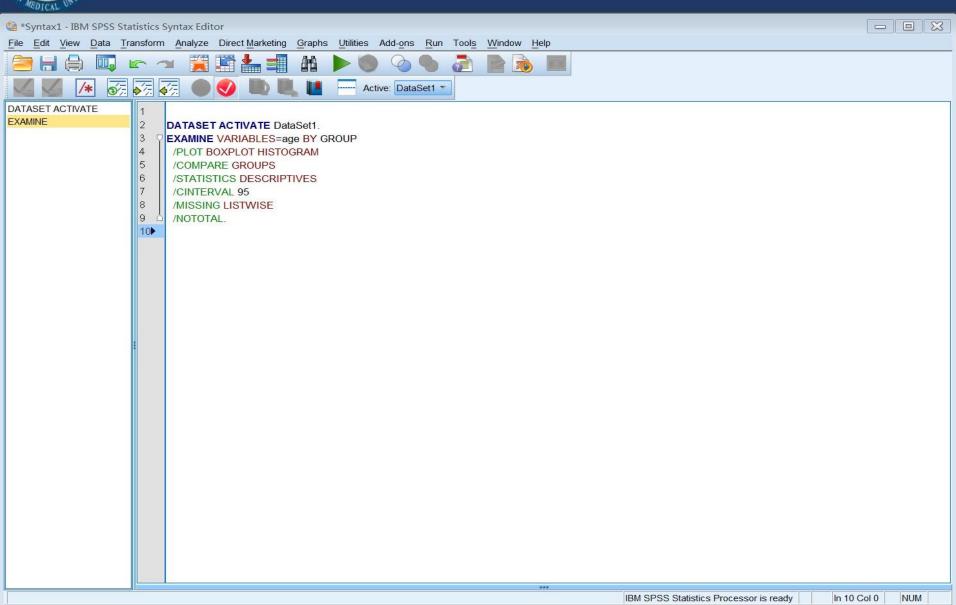
Syntax1

文件类型: SPSS Statistics Syntax File (.sps)

- ■满足专业人员使用
- ■避免重复劳动
- ■个别分析需编程实现(裂区分析,典则相关等)



程序编辑窗





显示菜单(View)

状态栏

- ■Status Bar 运行状况显示条
 - IBM SPSS statistics Processor is ready 正常运行
 - IBM SPSS statistics Processor is unavailable 不能正常运行(一般是因为安装的破解版本有问题,需要重新安装)



显示菜单(View)

• 菜单栏

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

• 数据编辑栏





第二章

数据文件的建立、导入和导出

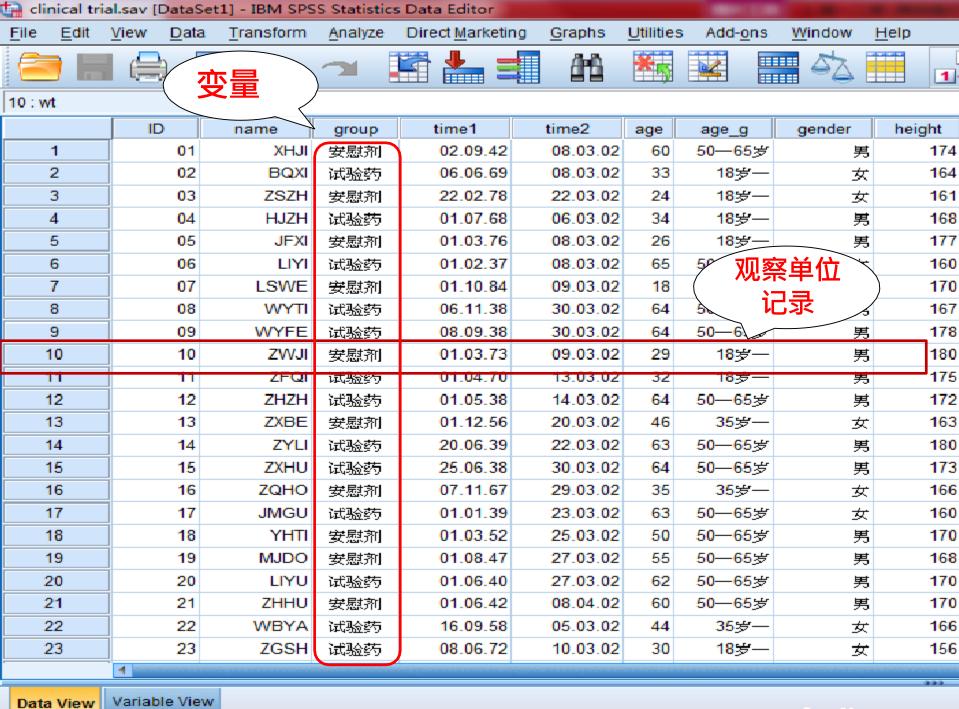


数据文件的建立—数据窗口

- ■数据文件结构
 - ■二维行列结构
 - ■行 cases、record (观察单位、记录):

个体、患者、动物

■列 — variables (变量): 指标

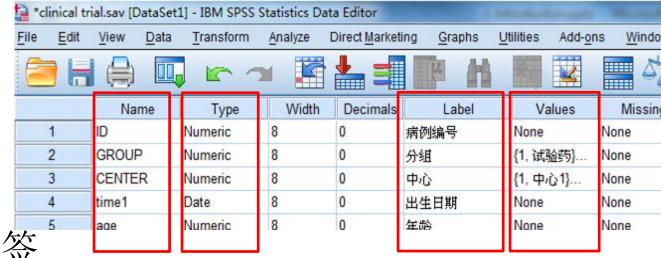


Variable View



数据文件的建立——变量窗口

- ■定义变量(变量窗口)
 - ■Name 变量名
 - ■Type 变量类型
 - ■数值型
 - ■字符型
 - ■日期型
 - ■Label 变量标签
 - ■Values 变量值标签





数据文件的建立

• 变量名 (Name)

不允许有两个相同的变量名

不超过64字节或者32个汉字

必须以字母或者汉字开头

对英文字母不区分大小写

- 变量标签 (Label) ----长度不限
- · 变量值标签 (Value) ----分类变量或等级变量



练习:

- 测得某班5名男生和5名女生身高(单位:米) 数据如下
 - -男: **1.70**, **1.68**, **1.72**, **1.75**, **1.77**
 - -女: 1.65, 1.60, 1.58, 1.55, 1.62
- 试建立数据文件
- 提示: 一个ID变量; 一个分组变量(男/女);
 - 一个指标变量(身高)



数据文件的导入和导出

➤可导入、导出的文件 件类型 SPSS 数据文件 Excel 数据文件 SAS 数据文件 STATA 数据文件 ASCII数据文件 ■将数据文件导入SPSS

File → Open →

Data



数据文件的导入和导出

	REDICAL DAY											
clinical trial.sav [DataSet1] - IBM SPSS Statistics Data Editor								_	_	_	_	
<u>F</u> ile	<u>E</u> dit	⊻iew	<u>D</u> ata	Transform	<u>A</u> nalyze D	irect <u>M</u> arketing	<u>G</u> rapi	hs <u>U</u> tili	ties	Add- <u>o</u> ns	<u>W</u> indo	w <u>H</u> elp
	<u>N</u> ew				>		26	8-8	-34-			<u>△</u>
	<u>O</u> pen				▶	Data		-		J		
	Open Da	ta <u>b</u> ase			▶	Syntax						
a	Rea <u>d</u> Tex	t Data				Output	ige	AGE_G	SEX	HT	WT	SBP
	Read Cog	gnos Da	ata			⊕ S <u>c</u> ript	27	<35岁	女	164	45	14.0
	Close			C	Ctrl+F4		60	50岁—	男	174	65	17.0
	Save				CtrI+S	08.03.02	64	50∌/—	女	158	49	20.0
	- Save As					08.03.02	33	<35≸	女	164	54	13.0
	Save All D					22.01.02	57 24	50岁—	男	177 161	75 53	16.0
	Export to I		22.03.02 05.03.02	26	<35岁	女 女	170	65	15.0 16.0			
	_		06.03.02	34	<35岁	男	168	68	16.0			
	Mar <u>k</u> File					08.03.02	26	<35₩	男	177	80	18.0
	Rena <u>m</u> e I					08.03.02	63	50岁—	女	162	53	18.0
			e Informatio	on	•	08.03.02	65	50岁—	女	160	70	18.0
(2)	Cac <u>h</u> e Da	ata				09.03.02	17	<35岁	男	170	60	17.0
4	Collect Va	ariable I	nformation	ı		30.03.02	63	50岁—	男	167	51	22.0
	Stop Proc	essor		C	Ctrl+Period	02.03.02	56	50岁—	男	174	70	17.0
***	S <u>w</u> itch Se	rver				09.03.02	21	<35岁	女	166	61	14.4
	<u>R</u> eposito	ry			▶	03.03.02	26	<35岁	女	162	60	13.0
	Set Viewe	er Outpu	ıt Options (Syntax)		30.03.02	64	50岁—	男	178	75	21.0
3	Print Pre <u>v</u>		23.03.02	64	50岁—	男	168	53	18.5			
ا 🖴	Print Ctrl+P					09.03.02	29	<35岁	男	180	80	20.0
	Welcome Dialog					13.03.02	32	<35岁	男	175	74	18.0
_	Recently Used Data					14.03.02	64 19	50岁—	男	172 170	75 65	21.0
	Recently Used Files					18.03.02 18.03.02	21	<35岁	男	170	70	17.0 17.0
	Exit		•			20.03.02	45	35岁—	男女	163	58	16.0
	25	25	安慰剂	中心1	20.09.7		27	<35≸	男	175	68	15.0
	26	26	安慰剂	中心1	26.10.3		63	50岁—	男	175	70	16.5
	27	27	试验药	中心1	20.06.3		63	50岁—	男	180	80	20.0
	28	28	安慰剂	中心1	06.05.5		44	35岁—	男	173	68	16.0
	29	29	试验药	中心1	25.06.3	8 30.03.02	64	50岁—	男	173	64	15.0
	30	30	试验药	中心1	07.11.6	7 29.03.02	34	<35岁	女	166	65	15.0
	31	31	试验药	中心1	01.01.3	9 23.03.02	63	50岁—	女	160	63	19.0
	32	32	安慰剂	中心1	01.03.5	25.03.02	50	50岁—	男	170	70	20.0
200									-			

Southern Medical University



数据文件的导入和导出

▶ SPSS文件类型

数据文件 *. sav

结果文件 *. spo

图形文件 *. cht

程序命令文件 *. sps

■将数据文件导出

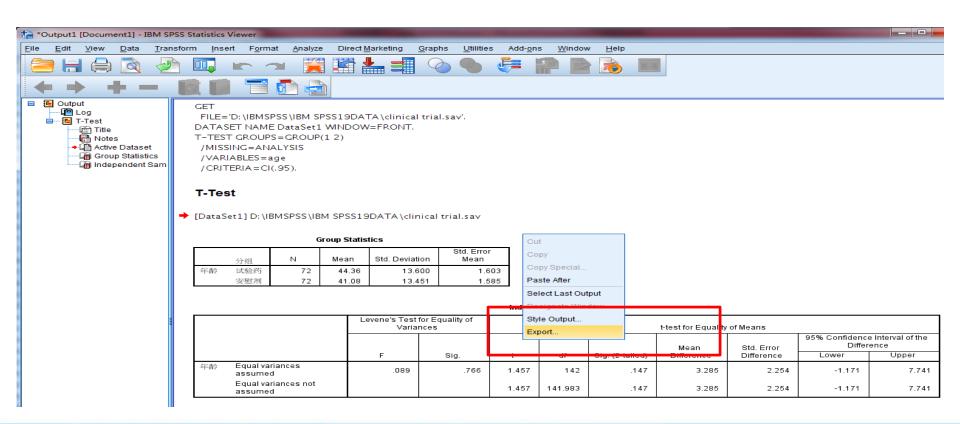
File →Save as



结果文件的导出

■将SPSS结果文件导出

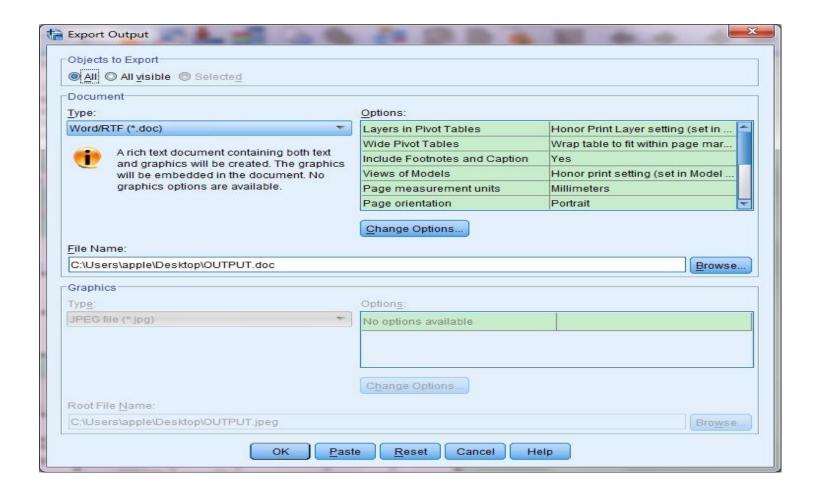
结果窗口:右键—Export…





结果文件的导出

■将SPSS结果文件导出为Word文档





第三章

数据文件的整理



主要内容

- ➤ 插入/删除变量(insert variable)/观测(insert case)
- ➤ 辨识重复观察单位(identify duplicate cases)
- ➤ 观测值排序(sort)
- ➤ 数据转置(transpose)
- ➤ 合并文件(merge files)
- ➤ 数据拆分(split files)
- ➤ 选择观察单位(select cases)
- > 变量值加权(weight cases)



1.插入/删除变量(insert variable)/观测(insert case)

- ■插入观测
- ■删除观测
- ■插入变量
- ■删除变量

1. Edit

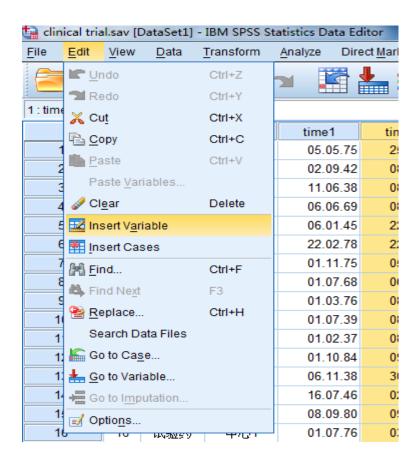
Insert variable/inset case

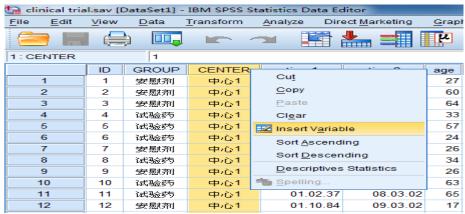
2. 选中-右键

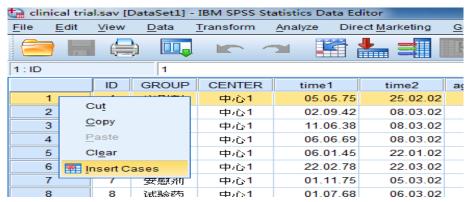
Insert variable/inset case



1. 插入/删除变量(insert variable)/观测(insert case)









2. 辨识重复观察单位(identify duplicate cases)

Data → Identify Duplicate Cases

例: 核查clinical trial. sav的数据中重

复的观测?



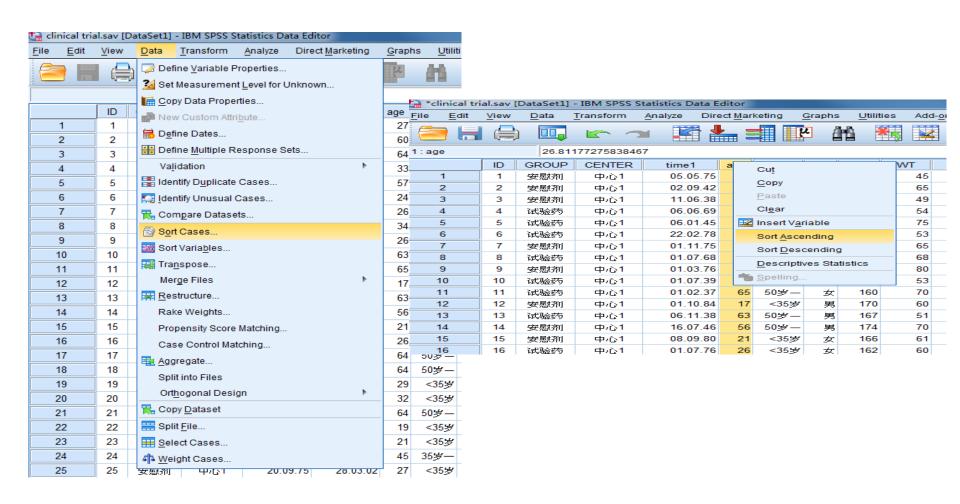
3. 观测值排序(sort)

- Data (Sort Cases) /右键 (Sort ascending /sort descending)
 - ■可对数值型和字符型变量进行排序
 - ■可同时按照多个变量进行排序
- 练习: clinical trial.sav, 练习按照年龄升序排列、

身高降序排列?



3. 观测值排序(sort)





4. 数据转置(transpose)

Data->Transpose

对原始数据的行和列进行互换(产

生新数据文件)

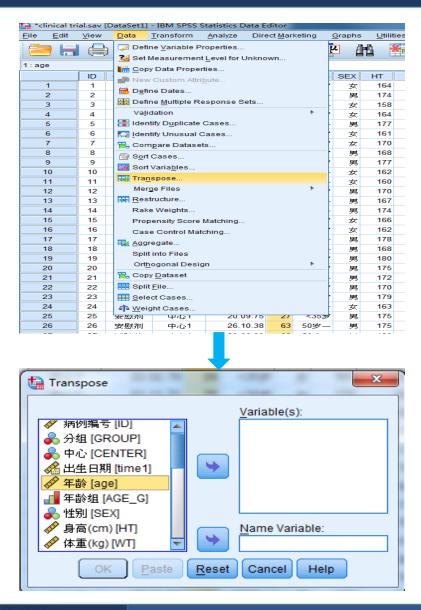
Variables: 选入要转置的变量

Name Variable:设定转置之后的

新变量名

▶ 练习: 对clinical

trial.sav进行数据转置操作





5. 合并文件(Merge Files)

>添加观察单位

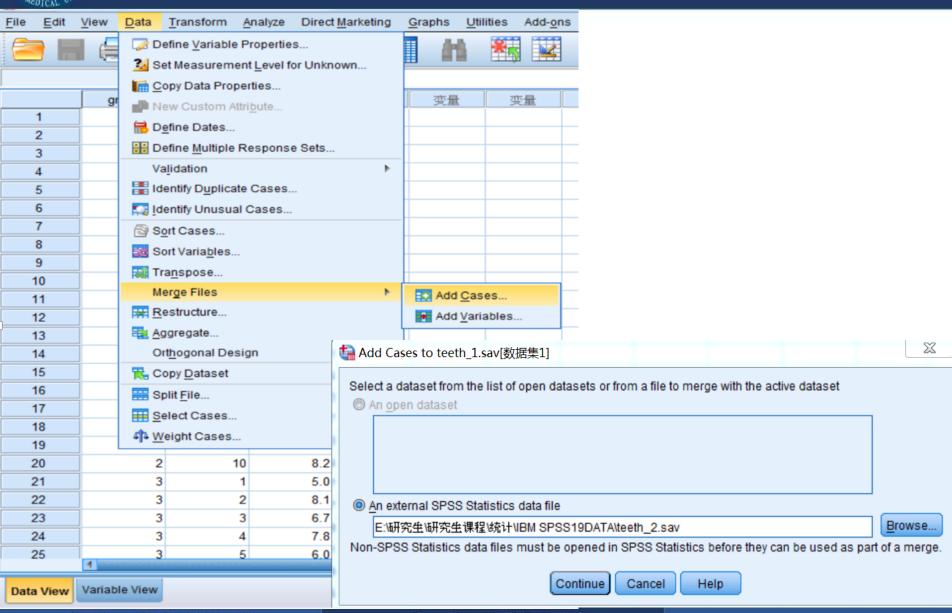
Data -> Merge Files -> Add Cases

将外部文件的观察单位添加到当前数据文件中,合并后新数据文件的观察单位数为两个文件观察单位数之和。

▶ 练习: 在teeth 1. sav基础上添加teeth 2. sav的变量



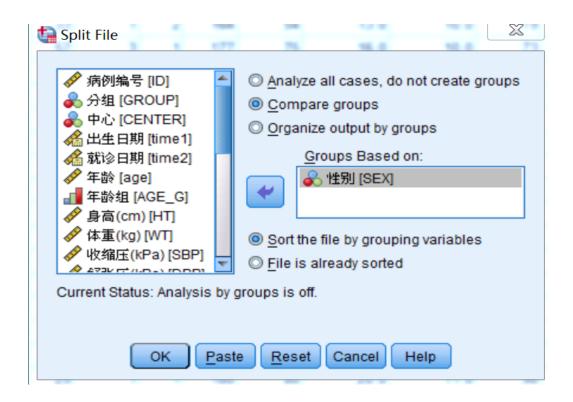
5. 合并文件(Merge Files)





6.拆分文件(split files)

- ➤ 拆分文件 (Split file): 对某些分类变量进行分层分析。避免重复分析, 通过Split File简化操作。 **Data -> split Files**
- ▶ 练习: 对clinical trial.sav的性别进行拆分.





- ➤ 选择观察单位(Select Cases)
 - ➤ If condition is satisfied 选择满足某条件的观察单位

数学运算		关系运算		逻辑运算	
+	加	<	小于	&	与
*	乘	~ =	不等于		或
**	幂	≤	小于等于	~	非



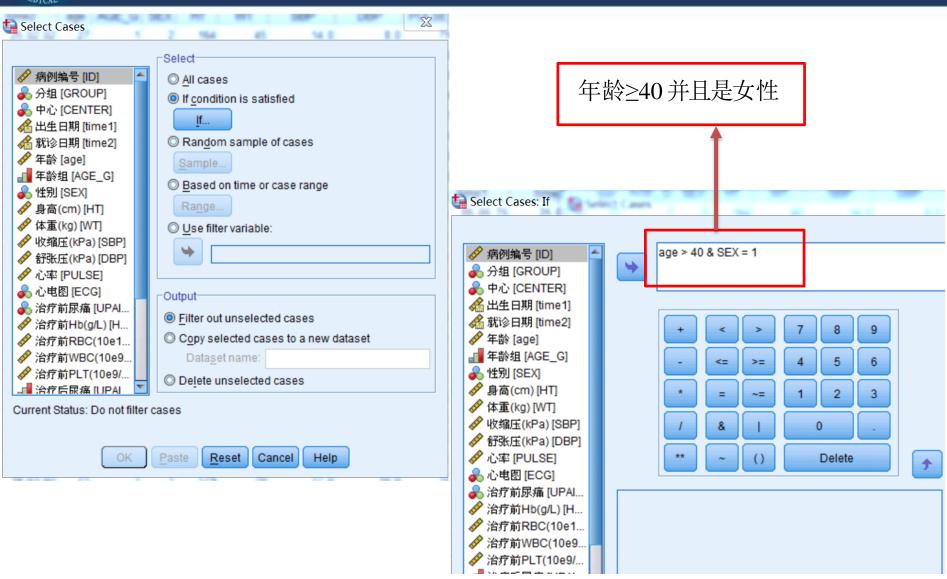
选择观察单位(Select Cases)

- All cases (所有数据)
- If condition is satisfied (满足if条件的数据)
- Random sample of cases (随机选择一定比例的数据)
- Based on time or case range(选择指定范围的数据)
- Use filter variable (选择滤过变量不为0 的数据)



- ▶ 练习:选择符合下列条件的记录
- ➤ 数据: clinical trial.sav
 - ■年龄≥40的女性
 - 收缩压大于15或舒张压大于11
 - ■随机选择70%的观察单位
 - ■从前30例中随机选择10例
 - ■选择第10-40例

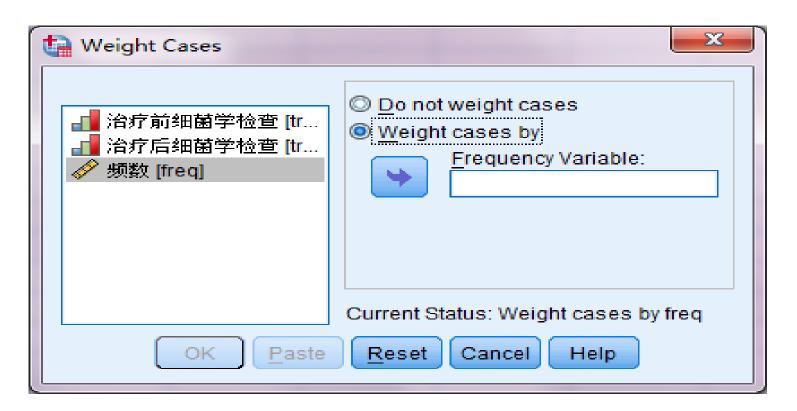






8. 变量值加权(weight cases)

■ **变量加权(weight cases**)是指以某一变量为权重对数据加权,常用于计数资料,当数据以频数形式录入时,以频数为权重,如列联表和等级资料频数表。





- ■数据的输入
 - ■不要使用输入法!

- ■变量
 - ■尽量用英文命名
 - ■最好加Label
 - ■注意变量类型 (numeric/string/date)



- ■分析方法的选择
 - ■结合生物统计学课程的知识
- ■数据格式
 - ■做什么分析,选择对应的格式



- ■正确的操作软件
 - ■操作顺序,选择那些需要的统计量
 - ■特別是需要进行拆分,加权等的时候
 - ■结果的解释
 - ■描述统计量的判读
 - 统计量、P值的判读
 - ■统计图的引用



- ■课程难点:
 - ■用什么方法?
 - ■怎么录数据?
 - ■怎么读结果?
 - ■Way Out ? 练!



Questions

&

Answers



Thank you