



第四章 数据转换(Transform)

南方医科大学生物统计学系

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

黄兴辉

1409938037@qq.com



复习1: 创建SPSS数据

➤ 为研究不同药物对肾病患者的一個治疗效果，收集了以下指标。请根据下表创建SPSS数据。

病例编号	分组	出生日期	性别	治疗前 红细胞 (10e12/L)	治疗后 红细胞 (10e12/L)
1	安慰剂	05.05.75	女	4.46	4.75
2	安慰剂	02.09.42	男	5.77	4.44
3	安慰剂	11.06.38	女	3.09	3.32
4	试验药	06.06.69	女	5.24	4.18
5	试验药	06.01.45	男	3.34	3.6
6	试验药	22.02.78	女	3.93	3.69

1 = "试验药"

2 = "安慰剂"

1 = "男"

2 = "女"



复习1：创建SPSS数据

- 删除出生日期这个变量，添加年龄这个指标；
- 删除ID为3的个体，添加ID为9的记录信息；
- 保存前3个变量到桌面并命名为“kidney.sav”。

病例编号	分组	性别	治疗前 红细胞 (10e12/ L)	治疗后 红细胞 (10e12/ L)	年龄
1	安慰剂	女	4.46	4.75	27
2	安慰剂	男	5.77	4.44	60
4	试验药	女	5.24	4.18	33
5	试验药	男	3.34	3.6	57
6	试验药	女	3.93	3.69	24
9	安慰剂	男	5.31	9.9	26



复习2: 数据文件的整理

打开文件” clinical trial.sav”

- ① 选择年龄 ≥ 40 的男性
- ② 对clinical trial.sav的分组(GROUP)和性别(SEX)进行分层, 分别求出不同分组不同性别的身高和体重的描述性统计量

描述统计量的运算过程:

Analyze--Descriptive statistic--Descriptives--Options



为什么做数据转换

- 在许多情况下，原始数据难以满足数据分析的全部要求，需要将原始数据进行适当的转换。

$$\text{体重指数 (BMI)} = \text{体重 (kg)} \div \text{身高 (m)}^2$$

- SPSS具有强大的数据转换功能
- 可以进行简单的变量变换和重新建立分类变量
- 可以进行复杂的统计函数运算以及逻辑函数运算



主要内容

- 计算产生变量(Compute)
- 计数统计(Count Values within Cases)
- 重新赋值(Recode)
- 自动重新赋值(Automatic Recode)
- 观察单位排秩(Rank Cases)
- 可视分类器(Visual Binning)
- 缺失值替代(Replace Missing Values)



计算产生变量(Compute)

- 功能：根据已存在的变量，经函数计算后，建立新变量或替换原变量。
- Compute中包含了各种函数，如：算术函数、统计函数、缺失值函数、字符型函数、逻辑函数、连续型变量的反分布函数、离散型随机变量函数等多种类型函数。



计算产生变量(Compute)

1. 函数计算

——BMI的计算

2. 时间函数

——年龄的计算

3. 随机数字的设立

——实验分组

4. 累积分布&非中心分布的累计分布函数

——计算 P 值和统计量



计算产生变量(Compute)

1. 函数计算

——BMI的计算

Transform

Compute Variable

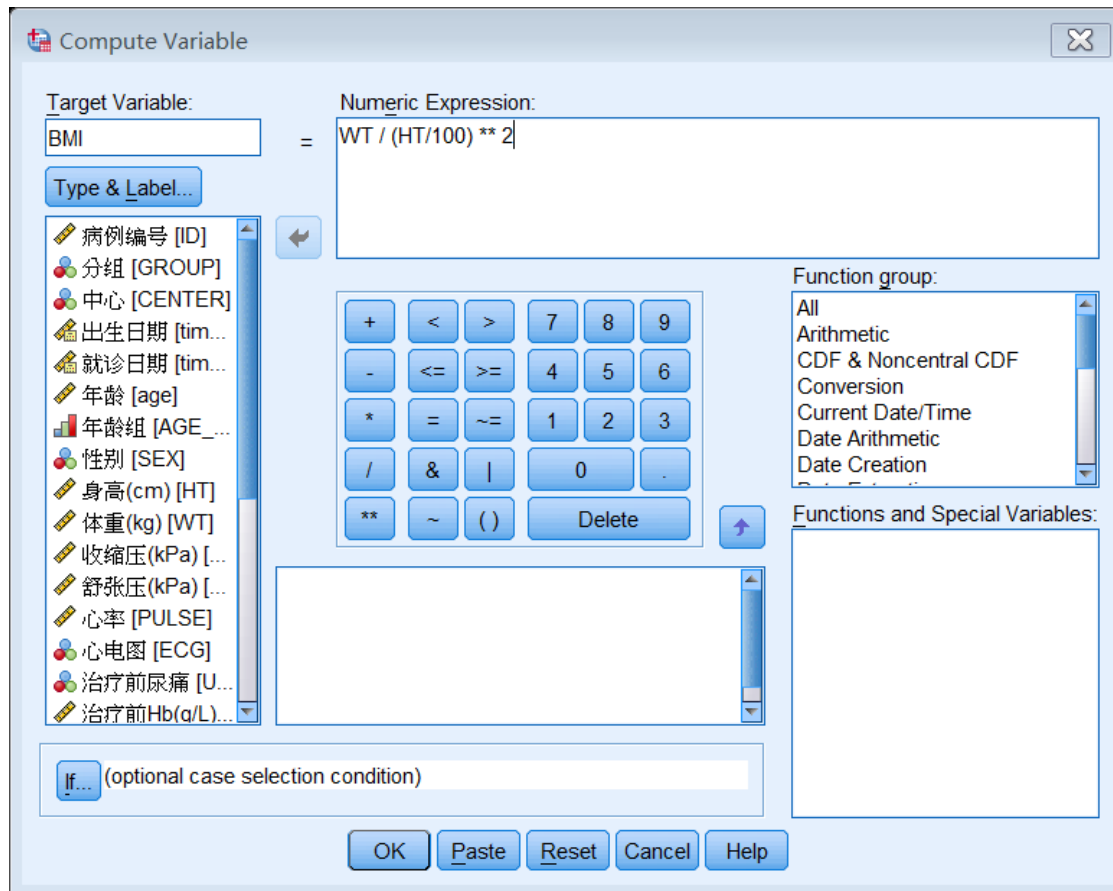
直接输入公式

保存lable(公式)



计算产生变量(Compute)

- 实例操作：打开文件” clinical trial.sav”，计算BMI。
- 操作步骤：**Transform**→**Compute Variable**





计算产生变量(Compute)

2. 时间函数

——年龄的计算

Transform

Compute Variable

函数族（日期运算Data Arithmetic P385）

函数DATEDIFF(参考说明进行)

保存lable(公式)



计算产生变量(Compute)

Compute Variable

Target Variable: = Numeric Expression:

Type & Label...

Function group:

- All
- Arithmetic
- CDF & Noncentral CDF
- Conversion
- Current Date/Time
- Date Arithmetic
- Date Creation

Functions and Special Variables:

- Datediff
- Datesum(3)
- Datesum(4)

Calculator:

Operators: +, -, *, /, **, ~, &, |, <, <=, >, >=, =, ~=

Numbers: 0-9, .

Functions: DATEDIFF(datetime2, datetime1, "unit"). Numeric. Calculates the difference between two date/time values and returns an integer (with any fraction component truncated) in the specified date/time units, where datetime2 and datetime1 are both date or time format variables (or numeric values that represent valid date/time values), and "unit" is one of the following string literal

If... (optional case selection condition)

OK Paste Reset Cancel Help



计算产生变量(Compute)

3. 随机数字的设立

——实验分组（随机分组，均衡分组）

种子数&随机化

——设定种子数为12345，产生10个服从均匀分布
(1,100) 的随机数，前5个小的为A组，剩下的为B组

说明：设立随机数字（按照自己需要，不一定是12345），
确保每次都一样，重现分组过程，保留分组证据。随机
是对患者来说，不是对随机数来说，所以种子数虽然是
固定的，对于患者来说还是随机分组的



计算产生变量(Compute)

3. 随机数字的设立

给每一个患者分配唯一ID号，建立ID号

Transform

Random Number Generators(种子数12345)

Compute Variable

函数族 (Random Numbers P389)

函数RV.UNIFORM(参考说明进行)

保存lable(公式)，产生数据

排序sort cases,按照规则分组

按照ID号再给患者分组



计算产生变量(Compute)

Random Number Generators

Active Generator

☐ Set Active Generator

☒ SPSS 12 Compatible

☐ Mersenne Twister

Active Generator Initialization


☒ Set Starting Point

☐ Random

☒ Fixed Value

Value: 12345

Current Active Generator: SPSS 12 Compatible

 The active generator setting applies immediately and to future sessions.

Some procedures have internal random number generators. See the help for a complete list.

OK Paste Reset Cancel Help



计算产生变量(Compute)

Compute Variable

Target Variable: = Numeric Expression:

Type & Label...

ID

Function group: Random Numbers

Functions and Special Variables: Rv.Uniform

RV.UNIFORM(min, max). Numeric. Returns a random value from a uniform distribution with specified minimum and maximum. See also the UNIFORM function.

If... (optional case selection condition)

OK Paste Reset Cancel Help



计算产生变量(Compute)

ID	number	group
7.00	8.71	A
5.00	19.20	A
4.00	28.35	A
1.00	36.93	A
9.00	71.02	A
6.00	73.16	B
8.00	73.70	B
2.00	74.77	B
10.00	76.64	B
3.00	83.27	B



计算产生变量(Compute)

4. 累积分布&非中心分布的累计分布函数 ——计算 P 值和统计量

对于自由度20的 t 分布，请问 $t=2.086$ 所对应的双侧尾部面积 P 为多少？



计算产生变量(Compute)

4. 累积分布&非中心分布的累计分布函数 ——计算P值和统计量

Transform

Compute Variable

函数族 (CDF&NoncentralCDF P384)

函数CDF.T(2.086,20)

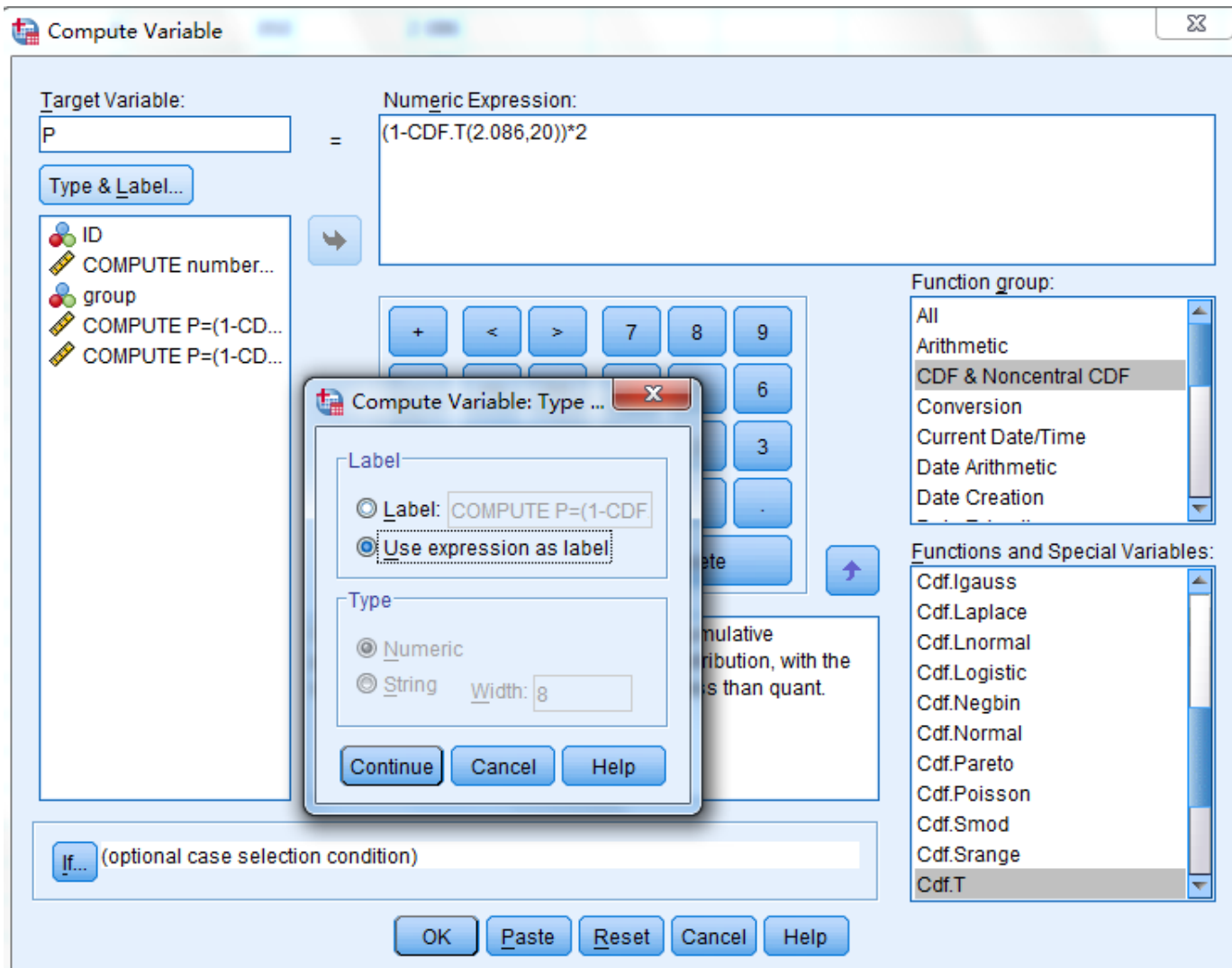
双侧尾部面积 $2*(1-\text{CDF.T}(2.086,20))$

保存lable(公式)

结果为0.050

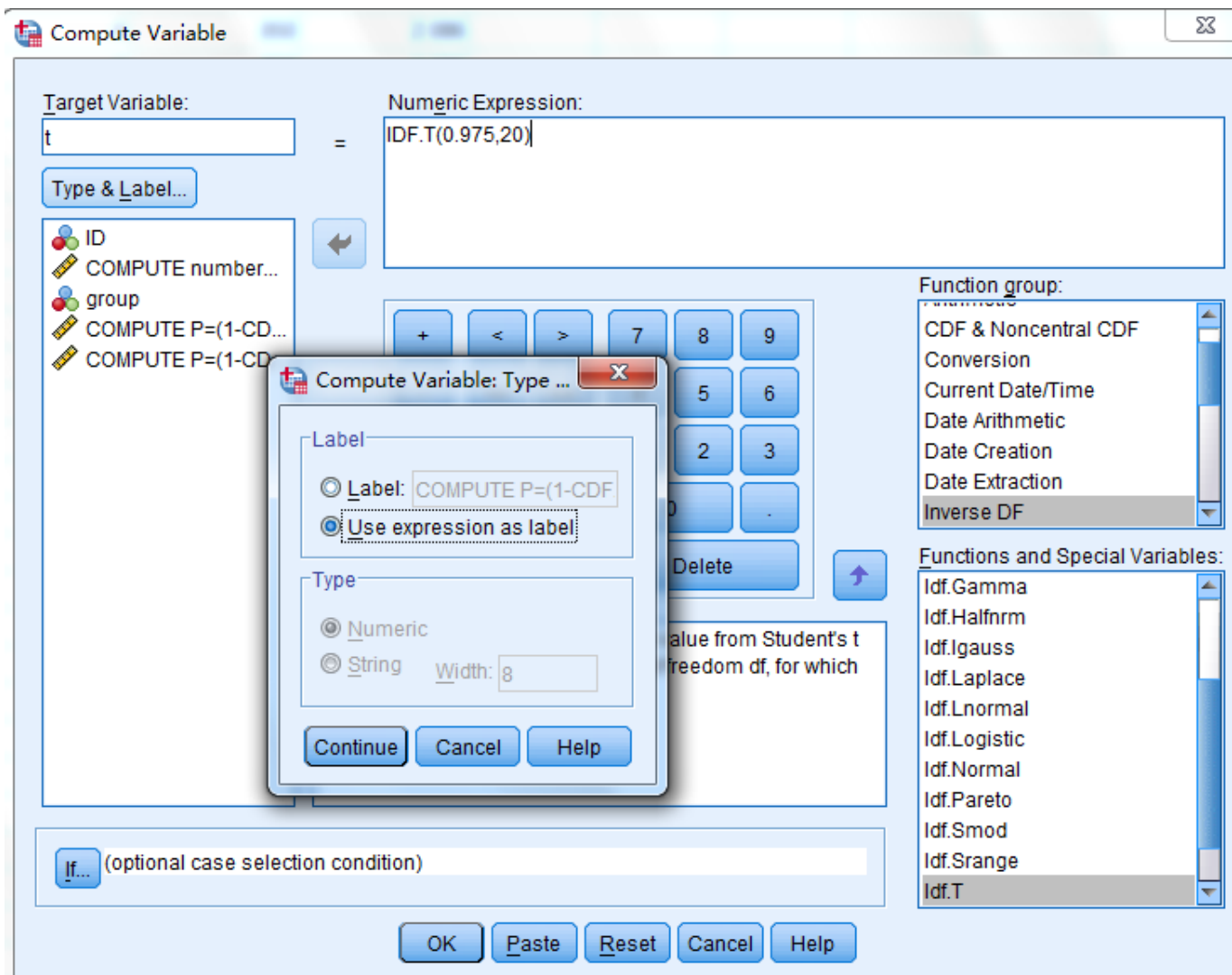


计算产生变量(Compute)





计算产生变量(Compute)





计数统计 (Count Values within Cases)

思考：黄老师录入了某班同学的医学统计学考试的**20**道选择题的答案。这里选择题一共有**A、B、C、D**四个选项。黄老师想统计一下**每个同学**在**20**道选择题里面，**A、B、C、D**分别选了多少个。他应该怎样做？



计数统计 (Count Values within Cases)

- ▶ 功能：可产生新的变量，记录**各观察单位**中某一（些）观察值出现的频数（个数）。

- ▶ 操作：从菜单中选择

Transform

Count Values within Cases



计数统计 (Count Values within Cases)

实例操作：数据文件“**IPSQ.sav**”是关于**20**个住院病人的满意度调查结果。调查表共有**39**个问题，即文件中的**39**个变量（含字母**acc, c, d, f, h, n, o, s**的变量名）。每个问题有**5**个可供选择的答案，即“很不满意，不太满意，一般，比较满意，很满意”，相应的赋值从**1**到**5**分。试统计每个调查对象的赋值分从**1**到**5**的分布。

步骤：**Transform/Count values within cases**



计数统计 (Count Values within Cases)

Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

Count Occurrences of Values within Cases

Target Variable: count1

Target Label: score=1

Numeric Variables: c4, acc4, h5, h1, acc1

Define Values... If... (optional case selected)

Count Values within Cases: Values to Count

Value: ☐ Value: 1 ☐ System-missing ☐ System- or user-missing ☐ Range:

through: Add Change Remove

☐ Range, LOWEST through value: ☐ Range, value through HIGHEST:

Continue Cancel Help

	var	var	var	var	var
4	3	2.00			
4	3	2.00			
4	3	1.00			
4	5	5.00			
5	3	1.00			
5	3	2.00			
4	5	2.00			
2	.	5.00			



重新赋值(Recode)

思考：某同学用**1、2、3、4**分别表示血型**A、B、AB、O**。
可是当他录入数据的时候，却用**1、2、3、4**分别表示血型**O、AB、B、A**，他应该如何把数据改过来？



重新赋值(Recode)

- 功能：根据已有的变量，按照一定的规则，更换原变量赋值或者产生新的变量。

应用：

- 修改原变量的赋值(Recode into Same Variables)
- 产生需要的新变量(Recode into Different Variables)



重新赋值(Recode)

- ▶ 实例操作：将数据文件“IPSQ.SAV”中变量“n4”的观察值“1,2,3,4,5”，相应的重新赋值为“5,4,3,2,1”，即用新赋值替代原观察值，不更改变量名。
- ▶ 步骤：**Transform/Recode into Same Variables**

IPSQ.sav [DataSet1] - IBM SPSS Statistics Data Editor

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

Compute Variable...
 Count Values within Cases...
 Shift Values...
Recode into Same Variables...
 Recode into Different Variables...
 Automatic Recode...
 Visual Binning...
 Optimal Binning...
 Prepare Data for Modeling
 Rank Cases...
 Date and Time Wizard...
 Create Time Series...
 Replace Missing Values...
 Random Number Generators...
 Run Pending Transforms Ctrl+G

	ID		h1	acc1
1	1		4	5
2	8		4	3
3	21		3	1
4	24		4	3
5	37		4	3
6	48		3	3
7	60		4	3
8	67		4	3
9	85		5	1
10	93		3	1
11	98		5	4
12	107		3	4
13	109		5	3
14	113	4 5 1	1	3
15	114	2 5 .	5	3



4							
---	--	--	--	--	--	--	--

Recode into Same Variables	Σ	d6	h4	d3	s4	s1	f3	f1
		4	4	4	3	3	1	1
		5	3	5	4	3	1	2

Recode into Same Variables

Numeric Variables:

ID
c4
acc4
h5
h1
acc1
c3
d6
h4
d3

Old and New Values

If... (optional case)

OK Paste Reset Cancel

Recode into Same Variables: Old and New Values

Old Value

☒ Value: 5

☐ System-missing

☐ System- or user-missing

☐ Range:

through

☐ Range, LOWEST through value:

☐ Range, value through HIGHEST:

☐ All other values

New Value

☒ Value: 1

☐ System-missing

Old -> New:

1 -> 5
2 -> 4
3 -> 3
4 -> 2

Add
Change
Remove

Continue Cancel Help

5	5	5	3
5	1	1	3
5	.	5	3
5	3	5	1
5	.	4	2
5	.	.	1
5	4	4	3
5	.	.	1



重新赋值(Recode)

实例：

- ▶ 血型数据：把血型变量的1、2、3、4（分别表示O、AB、B、A血型）改为4、3、2、1（赋值到相同变量）。
- ▶ 年龄数据：根据“age”变量，建立变量名为“age_g”的年龄组新变量。其对应规则为， $\text{age} \leq 35$ ， $\text{age_g} = 1$ ； $36 \leq \text{age} \leq 50$ ， $\text{age_g} = 2$ ； $\text{age} > 50$ ， $\text{age_g} = 3$ （赋值到不同变量）。



重新赋值(Recode)

Recode into Different Variables

Numeric Variable -> Output Variable:

age --> age_g

Output Variable

Name: age_g

Label: 年龄组

Change

病人编号 [ID]
姓名 [name]
分组 [group]
出生日期 [time1]
就诊日期 [time2]
年龄组 [age_g]
性别 [gender]
身高(cm) [height]
体重(kg) [wt]
收缩压(kPa) [SBP]
舒张压(kPa) [DBP]
心率 [pulse]
心电图 [ECG]
治疗前尿痛 [upai...]

wbc1	plt1	upain2	hb2	rbc2
9.41	242	轻度	134	4.1
9.60	213	无痛	138	4.1
4.81	170	无痛	126	4.1
4.50	262	无痛	117	3.1
3.90	177	中度	113	3.1

Recode into Different Variables: Old and New Values

Old Value

☒ Value:

☐ System-missing

☐ System- or user-missing

☐ Range:

☐ Range, LOWEST through value:

☒ Range, value through HIGHEST:

51

☐ All other values

New Value

☒ Value: 3

☐ System-missing

☐ Copy old value(s)

Old --> New:

Lowest thru 35 --> 1

36 thru 50 --> 2

Add

Change

Remove

☐ Output variables are strings Width: 8

☒ Convert numeric strings to numbers ('5'-->5)

Continue Cancel Help

35岁—	男	178
35岁—	男	172
35岁—	男	178
18岁—	女	162
1—65岁	女	165
18岁—	女	167
18岁—	男	168
35岁—	女	159
18岁—	男	177
1—65岁	女	160



自动重新赋值(Automatic Recode)

思考：某班同学的**SPSS**考试成绩已经按学号输入到一个**SPSS**文件中。现在，黄老师想知道每个同学成绩的排名情况，但是又想保持成绩是按学号的顺序排列的。他应该怎样做？



自动重新赋值(Automatic Recode)

功能：将数值变量或字符变量值转换为从1开始的顺序整数，并存为新变量。

实例操作：数据文件“**clinical trial. sav**”中根据年龄变量“**age**”进行排秩，产生新变量**age_rank**。

步骤：**Transform/ Automatic Recode**



	ID	GROUP	CENT
13	13	试验药	中心
14	14	安慰剂	中心
15	15	安慰剂	中心
16	16	试验药	中心
17	17	试验药	中心
18	18	安慰剂	中心
19	19	安慰剂	中心
20	20	试验药	中心
21	21	试验药	中心
22	22	试验药	中心
23	23	安慰剂	中心
24	24	安慰剂	中心
25	25	安慰剂	中心
26	26	安慰剂	中心
27	27	试验药	中心
28	28	安慰剂	中心
29	29	试验药	中心
30	30	试验药	中心
31	31	试验药	中心
32	32	安慰剂	中心
33	33	安慰剂	中心

Automatic Recode

Variable-> New Name

age-->age_rank

New Name:

Recode Starting from

☒ Lowest value ☐ Highest value

☐ Use the same recoding scheme for all variables

☐ Treat blank string values as user-missing

Template

☐ Apply template from:

☐ Save template as:

SBP
22.0
17.0
14.4
13.0
21.0
18.5
20.0
18.0
21.0
17.0
17.0
16.0
15.0
16.5
20.0
16.0
15.0
15.0
19.0
20.0
20.0



可视分类器(Visual Binning)

- ▶ 功能：将连续型变量按一定规则转换为分类变量
- ▶ 例：数据文件“clinical trial. sav”中根据年龄变量“age”建立变量名为“age_g2”的年龄组新变量，划分规则为： ≤ 35 岁的为1；35~50岁的为2； > 50 岁的为3。
- ▶ 步骤：Transform/ Visual Binning
 - ▶ 在variables to bin 中选入age
 - ▶ 手动定义value的分界值

可视分类器(Visual Binning)

et3] - IBM SPSS Statistics Data Editor

Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

name	group	time1	time2	age	age
XHJI	安慰剂	02.09.42	08.03.02	60	50-6
BQXI	试验药	06.06.69	08.03.02	33	18
ZSZH	安慰剂	22.02.78	22.03.02	24	18
HJZH	试验药	01.07.68	06.03.02	34	18
JFXI	安慰剂	01.03.76	08.03.02	26	18
LIYI	试验药	01.02.37	08.03.02	65	50-6
LSWE	安慰剂	01.10.84	09.03.02	18	18
WYTI	试验药	06.11.38	30.03.02	64	50-6
WYFE	试验药	08.09.38	30.03.02	64	50-6
ZWJI	安慰剂	01.03.73	09.03.02	29	18
ZFQI	试验药	01.04.70	13.03.02	32	18
ZHZH	试验药	01.05.38	14.03.02	64	50-6
ZXBE	安慰剂	01.12.56	20.03.02	46	35
ZYLI	试验药	20.06.39	22.03.02	63	50-6
ZXHU	试验药	25.06.38	30.03.02	64	50-6
ZQHO	安慰剂	07.11.67	29.03.02	35	35
JMGU	试验药	01.01.39	23.03.02	63	50-65岁

Visual Binning

Select the variables whose values will be grouped into bins. Data will be scanned when you click Continue.

The Variables list below contains all numeric ordinal and scale variables.

Variables:

- 分组 [group]
- 出生日期 [time1]
- 就诊日期 [time2]
- 年龄组 [age_g]
- 性别 [gender]
- 身高(cm) [height]
- 体重(kg) [wt]
- 收缩压(kPa) [SBP]
- 舒张压(kPa) [DBP]
- 心率 [pulse]
- 心电图 [ECG]

Variables to Bin:

- 年龄 [age]

☐ Limit number of cases scanned to:

Continue Cancel Help

可视分类器(Visual Binning)

Make Cutpoints

☒ **Equal Width Intervals**

Intervals - fill in at least two fields

First Cutpoint Location: 35

Number of Cutpoints: 2

Width: 15

Last Cutpoint Location: 50

☐ **Equal Percentiles Based on Scanned Cases**

Intervals - fill in either field

Number of Cutpoints:

Width(%):

☒ **Cutpoints at Mean and Selected Standard Deviations Based on Scanned Cases**

☐ +/- 1 Std. Deviation

☐ +/- 2 Std. Deviation

☐ +/- 3 Std. Deviation

i Apply will replace the current cutpoint definitions with this specification. A final interval will include all remaining values: N cutpoints produce N+1 intervals.

Apply Cancel Help

None	4	Right	Scale
None	6	Right	Scale

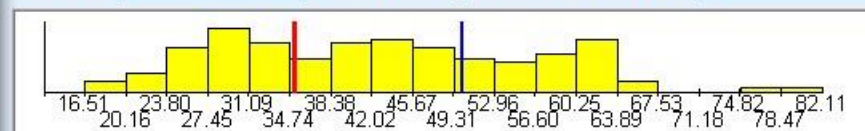
Visual Binning

Name: Label:

Current Variable: age 年龄

Binned Variable: 年龄 (Binned)

Minimum: 17 Nonmissing Values Maximum: 78



Enter interval cutpoints or click Make Cutpoints for automatic intervals. A cutpoint value of 10, for example, defines an interval starting above the previous interval and ending at 10.

Grid:

	Value	Label
1	35.0	<= 35
2	50.0	36 - 50
3		HIGH 51+
4		

Upper Endpoints

☒ Included (<=)

☐ Excluded (<)

Make Cutpoints...

Make Labels

☐ Reverse scale

OK Paste Reset Cancel Help


可视分类器(Visual Binning)

Make Cutpoints

☒ **Equal Width Intervals**
 Intervals - fill in at least two fields:
 First Cutpoint Location:
 Number of Cutpoints:
 Width:
 Last Cutpoint Location:

☐ **Equal Percentiles Based on Scanned Cases**
 Intervals - fill in either field:
 Number of Cutpoints:
 Width(%):

☐ **Cutpoints at Mean and Selected Standard Deviations Based on Scanned Cases**
☐ +/- 1 Std. Deviation
☐ +/- 2 Std. Deviation
☐ +/- 3 Std. Deviation

 Apply will replace the current cutpoint definitions with this specification. A final interval will include all remaining values: N cutpoints produce N+1 intervals.

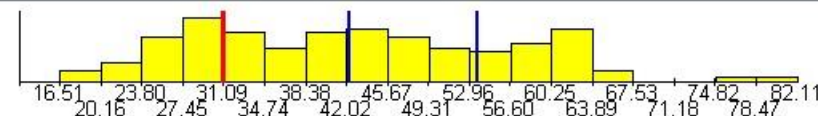
Visual Binning

Name: Label:

Current Variable:

Binned Variable:

Minimum: Nonmissing Values Maximum:



Enter interval cutpoints or click Make Cutpoints for automatic intervals. A cutpoint value of 10, for example, defines an interval starting above the previous interval and ending at 10.

Grid:

	Value	Label
1	30.8	<= 31
2	42.0	32 - 42
3	53.5	43 - 53
4		HIGH 54+
5		

Upper Endpoints
☒ Included (<=)
☐ Excluded (<)

☐ Reverse scale

可视分类器(Visual Binning)

Make Cutpoints

☒ Equal Width Intervals

Intervals - fill in at least two fields

First Cutpoint Location:

Number of Cutpoints:

Width:

Last Cutpoint Location:

☐ Equal Percentiles Based on Scanned Cases

Intervals - fill in either field

Number of Cutpoints:

Width(%):

☒ Cutpoints at Mean and Selected Standard Deviations Based on Scanned Cases

☒ +/- 1 Std. Deviation

☐ +/- 2 Std. Deviation

☐ +/- 3 Std. Deviation

i Apply will replace the current cutpoint definitions with this specification.
A final interval will include all remaining values: N cutpoints produce N+1 intervals.

Apply Cancel Help

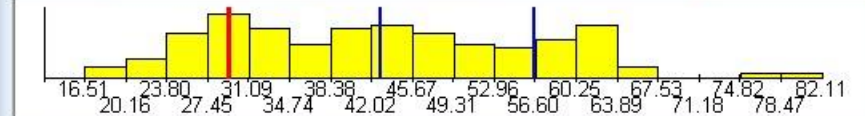
Numeric	8	0	治疗已显痛	无痛}...
None	5	Right	Ordinal	
None	4	Right	Scale	
None	6	Right	Scale	

Name: Label:

Current Variable:

Binned Variable:

Minimum: Nonmissing Values Maximum:



Enter interval cutpoints or click Make Cutpoints for automatic intervals. A cutpoint value of 10, for example, defines an interval starting above the previous interval and ending at 10.

Grid:

	Value	Label
1	29.1	<= 29
2	42.7	30 - 43
3	56.3	44 - 56
4		HIGH 57+
5		

Upper Endpoints

☒ Included (<=)

☐ Excluded (<)

Make Cutpoints...

Make Labels

☐ Reverse scale

OK Paste Reset Cancel Help



观察单位排秩(Rank Cases)

思考：现有一组病人的血压数据，需要获得每位病人血压值的秩次，应该怎样使用**SPSS**进行处理？



观察单位排秩(Rank Cases)

- ▶ 功能：根据某变量观测值的大小，按一定顺序排秩，生成一个代表其秩次的新变量，数据**本身顺序不改变**。另外，对于相同值，系统默认条件下，会得到**平均秩**。
- ▶ 与Sort Cases比较
- ▶ Sort Cases是根据某变量观测值的大小将数据重新排序（**本身顺序改变**），并不生成新变量。
- ▶ 操作：从菜单中选择

Transform

Rank Cases



观察单位排秩(Rank Cases)

实例操作：数据文件“**clinical trial. sav**”中根据年龄变量“**age**”进行排秩。

步骤：**Transform/ Rank Cases**



	ID	GROUP									BP	DBP
13	13	试验药									22.0	8.0
14	14	安慰剂									17.0	11.0
15	15	安慰剂									14.4	8.0
16	16	试验药									13.0	9.0
17	17	试验药									21.0	9.0
18	18	安慰剂									18.5	12.0
19	19	安慰剂									20.0	11.0
20	20	试验药									18.0	10.0
21	21	试验药									21.0	12.0
22	22	试验药									17.0	10.0
23	23	安慰剂									17.0	10.0
24	24	安慰剂									16.0	10.0
25	25	安慰剂									15.0	9.0
26	26	安慰剂	中心1	20.06.39	22.03.02	63	50岁—	男	173	70	16.5	10.5
27	27	试验药	中心1	20.06.39	22.03.02	63	50岁—	男	180	80	20.0	11.0
28	28	安慰剂	中心1	06.05.58	29.03.02	44	35岁—	男	173	68	16.0	9.0
29	29	试验药	中心1	25.06.38	30.03.02	64	50岁—	男	173	64	15.0	8.9
30	30	试验药	中心1	07.11.67	29.03.02	34	<35岁	女	166	65	15.0	9.0

Rank Cases

Variable(s): 年龄 [age]

By:

☒ Display summary tables

Assign Rank 1 to

☒ Smallest value

☐ Largest value

OK Paste Reset Cancel Help

Rank Types... Ties...



观察单位排秩(Rank Cases)

与Sort Cases比较

Sort Cases是根据某变量观测值的大小将数据重新排序（**本身顺序改变**），并不生成新变量。



观察单位排秩(Rank Cases)

Rank Cases与自动重新赋值(Automatic Recode)的主要区别：

- ① Rank Cases不可用于字符型变量，而Automatic Recode可以。
- ② Rank Cases在默认选项下，遇相同数求平均秩，Automatic Recode不求平均秩。



缺失值替代(Replace Missing Values)

思考：某位同学用60只小白鼠进行了一项实验，但是由于操作不当，有4只小白鼠的体重还没有进行测量就死亡了。除了这4只小白鼠的体重数据之外，其他指标值都有记录。在进行统计分析的时候，有一项分析对缺失值不纳入分析，而该同学又不想直接删除4只小白鼠的数据。他应该怎么办？



缺失值替代(Replace Missing Values)

- ▶ 功能：可根据研究目的和数据分布特征，选用不同的处理方法估计并替代缺失值。
- ▶ 操作步骤：从菜单中选择

Transform

Replace Missing Values

- ▶ **(不建议)**



练习题

1. 已知 $x=3$,求 $\sin(x)+\cos(x)+\exp(x)=$ ____（保留三位小数）。
2. 已知某人于1990年9月15日出生，请计算到2016年11月28日他来到这个世上有____天。
3. 已知 X 服从均值为5标准差为3的正态分布: $X \sim N(5, 3^2)$, 请问 $X > 10.88$ 的概率为____（保留三位小数）。



Questions

&

Answers



Thank you