

# 市场调查与分析

实验课——市场调查资料的整理

谢光明

成都信息工程大学物流学院



## 内容

- 1、编码系统的建立与应用
- 2、spss的基本操作
- 3、数据变量详解
- 4、数据导入和结果导出



# 1、编码系统的建立与应用



- ❖ 编码的概念：就是将市场调查回来的信息转化为统一设计的计算机可识别的代码，以便对其进行数据整理和分析。
- ❖ 编码的作用：将市场调查回来的信息转化为标准化的数据库，便于分析；建设录入和统计分析的工作量；降低误差。



## 编码的原则

- ❖ 准确性原则：代码能准确的代表原有信息
- ❖ 完整性原则：信息转化时尽量不要丢失信息，减少信息的损失和浪费；
- ❖ 效率性原则：易于操作
- 标准性原则：一般一个代码只代表一个数据，避免混淆和误解；
- ❖ 兼容性原则：便于与其他的数据分析软件接轨。



## ❖ 原始问卷

1、请问您购买海尔洗衣机的原因是（多选）（     ）

A:价格， C功能， D服务

2、请问您购买海尔洗衣机的原因是（可以多选，并按重要性排序）

（     ）

A:价格， C功能， D服务



## ❖ 编码表

原始题项	变量名称	变量标记	取值范围	答案编码
1	A1-1	价格	1-2	1勾选 2不勾选
	A1-2	功能	1-2	1勾选 2不勾选
	A1-3	服务	1-2	1勾选 2不勾选
2	A2-1	价格	1-3	1排序1, 2排序2, 3排序3
	A2-2	功能	1-3	1排序1, 2排序2, 3排序3
	A2-3	服务	1-3	1排序1, 2排序2, 3排序3



## 特殊编码（**开放式问题**）

- 1、将答案进行归类，然后逐一编码（答案分类简洁时）；
- 2、将答案进行归类，然后转化为几个维度，逐一编码（答案存在‘等等’）。



## ❖ 原始问卷:

1、你对改善小区生活环境的建议是:

---

答案可能是:

种树; 修草坪; 种花; 建儿童乐园; 保持清洁; 安装跷跷板等等。

## ❖ 编码表

原始题项	变量名称	变量标记	取值范围	答案编码
1	A1	环境建议	1-3	1绿化; 2儿童设施; 3小区清洁





# 课堂练习2



原始题项	变量名称	变量标记	取值范围	答案编码

您好！我是某公司的访问员，我们正在进行一项有关手机购买方面的调研，此问卷为匿名，请放心填写，感谢您的合作！↵

1、请问您最近三个月内是否有购买新手机的打算？【单选】↵

A:想 B 不想【跳答 11 题】↵

2、您喜欢的手机颜色有哪些？【至多选三项】↵

A 紫色 B 黄色 C 绿色 D 橙色 E 混合色 F 其他↵

3、在影响您选购手机的下列因素中，您认为最重要的前三位因素依次是第一重要（ ），第二重要（ ）第三重要（ ）↵

A 价格 B 功能 C 品牌 D 外观 E 性价比 F 其他↵

4、除了打电话、发短信以外，您还会经常用手机的什么功能？【多选】↵

A 音乐 B 上网 C 拍照 D 游戏 E 应用性 F 其他↵

5、请问您对手机 GPS 定位功能的需求程度如何？【单选】↵

A 非常需要 B 不太需要 C 一般 D 比较需要 E 非常需要↵

6、你愿意为您的手机支付多少钱？【单选】↵

A500 元以下 B501-1000 元 C1001-2000 元 D2001-3000 元 E3001-4000 元 F4000 以上↵

7、您的性别\_\_\_\_\_↵

8、您的年龄\_\_\_\_\_↵

9、您的专业\_\_\_\_\_↵

10、您的年级\_\_\_\_\_↵



## 课堂练习3——后编码



除了具体问卷调查前的编码，有时也可以实际进行问卷调查后，根据结果信息进行编码，称为后编码。

练习30分钟，依据问卷调查表、问卷调查原始数据，将其转换为问卷编码状态下的数据结果。



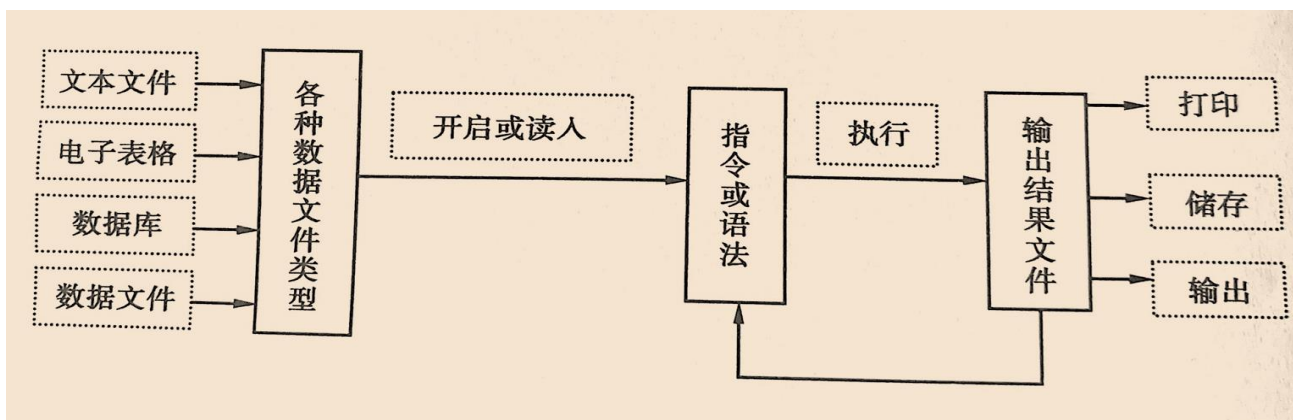
## 2、spss基本操作



- ❖ SPSS是“社会科学统计软件包”（Statistical Package for the Social Science）的简称，是一种集成化的计算机数据处理应用软件。1968年，美国斯坦福大学三位大学生开发了最早的SPSS统计软件，并于1975年在芝加哥成立了SPSS公司，已有30余年的成长历史，全球约有25万家产品用户，广泛分布于通讯、医疗、银行、证券、保险、制造、商业、市场研究、科研、教育等多个领域和行业。
- ❖ SPSS是世界上公认的三大数据分析软件之一（SAS、SPSS和STATA）

## 2、spss基本操作

- ❖ 视窗界面的 SPSS 软件，不像早期 PC 的 DOS 系统，要撰写语法程序，才能输出统计结果，视窗界面的改良，使研究者的操作只需开启数据文件，以点选鼠标为主，辅以键盘输入，即可进行统计分析，而其操作过程，也可全部转换为程序语法文件，加以储存，以便日后进行编辑或执行相类似的统计分析。
- ❖ 数据文件既可以以传统文本处理的方式创建，也可以数据库或电子表格方式创建，视窗界面的 SPSS 软件均能读取，目前以 EXCEL 的数据文件最为方便与普遍。
- ❖ SPSS 对于数据处理的流程如图所示：



本门课 SPSS 操作示例，将以 SPSS 20.0 版本为例



## 2、spss基本操作——步骤

- ① **把数据文件读入 SPSS:** 可以开启先前存储的 SPSS 数据文件（拓展名为.sav）；**可读取 Excel电子表格**；数据库建立的各种 SQL 格式文件（扩展名为.dat）；一般的文本文件（拓展名为.txt）；**直接在[数据编辑窗口]中输入数据。**
- ② **选取菜单程序:** 选取菜单栏的程序操作，如**文件**：新建、保存、打开等；**数据**：变量进行运算，选择个案等；**转化**：对变量进行重新编码等；**分析**：各种统计分析；**图形**：统计绘图功能；帮助等。
- ③ **选取分析的变量:** 数据文件中的所有变量，均会出现在各程序的对话框中，间想要分析的变量选入程序方块中，包含自变量、因变量（依变量）及可能干扰因变量的协变量等。
- ④ **执行程序:** 执行特定程序后，可查看结果，结果会出现在[SPSS 浏览器]（SPSS Viewer）输出窗口中。执行程序在于数据的数字运算及选择适当的统计方法。
- ⑤ **结果存档或打印:** 运行结果可输出为网页文件、Word 文件、Excel 文件、文本文件等，便于继续编辑修改；也可直接存为 Spss 结果文件（扩展名为.spo），中文版的 SPSS Viewer输出窗口下呈现的结果表格，可直接复制粘贴到 Word 里面进行编辑，不会出现乱码。



## 2、spss基本操作——窗口



❖ 多窗口软件，种类有四种，数据窗口，输出窗口，语法窗口，脚本窗口。

使用最多的是**数据窗口**和**输出窗口**。

□ **数据窗口**：主要用于数据的处理工作，分为**数据视图**和**变量视图**

● **数据视图**（data view）为原始数据键入的窗口；

● **变量视图**（variable view）为设定变量内容的窗口，包括：变量名称（name）、变量类型（type）、宽度（位数或字符数）（width）、小数位数（decimals）、变量标签（lable）、值（values）、缺失值（misssing）、变量列宽（columns）、度量尺度（测量尺度 measure）等。

□ **输出窗口**：用于输出统计分析的结果。

示例：各城市购买因素



## 2、spss基本操作——数据窗口



\*未标题2 [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

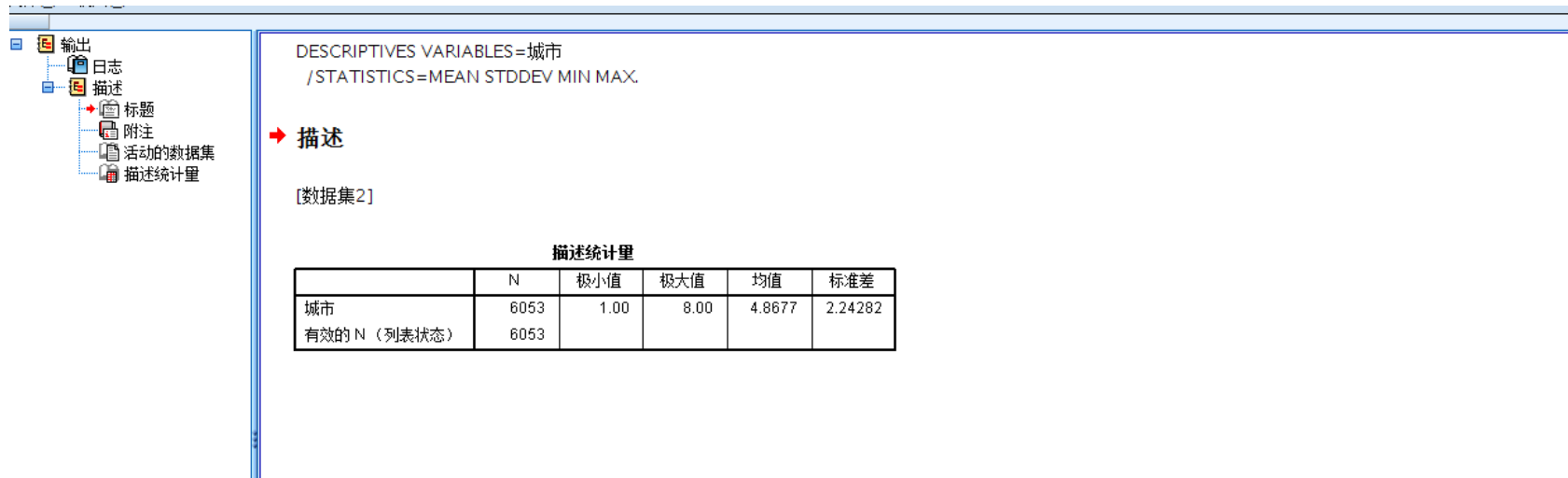


	名称	类型	宽度	小数	标签	值	缺失	列	对齐	度量标准	角色
1	城市	数值(N)	8	2		{1.00, 重庆}...	无	8	右	度量(S)	输入
2	购买因素	数值(N)	8	2		{1.00, 品种}...	无	8	右	度量(S)	输入
3	频数	数值(N)	8	2		无	无	8	右	度量(S)	输入
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											
25											

数据窗口

数据视图 变量视图

## 2、spss基本操作——输出窗口



SPSS Output Window Screenshot:

Output Window Structure:

- 输出
  - 日志
  - 描述
    - 标题
    - 附注
    - 活动的数据集
    - 描述统计量

DESCRIPTIVES VARIABLES=城市  
/STATISTICS=MEAN STDDEV MIN MAX.

→ 描述

[数据集2]

描述统计量

	N	极小值	极大值	均值	标准差
城市	6053	1.00	8.00	4.8677	2.24282
有效的 N (列表状态)	6053				

输出窗口





### 3、变量视图的详细解释



❖ 一个完整的SPSS数据结构包括:

变量名称、变量类型、变量名标签、变量值标签、缺失值的定义、度量尺度、及数据的显示属性（显示宽度、列宽度、对齐方式）等。



## ■ 变量名称



- ❖ 以中英文字母作为变量起始字符
- ❖ 不能以句点 “.” 作为变量名称结束字符
- ❖ 变量名的长度不能超过64个字符，太长了会不方便，要斟酌！
- ❖ 变量名必须唯一，并且不区分大小写
- ❖ 不能使用空格或特殊字符（如！？\*）
- ❖ 不能以SPSS保留字作为变量名（ALL、NE、EQ、GT AND、NOT、WITH等），如果在键入变量名时提及保留字，系统会提示
- ❖ 如不指定变量名，则系统默认变量名以 ‘VAR’ 开头后面跟5个数字

为了便于记忆和理解，建议变量名与它所代表的具体含义一致。



## ■ 变量名称



### 查看数据文件变量名的三种方法

- ❖ 在变量视图中浏览
- ❖ 在SPSS菜单中选择：实用程序—变量
- ❖ 在SPSS主界面选择：文件—显示数据文件信息—工作文件

## ❖ 数值型 (N) :

由0-9的阿拉伯数字和其他特殊符号，如美元，逗号或圆点组成。

❖ 字符型：变量的取值含有字母或者汉字（民族），使用该类型。

❖ 日期型：表示日前或者时间。

\*未标题1 [数据集0] - 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

	名称	类型	宽度
1	name	数值(N)	8
2	sex	数值(N)	8
3	VAR00003	数值(N)	8
4	VAR00004	数值(N)	8
5	VAR00005	数值(N)	8
6	VAR00006	数值(N)	8
7	VAR00007	数值(N)	8
8	VAR00008	数值(N)	8
9	VAR00009	数值(N)	8
10	VAR00010	数值(N)	8
11	VAR00011	数值(N)	8
12	VAR00012	数值(N)	8
13			
14			

变量类型

☒ 数值(N)

☐ 逗号(C) 宽度(W): 8

☐ 点(D) 小数位(P): 0

☐ 科学计数法(S)

☒ 日期(A)

☐ 美元(L)

☐ 设定货币(U)

☐ 字符串(R)

确定 取消 帮助

系统默认为数值型，呼应数据整理编码，一般不改变

# ■ 变量类型——数值型变量

## ❖ 数值型 (N) :

是SPSS默认的类型，  
宽度8、小数点2、  
小数点用圆点。

\*未标题1 [数据集0] - 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

	名称	类型	宽度
1	name	数值(N)	8
2	sex	数值(N)	8
3	VAR00003	数值(N)	8
4	VAR00004	数值(N)	8
5	VAR00005	数值(N)	8
6	VAR00006	数值(N)	8
7	VAR00007	数值(N)	8
8	VAR00008	数值(N)	8
9	VAR00009	数值(N)	8
10	VAR00010	数值(N)	8
11	VAR00011	数值(N)	8
12	VAR00012	数值(N)	8
13			
14			

变量类型

☒ 数值(N)

☐ 逗号(C)

☐ 点(D)

☐ 科学计数法(S)

☐ 日期(A)

☐ 美元(L)

☐ 设定货币(U)

☐ 字符串(R)

宽度(W): 8

小数位(P): 0

确定 取消 帮助

# ■ 变量类型——日期型变量(Date)

- ❖ 用于表示日期和时间的变量类型SPSS提供了多达29种日期型变量的格式。

\*未标题1 [数据集0] - 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)

	名称	类型
1	name	数值(N)
2	sex	数值(N)
3	VAR00003	数值(N)
4	VAR00004	数值(N)
5	VAR00005	数值(N)
6	VAR00006	数值(N)
7	VAR00007	数值(N)
8	VAR00008	数值(N)
9	VAR00009	数值(N)
10	VAR00010	数值(N)
11	VAR00011	数值(N)
12	VAR00012	数值(N)
13		
14		

变量类型

- ☐ 数值(N)
- ☐ 逗号(C)
- ☐ 点(D)
- ☐ 科学计数法(S)
- ☒ 日期(A)
- ☐ 美元(L)
- ☐ 设定货币(U)
- ☐ 字符串(R)

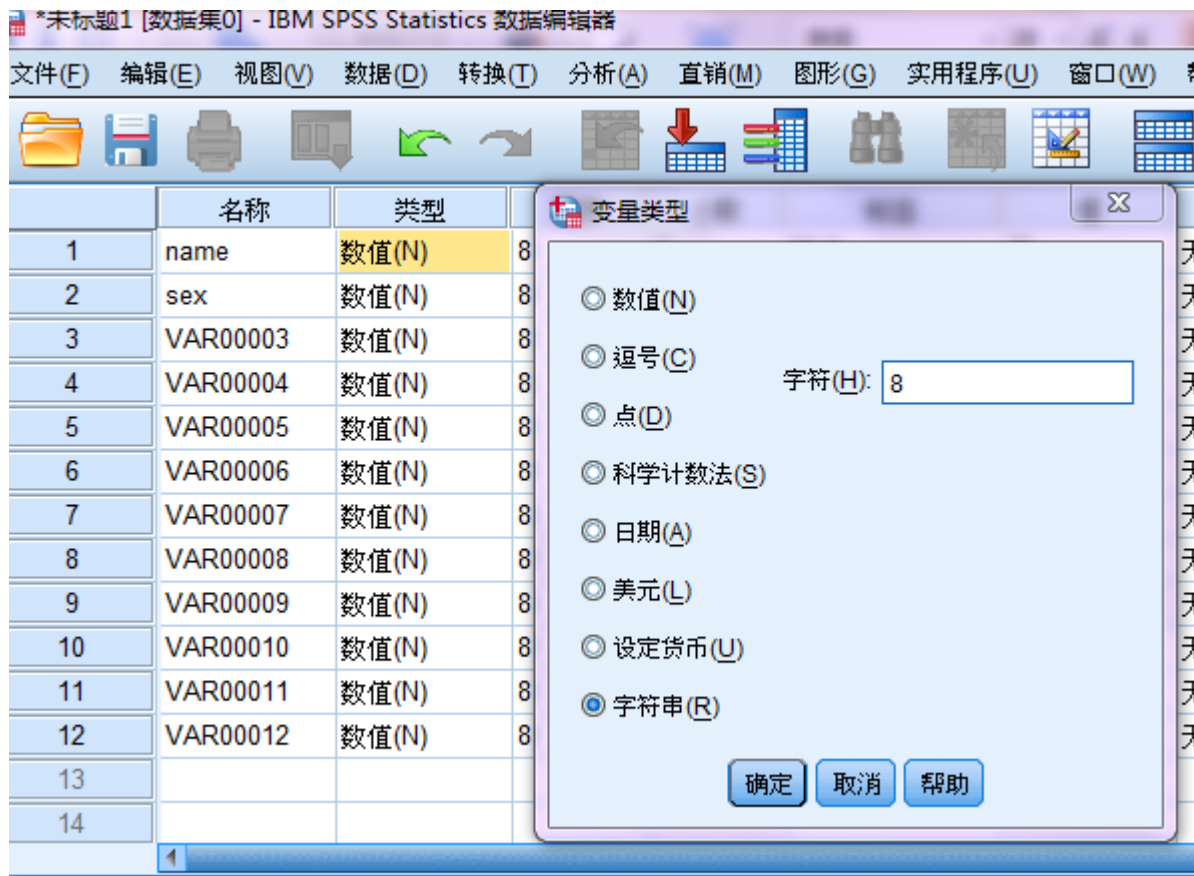
dd-mmm-yyyy  
dd-mmm-yy  
mm/dd/yyyy  
mm/dd/yy  
dd.mm.yyyy  
dd.mm.yy  
yyyy/mm/dd  
yy/mm/dd  
yyddd  
yyyyddd  
q Q yyyy  
q Q yy

确定 取消 帮助

数据视图 变量视图

# ■ 变量类型——字符型变量(String)

❖ 是非数值型变量类型,其值是由字符串组成的.对其定义的对话框只有一个输入项。



# ■ 变量类型——修改数据变量的默认值

❖ 对于SPSS变量的参数,系统都会自动给出默认值.也可通过编辑菜单中的选项对话框重新设置

\*未标题1 [数据集0] - IBM SPSS Statistics 数据编辑

文件(F) **编辑(E)** 视图(V) 数据(D) 转换(T) 分

	类型	宽度	小数
1	科学计数法	8	2
2	数值(N)	8	2
3	数值(N)	8	2
4	数值(N)	8	2
5	数值(N)	8	2
6	数值(N)	8	2
7	数值(N)	8	2
8	数值(N)	8	2
9	数值(N)	8	2
10	数值(N)	8	1
11	数值(N)	8	2
12	数值(N)	8	2
13			

选项

常规 查看器 数据 货币 输出标签 图表 枢轴表 文件位置 脚本 多重归因 语法编辑器

转换与合并选项

☒ 立即计算值(M)

☐ 使用前计算值(B)

显示新数值变量的格式

宽度(W): 8 小数位(D): 2

示例: 12345.67

设置用两位数字表示年号的年限全距

☒ 自动(U)

起始年: 1946  
结束年: 2045

☐ 设定(C)

起始年: 1946  
结束年: 2045

自定义变量视图...

更改字典...

随机数字生成器

☒ 与 SPSS 12 和先前版本兼容(O)

☐ 长期间 Mersenne 扭曲器(L)

读取外部数据

指定一个连续测量级别至数值字段, 使用至少(S) 24 唯一值。

数值取整与切断

RND 与 TRUNC 中使用的模糊位数: 6

确定 取消 应用(A) 帮助



## ❖ 变量名标签（variable label）

对变量名含义的进一步补充说明，定义变量名。

❖ 多个问项测量一个变量时，变量标签有很大的必要性。

\*未标题1 [数据集0] - IBM SPSS Statistics 数据编辑器

	名称	类型	宽度	小数	标签
1	name	数值(N)	8	2	姓名
2	sex	数值(N)	8	2	性别




## ❖ 变量值 (Value Label)

是仅对数值型变量各个取值的含义给予进一步的解释和说明，就是呼应我们最初的数据整理编码

\*未标题1 [数据集0] - 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)



	名称	类型	宽度	小数	标签	值
1	name	数值(N)	8	0	姓名	无
2	sex	数值(N)	8	0	性别	{1, 男}...

## ■ 变量缺失值

- ❖ 所谓的缺失值是受试者未填答时，研究者自行键入的数据，如五点式李克特量表，受试者在第 10 题未填答，研究者在键入数据时，可键入 ‘9’，在统计分析时，第 10 题的 ‘9’ 被设为了缺失值将不会纳入数据分析中。
- ❖ 系统默认无缺失值：如当前变量测试值、记录完全正确、无遗漏，则为 ‘无’ 缺失值状态。



The screenshot shows the IBM SPSS Statistics Data Editor window. The main table lists variables with their names, types, widths, and decimal places. The 'Missing' column for the 'name' variable is highlighted in yellow and contains the value '无' (None). A red box highlights the 'Missing' column header and the '无' value. A 'Missing Values' dialog box is open, showing the 'No missing values (N)' option selected.

	名称	类型	宽度	小数	标签	值	缺失
1	name	数值(N)	8	0	姓名	无	无
2	sex	数值(N)	8	0	性		
3	VAR00003	数值(N)	8	2			
4	VAR00004	数值(N)	8	2			
5	VAR00005	数值(N)	8	2			
6	VAR00006	数值(N)	8	2			
7	VAR00007	数值(N)	8	2			
8	VAR00008	数值(N)	8	2			
9	VAR00009	数值(N)	8	2			
10	VAR00010	数值(N)	8	1			
11	VAR00011	数值(N)	8	2			
12	VAR00012	数值(N)	8	2			
13							
14							

**缺失值**

☒ 没有缺失值(N)

☐ 离散缺失值(D)

☐ 范围加上一个可选离散缺失值(R)





低(L):  高(H):

离散值(S):

确定 取消 帮助





## ■ 度量尺度（测量尺度）

- ❖ 定距和定比型变量(scale)可以表示如温度、重量等含义的**连续性数值变量**,也可以表示年龄、次数等离散型变量.还可以表示时间的日期变量或者货币的货币型变量,**但不能是字符型变量。**（度量尺度）
- ❖ 定序型变量(ordinal)其值表示一种顺序的前后,如职称变量可分为高中低三个档次,可用A(1)、B(2)、C(3)表示, **定序型变量可以是数值型也可以字符型。**（序号）
- ❖ 定类型变量(nominal): 该变量不存在变量值之间的大小、顺序的前后等。只表示属于的类别。如性别中“1”表示男,“2”表示女等,可以是**数值型也可以字符型。**（名义）

度量标准	
	名义(N)
	序号(O)
	度量(S)
	度量(S)

# ■ 度量尺度（测量尺度）

- ❖ 为了便于后续的统计分析，研究者可以对变量的测量尺度加以正确的设定，尤其是名义尺度与序号尺度的区分。
- 以**独立样本 t检验与方差分析**为例，要进行上述分析，其要求**自变量**必须服从**名义尺度或序号尺度**，同时，**因变量**必须服从**度量尺度**（等距或等比变量）。若研究者未按要求设定，则后续统计分析无意义。

度量标准	
	名义(N)
	序号(O)
	度量(S)
	度量(S)



## □ 变量尺度的转换（呼应最初的数据整理编码）

- ❖ 连续变量是可以转换为分类变量。
- ❖ 如收入为连续变量，然后我们想把它转换为三分类：如1000元以下，1001-3000元，3001以上。
- ❖ 一种是重新编码为相同变量，对原始变量的取值进行直接的重新编码，替换原来的数据，原来的变量不存在了。
- ❖ 一种是“重新编码为不同变量”根据原始数据的取值生成一个新的变量，原来的变量还在
- ❖ (转换菜单中): “转换” - “重新编码为不同变量”，选择“旧值和新值”，选择“值”



# 课堂练习4



- 在软件中，自行熟悉消化一下刚才讲解过的变量视图下的各要素  
(5 分钟)



## 4、SPSS数据导入和结果导出



**SPSS数据导入**  
**SPSS结果导出**





# ■ SPSS数据导入



## ❖ 创建和读取SPSS数据文件

1. 直接通过SPSS ‘数据视图’ 窗口定义变量输入数据（如前所述）
2. 通过打开不同格式的数据文件，间接创建SPSS数据文件。

支持\*.doc, \*.xls, \*.txt等

操作：‘文件’ — ‘新建’ —— ‘数据’

‘文件’ — ‘打开’ — ‘数据’ — 选择文档类型——以excel表为例 ‘从第一行数据读取变量名’

## 数据导入练习（30 分钟）：

- （1）将前面课堂练习 2 的编码表导入SPSS中（也可直接在数据视图下键入数据），并检查变量视图下各要素；
- （2）将前面课堂练习 3 的编码表导入 SPSS 中，并检查变量视图各要素



SPSS结果的导出。

操作：‘文件’ — ‘导出’ — 可选择导出格式及导出路径

导出的对象

☒ 全部(A) ☐ 所有可见(V) ☐ 选定(D)

文档

类型(T):  
PowerPoint (\*.ppt)

将创建包含枢轴表及图形的 PowerPoint 演示文稿。图形将以标签图像文件 (TIF) 格式嵌入至文档中。不会导出文本项，只导出枢轴表。

选项(O):

Pivot Tables 中的层	按照打印层设置(在各表的表属性...
宽枢轴表	表格换行以适合页边距(W)
包括脚注及说明	是
使用查看内略图项作为幻灯片标...	否
模型视图	按照打印设置(在每个模型的模...
页面测量单位	毫米

更改选项(C)...

文件名(F):

浏览(B)...

图形

类型(E):  
JPEG 文件 (\*.jpg)

选项(S):

无可用选项	
-------	--

更改选项(H)...

根文件名(N):

C:\Users\吕刚\Documents\输出.jpeg

浏览(W)...

确定 粘贴(P) 重置(R) 取消 帮助