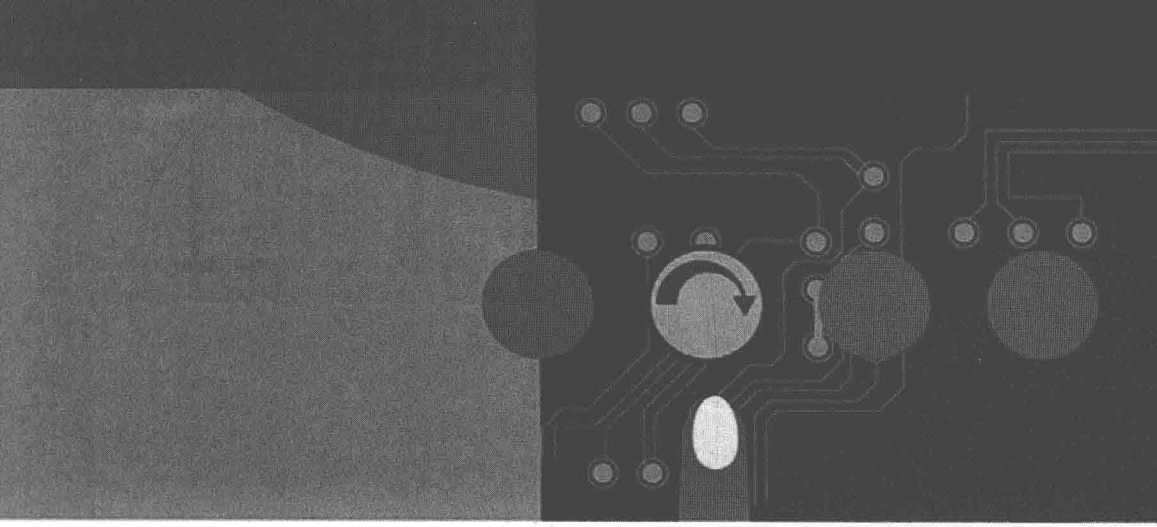
华南师范大学文科实验教材

# 语言研究统计学实验教程

主　编　周　榕

副主编　吴斯丹　冯　茵　刘巍巍

暨南大学出版社
JINAN UNIVERSITY PRESS



华南师范大学文科实验教材

# 语言研究统计学实验教程

主　编　周　榕

副主编　吴斯丹　冯　茵　刘巍巍

暨南大學出版社
JINAN UNIVERSITY PRESS

中国·广州

**图书在版编目（CIP）数据**

语言研究统计学实验教程/周榕主编. —广州：暨南大学出版社，2015.6
ISBN 978-7-5668-1447-0

Ⅰ.①语… Ⅱ.①周… Ⅲ.①语言统计—实验—教材 Ⅳ.①H0-05

中国版本图书馆 CIP 数据核字（2015）第 121058 号

出版发行：暨南大学出版社

地　址：中国广州暨南大学
电　话：总编室（8620）85221601
　　营销部（8620）85225284　85228291　85228292（邮购）
传　真：（8620）85221583（办公室）　85223774（营销部）
邮　编：510630
网　址：http://www.jnupress.com　http://press.jnu.edu.cn

排　版：广州联图广告有限公司
印　刷：佛山市浩文彩色印刷有限公司

开　本：787mm×1092mm　1/16
印　张：10.5
字　数：201 千
版　次：2015 年 6 月第 1 版
印　次：2015 年 6 月第 1 次

定　价：28.00 元

# 前　言

大学文科类学生的科研实验实践能力相对较弱，对此薄弱点，华南师范大学高度重视，出台了多项举措，如将各级大学生创新性实验计划项目都覆盖到文科类专业学生，并加强了文科类实验教学示范中心、文科类基础实验室、文科类教学实验教材的建设。本教程就是这些举措的产物，因而得到了学校的实验教材建设的立项和资助，也得到了学校教务处有关领导的大力支持，我们在此一并感谢。

本教程旨在提升学生的创新型实践能力，特别是运用统计软件 SPSS 开展语言学和应用语言学领域的科研实践的能力。SPSS 是国际上最著名、最通用的社会科学统计软件之一，广泛用于社会学、经济学、心理学、教育学、管理学等领域，也是语言学和应用语言学领域的实验性量化研究的重要工具。掌握了该统计软件的使用，无疑能使学生开拓学术视野，扩展科研手段，提升实践能力，因而有利于学生创新型实践能力和评判性思维能力的发展。

本教程将循序渐进地、系统有机地培养学生语言研究统计学的实践操作能力，配合“语言研究统计学”和“语言测试”等课程的教学，提供给学生在语言实验室实际操作和训练的指导，概言之，本教程有以下特点：

（1）根据语言研究统计学的需要，遵循实验室训练的特点，提供详细的 SPSS 操作指南，并配以截图，一步步指引学生由浅入深地上机操作，提升语言实验研究所需的统计学操作能力。

（2）以实验项目的形式编写，每一个实验项目既自成体系，训练一种统计分析的操作能力，又为下一个实验项目打下相关技能基础，逐级培养、提升全面的统计学实操能力。

（3）本教程尽量采用语言学和应用语言学研究中的实例，分析真实语言研究数据，使学生在掌握统计分析技能的同时，也提升语言研究的理论素养和科学研究的方法。

（4）语言尽量地浅显，操作尽量地直观，避免深奥的数理分析，让数学知识相对较弱的文科学生也能容易地操作和理解。

（5）每章后面都配有相关练习和参考答案，给予学生较多的实践操作练

习的机会，使学生没有教师指导时也能自主学习掌握。

笔者作为本教程主编，负责全书的设计、组织和统稿工作，并撰写了实验六、实验十一和实验十二。本书的实验四、实验八和实验九由吴斯丹老师撰写，实验一、实验五、实验七和实验十由冯茵老师撰写，实验二和实验三由刘巍巍老师完成。对本团队各位老师的辛勤工作，笔者在此表示衷心的感谢！

周　榕

2015 年 4 月

# 目录
contents

**前　言**　/001

**实验一**　数据初步整理能力的训练　/001

**实验二**　数据集中趋势分析能力的训练　/015

**实验三**　数据离散趋势分析能力的训练　/020

**实验四**　平均数差异显著性检验能力的训练　/025

**实验五**　方差分析能力的训练　/039

**实验六**　非参数检验分析能力的训练　/058

**实验七**　卡方检验分析能力的训练　/069

**实验八**　相关分析能力的训练　/088

**实验九**　因子分析能力的训练　/099

**实验十**　回归分析能力的训练　/114

**实验十一**　语言测试信度分析的训练　/132

**实验十二**　语言测试效度分析的训练　/148

**参考文献**　/161

# 实验一
# 数据初步整理能力的训练

要利用 SPSS 进行数据统计分析，首要的工作就是将研究数据导入到 SPSS 中，并建立相应的数据文件。但是实际收集到的数据往往是零散无序的，而且不能直接用于统计分析。因此，我们需要对原始数据进行科学的归纳、录入以及初步的整理，以便于更准确、更高效地进行统计分析。

实验一将详细介绍数据统计分析的前期工作，如数据文件的建立和导入，数据属性的设定，数据的整理、计算和变换等。这为日后采用其他的统计分析打下基础，为数据分析做好充分的准备。

## 1.1 SPSS 数据文件的建立

### 1.1.1 新建数据文件

打开 SPSS 软件后，选择菜单栏中的【File】，然后依次选择【New】、【Data】命令，创建新的空白数据文件，之后便可直接录入数据，如图 1－1 所示。

### 1.1.2 直接打开已有数据文件

打开 SPSS 软件后，选择菜单栏中的【File】，然后依次选择【Open】、【Data】命令，【Open Data】对话框弹出，如图 1－2 所示。选中需要打开的数据类型和文件名，双击左键即可打开该文件。

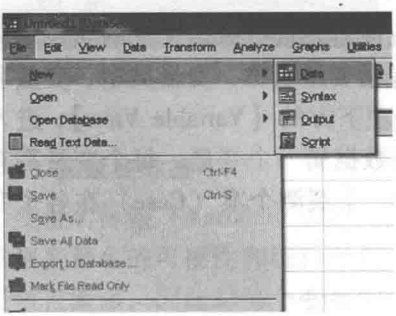



图 1－1 新建数据文件

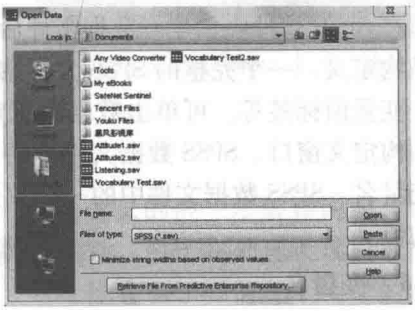
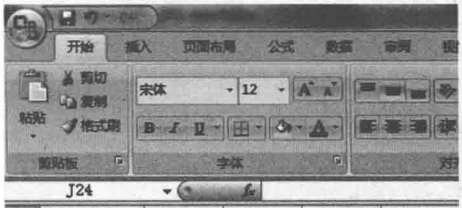



图 1－2 打开已有数据文件

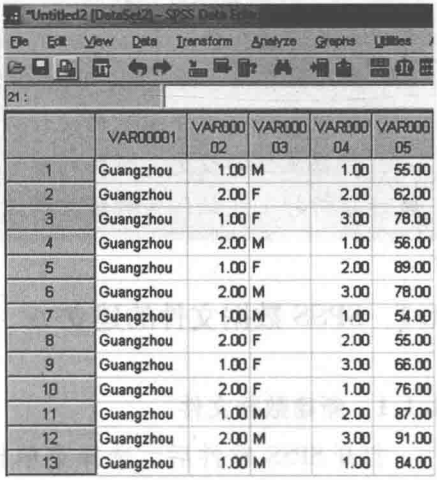
### 1.1.3 从 Excel 复制数据到新建文件

图 1-3 是三个城市在实施了不同的教学模式一年后的水平测试成绩表。复制 Excel 数据，首先用鼠标选取各列、行数据，拷贝并粘贴于新建的 SPSS 文件中，如图 1-4 所示。值得注意的是，这些可直接拷贝的原始数据结构应与 SPSS 文件的数据结构相一致，且不需要拷贝 Excel 工作表第一行的名称，名称可以在 SPSS 数据浏览窗口右下方的【Data View】通过重新定义变量名来确定。



| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | City | Method | Sex | Grade | Score |
| 2 | Guangzhou | 1 | M | 1 | 55 |
| 3 | Guangzhou | 2 | F | 2 | 62 |
| 4 | Guangzhou | 1 | F | 3 | 78 |
| 5 | Guangzhou | 2 | M | 1 | 56 |
| 6 | Guangzhou | 1 | F | 2 | 89 |
| 7 | Guangzhou | 2 | M | 3 | 78 |
| 8 | Guangzhou | 1 | M | 1 | 54 |
| 9 | Guangzhou | 2 | F | 2 | 55 |
| 10 | Guangzhou | 1 | F | 3 | 66 |
| 11 | Guangzhou | 2 | F | 1 | 76 |
| 12 | Guangzhou | 1 | M | 2 | 87 |

图 1-3 调查问卷的初步数据



| | VAR00001 | VAR00002 | VAR00003 | VAR00004 | VAR00005 |
|---|---|---|---|---|---|
| 1 | Guangzhou | 1.00 | M | 1.00 | 55.00 |
| 2 | Guangzhou | 2.00 | F | 2.00 | 62.00 |
| 3 | Guangzhou | 1.00 | F | 3.00 | 78.00 |
| 4 | Guangzhou | 2.00 | M | 1.00 | 56.00 |
| 5 | Guangzhou | 1.00 | F | 2.00 | 89.00 |
| 6 | Guangzhou | 2.00 | M | 3.00 | 78.00 |
| 7 | Guangzhou | 1.00 | M | 1.00 | 54.00 |
| 8 | Guangzhou | 2.00 | F | 2.00 | 55.00 |
| 9 | Guangzhou | 1.00 | F | 3.00 | 66.00 |
| 10 | Guangzhou | 2.00 | F | 1.00 | 76.00 |
| 11 | Guangzhou | 1.00 | M | 2.00 | 87.00 |
| 12 | Guangzhou | 2.00 | M | 3.00 | 91.00 |
| 13 | Guangzhou | 1.00 | M | 1.00 | 84.00 |

图 1-4 粘贴到 SPSS 文件的数据

思考题 1：数据导入到 SPSS 后能否马上开展统计工作呢？你将如何对图 1-4 的数据进行初步的整理呢？

## 1.2 SPSS 变量的定义

在 SPSS 软件中创建好数据文件后，应该首先进行数据文件的属性定义或者结构定义。一个完整的 SPSS 文件结构包括变量名称、变量类型、变量名标签、变量值标签等。可单击数据浏览窗口左下方的【Variable View】，进入数据结构定义窗口。SPSS 数据文件中的一列数据属一个变量，每个变量都有一个变量名。SPSS 数据文件中的一行数据为一个观测个案（Case）在各个变量上的值。

### 1.2.1 变量名

变量名（Name）是变量存取的标志，在定义 SPSS 数据属性时应首先定

好每个变量的名字。SPSS 默认的变量名为 VAR00001、VAR00002 等，如图 1－4 所示。我们可以根据需要来命名变量，如单击图 1－5 中【Name】那一列，分别输入“City”和“Score”等。变量的命名应遵循以下基本规则：

（1）SPSS 变量长度不能超过 64 个字符（32 个汉字）；

（2）首字母必须是字母或者汉字；

（3）变量名的结尾不能是圆点、句号或者下划线；

（4）变量名必须是唯一的，不区分大小写；

（5）变量名不能使用 SPSS 的保留字，例如 ALL，AND，OR，NOT，EQ，GE，GT，WITH，以及一些常用的函数符号等。

### 1.2.2　变量类型

设置变量类型，要先单击图 1－5 中的【Type】按钮，待出现定义变量类型的对话框后再进行编辑，如图 1－5 所示。在对话框中选择合适的变量类型，然后单击【OK】，即可定义变量类型。如果选择了【String】，则在图 1－5 的【Characters】键入字符串的长度；如果选择了【Numeric】，则在图 1－6 的【Width】和【Decimal Places】中键入相应的数值型的长度以及小数点位数，两者的系统默认值分别为 8 和 2。

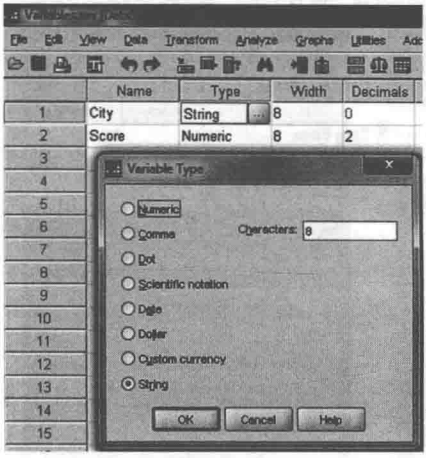



图 1－5　定义变量【String】时弹出的对话框

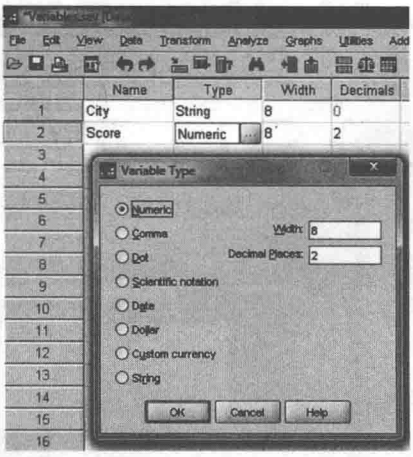



图 1－6　定义变量【Numeric】时弹出的对话框

### 1.2.3　变量名标签

变量名标签（Variable Label）是对变量名含义的进一步解释说明，它可以增强变量名的可视性和统计分析结果的可读性。由于在处理大规模数据时变量数目繁多，因此对每个变量的含义加以标注有利于弄清每个变量的实际含义。变量名标签一般用英文，如果是中文版的 SPSS，也可以用中文标记，

总长度可达 120 个字符。

### 1.2.4 变量值标签

变量值标签（Value Label）是对变量的每一个可能取值的进一步描述。当变量是定类或者定序变量时，不同的变量值一定要确定其相应的标签。设置变量值标签须单击图 1－7 中的【Values】下相应的  框，此时弹出【Value Labels】对话框。如果要把“Method”变量分为任务型方法（Task-Based Method）和教授型方法（Lecture-Based Method），可用“1”和“2”来分别代表这两种方法，以下是如何给出变量值标签的步骤。

（1）在【Value Labels】对话框中的【Value】右边的文本框输入“1”，在【Label】后输入“Task-Based Method”，代表第一种教学方法，再点击【Add】。

（2）以同样的方法，在【Value】右边的文本框输入“2”，在【Label】后输入“Lecture-Based Method”，代表第二种教学方法，再点击【Add】。如果有更多的类别，可按同样的方法定义变量值。如果需要改变和删除变量值标签，可以单击相应项目，按【Change】改变或者【Remove】进行删除，最后按【OK】。

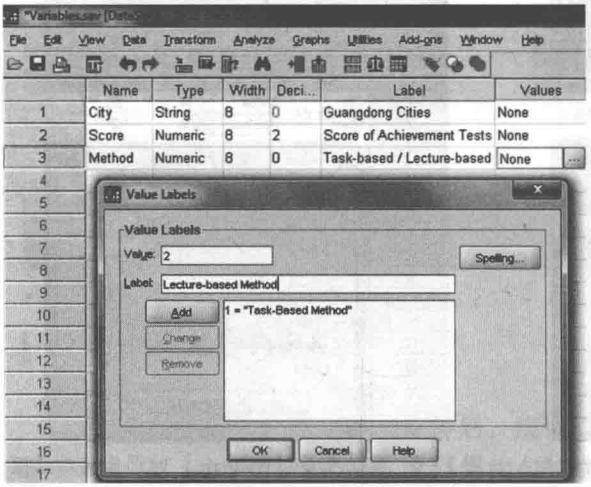



图 1－7 【Value Labels】对话框

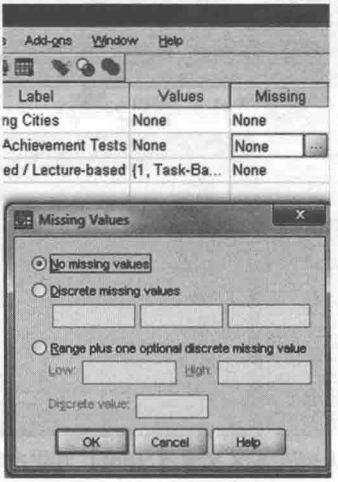



图 1－8 【Missing Values】对话框

### 1.2.5 变量缺失值

在统计分析中，收集到的数据可能会出现这样的情况：数据中出现明显的错误和不合理的情形，或者某些数据项的数据填错、漏填了。如果文件中的所有数据在分析中都可以使用，就会使错误数据参与统计，致使分析结果

不准确。为避免出现这种情况，可通过自行定义缺失值来筛选数据，使缺失值不参与统计分析。

如图 1 -8 所示，在数据定义窗口的【Missing】列中任意选择一个单元格，单击其右侧的 ...，【Missing Value】对话框弹出，框中有三种缺失值的定义方式可以选择。

（1）No missing values：不指定缺失值，这是系统默认项。

（2）Discrete missing values：对字符型或数据型变量，可填入 1 ~3 个特定的离散值作为定义的缺失值。

（3）Range plus one optional discrete missing value：缺失值可定义在一个由下限和上限组成的连续区间里面，并可输入一个该区间以外的离散值。

### 1.2.6　变量列宽

如图 1 -9 所示，【Data View】窗口中的【Columns】列，主要用于定义列宽，单击向上和向下的箭头按钮可以调整列宽，系统默认宽度为 8。

*Variables.sav [DataSet1] - SPSS Data Editor

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

| | Name | Type | Width | Decimals | Label | Values | Missing | Columns | Align | Measure |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | City | String | 8 | 0 | Guangdong Cities | None | None | 8 | Left | Nominal |
| 2 | Method | Numeric | 8 | 0 | Task-based / Lecture-b... | {1, Task-ba... | None | 3 | Left | Nominal |
| 3 | Sex | String | 8 | 0 | Male Femal | None | None | 5 | Left | Nominal |
| 4 | Grade | Numeric | 8 | 0 | Grade 7, 8, 9 | None | None | 5 | Left | Ordinal |
| 5 | Score | Numeric | 8 | 2 | Score of Achivement Test | None | None | 8 | Left | Scale |
| 6 | | | | | | | | | | |

图 1 -9　【Data View】对话框

### 1.2.7　变量对齐方式

如图 1 -9 所示，【Data View】窗口中的【Align】列，主要用于定义数据对齐方式，可选择“Left”左对齐、“Right”右对齐，或者“Center”居中对齐。

### 1.2.8　变量测量类型

如图 1 -9 所示，【Data View】窗口中的【Measure】列，主要用于定义变量的测量类型。根据情况可以有三种选择：对于等距型变量或比率变量可选择“Scale”；对于顺序型变量可以选择“Ordinal”，可以是数值型变量，也可以是字符型变量；对于称名型变量可以选择“Nominal”，可以是数值型变量，也可以是字符型变量。

思考题 2：根据图 1 -9，解释什么是等距型变量数据、顺序型变量数据和称名型变量数据。

## 1.3 SPSS 数据的整理

在 SPSS 数据编辑窗口里，表格的顶部标有经过定义的变量名，左侧有观测个案（Case）的序号。一个变量名和一个观测个案对应的单元格的数据便是该个案在该变量上的值。

通常情况下，刚刚建立的 SPSS 数据文件不能立即进行统计分析，这是因为输入的数据还是原始数据，可能存在错误、缺失等。此时，需要对原始数据进行进一步的加工、整理，使统计工作更科学、系统和合理。这部分将介绍如何对 SPSS 数据进行增减、排序、转置和分类汇总等。

### 1.3.1 数据的增减

增加一个观测个案有两步：① 如图 1－10 所示，点击选定要插入行的最左边的序号单元格，如第四行，被选中的观测个案所在行的全部单元格都被选中。②单击鼠标右键，选【Insert Cases】命令，便可在选中的行之上增加一个空行，此空行可输入观测个案的各个变量值。原来的测量个案及其后的观测个案都自动下移一行。

删除一个观测个案也是两步：① 如图 1－10 所示，点击选定要删除行的最左边的序号单元格，如第四行，被选中的观测个案所在行的全部单元格都被选中。② 单击鼠标右键，选【Clear】命令，此行的观测个案消失，下面的观测个案都自动上移一行。

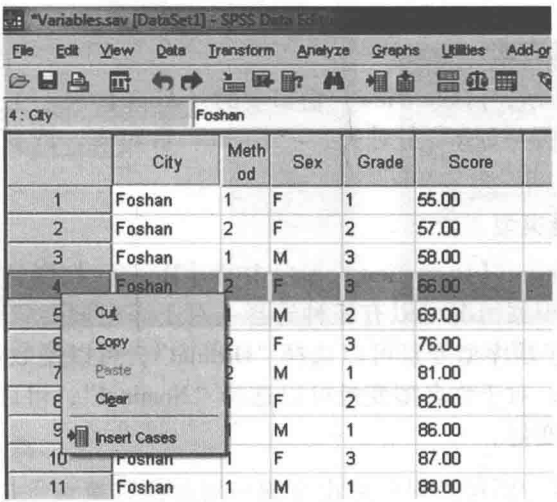



图 1－10　增减观测量

### 1.3.2　数据的排序（相当于 Excel 里面的排序）

在进行数据处理时，需要按某个变量值重新排列各个观测个案在文件中的先后顺序。以下是为数据排序的步骤：

（1）点击菜单栏【Data】，选取下拉菜单的【Sort Cases】命令，弹出【Sort Cases】对话框，如图 1－11 所示。

（2）假设我们选取“Score...”作为排序变量。在对话框左侧的候选变量列表框中选择排序变量“Score...”，单击右向箭头按钮，将其移至【Sort by】列表框中。

（3）在【Sort Order】选项组中可点击【Ascending】，即按变量的大小升序排序；若点击【Descending】，则按变量的大小降序排列。最后单击【OK】，结果如图 1－12 所示。

思考题 3：如果排序的变量不止一个怎么办？

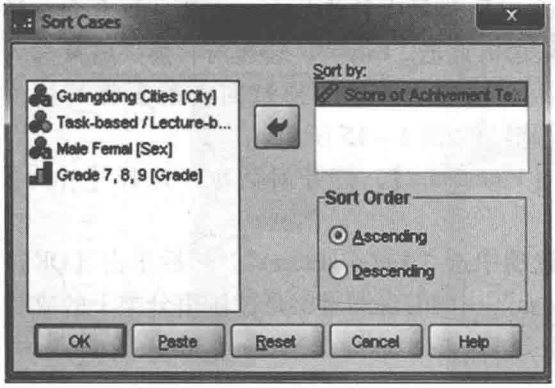
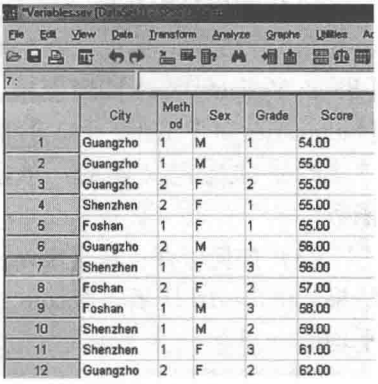



图 1－11　【Sort Cases】对话框



| | City | Method | Sex | Grade | Score |
|---|---|---|---|---|---|
| 1 | Guangzho | 1 | M | 1 | 54.00 |
| 2 | Guangzho | 1 | M | 1 | 55.00 |
| 3 | Guangzho | 2 | F | 2 | 55.00 |
| 4 | Shenzhen | 2 | F | 1 | 55.00 |
| 5 | Foshan | 1 | F | 1 | 55.00 |
| 6 | Guangzho | 2 | M | 1 | 56.00 |
| 7 | Shenzhen | 1 | F | 3 | 56.00 |
| 8 | Foshan | 2 | F | 2 | 57.00 |
| 9 | Foshan | 1 | M | 3 | 58.00 |
| 10 | Shenzhen | 1 | M | 2 | 59.00 |
| 11 | Shenzhen | 1 | F | 3 | 61.00 |
| 12 | Guangzho | 2 | F | 2 | 62.00 |

图 1－12　【Sort Cases】排序结果

### 1.3.3　数据的转置

利用 SPSS 的转置功能可以将原数据文件中的行、列进行互换，即将观测个案转变为变量。当转置结束后，系统会创建一个新的数据文件。进行行列转置的步骤如下：

（1）在菜单栏【Data】选取下拉菜单的【Transpose】命令，弹出【Transpose】对话框，如图 1－13 所示。

（2）选取所有变量作为转置对象，单击右向箭头按钮，将其移至【Variable（s）】列表框中，最后单击【OK】，结果如图 1－14 所示。由于数据文件转置后，数据属性的定义会丢失，因此要慎用此功能。

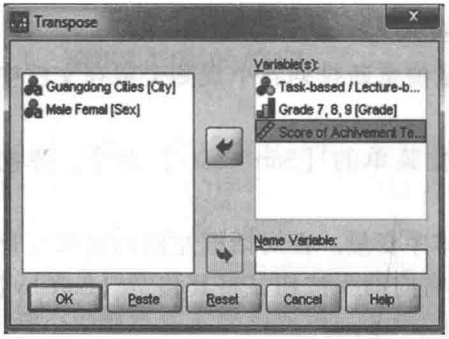
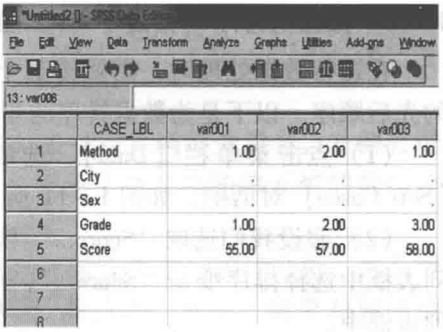



图 1 - 13 【Transpose】对话框




| | CASE_LBL | var001 | var002 | var003 |
|---|---|---|---|---|
| 1 | Method | 1.00 | 2.00 | 1.00 |
| 2 | City | . | . | . |
| 3 | Sex | . | . | . |
| 4 | Grade | 1.00 | 2.00 | 3.00 |
| 5 | Score | 55.00 | 57.00 | 58.00 |
| 6 | | | | |
| 7 | | | | |
| 8 | | | | |

图 1 - 14 【Transpose】转置结果

### 1.3.4 数据的分类汇总

SPSS 可对数据按照指定变量的数值进行归类，分组汇总。例如要了解不同城市的学生水平测试平均成绩，需要对数据按照所在城市进行分类，步骤如下：

（1）如何分配变量：单击【Data】菜单中的【Aggregate Data】命令，弹出如图 1 - 15 的对话框。在左侧候选框点击“City”，通过向右箭头将其移入【Break Variable（s）】列表框，再点击“Score”，通过向右箭头将其移入【Summaries of Variable（s）】列表框，如图 1 - 15 所示。

（2）如何定义函数：单击【Function...】，打开对话框，选择【Mean】函数。

（3）如何保存：在【Save】选项中选“Add... dataset”，然后单击【OK】，结果如图 1 - 16 所示。“Score_ mean”一列的观测量就是按城市分类下的成绩平均值。

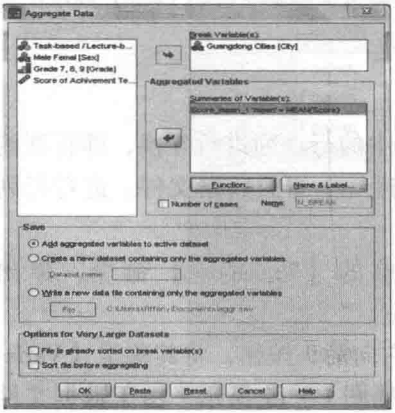
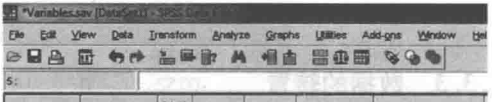



图 1 - 15 【Aggregate Data】对话框




| | City | Method | Sex | Grade | Score | Score_mean |
|---|---|---|---|---|---|---|
| 1 | Foshan | 1 | F | 1 | 55.00 | 76.57 |
| 2 | Foshan | 2 | F | 2 | 57.00 | 76.57 |
| 3 | Foshan | 1 | M | 3 | 58.00 | 76.57 |
| 4 | Foshan | 2 | F | 3 | 66.00 | 76.57 |
| 5 | Foshan | 1 | M | 2 | 69.00 | 76.57 |
| 6 | Foshan | 2 | F | 3 | 90.00 | 76.57 |
| 7 | Guangzho | 1 | M | 1 | 54.00 | 72.94 |
| 8 | Guangzho | 1 | M | 1 | 55.00 | 72.94 |
| 9 | Guangzho | 2 | F | 2 | 55.00 | 72.94 |
| 10 | Shenzhen | 2 | F | 1 | 55.00 | 74.72 |
| 11 | Shenzhen | 1 | F | 3 | 56.00 | 74.72 |
| 12 | Shenzhen | 1 | M | 2 | 59.00 | 74.72 |
| 13 | Shenzhen | 1 | F | 3 | 61.00 | 74.72 |

图 1 - 16 【Aggregate Data】分类汇总结果

## 1.4　SPSS 变量的处理

### 1.4.1　变量的增减处理

增加一个变量有两步：① 如图 1－17 所示，若要在第四列前增加新的一列，先单击第四列列首，此时整个列被选。② 单击鼠标右键后，点击【Insert Variable】命令。系统将在第四列前插入一个新的变量列，可以在该列输入观测值，原第四列自动向右移动成为第五列。

删除一个变量也有两步：①如图 1－17 所示，若要删除第四列，先单击第四列的列头，此时整个列被选。②单击鼠标右键后，点击【Clear】命令，此列所有观测值消失，右边的变量左移一列。

思考题 4：如何通过【Variable View】增加和删除变量？

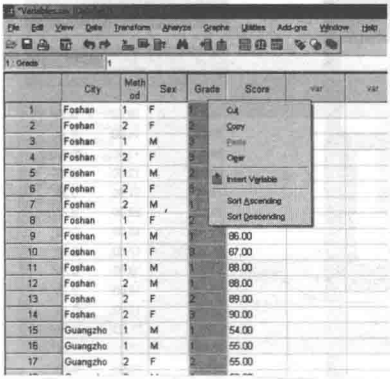



图 1－17　如何增减变量

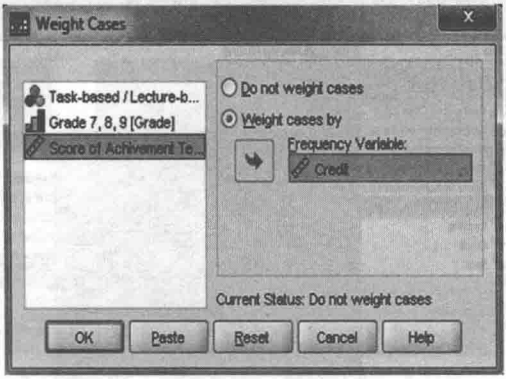



图 1－18　【Weight Cases】对话框

### 1.4.2　变量的加权处理

在实际统计中常常需要计算数据的加权平均值。例如，如果单靠期末考试成绩的算术平均数去评定某学生的学习成绩，这是不合理的，还应该考虑各学科的学分对平均成绩的影响。因此，以某学科的学分作为权重来计算一个学生考试成绩的平均数，更为合理。进行变量的加权处理的步骤如下：

（1）单击【Data】菜单中的【Weight Cases】命令，弹出如图 1－18 对话框。

（2）选择“Weight cases by”。

（3）从左侧候选框点击“Credit”，通过向右箭头将其移入【Frequency

Variable】列表框，再点击【OK】完成。

### 1.4.3 变量的计算处理

有时候收集到的原始数据并不能直接提供有用的信息，此时需要将原数据进行变换，生成有用的新变量。例如，三个城市的英语水平测试里面包含了听力成绩、阅读成绩等，我们需要计算听力占总成绩的百分比，用这个比例作为一个新的变量。SPSS 变量计数处理的步骤如下：

（1）单击【Transform】菜单中的【Compute Variable】命令，弹出如图1－19的对话框。

（2）在【Target Variable】文本框定义目标函数名为“Rate”，表示听力占总成绩的百分比。

（3）在【Numeric Expression】文本框中输入计算表达式“Listening/Score”，如图 1－19 所示。最后单击【OK】，结果如图 1－20 所示，新增的“Rate”变量就是计算出来的比例值。

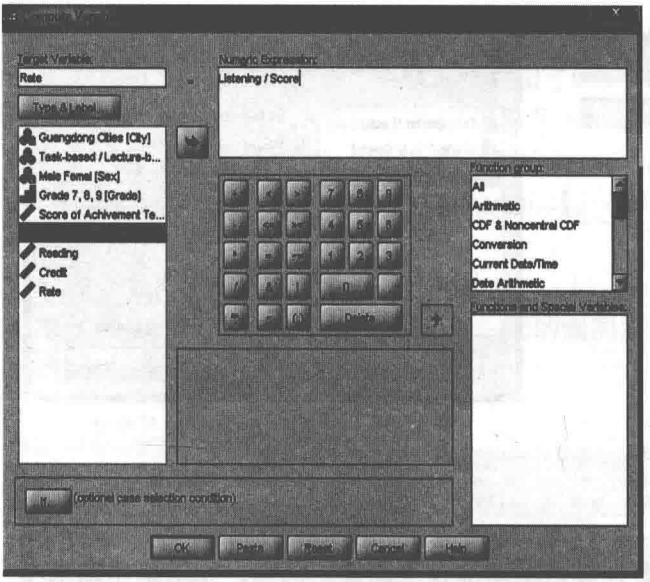



图 1－19 【Compute Variable】对话框

| Score | Listening | Rate |
|---|---|---|
| 55.00 | 12.00 | 0.22 |
| 57.00 | 13.00 | 0.23 |
| 58.00 | 14.00 | 0.24 |
| 66.00 | 13.00 | 0.20 |
| 69.00 | 12.00 | 0.17 |
| 76.00 | 21.00 | 0.28 |
| 81.00 | 22.00 | 0.27 |
| 82.00 | 18.00 | 0.22 |
| 86.00 | 17.00 | 0.20 |
| 87.00 | 18.00 | 0.21 |
| 88.00 | 13.00 | 0.15 |
| 88.00 | 12.00 | 0.14 |
| 89.00 | 21.00 | 0.24 |
| 90.00 | 22.00 | 0.24 |
| 54.00 | 18.00 | 0.33 |
| 55.00 | 17.00 | 0.31 |
| 55.00 | 18.00 | 0.33 |
| 56.00 | 20.00 | 0.36 |
| 62.00 | 21.00 | 0.34 |
| 65.00 | 22.00 | 0.34 |
| 66.00 | 21.00 | 0.32 |
| 76.00 | 18.00 | 0.24 |
| 76.00 | 15.00 | 0.20 |
| 78.00 | 24.00 | 0.31 |

图 1－20 变量计算结果

### 1.4.4 变量的自动分组处理

统计过程往往需要对某些变量进行分组，使其变成类别变量，即对变量进行重新赋值或者重新编码。例如在三个城市的英语水平测试中，将 80 分以上的成绩定为 A 水平等级，70～80 分为 B 水平等级，70 分以下为 C 水平等级。SPSS 分组变量的步骤如下：

（1）单击【Transform】菜单中的【Recode into Different Variables】命令，弹出如图 1－21 的对话框。

（2）在【Recode into Different Variables】对话框的左边的变量名列表中选定一个需要分组或重新赋值的变量，如“Score”，移入右边的【Numeric Variable→Output Variable】文本框内，同时在【Name】和【Label】输入新的变量名和标签，便于注释说明。

（3）单击【Old and New Values...】，打开相应文本框，如图 1－22 所示。在【Old Value】变量值范围的【Range，LOWEST through value】文本框中输入“70”，然后在【New Value】的最下面勾选【Output variables are strings】选项，在其【Value】文本框输入“C level”，这表示成绩在 70 分以下的范围在新变量中被定为 C 水平等级，然后单击【Add】，完成第一个变量赋值。以同样方式，定义旧变量的范围和新变量的变量值，得到如图 1－22【Old→New】文本框的结果。最后单击【Continue】回到图 1－21 的对话框。

（4）单击【Change】完成变量赋值，最后点击【OK】。结果如图 1－23“Grading”列所示。

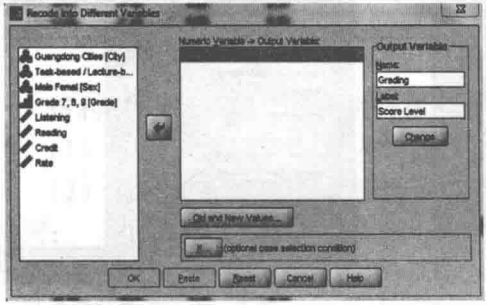



图 1－21　【Recode into Different Variables】对话框

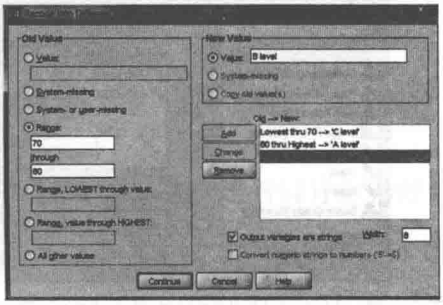



图 1－22　变量自动赋值处理

| | City | Method | Sex | Grade | Score | Listening | Reading | Credit | Rate | Grading |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Foshan | 1 | F | 1 | 55.00 | 12.00 | 30.00 | 2.00 | 0.22 | C level |
| 2 | Foshan | 2 | F | 2 | 57.00 | 13.00 | 32.00 | 3.00 | 0.23 | C level |
| 3 | Foshan | 1 | M | 3 | 58.00 | 14.00 | 28.00 | 4.00 | 0.24 | C level |
| 4 | Foshan | 2 | F | 3 | 66.00 | 13.00 | 27.00 | 5.00 | 0.20 | C level |
| 5 | Foshan | 1 | M | 2 | 69.00 | 12.00 | 29.00 | 3.00 | 0.17 | C level |
| 6 | Foshan | 2 | F | 3 | 76.00 | 21.00 | 30.00 | 4.00 | 0.28 | B level |
| 7 | Foshan | 2 | M | 1 | 81.00 | 22.00 | 40.00 | 2.00 | 0.27 | A level |
| 8 | Foshan | 1 | F | 2 | 82.00 | 18.00 | 42.00 | 3.00 | 0.22 | A level |

图 1－23　变量自动赋值结果

### 1.4.5　变量的计数处理

数据分析中，常常需要计算一些变量在同一个观测个案中满足要求的特