





销售分类 / 参考书

策划编辑 薛 琿 王祎星  
责任编辑 薛 琿 王祎星  
封面设计  北京华艺设计室  
郭 淼  
版式设计 李秋斋

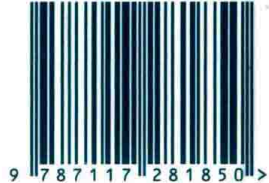
**人卫智网**  
[www.ipmph.com](http://www.ipmph.com)  
医学教育、学术、考试、健康，  
购书智慧智能综合服务平台

**人卫官网**  
[www.pmph.com](http://www.pmph.com)  
人卫官方资讯发布平台



关注人卫健康  
提升健康素养

ISBN 978-7-117-28185-0



9 787117 281850 >

定 价：65.00 元

# SPSS 轻松学

## 实战案例

主 审 吕 繁

主 编 汤后林

副主编 徐 鹏 琚腊红

编 委 (按姓氏笔画排序)

汤后林 许 娟 李 健 杨跃诚

徐 鹏 琚腊红 韩 晶 臧春鹏



人民卫生出版社

图书在版编目(CIP)数据

SPSS 轻松学: 实战案例 / 汤后林主编. —北京:  
人民卫生出版社, 2019

ISBN 978-7-117-28185-0

I. ①S… II. ①汤… III. ①医学统计—统计分析—  
软件包 IV. ①R195.1-39

中国版本图书馆 CIP 数据核字(2019)第 033391 号

人卫智网	<a href="http://www.ipmph.com">www.ipmph.com</a>	医学教育、学术、考试、健康, 购书智慧智能综合服务平台
人卫官网	<a href="http://www.pmph.com">www.pmph.com</a>	人卫官方资讯发布平台

版权所有, 侵权必究!

SPSS 轻松学——实战案例

主 编: 汤后林

出版发行: 人民卫生出版社(中继线 010-59780011)

地 址: 北京市朝阳区潘家园南里 19 号

邮 编: 100021

E - mail: [pmph@pmph.com](mailto:pmph@pmph.com)

购书热线: 010-59787592 010-59787584 010-65264830

印 刷: 三河市潮河印业有限公司

经 销: 新华书店

开 本: 787×1092 1/16 印张: 22

字 数: 535 千字

版 次: 2019 年 3 月第 1 版 2019 年 3 月第 1 版第 1 次印刷

标准书号: ISBN 978-7-117-28185-0

定 价: 65.00 元

打击盗版举报电话: 010-59787491 E-mail: [WQ@pmph.com](mailto:WQ@pmph.com)

(凡属印装质量问题请与本社市场营销中心联系退换)

# 序

SPSS (statistical product and service solution) 是国内外应用很广的一款统计分析软件, 由于好学易懂、窗口菜单操作方便, 深受广大用户的喜爱。2009 年 IBM 公司全资收购 SPSS 公司, 现使用的 SPSS 分析软件为 IBM SPSS Statistics。目前介绍 SPSS 统计分析的书籍有不少, 但基本上都是 SPSS 软件和统计学方法介绍, 基层工作人员在实际工作中的统计需求难以满足。

本书的主编汤后林为流行病与卫生统计学博士, 曾在中国疾病预防控制中心和加州大学洛杉矶分校学习。在基层和中国疾病预防控制中心工作近 20 年, 参与过多个国家级重大科技项目、国际合作项目和疾病信息系统的数据处理, 在运用 SPSS、SAS 等统计软件进行数据整理和分析方面积累了丰富的实践经验。本书从使用者的角度出发, 专门针对基层从事公共卫生工作人员和一线临床医生撰写了本书。本书内容全面、新颖独创、注重实用、可操作性强, 是目前市场上难得的一本专门为基层从事公共卫生工作人员和一线临床医生撰写的 SPSS 实用参考书。

本书集实用性和操作性于一体, 理论与实践应用紧密结合, 基础理论知识、知识点和针对每个知识点的实例互相融合穿插, 使读者不但知其然而且知其所以然, 可以使读者快速地学习、掌握 SPSS 软件操作及应用知识点解决实践中的问题, 从而达到学以致用和融会贯通的目的。

于石成

2018 年 6 月

# 前言

SPSS 是一款好学易懂的统计学软件,既可以通过窗口菜单,也可以通过语法程序进行数据整理和统计分析,是当前应用最为广泛的统计分析软件之一。

随着计算机信息技术发展,各行各业在工作和科研活动中产生了海量的原始数据,这些原始数据需要借用专门的数据分析软件进行整理、分析和利用,以用于指导开展工作或撰写报告和论文。作为一名公共卫生工作人员和一线临床医生,在日常的疾病监测、疾病控制、专题调查、临床实践等活动中,收集到了大量的数据,而这些数据整理和分析用一般计算器无法完成,需要借用专门的统计分析软件进行整理和分析。笔者在过去几年到各级医疗机构和疾病控制机构与工作人员进行沟通时发现,大家普遍感觉到在数据分析方面心有余而力不足,不能很好地分析利用现有的数据,甚至有时错误应用统计方法进行统计分析,迫切希望有一款通俗易懂、操作性和专业性强的参考书来指导进行现有数据分析利用,解决基层工作人员的实际问题。

结合多年 SPSS 使用和从事公共卫生和临床数据分析工作的经验,针对基层从事公共卫生工作人员和一线临床医生撰写了此本 SPSS 参考书,本书不求介绍高深复杂的统计方法,唯求简单、简洁和实用。第一章至第四章主要介绍 SPSS 数据整理方法,第五章主要介绍 SPSS 中的常用抽样方法,第六章主要介绍 SPSS 中常用图的作图方法,第七章主要介绍常用检验方法和回归模型,第八章和第九章主要介绍工作数据和论文数据的数据整理和统计分析方法,第十章主要介绍数据整理和数据分析中的一些技巧,第十一章主要介绍了近年来公开发表论文中的一些统计学方法误用及如何纠正。本书主要采取实际工作中的案例进行剖析讲解,力求达到举一反三的效果。

本书在编写中存在一定的不足,希望大家在使用过程中提出更多的宝贵意见和建议。

汤后林

2018年5月

# 目 录

第一章 初步认识 SPSS .....	1
第一节 SPSS 运行窗口 .....	1
第二节 SPSS 运行方式 .....	10
第二章 SPSS 数据录入和获取 .....	13
第一节 数据属性及定义 .....	13
第二节 直接录入数据 .....	19
第三节 外部导入数据 .....	24
第三章 SPSS 数据管理 .....	36
第一节 查重个案 .....	36
第二节 排序个案 .....	39
第三节 选择个案 .....	41
第四节 加权个案 .....	46
第五节 数据转置 .....	48
第六节 数据文件的合并 .....	49
第七节 数据的分类汇总 .....	58
第八节 拆分文件 .....	61
第四章 SPSS 数据转换 .....	63
第一节 变量计算 .....	63
第二节 变量重新编码 .....	67
第三节 自动重新编码 .....	72
第四节 变量值转换 .....	73
第五节 个案排秩 .....	75
第六节 日期和时间向导 .....	77
第七节 替换缺失值 .....	90
第八节 随机数字生成 .....	91
第九节 数据格式重排 .....	93
第十节 变量值中字符的查找与替换 .....	102

第五章 SPSS 抽样方法	109
第一节 简单随机抽样	109
第二节 复杂抽样	111
第六章 SPSS 作图方法	138
第一节 SPSS 统计图概述	138
第二节 直方图	138
第三节 饼图	146
第四节 条图	150
第五节 线图	152
第六节 散点图	154
第七节 ROC 曲线	157
第七章 常用检验方法和回归模型	160
第一节 $t$ 检验	160
第二节 卡方检验	163
第三节 Logistic 回归模型	169
第四节 Cox 回归模型	175
第八章 实战案例分析一：工作数据	179
第一节 数据准备	180
第二节 数据整理	180
第三节 数据分析	242
第九章 实战案例分析二：论文数据	250
第一节 数据准备	250
第二节 数据整理	251
第三节 数据分析	284
第十章 实战案例分析三：实战技巧	295
第一节 外部数据导入技巧	295
第二节 多选题录入和分析技巧	298
第三节 四格表和列联表数据录入技巧	302
第四节 分类变量的哑变量设置技巧	304
第五节 批量导出数据技巧	308
第六节 配对 Logistic 回归模型的分析技巧	312
第七节 生存曲线及生存率分析技巧	314

第十一章 论文常见错误分析.....	322
第一节 多个样本率的多重比较问题.....	322
第二节 多因素回归分析模型中的自变量设置问题.....	324
第三节 卡方检验应用问题.....	327
第四节 $t$ 检验应用问题 .....	334
第五节 重复测量资料应用问题.....	336
参考文献.....	341

# 第一章 初步认识 SPSS



社会科学统计软件包 (statistical package for the social science, SPSS) 是世界上应用最广泛的专业统计和数据模型软件之一。经过四十多年的发展, 在全球已拥有大量的用户。目前, SPSS 使用 Windows 的窗口方式展示各种管理和分析数据的方法, 可方便地用于各类商业管理和科研数据管理及统计分析工作。本章概要介绍 SPSS 的发展、主要版本、运行方式和启动与退出的使用。

1968 年, 3 位美国斯坦福大学的学生开发了最早的 SPSS 系统, 并基于这一系统于 1975 年在芝加哥合伙成立了 SPSS 公司。SPSS 的基本功能包括数据管理、统计分析、图表分析、输出管理等。2009 年 IBM 收购了 SPSS 公司。SPSS 被 IBM 全资收购后, 成为 IBM 软件部的一个产品, 现该产品归属在大数据与分析集团 Analytics 下。

到目前为止, SPSS 已成为适合于 DOS、Windows、UNIX 等多种操作系统的产品, 国内常用的是其适用于 DOS 和 Windows 的版本。本书以运行于 Windows 操作系统上的 SPSS 24.0 for Windows 标准版为例, 并在本书后面的内容中简称为 SPSS。

学习软件, 得先掌握软件的基本模块及其逻辑关系。因此, 要想熟练掌握 SPSS 软件并在数据分析中熟练运用, 先要掌握 SPSS 软件的基本模块及其基本步骤。

## 第一节 SPSS 运行窗口

SPSS 软件的安装非常简单方便, 直接根据安装盘的提示进行安装即可, 市场上各类 SPSS 软件教材均有介绍, 本文不再赘述。本书中有关操作步骤和界面是以“IBM SPSS Statistics 24”为例, 不同版本的 SPSS 软件之间显示界面稍微有些差别, 但不影响读者使用。

### 一、SPSS 如何打开

1. 打开 SPSS 软件的方法, 既可以直接双击桌面快捷图标, 也可以从电脑桌面通过“开始—所有程序—IBM SPSS Statistics”的目录下, 点击“IBM SPSS Statistics 24”, 如图 1-1 所示。

2. 打开 SPSS 后, 如图 1-2 所示界面。这时, SPSS 浮现窗口会问“你希望做什么”, 对于初学者, 可以不用管它, 直接点击“取消”按钮(第二章具体讲如何读取数据至 SPSS 软件中), 这样就打开了 SPSS 主界面。

3. SPSS 主界面的数据编辑窗口由菜单栏、工具栏、编辑栏、变量名栏、内容区、窗口切换标签页和状态栏组成, 如图 1-3 所示。



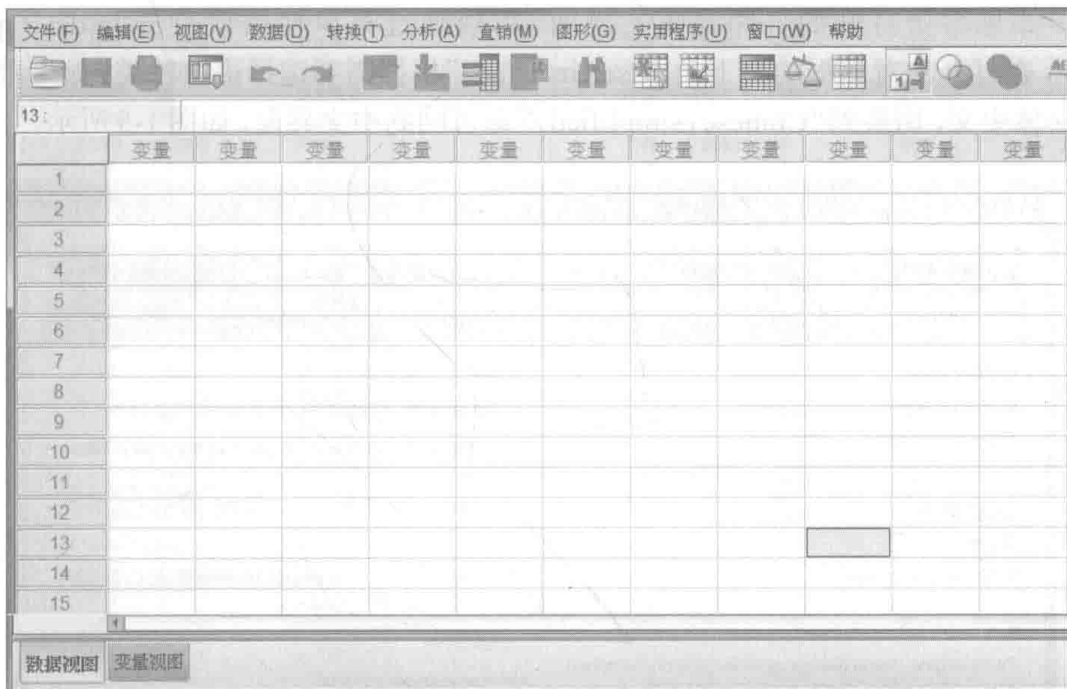


图 1-3 SPSS 数据编辑窗口

4. SPSS 提供了多种语言的用户界面, 可以根据需要进行切换。

(1) 如果希望打开的界面是中文界面, 可以通过菜单栏点击“编辑—选项—常规”, 在打开的对话框的右下角“用户界面”, 根据需要选择语言种类, 如果希望用户界面为英文, 选择“英文”即可。如图 1-4 所示。

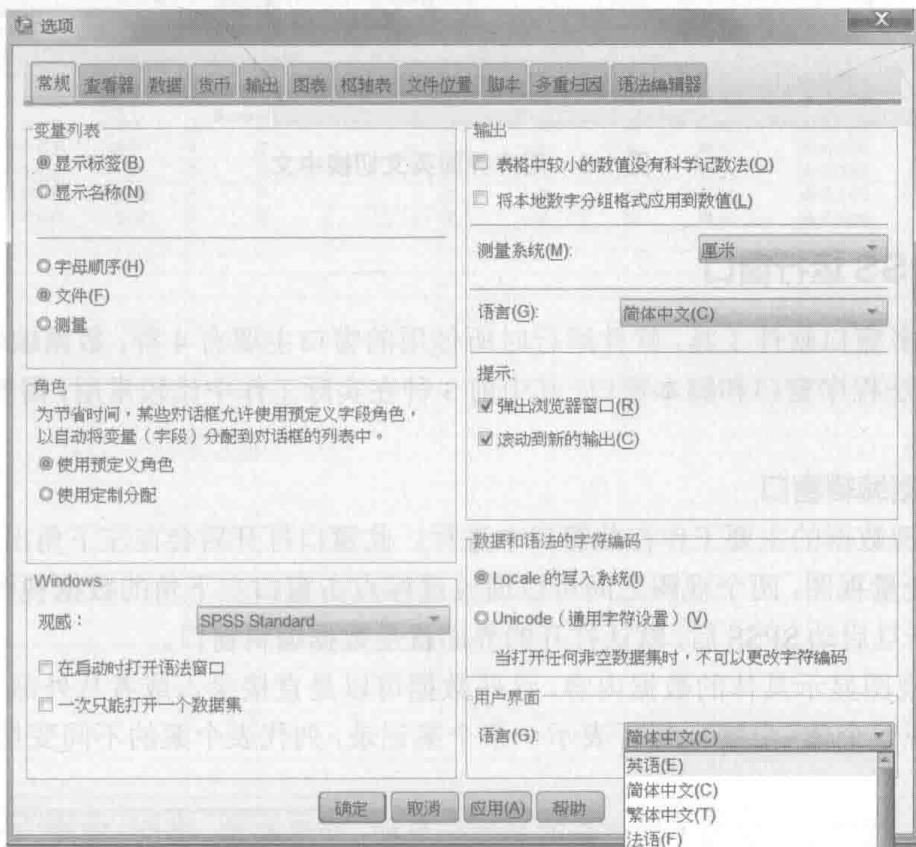


图 1-4 用户界面中文切换英文

(2) 如果希望打开的界面是英文界面,可以通过菜单栏依次点击“Edit-Options-General”,在打开的对话框的右下角“User interface”根据需求选择语言种类;如果希望用户界面为简体中文,切换到“Chinese (simplified)”即可回到中文界面,如图 1-5 所示。

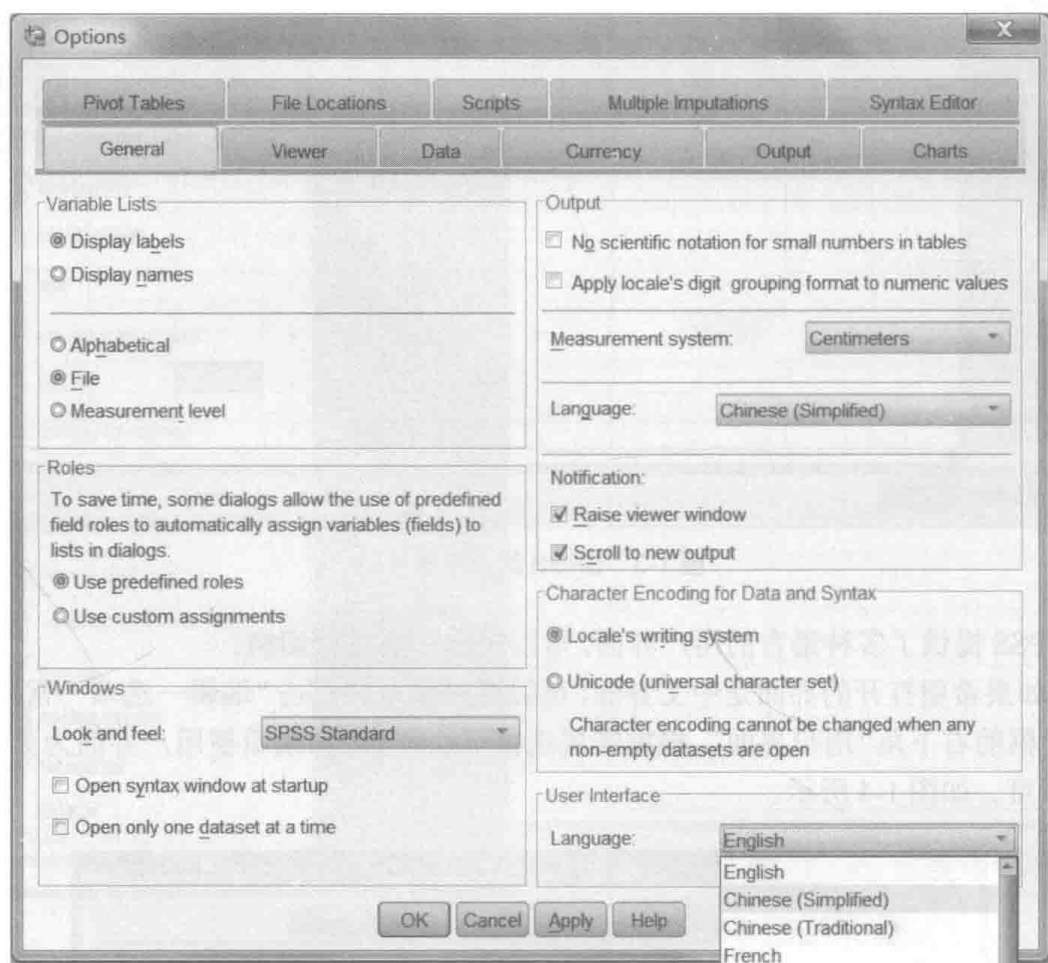


图 1-5 用户界面英文切换中文

## 二、SPSS 运行窗口

SPSS 是多窗口软件工具,软件运行时所使用的窗口主要有 4 种:数据编辑窗口、结果输出窗口、语法程序窗口和脚本窗口,其中前 3 种在实际工作中比较常用,需要大家熟练掌握,灵活运用。

### (一) 数据编辑窗口

SPSS 处理数据的主要工作在此窗口中进行。此窗口打开后会在左下角出现两个视图:数据视图和变量视图,两个视图之间可以通过鼠标点击窗口左下角的数据视图或变量视图进行切换。一旦启动 SPSS 后,默认打开的界面就是数据编辑窗口。

1. 数据视图显示具体的数据内容,这些数据可以是直接录入或者从外部数据导入,其中的行代表观察个体(个案),每行表示一条个案记录,列代表个案的不同变量值,如图 1-6 所示。

2. 变量视图显示个案的不同变量的具体信息如:变量名称、类型、宽度、小数、标签、值等信息,如图 1-7 所示。

	地区编码	地区名称	地区全称	地区级别	行政级别	城市农村标记	东中西部标记
1	1100	北京市	北京市	省级	直辖市	城市	东部
2	1101	市辖区	北京市市辖区	市级	无	城市	无
3	1102	县	北京市县	市级	无	农村	无
4	1199	不详市	北京市不详市	市级	无	无	无
5	1200	天津市	天津市	省级	直辖市	无	东部
6	1201	市辖区	天津市市辖区	市级	无	无	无
7	1202	县	天津市县	市级	无	无	无
8	1299	不详市	天津市不详市	市级	无	无	无
9	1300	河北省	河北省	省级	省	城市	中部
10	1301	石家庄市	河北省石家庄市	市级	市	城市	无
11	1302	唐山市	河北省唐山市	市级	市	城市	无
12	1303	秦皇岛市	河北省秦皇岛市	市级	市	城市	无
13	1304	邯郸市	河北省邯郸市	市级	市	城市	无
14	1305	邢台市	河北省邢台市	市级	市	城市	无
15	1306	保定市	河北省保定市	市级	市	城市	无
16	1307	张家口市	河北省张家口市	市级	市	城市	无
17	1308	承德市	河北省承德市	市级	市	城市	无
18	1309	沧州市	河北省沧州市	市级	市	城市	无
19	1310	廊坊市	河北省廊坊市	市级	市	城市	无
20	1311	衡水市	河北省衡水市	市级	市	城市	无
21	1399	不详市	河北省不详市	市级	无	城市	无
22	1400	山西省	山西省	省级	省	无	中部
23	1401	太原市	山西省太原市	市级	市	城市	无
24	1402	大同市	山西省大同市	市级	市	城市	无
25	1403	阳泉市	山西省阳泉市	市级	市	城市	无
26	1404	长治市	山西省长治市	市级	市	城市	无
27	1405	晋城市	山西省晋城市	市级	市	城市	无
28	1406	朔州市	山西省朔州市	市级	市	城市	无
29	1407	晋中市	山西省晋中市	市级	市	城市	无
30	1408	运城市	山西省运城市	市级	市	城市	无

数据视图 变量视图

图 1-6 数据视图窗口

名称	类型	宽度	小数	标签	值	缺失	列	对齐	度量标准	角色
1	地区编码	字符串	6	0	无	无	6	左	名义(N)	输入
2	地区名称	字符串	12	0	无	无	12	左	名义(N)	输入
3	地区全称	字符串	30	0	无	无	30	左	名义(N)	输入
4	地区级别	字符串	4	0	无	无	6	左	名义(N)	输入
5	行政级别	字符串	6	0	无	无	6	左	名义(N)	输入
6	城市农村	数值(N)	4	0	无	无	11	右	名义(N)	输入
7	东中西部	数值(N)	4	0	无	无	15	右	名义(N)	输入
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										
25										
26										
27										
28										
29										
30										
31										
32										

数据视图 变量视图

图 1-7 变量视图窗口

## (二) 结果输出窗口

SPSS 结果输出窗口用于显示和编辑数据分析的输出结果。如果分析过程正确,显示分析结果。如果分析过程中发生错误使处理失败,则在窗口中显示出系统给出的错误信息,通过给出的错误信息,就可以知道分析过程中出现的具体问题。结果输出窗口分为两个区:左侧目录区和右侧内容区。

1. 目录区显示的是 SPSS 不同分析结果的目录。
2. 内容区显示的是与目录一一对应的具体分析结果。

3. 通过点击结果输出窗口中的不同级别的目录,在右侧的内容区中会显示该目录下的分析结果的具体内容,对于多次执行不同的分析,产生的分析内容比较多,通过点击目录,查找对应的分析结果就非常方便,如图 1-8 所示。

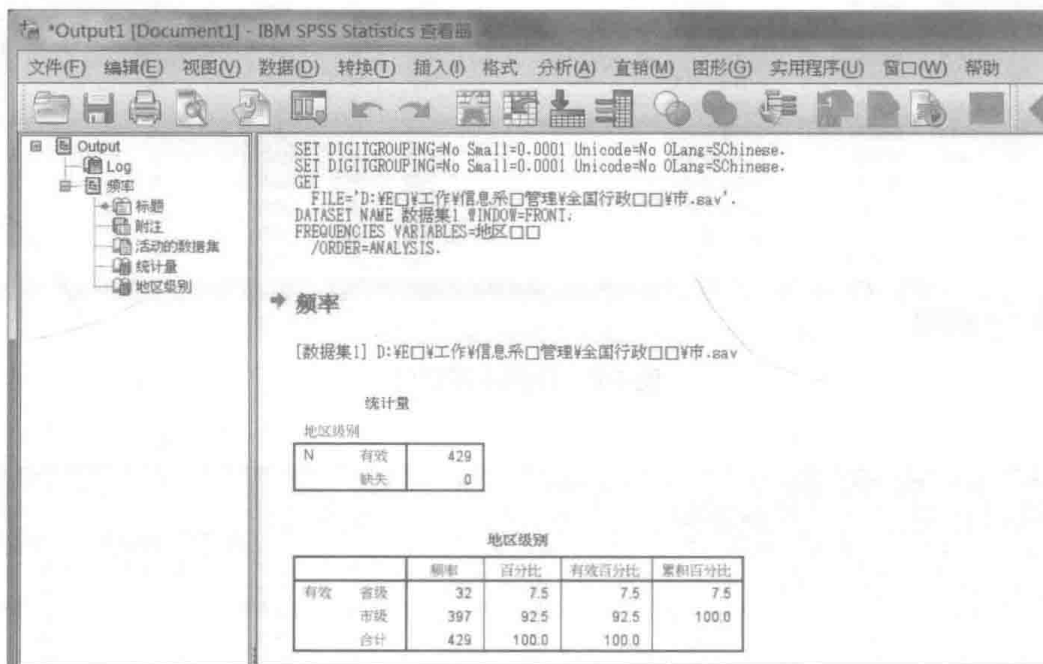


图 1-8 结果输出窗口

## (三) 语法命令窗口

也称命令语句窗口或语法编辑窗口,是编辑和运行程序命令的编辑器。与其他一些数据分析软件类似,SPSS 也提供了语法程序运行方式进行数据分析,一方面与菜单运行方式进行互补,另一方面便于对重复工作进行简化,一次编辑完成,多次使用。例如图 1-9 是对变量(地区级别)进行频数分析的程序语法语句。

1. 第 1~2 行是从计算机中读取 SPSS 数据库文件。
2. 第 3 行是临时生成的数据库并作为活动数据集(注:SPSS 读取数据库文件后会自动生成一个临时活动数据集,这样便于操作运行后保存为所需要的数据库文件)。
3. 第 4~5 行是对变量“地区级别”进行频数分析。

## (四) 脚本窗口

SPSS 脚本是用 Sax Basic 语言编写的程序。在脚本窗口也可以像 SPSS 宏一样构建和运行 SPSS 命令语句。可以将经常做的工作编辑成脚本文件,供使用时直接调用,这样可以大大提高工作效率,对于一般使用者来说,没必要去专门学习掌握。



图 1-9 语法命令窗口

### 三、SPSS 选项设置

打开 SPSS 后,各窗口的界面显示方式可以根据用户喜好或者需要进行设置。这里就好比装修房子,把需要的各类材料尺寸、颜色等规格都预先设置一下,这样就既省工又省料,界面看上去还很美观。那么怎么做到呢?在打开的 SPSS 数据视图界面的菜单中选择“编辑—选项”,如图 1-10 所示,以下主要介绍常用的功能。

1. 常规 对变量列表、角色、Windows 观感效果、输出语言方式、用户界面的语言界面等进行设置,设置完毕点击“应用”或者直接点击“确定”按钮。

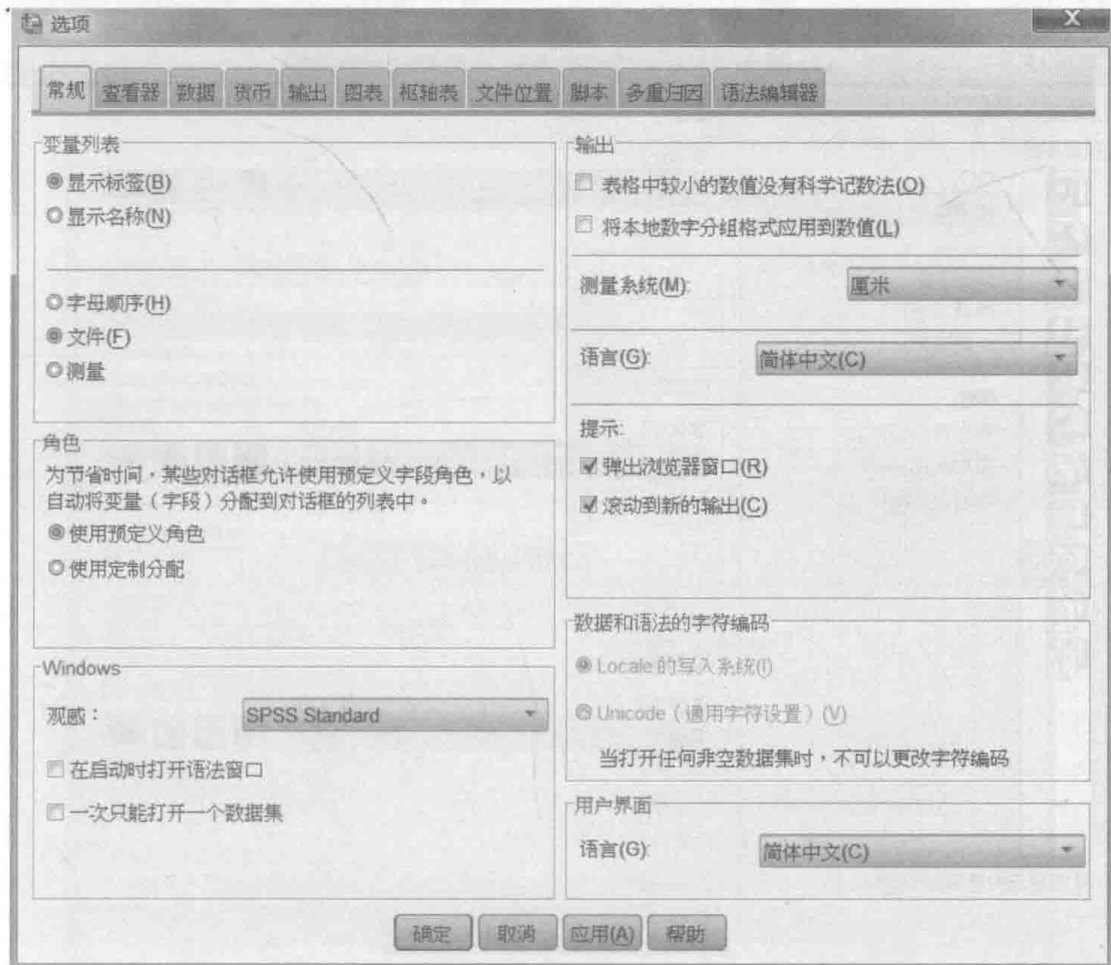


图 1-10 常规选项设置

2. 查看器 对各种结果的初始输出状态进行设置,包括输出的标题、页面标题和文本输出的字体、大小、颜色等进行设置。如图 1-11 所示。

3. 数据 通过预先设置数据的转换与合并、显示新数值变量的格式等等,在随后的数据管理过程中就会默认预先设置的结果。如图 1-12 所示。

4. 货币 预先设置货币型数值的显示格式。

5. 图表 对输出图表的宽高比、字体、样式、各条目的对比、线条等进行设置。

6. 枢轴表 主要是对输出的统计表格的外观、列宽等进行设置,以便符合统计要求。此项功能比较重要,默认为缺失,一般根据统计学习习惯要求,选择“Academic”。当然也可以根据大家的喜好,进行多种选择。如图 1-13 所示。

7. 语法编辑器 主要是对语法程序语句格式进行设置,不同语句的编码显示颜色,错误的语句编码显示颜色,以便于区分和识别。如图 1-14 所示。

● 请注意:“从对话框粘贴语法”默认为“自上一条命令之后”,就是说,每次粘贴一个操作的语法程序到语法程序窗口中,自动粘贴在上一次粘贴的语法程序之后,这样就不会乱,也便于后期保存连续操作后的全部语法程序,以便下一次重复该项操作时,直接运行语法程序即可得到预期的结果。这个在后面讲实战案例时,大家就更有体会。

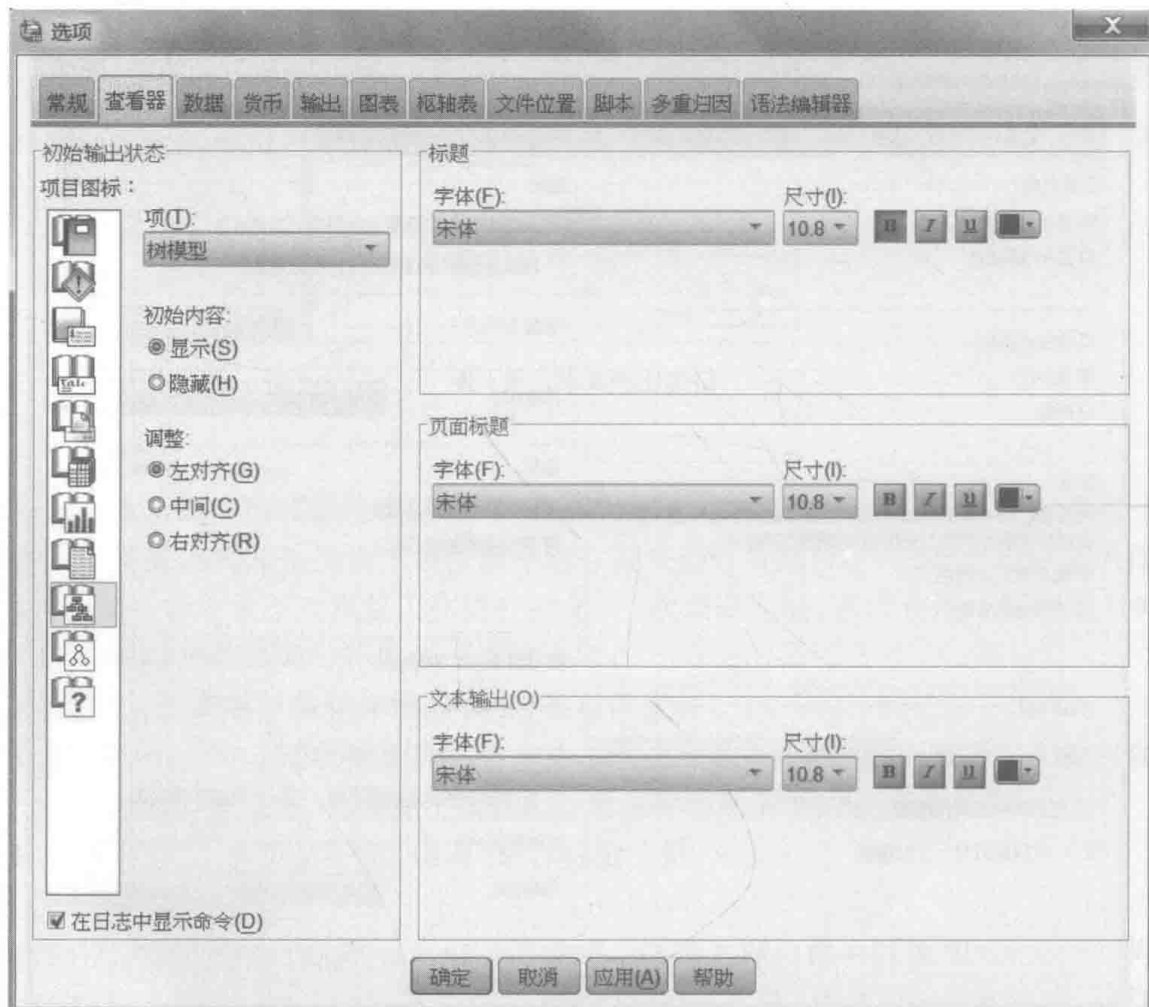


图 1-11 查看器选项设置