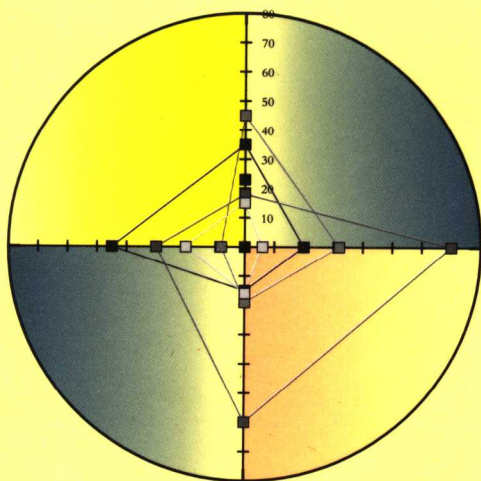


21世纪高等院校规划教材

统计实验

主编 林 亮 副主编 吴群英 王远清



国防工业出版社

National Defense Industry Press

21 世纪高等院校规划教材

统计实验

主 编 林 亮
副主编 吴群英 王远清
编 者 徐海云 涂雄苓 何宝珠 刘筱萍

国防工业出版社

· 北京 ·

内 容 简 介

本书是广西教育厅数学实验的探索与研究课题的研究成果,是专为统计学专业的实习课程所编的教材。本书内容分 SPSS 软件介绍,抽样调查实验,统计软件实验,统计综合实验,共 4 个部分。实验内容丰富,大部分都来自于真实的案例与分析,而且不少内容和大学生的现实问题密切相关,用以培养学生分析问题和解决实际问题的能力,有助于统计学专业学生对统计原理和数理统计基础知识的理解和正确应用,同时也能更好地掌握统计软件 SPSS 的操作与应用。

本书可作为高等学校统计专业的统计试验教材,也可以作为理工、财经、管理等专业本、专科学生的实验教材和相关工程技术人员的参考用书。

图书在版编目(CIP)数据

统计实验 / 林亮主编; 吴群英, 王远清副主编. —北京: 国防工业出版社, 2006. 9
21 世纪高等院校规划教材
ISBN 7-118-04595-0

I. 统... II. ①林... ②吴... ③王... III. 统计学—实验—高等学校—教材 IV. C8-33

中国版本图书馆 CIP 数据核字(2006)第 069493 号

※

国防工业出版社出版发行

(北京市海淀区紫竹院南路 23 号 邮政编码 100044)

天利华印刷装订有限公司印刷

新华书店经售

*

开本 710×960 1/16 印张 17¼ 字数 306 千字

2006 年 9 月第 1 版第 1 次印刷 印数 1—4000 册 定价 26.00 元

(本书如有印装错误,我社负责调换)

国防书店:(010)68428422

发行邮购:(010)68414474

发行传真:(010)68411535

发行业务:(010)68472764

前 言

本书是广西教育厅数学实验的探索与研究课题的研究成果。作者长期从事数学、统计学和统计学实验的教学,同时组织了多届数学建模比赛,有丰富的实践经验。

本书是专为统计学专业的实习课程所编的教材,由 SPSS 软件介绍,抽样调查实验,统计软件实验,统计综合实验,共 4 个部分组成。统计实验内容丰富,大部分都来自于真实的案例与分析,而且不少内容和大学生的现实问题密切相关,用以培养学生分析问题和解决实际问题的能力,有助于统计学专业学生对统计原理和数理统计基础知识的理解和正确应用。

统计学是 20 世纪迅速发展的学科,是研究各种随机现象的本质与内在规律性以及自然科学、社会科学等各个学科中各种类型数据的科学的综合处理及统计推断方法。随着人类社会各种体系的日益庞大、复杂、精密,以及计算机的广泛使用,统计学的重要性越来越大。

我国原有的统计学专业基本上以社会经济统计为主,是与计划经济体制相适应的专业。尽管很多高校的统计学专业都做了不同程度的教学改革,但未从根本上改变学生的知识结构与技能结构,其数理统计理论与方法及运用计算机能力均相当薄弱,难以适应社会主义市场经济的需要。统计学权威人士认为,对各级政府来说,今后需要的统计人才必须是具有掌握国民经济核算体系、宏观经济统计分析、抽样调查理论与方法、社会问题统计研究理论与方法及计算机知识与技能的高级人才。对企业来说,还要求具有对经营管理所需要的各种信息进行搜集、整理、分类的知识与技能,能开展统计预测,参与企业决策。因此,建立数理统计与社会经济统计相结合的“大统计”成为时代的要求,是统计界的主流呼声。

统计学随着计算机的发展其应用越来越重要,如何培养学生分析问题与解决实际问题的能力,是许多高校统计学教学改革的重点。其中一个重要的方面就是增加实验课程及实习时间,相应地,教材改革也应增加这方面的内容。

本书除了可作为高等学校统计专业的统计试验教材外,还可作为理工、财经、管理等专业本、专科学生的实验教材用书,也可供有关工程技术人员参考。

本书主要针对统计的实际问题提供一些实验的参考资料和数据,使学生在实验与实习中得以借鉴与参考。

本书有的实验是我院统计专业 2002 级学生完成的实验作业,在编写过程中得到该班同学的不少帮助,校稿还得到了我院年轻教师封全喜的帮忙,在此特向他们表示感谢。

由于时间仓促,加上我们的水平有限,书中错误在所难免,恳请读者指正。

编者

2006 年 7 月

目 录

第一部分 SPSS 软件介绍

第一章 SPSS 12.0 for Windows 简介	1
1.1 SPSS 12.0 for Windows 的基本特点	1
1.2 SPSS 12.0 for Windows 的运行环境	2
1.3 SPSS 12.0 for Windows 的功能	3
1.4 SPSS 的主要窗口	4
第二章 信息的输入和输出	13
2.1 数据的输入	13
2.2 数据保存与统计结果输出	21
第三章 数据文件的编辑	33
3.1 数据的排序	33
3.2 数据的转置	35
3.3 数据的聚合	37
3.4 数据文件的拆分	41
3.5 数据文件的合并	45
3.6 数据的选择	49
3.7 统计软件 SPSS 12.0 for Windows 应用及开发指南	54

第二部分 抽样调查实验

实验一 大学生兼职状况调查报告	63
一、调查方案设计	63
二、数据分析	65
实验二 桂林工学院大学生生活状况调查分析	71
一、调查方案	71

二、问卷设计	71
三、调查报告	73
四、总结	80
实验三 大学生的人生价值取向调查	82
一、调查方案设计	82
二、数据分析	84
三、总结	88
实验四 关于外宿大学生性行为问题的调查	90
一、调查方案设计	90
二、数据分析	93
三、总结	109
实验五 关于逃课问题的调查	111
一、调查方案设计	111
二、调查分析报告	112
三、总结	118
附录 调查问卷结果汇总	118
实验六 关于毕业去向问题的调查报告	119
一、方案设计	119
二、数据整理	119
三、数据分析	119
附录	125
一、调查问卷	125
二、原始数据统计表	126
三、原始数据	127

第三部分 统计软件实验

实验一 参数估计与假设检验实验	133
一、参数估计	133

二、假设检验	134
三、非参数假设检验	137
实验二 单因素方差分析与协方差分析实验	143
一、单因素完全随机试验的方差分析	143
二、单因素随机区组试验的方差分析	148
三、单因素系统分组设计试验的方差分析	152
四、拉丁方试验的方差分析	155
五、协方差分析	159
实验三 回归分析与相关分析实验	164
一、一元直线回归与简单相关	164
二、相关分析与偏相关分析	167
三、多元线性回归与复相关	173
四、非线性回归分析	178
实验四 析因试验的方差分析实验	185
一、多因素完全随机试验的方差分析	185
二、多因素随机区组试验的方差分析	189
三、裂区与条区试验设计的方差分析	193
四、多年、多地点试验的方差分析	200
五、析因试验的方差分析	202
六、正交试验的方差分析	205

第四部分 统计综合实验

实验一 中国股民股票投资状况问卷调查与分析	211
一、实验简介	211
二、学习本实验的目的	211
三、调查表表式	212
四、数据文件说明	213
五、变量及代码说明	214
六、本实验中问卷调查的程序	217
七、需要学生讨论和解决的问题	218

附录	219
一、教学目的	219
二、本实验所采用的分析方法	219
三、实验教材中讨论题的解答要点	220
四、需要进一步讨论的问题	221
实验二 中国工业生产发展的短期预测分析	223
一、实验目的和要求	223
二、问题的提出	223
三、方法的选择	224
四、统计指标和数据的收集	224
五、方法推荐	224
六、讨论	226
实验三 财政收入多元回归分析	227
一、论文摘要	227
二、基本假设	227
三、符号说明	227
四、问题的分析和解答	228
实验四 中国城镇居民消费结构分析和需求函数估计	234
一、实验简介	234
二、关于实验数据的说明	234
三、实验所需要的预备知识	236
四、实验分析的具体步骤和需要讨论的问题	238
实验五 基尼系数——一种收入分配平均程度的测度方法	239
一、实验简介	239
二、实验数据	239
三、学习目的和要求	241
四、洛伦兹曲线	241
五、基尼系数	242
六、本实验需要讨论的几个问题	245

实验六 年龄、文化程度、婚姻状况 and 职业对人口死亡率的影响·····	246
一、实验简介·····	246
二、问题解答·····	247
三、关于本实验教学的一些建议·····	254
附录 I 上海市第四次人口普查(1990 年 7 月 1 日零时)的居民人数分类 数据·····	256
附录 II 上海市第四次人口普查(1990 年 7 月 1 日零时)的死亡人数分类 数据·····	260
参考文献·····	265

第一部分 SPSS 软件介绍

第一章 SPSS 12.0 for Windows 简介

SPSS 公司自 1972 年正式成立以来,不断推出 SPSS 软件的新版本,从最初的 SPSS/PC' for DOS 到 SPSS 6.0、7. x、8. x、9.0、10. x、11. x、12. x for Windows。随着版本的不断更新,软件功能不断完善,操作越来越简便,与其他软件的接口也越来越多。现在的 SPSS 软件,不仅能实现统计功能,还能将分析结果用图表和地图来表示,真正做到了实用与美观的统一。

1.1 SPSS 12.0 for Windows 的基本特点

SPSS 12.0 for Windows 具有下面一些显著的特点。

- (1)使用该软件,只要给出分析指令,系统便自动进行数据处理,得到相应的结果,从而使数理统计的计算过程大大简化。
- (2)统计功能强大,选用不同的模块可以完成基本统计任务和专业统计任务。
- (3)完全的 Windows 风格界面,输入数据文件以后,只需用鼠标结合简单的数据输入便可完成操作。
- (4)完善的帮助系统(包括图解帮助、在线帮助和联机帮助等),使用户能更好地自学该软件。
- (5)具有丰富的图表表达功能,可以根据需要选择不同风格、不同功能的表和图来整理数据,使报告、论文等更加简练、美观。
- (6)可以与很多其他软件进行数据传输。该软件可以打开扩展名为 .dat、.xls、.slk、.dbf 和 .wk3 等多种数据文件,该软件生成的图形可以保存为多种图形格式。
- (7)可以生成主题地图,表示不同地理区域的数据分布特征。
- (8)单击对话框中的“Paste”按钮,可以自动生成“Syntax”程序代码,经过简

单的编辑或直接运行,就可以实现指定的功能。该功能使熟悉编程的用户节省很多工作量。

(9)与 Excel、Word 等软件一样,该软件具有内置的 VBA 客户语言——SaxBasic 语言。该语言与 Syntax 命令行语言混合编程,可以提高工作效率。利用 SPSS 的类库对象,可以实现对 SPSS 主要窗口和图表对象的操作。

(10)该软件支持 DLL、DDE、OLE 和 ActiveX 技术等多种接口方式,可以通过接口与其他软件协同工作来增加 SPSS 的功能。

1.2 SPSS 12.0 for Windows 的运行环境

安装 SPSS 12.0 for Windows 以后,在 Windows 界面下按照“开始”→“程序”SPSS for Windows→SPSS 12.0 for Windows 的顺序选择菜单选项,进入 SPSS 12.0 for Windows 主界面,如图 1-1-1 所示。

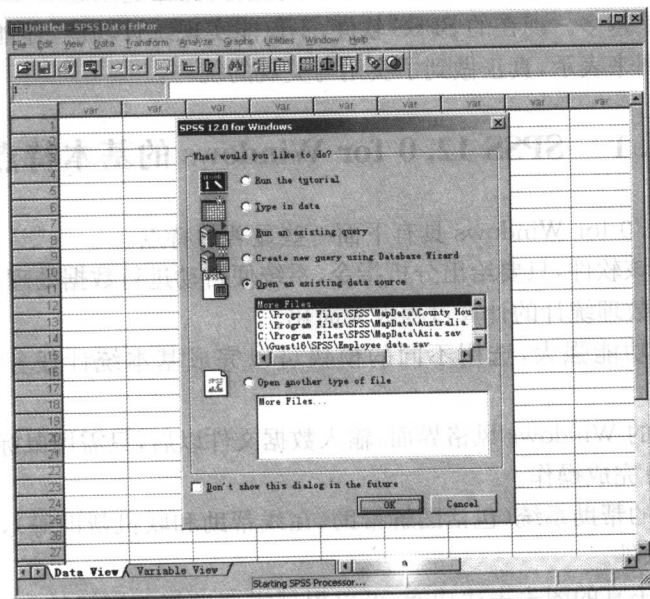


图 1-1-1 SPSS 12.0 for Windows 主界面

刚安装 SPSS 软件时,如果不是直接从“开始”菜单的“文档”次级菜单中打开 SPSS 数据文件(扩展名为“.sav”),进入 SPSS 主环境时会显示一个对话框,提示下一步准备什么,进行选择可以快捷地完成任务。程序打开时显示的是数据编辑器主界面(如图 1-1-1 所示),界面比较简洁,主要包括菜单、工具条、数据编辑区、变量定义区和状态栏。在该界面中,在变量定义区定义变量,在数据

编辑区输入数据。直接输入数据或将已经存在的数据导入到数据编辑器以后，就可以进行绘图和统计分析等操作了。

1.3 SPSS 12.0 for Windows 的功能

SPSS 12.0 for Windows 可以实现数据编辑、图表生成和编辑、主题地图制作、与其他程序接口以及各种统计分析功能。

1.3.1 SPSS 的数据编辑功能

在 SPSS 的数据编辑器窗口中，不仅可以对已打开的数据文件进行增加、删除、复制、剪切和粘贴等常规操作，还可以对数据文件中的数据进行排序、转置、拆分、聚合、加权等操作，对多个数据文件可以根据变量或个案进行合并。根据需要还可以把将要分析的变量集中到一个集合中，打开时指定打开该集合，而不必打开整个数据文件。

1.3.2 表格的生成和编辑

利用 SPSS 可以生成数十种风格的表格，在专门的编辑窗口或直接在查看器中编辑所生成的表格。

1.3.3 图形的生成和编辑

利用 SPSS 可以生成数十种基本图和交互图。其中基本图包括条形图、线图、面积图、饼图、高低图、帕累托图、控制图、箱图、误差条图、散点图、直方图、ROC 曲线图、P-P 概率图、Q-Q 概率图、序列图和时间序列图等，有的基本图中又可进一步细分。交互图比基本图更漂亮，可以对组成交互图的图形元素进行交互编辑并有不同风格的二维图、三维图。交互图包括条形交互图、点形交互图、线形交互图、带形交互图、饼形交互图、箱形交互图、误差条形交互图、直方交互图和散点交互图等。图形生成以后，可以进行编辑。除了基本图和交互图外，SPSS 12.0 for Windows 还提供了主题地图的制作和编辑功能，利用 SPSS 可以生成范围值地图、点密度地图、分级标记地图、单值地图、条形图地图、饼图地图和多主题地图等。

1.3.4 与其他程序交换数据

SPSS 能打开 Excel、dBASE、FoxBase、Lotus 1-2-3、Access 和文本编辑器等生成的数据文件。SPSS 生成的图形可以保存为 BMP、JPG 等多种图形

格式。

1.3.5 SPSS 的统计功能

SPSS 的统计功能是 SPSS 的核心部分,其基本统计功能包括以下几种:

(1)概率分布;(2)样本数据的描述和预处理;(3)假设检验(包括参数检验、非参数检验及其他检验);(4)方差分析(包括一般的方差分析和多元方差分析);(5)列联表分析;(6)相关分析;(7)回归分析;(8)聚类分析;(9)判别分析;(10)因子分析。

除了这些基本功能外,SPSS 还提供了其他高级和专门分析模块,可以实现更为复杂的统计任务,如混合模型、对数线性模型、信度分析和生存分析等。

1.3.6 二次开发功能

SPSS 12.0 for Windows 有比较完备的二次开发功能。它不但有源于 DOS 时期的 Syntax 命令集,还有内置的 VBA 客户语言——SaxBasic 语言。在 SPSS 的脚本编辑器中混合使用 SaxBasic 和 Syntax,可以很方便地实现脚本的编制,从而能提高统计分析的效率。除了能编制内部脚本,SPSS 12.0 for Windows 还支持 DLL 技术,调用 Windows 核心库或外部库中的 API 函数。它还支持 DDE、OLE 和 ActiveX 等技术,可以与其他同样支持该技术的程序进行接口。

1.4 SPSS 的主要窗口

SPSS 中的主要窗口有数据编辑器、语法编辑器、输出查看器、草稿输出器和脚本编辑器等 5 种。

1.4.1 数据编辑器

进入 SPSS,系统自动打开数据编辑器,如图 1-1-2 所示。数据编辑器包括 Data View 和 Variable View 两个视区,分别定义变量的值(即数据)和变量(名称、类型等)。打开数据编辑器以后,首先在“Variable View”视区中定义变量信息(如图 1-1-2 所示),在该视区中定义变量包括 Name、Type 和 Width 等方面的内容。具体说明如下。

(1)Name:为变量名。SPSS 中变量名的定名规则与其他软件大同小异,具体而言,有以下一些规定。

①变量名的第一个字符必须为字母,后面可跟任意字母、数字、句点或@、#、\$ 等符号;

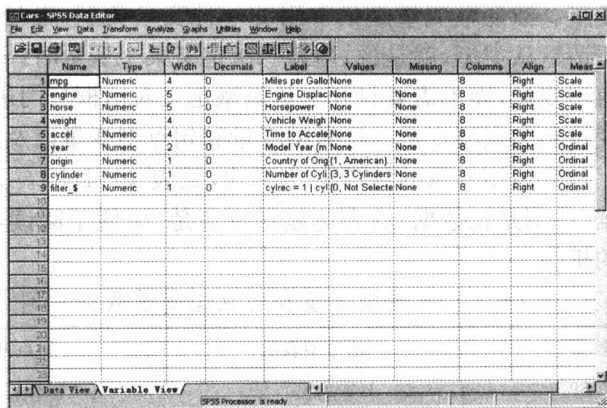


图 1-1-2 数据编辑器

②变量名不能以句点结尾；

③定名时应避免最后一个字符为下划线“_”(因为某些过程运行时自动创建的变量名的最后一个字符有可能为下划线,这样有可能造成冲突)；

④变量名的长度一般不能超过 8 个字符；

⑤空格和特殊字符(如!、?、’和*等)不能用于变量名；

⑥每个变量名必须保证是唯一的,不区分大小写；

⑦下面的关键词不能用作变量名: ALL、NE、EQ、TO、LE、LT、GE、BY、OR、GT、AND、NOT、WITH。

(2)Type:为变量类型。单击该列中的单元,将弹出“Variable Type”对话框,如图 1-1-3 所示。对话框中列出了 8 种可选的基本变量类型:数值型(Numeric)、逗号型(Comma)、句点型(Dot)、科学计数型(Scientific notation)、日期格式型(Date)、美元型(Dollar)、定制货币型(Custom currency)和字符串型(String)。只要单击要定义的变量类型的标签或标签前面的单选按钮,便可进行变量类型的定义。

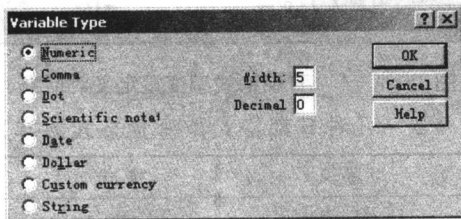


图 1-1-3 变量类型定义对话框

现对 8 种变量类型分述如下。

①数值型变量。数值型变量是最常用的一种变量类型,也是 SPSS 中变量的默认类型。所谓默认类型,就是说,当第一次定义某个变量时,SPSS 自动将该变量设置为数值型变量,而不需要打开变量类型定义对话框专门进行设置。

②逗号型变量和句点型变量。逗号型变量和句点型变量以及后面的科学计数型变量实际上是数值型变量的特殊表达形式,见表 1-1-1。采用这些表达形式可以使数据(特别是长数据)的表达显得更清楚、更便于阅读,因此,在应用中可减少出错的机会。

表 1-1-1 逗号型变量和句点型变量数据示例

原变量数据	逗号型表示	句点型表示
1234567.89	1,234,567.89	1.234.567,89

逗号型变量用句点来分隔数据的整数部分和小数部分,而整数部分从个位开始往前逢 3 的进位使用逗号分隔开;句点型变量则反过来,用逗号分隔数据的整数部分和小数部分,用句点来分隔整数部分(逢 3 的进位)。

③科学计数型变量。科学计数型数据实际上是用科学计数法来记录的数据。它是使用 a.bEc 的形式来表达的,其中 a 为原数据的最高位数字,b 为原数据除去最高位数据以外余下的数字,c 为原数据最高位 10 的方次数。

④日期格式型变量。日期格式型变量以及后面的几种变量的定义对话框均与前面的有所不同。图 1-1-4 是日期格式型变量定义对话框,框中列出了各种日期或时间格式,用户只要用鼠标进行选择就行了。表 1-1-2 列出了几种日期格式型变量格式的具体意义。

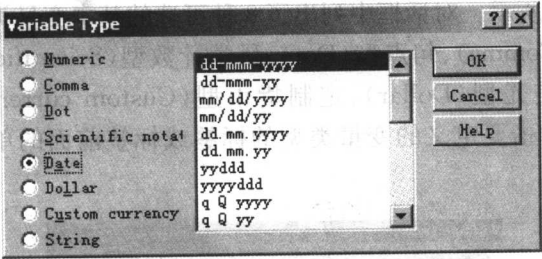


图 1-1-4 日期格式型变量定义对话框

表 1-1-2 几种日期格式型变量格式的意义说明

格式	意义	格式	意义
dd-mmm-yyyy	日-月(英)-年	dd.mm.yy	日.月.年(简称)
dd-mmm-yy	日-月(英)-年(简称)	yyddd	年(简称)天数
mm/dd/yy	月/日/年(简称)	QQyy	季度 Q 年(简称)

(续)

格式	意义	格式	意义
mmmyyyy	月(英)年	hh:mm	时:分
wwWKyyyy	周数 WK 年	hh:mm:ss	时:分:秒
MANDA,TUESDAY...	用英文单词表示的星期几	ddd hh:mm	日(英) 时:分
JAN,FEB,MAR...	用英文单词表示的月份	ddd hh:mm:ss	日(英) 时:分:秒
dd-mm-yyyy-hh:mm	日-月(英)-年时:分	ddd hh:mm:ss.ss	日(英) 时:分:秒. 百分秒
注:上表中括号内的注释项中,英——用英文单词表示;简称——只取年号的最后两个数字,如 1999 年简称为 99 年			

⑤美元型变量。美元型变量是在逗号型变量前加上美元符号“\$”。如图 1-1-5 所示,在美元型变量定义对话框中可以选择格式,也可以自己定义格式。列表框中提供了多种固有格式,可以根据需要进行选择。比如说,某职员年薪为 5678.91 美元,如果记录时希望保留两位小数,可以选择“\$#,###.##”格式;如果不希望保留小数,可以选择“\$#,###”格式。

如果列表框中没有满意的数据格式,可以自己设置。设置的方法与数值型变量大致相同,即在“Width”窗口和“Decimal”窗口中输入数字;但“Width”窗口中的数字必须比“Decimal”窗口中的数字大 2,否则出错。

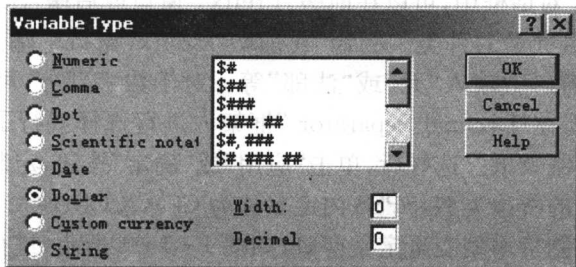


图 1-1-5 美元型变量定义对话框

⑥定制货币型变量。尽管 SPSS 只提供了美元型变量来作为现成的货币描述量,但可以通过自己定制来定义自己的货币型变量。在进行定制之前,先在数据编辑器中增加一个变量,如图 1-1-6 所示,在编辑器的最后一列增加“price”变量,用来描述汽车的价格。在这里,数据类型是按照默认类型——数值型定义的。

	mpg	engine	horse	weight	price
1	18	307	130	3504	1790000
2	15	350	165	3693	1900000
3	18	318	150	3436	1890000

图 1-1-6 增加“price”变量