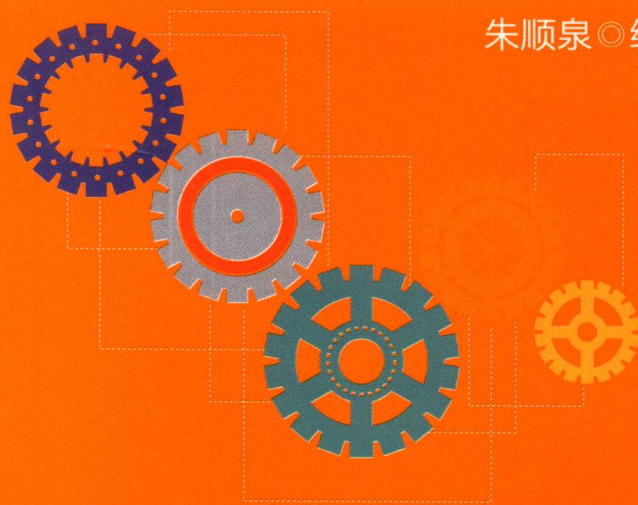


21世纪经济与管理精编教材·经济学系列

# Stata数据 分析应用

Data Analysis and  
Application Using Stata

朱顺泉◎编著



北京大学出版社  
PEKING UNIVERSITY PRESS

21世纪经济与管理精编教材·经济学系列

# Stata数据 分析应用

Data Analysis and  
Application Using Stata

朱顺泉◎编著



北京大学出版社  
PEKING UNIVERSITY PRESS

## 图书在版编目(CIP)数据

Stata 数据分析应用/朱顺泉编著. —北京:北京大学出版社,2015.9

(21世纪经济与管理精编教材·经济学系列)

ISBN 978-7-301-25932-0

I. ①S… II. ①朱… III. ①统计分析—应用软件—高等学校—教材 IV. ①C819

中国版本图书馆CIP数据核字(2015)第125008号

书 名 Stata 数据分析应用  
Stata Shuju Fenxi Yingyong

著作责任者 朱顺泉 编著

责任编辑 周莹

标准书号 ISBN 978-7-301-25932-0

出版发行 北京大学出版社

地 址 北京市海淀区成府路205号 100871

网 址 <http://www.pup.cn>

电子信箱 [em@pup.cn](mailto:em@pup.cn) QQ:552063295

新浪微博 @北京大学出版社 @北京大学出版社经管图书

电 话 邮购部 62752015 发行部 62750672 编辑部 62752926

印刷者 北京大学印刷厂

经 销 者 新华书店

787毫米×1092毫米 16开本 19.25印张 480千字

2015年9月第1版 2015年9月第1次印刷

印 数 0001—3000册

定 价 39.00元

---

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究

举报电话:010-62752024 电子信箱: [fd@pup.pku.edu.cn](mailto:fd@pup.pku.edu.cn)

图书如有印装质量问题,请与出版部联系,电话:010-62756370

# 前 言

在当今大数据时代的背景下,数据已经成为商务决策最为重要的参考依据之一,数据分析行业已经迈入了一个全新的阶段。《Stata 数据分析应用》侧重于 Stata 软件的数据存取、图形展示和统计与计量经济数据分析,重点介绍了 Stata 软件在统计学与计量经济分析中的应用,同时结合大量的实例,通过大量的丰富的实例,对 Stata 软件进行科学、准确和全面的介绍,以便使读者能深刻理解 Stata 的精髓和灵活、高效的使用技巧。

Stata 软件是金融数据分析、宏观经济预测、销售预测和成本分析等领域应用非常广泛的统计与计量软件之一,它具有易操作、运算速度快、绘图功能强、更新和发展速度快等特点,因而受到广大用户的欢迎和喜爱。本书将通过丰富的实例,详细介绍 Stata 12.0 在现实生活中的应用。

本书注重理论与应用相结合,实例丰富且通俗易懂,实用性、实验和可操作性特别强,详细介绍了各种统计和计量经济方法在 Stata 软件中的实现过程。

本书的内容安排如下:第 1 章介绍 Stata 数据处理;第 2 章介绍 Stata 图形的绘制;第 3 章介绍 Stata 描述统计;第 4 章介绍 Stata 参数假设检验;第 5 章介绍 Stata 相关分析;第 6 章介绍 Stata 最小二乘线性回归分析;第 7 章介绍 Stata 因变量受限回归分析;第 8 章介绍 Stata 异方差计量检验与应用;第 9 章介绍 Stata 自相关计量检验与应用;第 10 章介绍 Stata 多重共线性计量检验与应用;第 11 章介绍 Stata 时间序列分析;第 12 章介绍 Stata 面板数据分析。

本书实例经典,内容丰富,有很强的针对性,书中各章详细地介绍了实例的 Stata 具体操作过程,读者只需按照书中介绍的步骤进行操作,就能掌握全书的内容。为了帮助读者更加直观地学习本书,我们将书中实例的全部数据文件制成电子包,读者可来信索取。

本书作为广东省研究生教育创新计划项目、广东财经大学 MBA 中心数据模型与决策示范课程的阶段性成果之一,适合作为经济学、金融学、统计学等相关专业的本科生或研究生学习统计学、计量经济学等课程的教材或实验参考用书,同时对从事数据分析的实际工作者也大有裨益。

本书的出版,得到了北京大学出版社周莹编辑的支持、帮助,应该感谢她为读者们提供了这么好的一个工具!由于时间和水平的限制,书中疏漏之处,恳请读者批评指正。

作 者

2015 年 5 月于广州

# 目 录

## 第 1 篇 Stata 应用基础

第 1 章 Stata 基本窗口、变量与数据处理 .....	(3)
1.1 Stata 12.0 基本窗口说明 .....	(3)
1.2 Stata 12.0 数据文件的创建与读取 .....	(4)
1.2.1 Stata 12.0 数据文件的创建 .....	(4)
1.2.2 Stata 12.0 数据文件的读取及转换 .....	(8)
1.3 创建和替代变量 .....	(10)
1.4 分类变量和定序变量操作 .....	(15)
1.5 数据的基本操作 .....	(19)
1.6 定义数据的子集 .....	(25)
第 2 章 Stata 图形的绘制 .....	(29)
2.1 直方图的绘制 .....	(29)
2.2 散点图的绘制 .....	(33)
2.3 曲线标绘图的绘制 .....	(36)
2.4 连线标绘图的绘制 .....	(40)
2.5 箱图的绘制 .....	(43)
2.6 饼图的绘制 .....	(46)
2.7 条形图的绘制 .....	(49)
2.8 点图的绘制 .....	(52)

## 第 2 篇 Stata 统计应用

第 3 章 Stata 描述统计 .....	(57)
3.1 描述统计基本理论 .....	(57)
3.1.1 数据分析的基本描述统计 .....	(57)
3.1.2 常用的统计分布 .....	(64)

3.2	定距变量的描述性统计 .....	(67)
3.3	正态性检验和数据转换 .....	(71)
3.4	单个分类变量的汇总 .....	(76)
3.5	两个分类变量的列联表分析 .....	(77)
3.6	多表和多维列联表分析 .....	(80)
<b>第4章</b>	<b>Stata 参数假设检验 .....</b>	<b>(85)</b>
4.1	参数假设检验的基本理论 .....	(85)
4.1.1	区间估计 .....	(85)
4.1.2	假设检验 .....	(86)
4.2	单个样本 $t$ 检验应用 .....	(95)
4.3	两个独立样本 $t$ 检验应用 .....	(97)
4.4	配对样本 $t$ 检验应用 .....	(100)
4.5	单一样本方差的假设检验应用 .....	(103)
4.6	双样本方差的假设检验应用 .....	(105)
<b>第5章</b>	<b>Stata 相关分析 .....</b>	<b>(108)</b>
5.1	简单相关分析的基本理论 .....	(108)
5.2	简单相关分析的基本应用 .....	(109)
5.3	偏相关分析的基本理论 .....	(112)
5.4	偏相关分析的基本应用 .....	(112)
<b>第6章</b>	<b>Stata 最小二乘线性回归分析 .....</b>	<b>(115)</b>
6.1	一元线性回归分析基本理论 .....	(115)
6.1.1	一元线性回归分析模型 .....	(115)
6.1.2	一元线性回归的假设 .....	(115)
6.1.3	方差分析 .....	(116)
6.1.4	决定系数 .....	(116)
6.1.5	估计的标准误 .....	(116)
6.1.6	回归系数的假设检验 .....	(117)
6.1.7	回归系数的置信区间 .....	(117)
6.2	一元线性回归分析的应用 .....	(118)
6.3	多元线性回归分析基本理论 .....	(126)
6.3.1	多元线性回归模型 .....	(126)
6.3.2	方差分析 .....	(127)
6.3.3	决定系数 .....	(127)
6.3.4	估计的标准误 .....	(127)
6.3.5	回归系数的 $t$ 检验和置信区间 .....	(127)
6.3.6	回归系数的 $F$ 检验 .....	(128)

6.3.7	虚拟变量	(129)
6.4	多元线性回归分析的应用	(129)
<b>第7章</b>	<b>Stata 因变量受限回归分析</b>	(139)
7.1	断尾回归分析	(139)
7.1.1	断尾回归分析的概念	(139)
7.1.2	在简单线性回归分析中的应用	(139)
7.2	截取回归分析	(146)
7.2.1	截取回归分析的概念	(146)
7.2.2	截取回归分析的应用	(146)

## 第3篇 Stata 计量经济应用

<b>第8章</b>	<b>Stata 异方差计量检验与应用</b>	(157)
8.1	回归模型的异方差计量检验基本理论	(157)
8.1.1	异方差的概念	(157)
8.1.2	异方差产生的原因	(159)
8.1.3	异方差的后果	(160)
8.1.4	异方差的识别检验	(160)
8.1.5	消除异方差的方法	(162)
8.2	回归模型的异方差计量检验的应用	(164)
<b>第9章</b>	<b>Stata 自相关计量检验与应用</b>	(180)
9.1	回归模型的自相关计量检验基本理论	(180)
9.1.1	自相关的概念	(180)
9.1.2	产生自相关的原因	(180)
9.1.3	自相关的后果	(182)
9.1.4	自相关的识别和检验	(182)
9.1.5	自相关的处理方法	(184)
9.2	回归模型的异方差计量检验的应用	(186)
<b>第10章</b>	<b>Stata 多重共线性计量检验与应用</b>	(198)
10.1	回归模型的多重共线性计量检验基本理论	(198)
10.1.1	多重共线性的概念	(198)
10.1.2	多重共线性的后果	(199)
10.1.3	产生多重共线性的原因	(199)
10.1.4	多重共线性的识别和检验	(200)
10.1.5	消除多重共线性的方法	(201)
10.2	回归模型的多重共线性计量检验的应用	(203)

<b>第 11 章 Stata 时间序列分析</b> .....	(215)
11.1 时间序列分析的基本理论 .....	(215)
11.1.1 平稳、协整、因果检验的基本概念 .....	(215)
11.1.2 单位根检验 .....	(215)
11.1.3 协整检验 .....	(217)
11.1.4 误差修正模型 .....	(218)
11.2 时间序列分析的基本应用 .....	(218)
11.3 单位根检验 .....	(227)
11.4 协整检验 .....	(236)
11.5 格兰杰因果关系检验 .....	(244)
<b>第 12 章 Stata 面板数据分析</b> .....	(253)
12.1 面板数据分析的基本理论 .....	(253)
12.2 短面板数据分析的基本应用 .....	(254)
12.3 长面板数据分析的基本应用 .....	(277)
<b>参考文献</b> .....	(298)



# 第 1 篇

Stata 应用基础



# 第1章 Stata基本窗口、变量与数据处理

Stata 是一种功能全面的统计与计量软件包,是目前最为流行的统计与经济计量软件包之一。它具有易操作、运算速度快、绘图功能强等特点。它不仅包括一套事先编制好的数据分析功能,同时还允许用户根据自己的需要来创建程序。该软件被引入中国后,得到了广大学者的认可,其适用范围越来越广。本章主要对 Stata 12.0 的基本窗口、变量与数据管理作一些初步介绍。

## 1.1 Stata 12.0 基本窗口说明

将 Stata 12.0 安装在 E 盘(也可以是其他盘)的 stata12 目录上,在 stata12 上创建一个 zsq 目录(命名可随意),将光盘上的所有文件复制在 E:\stata12\zsq 目录中。

在正确安装 Stata 12.0 后,双击 Stata 主程序的图标文件,即可打开 Stata 的主界面,如图 1-1 所示。

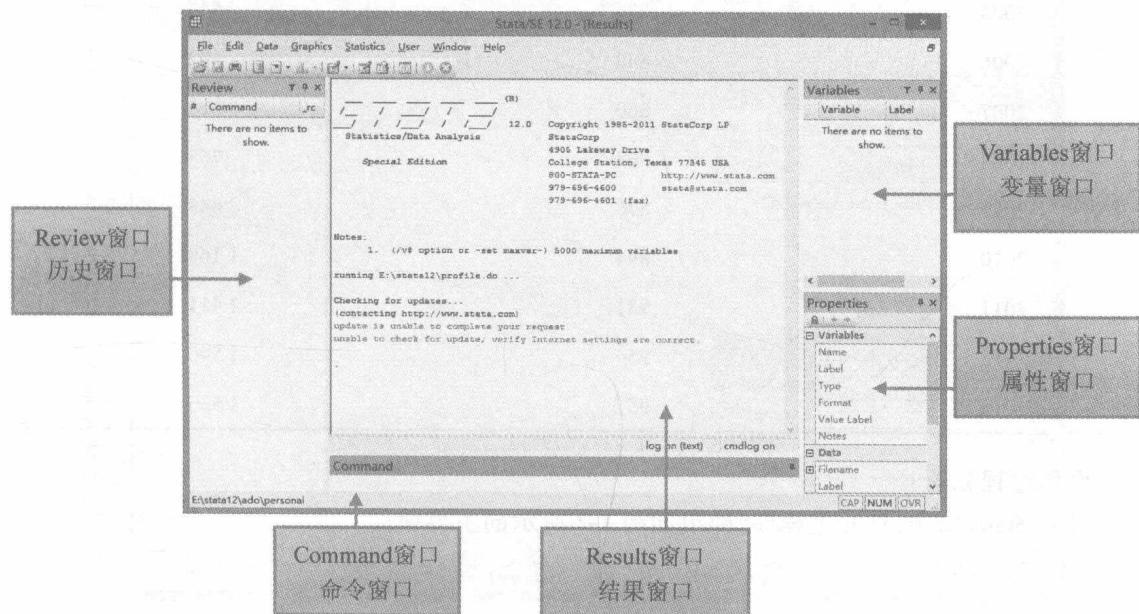


图 1-1 Stata 12.0 主界面

在图 1-1 中,有 Review、Command、Results、Variables、Properties 5 个区域:Review 窗口(历史窗口)显示的是自本次启动 Stata 12.0 以来执行过的所有命令;Command 窗口(命令窗口)是最重要的窗口,它的作用是在本窗口内输入准备执行的命令;Results 窗口(结果窗口)显示的是每次执行 Stata 命令后的执行结果,无论成功还是失败;Variables 窗口(变量窗口)显示的是当前 Stata 数据文件中的所有变量;Properties 窗口(属性窗口)显示的是当前数据文件中设定的变量及数据的性质。

## 1.2 Stata 12.0 数据文件的创建与读取

### 1.2.1 Stata 12.0 数据文件的创建

**例 1-1** 表 1-1 是我国 2000—2013 年上市公司数量的数据,创建 Stata 12.0 格式的数据文件并保存。

表 1-1 我国 2000—2013 年上市公司数量

年份	上交所上市公司数	深交所上市公司数
2000	572	516
2001	646	514
2002	715	509
2003	780	507
2004	837	540
2005	834	547
2006	842	592
2007	860	690
2008	864	761
2009	870	848
2010	894	1 169
2011	931	1 411
2012	954	1 540
2013	957	1 536

操作过程如下:

进入 Stata 12.0,打开主程序,弹出如图 1-2 所示的主界面。

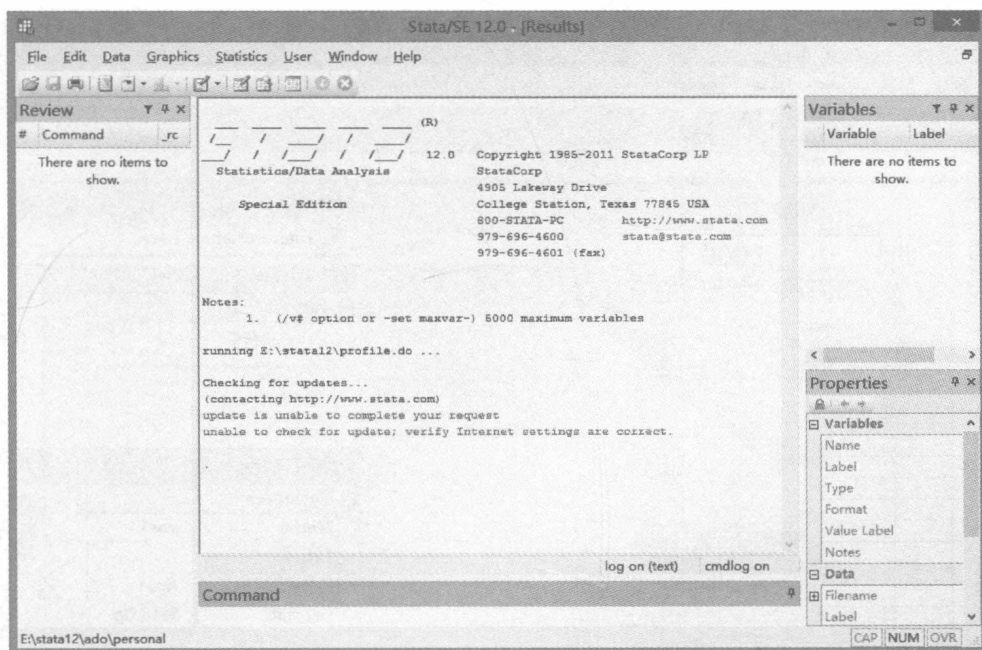


图 1-2 程序主界面

选择“Data”|“Data Editor”|“Data Editor( Edit)”,弹出如图 1-3 所示的“Data Editor( Edit)”对话框。

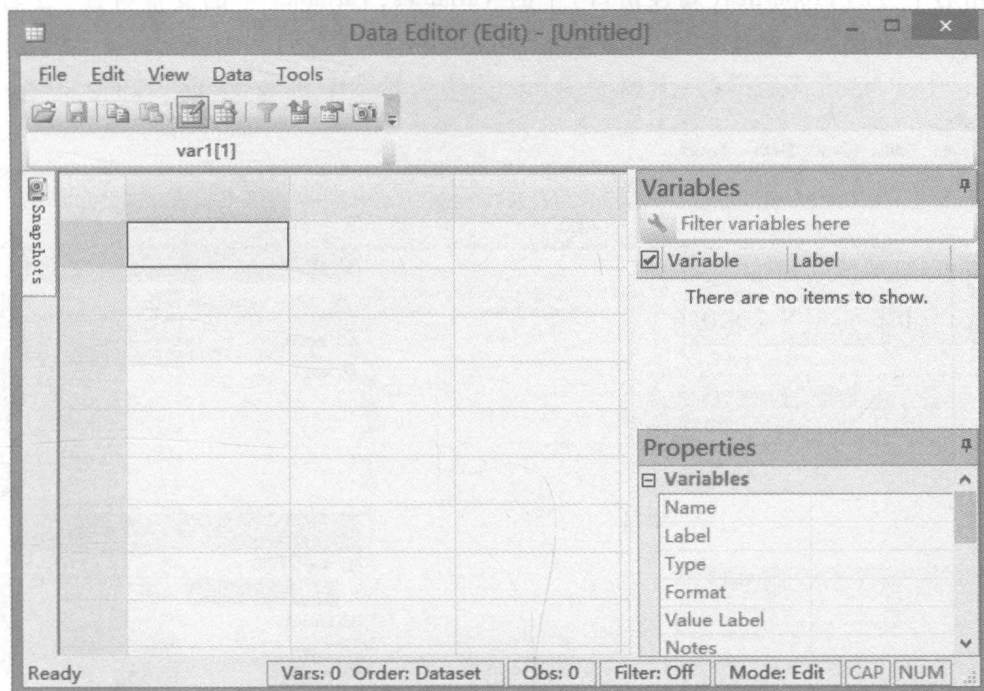


图 1-3 “Data Editor( Edit)”对话框(1)

在“Data Editor( Edit)”对话框左上角的单元格中输入第 1 个数据“2 000”,系统即自动创建“var1”变量,如图 1-4 所示。

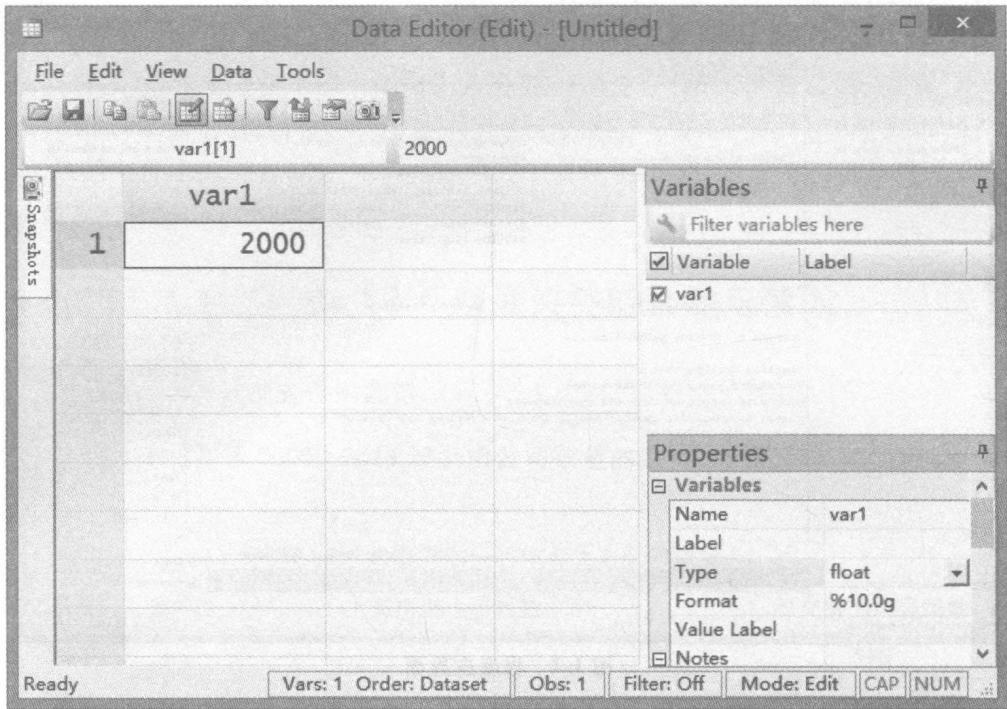


图 1-4 “Data Editor (Edit)”对话框 (2)

单击右下方的 Properties (属性窗口) 中的 Variables, Variables 中的变量特征 (含名称、类型、长度等) 即可进入可编辑状态, 如图 1-5 所示。

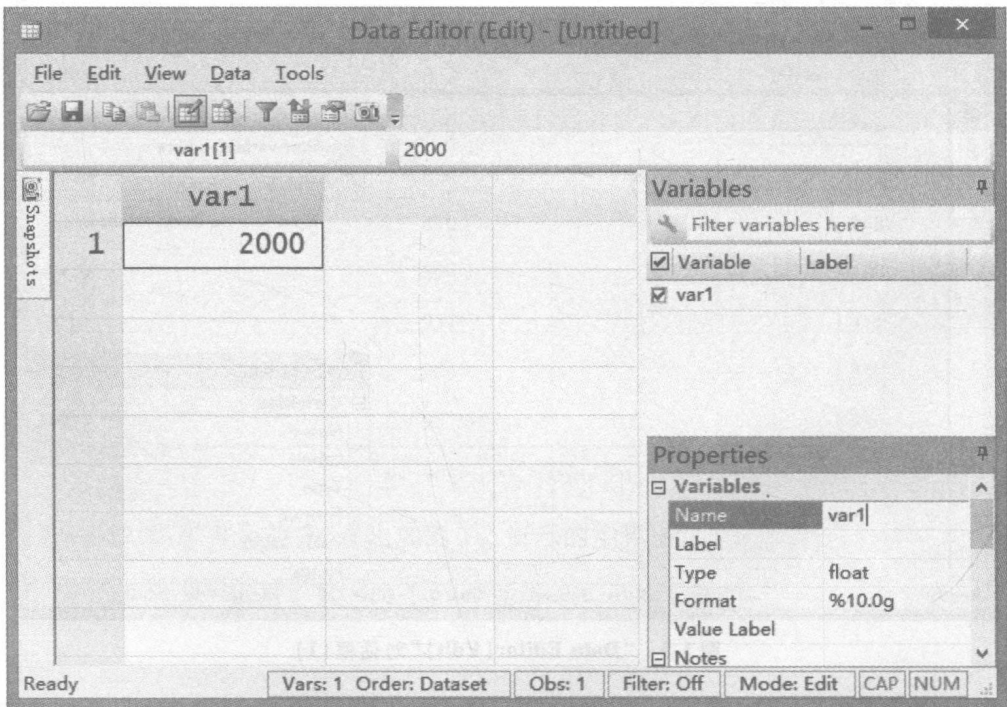


图 1-5 “Data Editor (Edit)”对话框 (3)

对变量名称进行必要的修改,第1个变量是年份,可把“var1”修改为“year”,其他采取系统默认设置,修改完成后在左侧数据输入区域单击,即可弹出如图1-6所示的对话框。

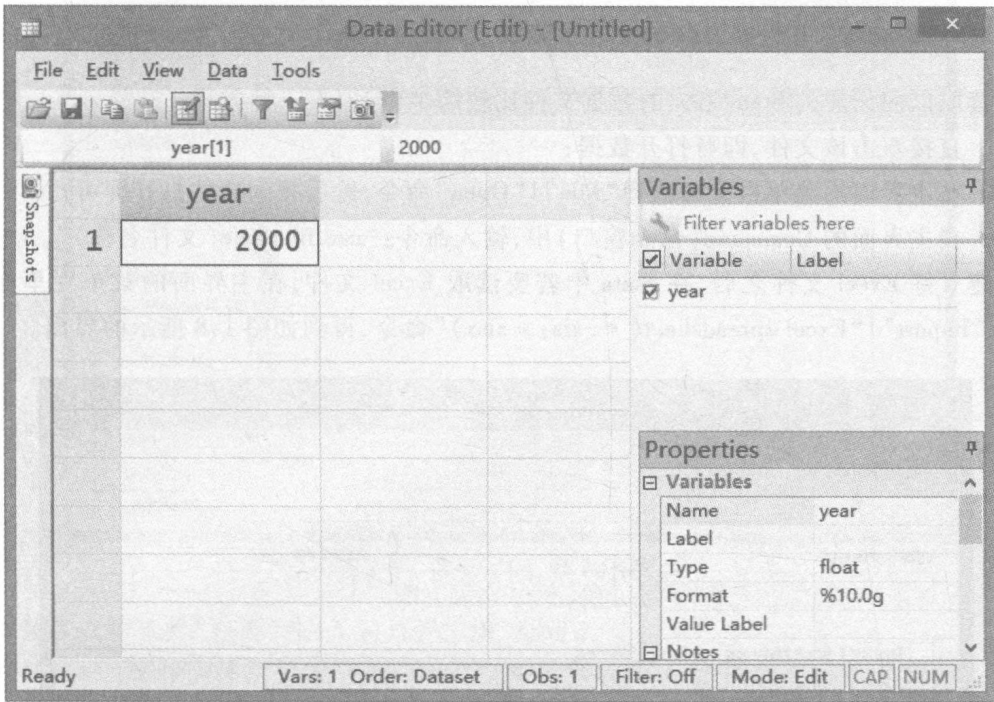


图 1-6 “Data Editor (Edit)”对话框(4)

逐个录入其他数据,其他两个变量参照“year”进行设置,将其定义为“shangjiao”“shen-jiao”,数据录完后,如图1-7所示。

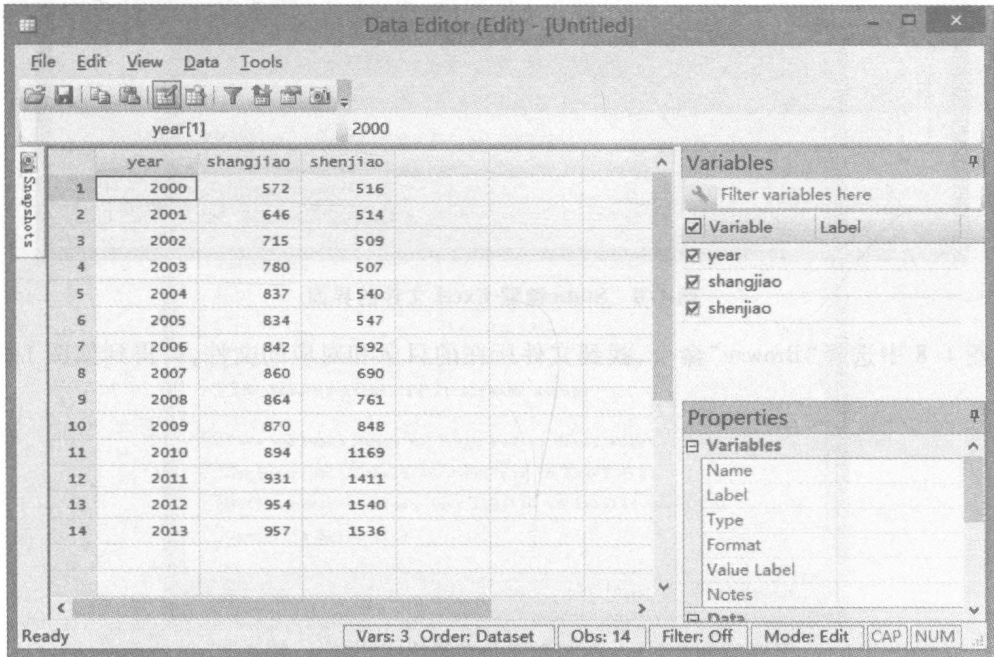


图 1-7 “Data Editor (Edit)”对话框(5)

关闭“Data Editor(Edit)”对话框,在主界面的工具栏里单击 Save 把数据保存在 E:\stata12\zsq\chap01 中,文件名为 all-1. dta。

### 1.2.2 Stata 12.0 数据文件的读取及转换

要读取以前创建的 Stata 格式的数据文件比较简单,有以下三种方式:

- (1) 直接双击该文件,即可打开数据;
- (2) 在主界面的菜单栏里面选择“File”|“Open”命令,然后找到文件打开即可;
- (3) 在主界面的 Command(命令窗口)中,输入命令:“use filename(文件名称)”。

在建立好 Excel 文件之后,在 Stata 中若要读取 Excel 文件,在主界面的菜单栏里面选择“File”|“Import”|“Excel spreadsheet(\*.xls;\*.xlsx)”命令,得到如图 1-8 所示的界面。

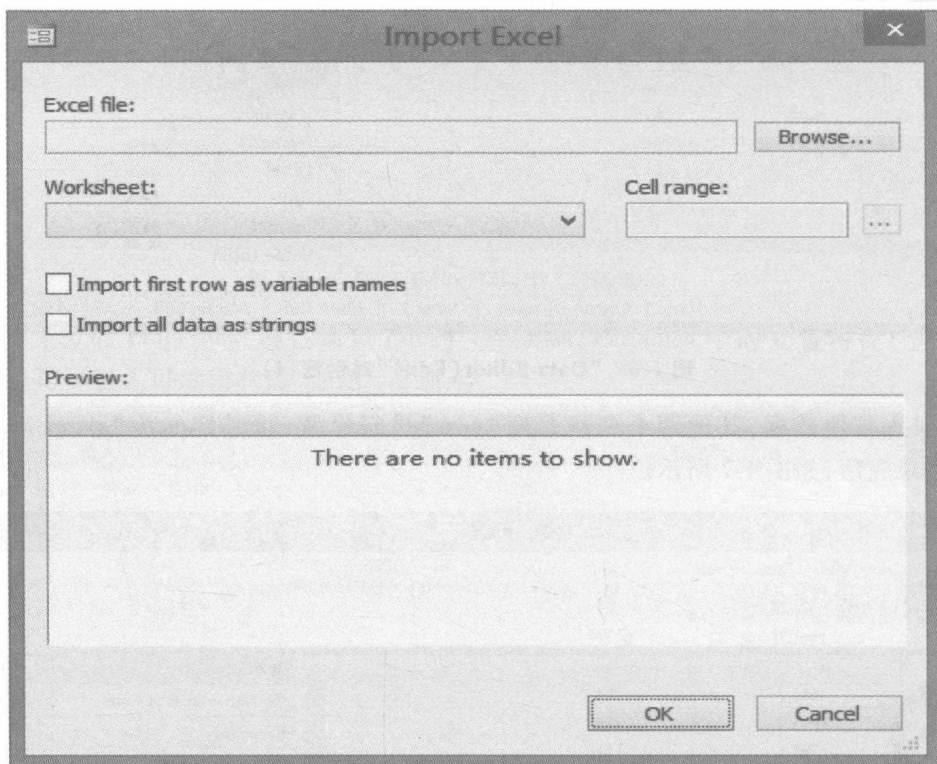


图 1-8 Stata 读取 Excel 文件的界面

在图 1-8 中选择“Browse”命令,找到文件所在的目录和对应的文件,可得到如图 1-9 所示的界面。



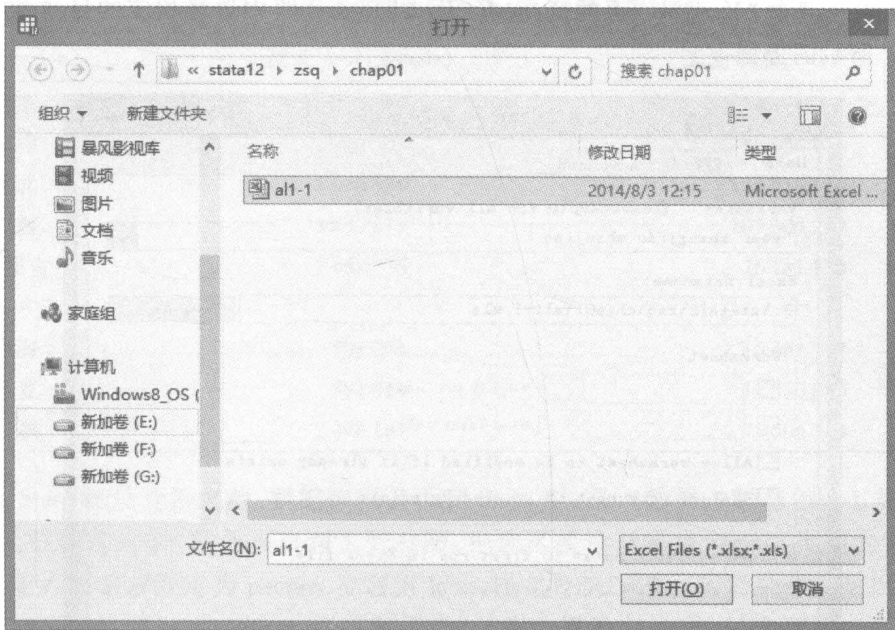


图 1-9 读取 al1-1.xls 文件数据

在图 1-9 中选择“打开”按钮,再选择“OK”即可。

在建立好 Excel 文件之后,若要把 Excel 文件转换成 Stata 数据文件也可用如下命令实现:

```
import excel "E:\stata12\zsq\chap01\al1-1.xls", sheet("Sheet1") clear
```

若要把 Stata 数据文件转换成 Excel 文件,先把 Stata 文件打开,命令如下:

```
use "E:\stata12\zsq\chap01\al1-1.dta", clear
```

然后在主界面的菜单栏里面选择“File”|“Export”|“Excel spreadsheet(\*.xls;\*.xlsx)”命令,得到如图 1-10 所示的界面。

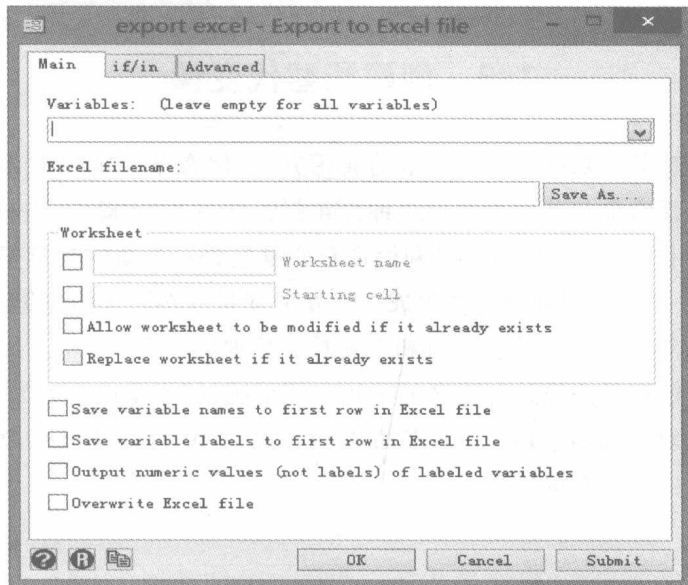


图 1-10 把 Stata 数据文件转换成 Excel 文件初始界面