

[lianxh.cn](http://lianxh.cn)

# Stata 101 - 入门指南


**连玉君 (中山大学)**

[arlionn@163.com](mailto:arlionn@163.com)

最新版: [网页版](#) | [PDF 版](#)



## 在线课程和资料

- 连玉君, Stata 33 讲: b 站视频, 课件
- 连玉君老师经验分享: 一个资深 Stata 用户的若干思考 
- 🍏 连享会推文: Stata入门 | Stata资源 | Stata程序
- DATA ANALYSIS NOTES: LINKS AND GENERAL GUIDELINES
- Internet Guide to Stata
- SSCC 在线课程
- GitHub
  - 提供了大量用 Stata 完成的项目的完整数据和 dofile 等。
  - 可以在 [github.com](https://github.com), 中搜索 stata 。
  - 点击 `fork` 按钮, 克隆之 → Clone 到国内 码云 ([gitee.com](https://gitee.com)) 账户

## Stata 基本功能:

[stata.com/features](https://www.stata.com/features) 🍏 🍓 务必通读 [U] User Guide

- `help help` // [R] help 🍏 `help search` // [R] search
- `help language` // [U] 11 Language syntax
- `help data_management` // [D] data\_management
- `help program` // [P] program
- `help graph` // [G-2] graph 🍏 `help graph intro` // [G-1] graph intro
- `help estcom` // 统计和回归分析 [U] 27 Overview of Stata estimation commands
- [U] 28 Commands everyone should know, 你必须知道的 Stata 命令

Stata/MP 17.0 - D:\stata17\ado\base\a\auto.dta

文件(F) 编辑(E) 数据(D) 图形(G) 统计(S) 用户(U) 窗口(W) 帮助(H)

可以在这里查看手册目录, 搜索等

历史窗口

过滤命令

# 命令 | \_rc

1.. cls

1.. dis \_n \_n \_n \_n

1.. sysuse "auto.dta", clear

1.. reg mpg wei len price ...

1.. sum price wei len

已经执行过的命令  
双击可再次执行

当前工作路径  
输入 cd 命令可以查看  
输入 cd D:/data 可更改路径

dofile 浏览数据

显示/关闭特定窗口

鼠标右击任何窗口, 都可以调整字号、颜色等

```

. sysuse "auto.dta", clear
(1978 automobile data)

. reg mpg wei len price foreign, noheader

```

	mpg	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
weight		-0.004	0.002	-2.09	0.040	-0.008	-0.000
length		-0.086	0.058	-1.48	0.143	-0.202	0.030
price		-0.000	0.000	-0.19	0.853	-0.000	0.000
foreign		-1.574	1.292	-1.22	0.227	-4.152	1.003
_cons		50.718	6.364	7.97	0.000	38.022	63.414

```

. sum price wei len

```

Variable	Obs	Mean	Std. dev.	Min	Max
price	74	6165.257	2949.496	3291	15906

命令窗口

这里输入新命令

变量窗口

过滤变量

名称	标签
make	Make and n
price	Price
mpg	Mileage (m
rep78	Repair reco
headroom	Headroom
trunk	Trunk space
weight	Weight (lbs

双击变量名  
呈现于命令窗口

属性窗口

变量

名称	make
标签	Make and mo
类型	str18
格式	%-18s
值标签	英
注释	

数据

D:\stata\personal\PX\_A\_2022b\A1\_intro\data

CAP NUM INS

## Useful shortcuts

**F2** — keyboard buttons

describe data

**Ctrl** + **8**

open the data editor

**clear**

delete data in memory

AT COMMAND PROMPT

**PgUp** **PgDn**

scroll through previous commands

**Tab**

autocompletes variable name after typing part

**cls**

clear the console (where results are displayed)

**Ctrl** + **9**

open a new do-file

**Ctrl** + **D**

highlight text in do-file,  
then ctrl + d executes it  
in the command line



## Set up

**pwd**

print current (working) directory

**cd** "C:\Program Files\Stata16"

change working directory

**dir**

display filenames in working directory

**dir \*.dta**

List all Stata data in working directory

**capture log close**

close the log on any existing do-files

underlined parts  
are shortcuts –  
use "capture"  
or "cap"

**log using "myDoFile.txt", replace**

create a new log file to record your work and results

**search mdesc**

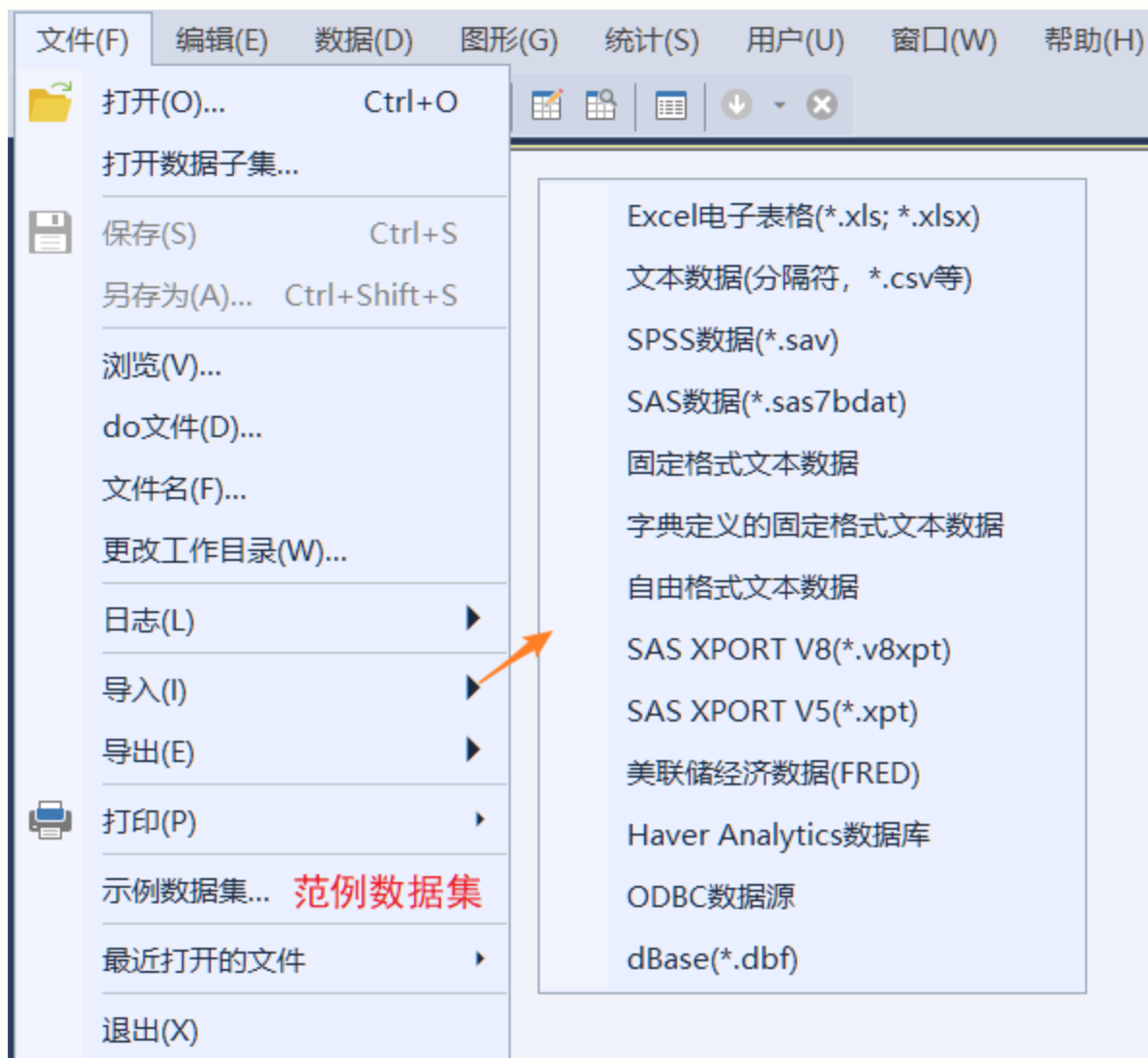
find the package mdesc to install

**ssc install mdesc**

install the package mdesc; needs to be done once

packages contain  
extra commands that  
expand Stata's toolkit

# 文件[File] 菜单



# 界面风格设置及保存

The image shows a software interface with a menu bar (文件(F), 编辑(E), 数据(D), 图形(G), 统计(S), 用户(U)) and a '编辑(E)' dropdown menu. The '首选项(P)' option is highlighted in yellow. A red arrow points from '首选项(P)' to a '常规首选项' dialog box. In this dialog, the '常规' tab is selected, and the '整体颜色方案' is set to '蓝色'. A red arrow points from the '常规首选项' dialog to a '常规首选项(E)...' menu item. Another red arrow points from this menu item to a '新建首选项设置...' dialog box, which lists '2021A' and '2022A'. A final red arrow points from the '新建首选项设置...' dialog to a '出厂窗口设置' dialog box, which lists various window layouts like '宽广布局', '组合布局', '紧凑布局', and '演讲布局(1920x1080) 1080p'.

常规首选项

常规

结果

浏览器

命令

数据编辑器

整体颜色方案：  
蓝色

注意：更改整体方案将同时改变其他颜色设置。

对现有dta文件保存覆盖之前提示

启用自动更新检查

出厂窗口设置

宽广布局

组合布局

紧凑布局

演讲布局(1920x1080) 1080p

演讲布局(1024x768)

演讲布局(1024x768) 大字体

演讲布局(800x600)

2021A

2022A

常规首选项(E)...

图形首选项(G)

用户界面语言...

加载首选项设置(L)

保存首选项设置(S)

删除首选项设置(D)

重置文件关联(A)

新建首选项设置...

2021A

2022A

你可以根据喜好设定字体、配色等，并将这些设置保存下来，如 2021A

## 数据/图形/统计 菜单

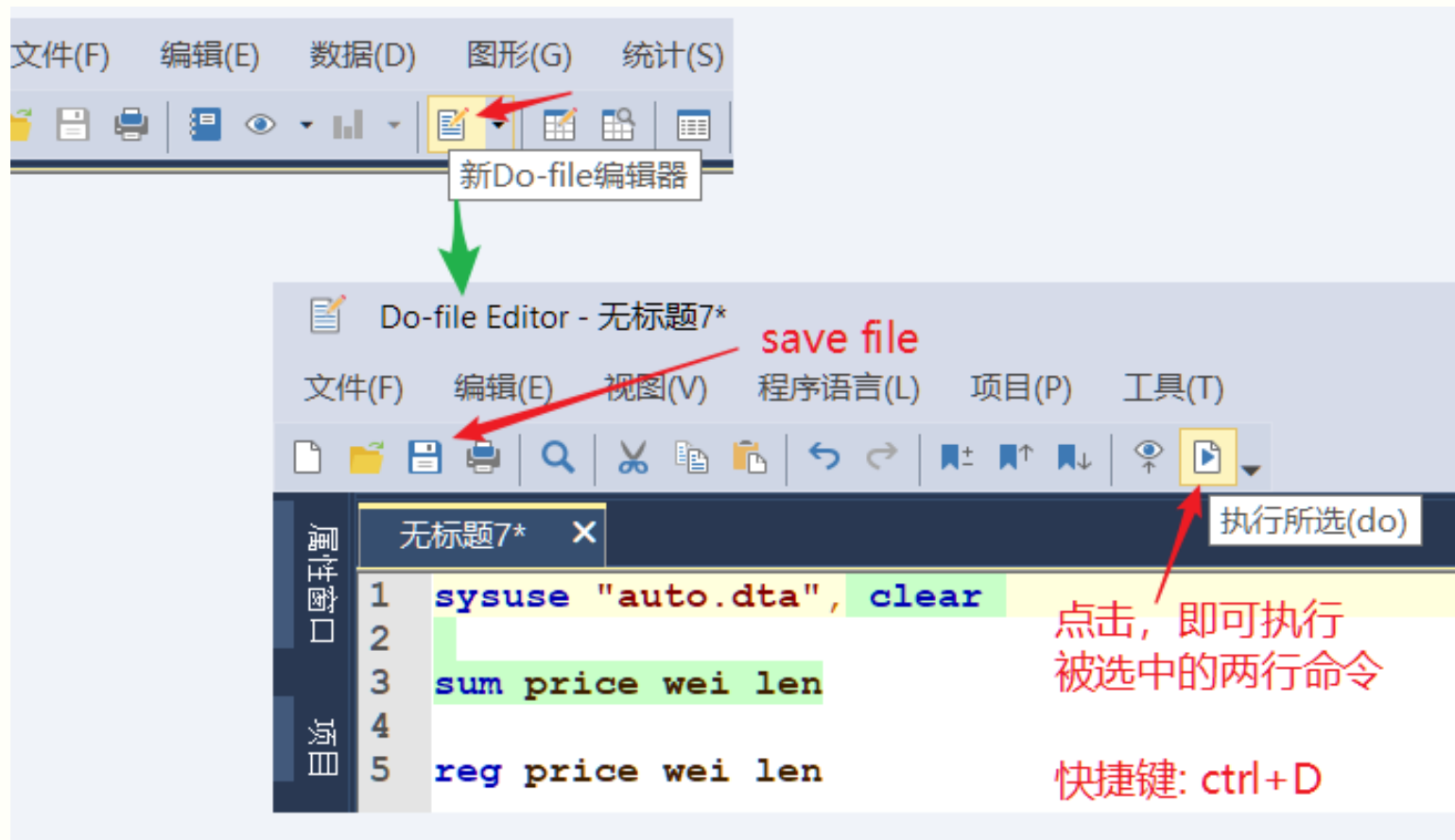




在 dofile 中编写代码  
是 Stata 的标准动作

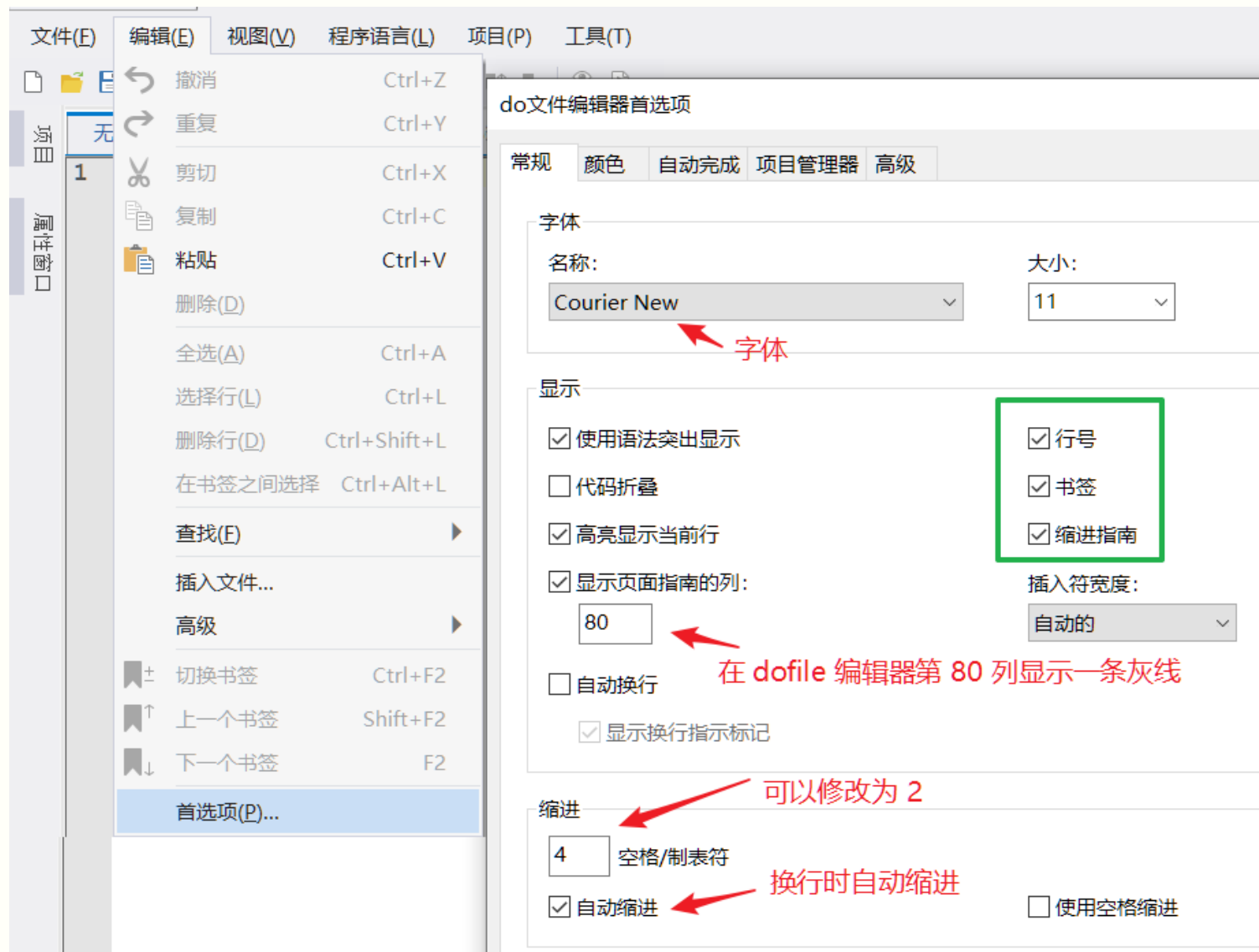
## dofile: 新建和执行

- Dofile 编辑器的使用
- dofile 编辑器的配置
- dofile 模板 DIY
- dofile 模板
- dofile-转-PDF-讲义



# dofile 编辑器设定

## 编辑 → 首选项



```

*-----
*-基本回归分析 see [A4_Regress.do]
  global path "D:/paper01/" // 设置路径
  global D "$path/data"
  global Out "$path/out"
  cd "$D" // 当前工作路径

  sysuse "nlsw88", clear
  global y "wage" //存放被解释变量的全局暂元
  global x "hours tenure married collgrad age" //help macro, 存放解释变量

*-去除缺漏值
  qui reg $y $x i.race i.industry i.occupation
  keep if e(sample)

*-呈现并输出基本统计量(`////' 是换行命令)
  logout, save("$Out/Table1_sum01") excel replace: ///
  tabstat $y $x, column(stats) format(%6.2f) ///
  stats(mean sd min p50 max)

```

\*-呈现相关系数矩阵

```
logout, save("$Out/Table2_corr") excel replace: ///  
    pwcorr_a $y $x
```

\*-回归分析：基本版本

```
reg $y $x //OLS 回归, basic model  
    est store m1 //存储回归结果  
reg $y $x i.race //i.race表示种族虚拟变量, help fvvarlist  
    est store m2  
reg $y $x i.race i.industry  
    est store m3  
reg $y $x i.race i.occupation  
    est store m4
```

\*-列表呈现回归结果 see [A4\_Regress, 4.7 回归结果呈现]

```
local s "using $Out/Table3.csv" //指定存储结果的 Excel 文档名称  
local m "m1 m2 m3 m4"  
esttab `m' `s', replace nogap compress ///  
    ar2 scalar(N F rss) b(%6.3f) /////  
    star(* 0.1 ** 0.05 *** 0.01) /////  
    noomit nobase /////  
    indicate("行业效应=*.industry" "职业效应 =*.occupation")
```

## Stata 代码习惯: dofile 模板等

命令: `. lianxh 码农 重现代码 可重现`

- 专题: 论文写作
  - Stata论文复现: 做一个优雅的码农
  - Stata: 论文重现代码模板
  - 可重复性研究: 如何保证你的研究结果可重现?
- 专题: Stata资源
  - 会计期刊论文的结果可重现吗?
- Stata数据处理: 快来更新你的 Stata-Workflow

## Project 功能

Stata 官方 video, -PDF 实操-

- **目的:** 将同一个 project 相关的文档汇总到一个文档下, 便于管理
- **用途:** 多个 dofile/data 的管理, 如论文, 课题或课程资料
- **新建**一个 Project:
  - [1] 打开 Do-file Editor,
  - [2] 依次点击 File → New → Project..., **保存**
- **打开**一个 Project:
  - [1] 打开 Do-file Editor,
  - [2] 依次点击 File → Open → Project....

Source:  
[stata.com](http://stata.com)

The screenshot shows the Do-file Editor interface. The Project pane on the left displays a folder named 'gsem' containing subfolders 'builder', 'cert', 'data', 'gllamm', and 'xtme', along with do-files 'base01.do' through 'base06.do' and 'bernoulli.do'. The Properties pane below shows details for the 'gsem' project, including its name, type, location, relative path, and full path (C:\stata\gsem\gsem.stp). The main editor window displays the content of 'test.do', which is a list of do-files to be executed, including 'tools', 'base01' through 'base06', 'cns01', 'dnumerical', 'models', 'flatent', 'ilatent', 'scores', 'svyset', 'models\_quad' (with 'intmethod(mv)' and 'intmethod(mc)'), 'models\_lap', 'models\_me', and 'gauss\_exteas'. The status bar at the bottom indicates 'Line: 33, Col: 1 CAP NUM OVR'.

```
33  *** Start of test scripts
34
35  do tools                `noisily' `debug'
36
37  do by                   `opts'  show
38  do base01              `opts'  show
39  do base02              `opts'  show
40  do base03              `opts'  show
41  do base04              `opts'  show
42  do base05              `opts'  show
43  do base06              `opts'  show
44  do base06              `opts'  show fvstand
45  do cns01               `opts'  show
46  do dnumerical          `opts'  show
47
48  do models              `opts'  show
49
50  do flatent
51  do ilatent
52
53  do scores
54  do svyset
55
56  do models_quad  intmethod(mv)    `opts'  show
57  do models_quad  intmethod(mc)    `opts'  `show'
58
59  do models_lap      `opts'  `show' `superall'
60
61  `DA' models_me    `opts'  show
62
63  do gauss_exteas   `opts'  show
```

# Stata 学习资源



# Stata 主页资料

- 🍏 [Stata 小抄 - PDF](#)
- [Installation guide](#): These steps will guide you through the installation process.
- [FAQs](#): Search the frequently asked questions.
- [Documentation](#): Discover the inner workings of Stata's features.
- [Examples and datasets](#): Access third-party resources with examples in Stata and datasets in Stata format.
- [Web resources for learning Stata](#): Explore tutorials, programs, and other resources available on the web.
- [The Stata Blog](#): 推文
- [TEXTBOOK EXAMPLES](#), 经典教科书的 Stata 代码和数据

# Stata 主页：入门介绍

1. [Introducing Stata—sample session](#)
2. [The Stata user interface](#)
3. [Using the Viewer](#)
4. [Getting help](#)
5. [Opening and saving Stata datasets](#)
6. [Using the Data Editor](#)
7. [Using the Variables Manager](#)
8. [Using the Do-file Editor—automating Stata](#)
9. [Learning more about Stata](#)
10. [Subject index](#)

## 论坛：讨论和交流

- [StataList](#)
- [Stack overflow](#)
- [Stack Exchange](#)
- [Stack Exchange Superuser](#) 高级用户经常出没的地方
- [Economics Job market Rumors](#) 很多博士生出没

## Stata Textbooks

- [UCLA-主流计量教科书Stata实操](#)
- [经典计量经济学教材推荐](#)
- [Stata 参考书和经典教材!](#)
- Stata Press books
  - [书目列表](#)     [Data and Codes](#)

## 计量教材+ Stata实操

- Baum, C. An introduction to modern econometrics using Stata[M]. Stata Press, 2006.
- Cameron, A., P. Trivedi. Microeconometrics using stata[M]. Stata Press, 2009.
- Acock, A. C. A gentle introduction to stata (4ed)[M]. Stata Press, 2014.

## 数据处理/绘图/编程

- Kohler, U., F. Kreuter. Data analysis using stata (3th)[M]. Stata Press, 2012.
- Long, J. The workflow of data analysis using stata[M]. Stata Press, 2009.
- Mitchell, M. N. Data management using stata[M]. Stata Press, 2010.
- Mitchell, M. N. Interpreting and visualizing regression models using stata[M]. 2022.
- Weinberg, S. L., S. K. Abramowitz. Statistics using stata[M]. Cambridge: Cambridge University Press, 2020. [-Link-](#), [-PDF1-](#)

## 离散变量/Logit/Count Dtata

- Long, J., J. Freese. Regression models for categorical dependent variables using stata[M]. Stata Press, 2014, 3eds. [-Link-](#)

## 面板/多层次

- Rabe-Hesketh, S., A. Skrondal. Multilevel and longitudinal modeling using stata[M]. Stata press, 2012.
- Sul, D. Panel data econometrics: Common factor analysis for empirical researchers[M]. 2019. [-Link-](#), [-PDF1-](#)

## 金融/时序

- Boffelli, S., G. Urga. Financial econometrics using stata[M]. Stata Press, 2016.
- Levendis, J. D. Time series econometrics: Learning through replication[M]. Springer, 2019. [-Link-](#), [-PDF1-](#), [PDF2](#)

## 生存分析/结构方程

- Acock, A. C. Discovering structural equation modeling using stata[M]. 2013.
- Cameron, A., P. Trivedi. Microeconometrics using stata[M]. Stata Press, 2009.

## SFA

- Kumbhakar, S. C., H. J. Wang, A. P. Horncastle, 2015, A practitioner's guide to stochastic frontier analysis using stata, Cambridge University Press. [-Link-](#)

# Stata 数据处理

👉 Stata 命令: `. lianxh 数据处理`

- 专题: [Stata教程](#)
- 专题: [数据处理](#)
- 专题: [面板数据](#)



## 导入数据

```
help use           // 导入 Stata 格式数据

help sysuse       // Stata 自带数据
sysuse dir, all  // 数据列表

help webuse       // 电子手册数据
help dta_manuals // 数据列表

help import       // 导入 Excel, csv, txt 格式的数据
import excel data.xls, firstrow clear
import excel "D:/data.xlsx", cellrange(A1:A10) firstrow clear

copy "https://www.stata.com/examples/auto.csv" "auto.csv" // CSV
import delimited auto, rowrange(3:6)

help bcuse        // Wooldridge 书中配套数据
bcuse wagepan, clear
```

## 基本语法、运算符和函数

- **[U]** [Language syntax](#) Stata 基本命令的语法规则 (务必认真研读)
  - `help language`
- **[U]** [Functions and expressions](#) 函数和表达式
  - `help operators`
  - `help function`
  - `help expression`
- **[D]** [gen](#) 变量生成和函数 `help gen`
- **[D]** [egen](#) 大量扩展函数 `help egen`

```

sysuse "auto.dta", clear

gen wei2len = weight/length
gen lnPrice = ln(price)
gen bad = (rep78>=4 & rep78!=.)
bysort foreign: egen sd_price = sd(price)

regress mpg price weight##weight i.foreign i.foreign#weight

```

+	addition	&	and	>	greater than
-	subtraction		or	<	less than
*	multiplication	!	not	>=	greater than or equal
/	division	~	not	<=	less than or equal
^	power			==	equal
-	negation			!=	not equal
+	string concatenation			~ =	not equal

## 变量的简洁表示

- 因子变量 `reg wage hours i.race i.industry i.industry#c.year`
  - 详情: [Stata: 因子变量全攻略](#)
  - 帮助: `help fvvarlist` ; `help varlist`
- 时序变量 `reg D.wage D.hours D.L.(z1 z2 z3)`
  - 帮助: `help varlist` ; `help tsvarlist`

## 因子变量

详情: [Stata: 因子变量全攻略](#)    帮助: `help fvvarlist` ; `help varlist`

符号	含义	实例
<code>i.</code>	标示为类别变量	<code>reg y x i.id i.year</code> → $y_{it} = a_i + x_{it}\beta + \lambda_t + u_{it}$
<code>c.</code>	标示为连续变量	
<code>o.</code>	略去某个类别或变量	
<code>#</code>	交乘项	<code>reg y D x i.D#c.x</code> → $y_i = a + D_i\beta_1 + x_i\beta_2 + D_ix_i\gamma + u_i$ <code>reg y x c.x#i.year</code> → $y_{it} = a + x_{it}\beta_t + u_{it}$
<code>##</code>	两个变量及其交乘项	<code>reg y x##z</code> → $y_i = a + x_i\beta_1 + z_i\beta_2 + x_iz_i\beta_3 + u_i$

```

clear
input group x
      1   30
      1   50
      2   40
      2   60
      2   80
      3   70
end

```

```
. list group i.group i.group#c.x, clean
```

	1.	2.	3.	1.group#	2.group#	3.group#
group	group	group	group	c.x	c.x	c.x
1.	1	1	0	30	0	0
2.	1	1	0	50	0	0
3.	2	0	1	0	40	0
4.	2	0	1	0	60	0
5.	2	0	1	0	80	0
6.	3	0	0	0	0	70

```

clear
input group x
      1  30
      1  50
      2  40
      2  60
      2  80
      3  70
end

```

```
. reg x i.group, noheader
```

x	Coefficient	Std. err.	t	P> t
group				
2	20.000	16.667	1.20	0.316
3	30.000	22.361	1.34	0.272
_cons	40.000	12.910	3.10	0.053

```
. regfit
```

```

x = 40.00 + 0.00*1b.group + 20.00*2.group + 30.00*3.group
    (12.91) (0.00)          (16.67)          (22.36)
    N = 6, R2 = 0.43, adj-R2 = 0.05

```

## 因子变量：设定基准组的运算符及含义

基准组运算符	含义
<code>ib#.x</code>	使用*作为基准组，*为变量中其中一类的值
<code>ib(#*).x</code>	使用变量值中的第*位排序的值所对应的类别作为基准组
<code>ib(first).x</code>	使用变量的最小值所对应的类别作为基准组 (该项为 Stata 默认选项)
<code>ib (last).x</code>	使用变量的最大值所对应的类别作为基准组
<code>ib( freq).x</code>	使用变量值的频数最大的类别作为基准组
<code>ibn.x</code>	不设基准组



# 例 1: 种族与工资

模型设定:  $Wage_i = \alpha + black_i\beta_1 + hours_i\beta_2 + black_i \times hours_i\beta_3 + u_i$

```
. sysuse "nlsw88.dta", clear

*-传统方法
. gen black=1
. replace black=0 if race!=2
. gen black_x_hours = black*hours
. reg wage black hours black_x_hours

*-因子变量法
. reg wage 2.race c.hours 2.race#c.hours
. reg wage 2.race##c.hours //与上一行等价
```

	wage	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
race							
Black		-1.619	1.266	-1.28	0.201	-4.101	0.864
hours		0.089	0.012	7.24	0.000	0.065	0.113
race#c.hours							
Black		0.007	0.033	0.22	0.827	-0.057	0.071
_cons		4.810	0.474	10.14	0.000	3.880	5.741


模型设定:  $Wage_i = \alpha + B_i\beta_1 + h_i\beta_2 + B_i \times h_i\beta_3 + h_i^2\beta_4 + B_i \times h_i^2\beta_5 + u_i$

```
. sysuse "nlsw88.dta", clear
. reg wage 2.race##c.hours##c.hours
```

wage	Coeff	SE	t	P> t
race				
Black	-3.461	2.274	-1.52	0.128
hours	0.133	0.044	3.03	0.002
race#c.hours				
Black	0.123	0.127	0.97	0.331
c.hours#c.hours	-0.001	0.001	-1.05	0.296
race#c.hours#c.hours				
Black	-0.002	0.002	-0.95	0.344
_cons	4.165	0.778	5.35	0.000

# 时序变量的表示

帮助: `help varlist`; `help tsvarlist`

 `reg y L.y L(0/2).x D.L.z`  $\rightarrow y_t = a + \rho y_{t-1} + \sum_{s=0}^3 \theta_s x_{t-s} + \Delta z_{t-1} + u_t$

Operator	Meaning	
L.x	$x_{t-1}$	1-period lag
L2.x	$x_{t-2}$	2-period lag
...		
F.x	$x_{t+1}$	1-period lead
F2.	$x_{t+2}$	2-period lead
...		
D.	$\Delta x_t = x_t - x_{t-1}$	first difference
D2.	$\Delta^2 x_t = (x_t - x_{t-1}) - (x_{t-1} - x_{t-2})$	difference of difference
...		
L(0/2).x	$x_t, x_{t-1}, x_{t-2}$	several variables
L.D.x	$\Delta x_{t-1}$	can be nested

```

. use "https://www.stata-press.com/data/r17/gxmpl1", clear
. format gnp cpi %5.1f

. list year L(1/3).(gnp cpi) D.cpi, clean

```

	year	L. gnp	L2. gnp	L3. gnp	L. cpi	L2. cpi	L3. cpi	D. cpi
1.	1989	.	.	.	.	.	.	.
2.	1990	5837.9	.	.	124.0	.	.	6.7
3.	1991	6026.3	5837.9	.	130.7	124.0	.	5.5
4.	1992	6367.4	6026.3	5837.9	136.2	130.7	124.0	4.1
5.	1993	6689.3	6367.4	6026.3	140.3	136.2	130.7	4.2
6.	1994	7098.4	6689.3	6367.4	144.5	140.3	136.2	3.7
7.	1995	7433.4	7098.4	6689.3	148.2	144.5	140.3	4.2
8.	1996	7851.9	7433.4	7098.4	152.4	148.2	144.5	4.5

```

. use "https://www.stata-press.com/data/r17/invest2", clear
. keep if time<=5&company<=2
. format invest market %4.1f
. list company time invest L(1/2).invest market D1.market

```

	company	time	invest	L. invest	L2. invest	market	D. market
1.	1	1	317.6	.	.	3078.5	.
2.	1	2	391.8	317.6	.	4661.7	1583.2
3.	1	3	410.6	391.8	317.6	5387.1	725.4
4.	1	4	257.7	410.6	391.8	2792.2	-2594.9
5.	1	5	330.8	257.7	410.6	4313.2	1521.0
6.	2	1	40.3	.	.	417.5	.
7.	2	2	72.8	40.3	.	837.8	420.3
8.	2	3	66.3	72.8	40.3	883.9	46.1
9.	2	4	51.6	66.3	72.8	437.9	-446.0
1.	2	5	52.4	51.6	66.3	679.7	241.8

## 数据的长宽转换 long <--> wide

- 专题：Stata命令
  - Stata：宽数据到长数据的转换-tolong
- 专题：数据处理
  - Stata数据处理：纵横长宽转换-reshape命令一文读懂！（下）-sreshape-fastreshape-xpose
  - Stata数据处理：纵横长宽转换-reshape命令一文读懂！（上）
  - Stata数据处理：纵横长宽转换-reshape的兄弟-gather和spread.md

long				wide		
+	-----	+		+	-----	+
	i	j			var1	var2
	-----				-----	
	1	1			1	4.1
	1	2			4.1	4.5
	2	1			3.3	3.0
	2	2			3.3	3.0
	-----				-----	
+	-----	+		+	-----	+

reshape  
 <----->

// long --> wide:

j 旧变量名称  
/

reshape wide year, i(i) j(j)

// wide --> long:

reshape long stub, i(i) j(j)  
 \ j 新变量名称

```
. reshape long inc ue, i(id) j(year) //Note: [1] sex 不发生变化, 无需转换
// [2] j() 选项中填写新的变量名称
```

\* -wide- 数据

id	sex	inc1980	inc1981	inc1982	ue1980	ue1981	ue1982
1	0	5000	5500	6000	0	1	0
2	1	2000	2200	3300	1	0	0
3	0	3000	2000	1000	0	0	1

```
reshape long inc ue, i(id) j(year)
```

\* -long- 数据

id	year	sex	inc	ue
1	1980	0	5000	0
1	1981	0	5500	1
1	1982	0	6000	0
2	1980	1	2000	1
2	1981	1	2200	0
2	1982	1	3300	0
3	1980	0	3000	0
3	1981	0	2000	0
3	1982	0	1000	1



# Stata 数据处理：推文

 Stata 命令：

```
. lianxh 数据处理 面板
```

## 推文：Stata 数据处理 (1)

- 专题：Stata教程
  - 普林斯顿Stata教程(一) - Stata数据处理
- 专题：
  - Stata数据处理：数据框使用教程
  - Stata数据处理：将字符变量编码为数值变量-encoder
  - Stata数据处理：一文搞定CEIC数据库
  - Stata数据处理：清洗中国城市建设统计年鉴
  - Stata数据处理：清洗CFPS数据库
  - CFPS数据处理：少儿代答库与成人库匹配

## 推文：Stata 数据处理 (2)

- [Stata数据处理：批量处理被保护的年鉴数据-dxls-txls](#)
- [Stata数据处理：缺失值与多重补漏分析（一）](#)
- [Stata数据处理：缺失值与多重补漏分析（二）](#)
- [Stata数据处理：缺失值与多重补漏分析（三）](#)
- [Stata数据处理：一文读懂微观数据库清理（上）](#)
- [Stata数据处理：一文读懂微观数据库清理（下）](#)
- [Stata数据处理：iebalab和ieddtab命令介绍-T208](#)
- [Stata数据处理：超大Excel文档如何读入](#)

## 推文：Stata 数据处理 (3)

- [滚动吧统计量！ Stata数据处理](#)
- [Stata数据处理：各种求和方式一览](#)
- [Stata数据处理：字符型日期变量的转换](#)
- [Stata数据处理：统计组内非重复值个数](#)
- [Stata数据处理：赫芬达尔指数-\(hhi5\)-命令介绍](#)
- [Stata数据处理：面板数据的填充和补漏](#)
- [Stata数据处理：xtbalance-非平衡面板之转换](#)

# Stata 程序基础

- 专题: [Stata教程](#)
- 专题: [Stata程序](#)
- [\[U\] 18 Programming Stata](#)

 [连玉君 b 站视频公开课: Stata程序的编写和发布](#)

## Stata 程序：推文

- [普林斯顿Stata教程\(三\) - Stata编程](#)
- [Stata编程：暂元, local! 暂元, local!](#)
- [Stata程序：暂元-\(local\)-和全局暂元-\(global\)](#)
- [Stata小白编程：暂元及macrolists命令](#)
- [Stata：复合双引号在编程中的应用](#)
- [Stata程序：10 分钟快乐编写 ado 文件](#)
- [Stata程序：Monte-Carlo-模拟之产生符合特定分布的随机数](#)
- [Stata 程序：数值求解极值的基本思想](#)
- [statsby: 不用循环语句的循环](#)

# Stata 绘图

## Stata 绘图：基本资料

- Stata 官网：[图形范例及代码](#) → [各类图形](#)
- Stata 绘图：[PDF 手册](#)
  - [绘图简介](#) → [twoway](#) → [Options](#)
  - [\[G-1\] Graph Editor](#), 手动编辑图形
- **Book:** Mitchell, M. A visual guide to stata graphics (4e). Stata Press, 2022. [-Link-](#)
  - [Data-Codes](#) → 直接下载：

```
. net get vgsg4
```
- 普林斯顿Stata教程(二) - Stata绘图, [-Link-](#), [E-version](#)
- Denny, K., 2021, Basic Stata graphics for social science students, [-PDF-](#)
- An Introduction to Stata Graphics, [-Link-](#)



# Stata 绘图：图形模板和选项

. lianxh 模板 选项

## • 绘图模板

- Stata绘图极简新模板：plotplain和plottig-T251
- Stata绘图：一个干净整洁的-Stata-图形模板qlean
- Stata：图形美颜-自定义绘图模板-grstyle-palettes
- Stata黑白图形模板：中文期刊风格的纯黑白图形
- 史上最牛Stata绘图模版-schemepack：酷似R中的ggplot2
- Stata 绘图：用 Stata 绘制一打精美图片-schemes

## • 图形选项 PDF-Options

- 给你的图形化个妆：Stata绘图常用选项汇总-上篇
- 给你的图形化个妆：Stata绘图常用选项汇总-下篇

## Stata 绘图：系数可视化

`.lianxh` 可视

- Stata可视化：让他看懂我的结果！
- Stata绘图-可视化：组间差异比较散点图
- Stata：分仓散点图绘制-binscatter-binscatter2
- Stata绘图：世行可视化案例-条形图-密度函数图-地图-断点回归图-散点图
- Stata：边际处理效应及其可视化-mtefe-T309
- Stata绘图-可视化：组间差异比较散点图
- Stata可视化：biplot一图看尽方差、相关性和主成分
- Stata绘图-组间差异可视化：不良事件火山图、点阵图

## Stata绘图：直方图/柱状图/密度图

.lianxh 直方 柱状 密度

- Stata: 读懂直方图
- Stata绘图: 柱状图专题
- Stata绘图: 多维柱状图绘制
- Stata: 多个核密度函数图叠加-mkdensity
- Stata: 双变量联合核密度函数图-kdens2

## Stata 绘图：散点图

.lianxh 散点

- Stata绘图：绘制美观的散点图-superscatter
- Stata绘图-可视化：组间差异比较散点图
- Stata：分仓散点图绘制-binscatter-binscatter2
- Stata绘图：世行可视化案例-条形图-密度函数图-地图-断点回归图-散点图

## Stata 绘图：进阶

- 论文中因果推断的经典图形
- Stata绘图：如何更高效的绘制图形
- Stata绘图：addplot-层层叠加轻松绘图

Stata 绘图最难也是最灵活的部分就是选项的设定。

若将一幅图形的各个要素拆解开，了解对应的选项设定和帮助文件，积累一段时间，便可以熟练地绘制图形了。选项大致可以分成以下几类：

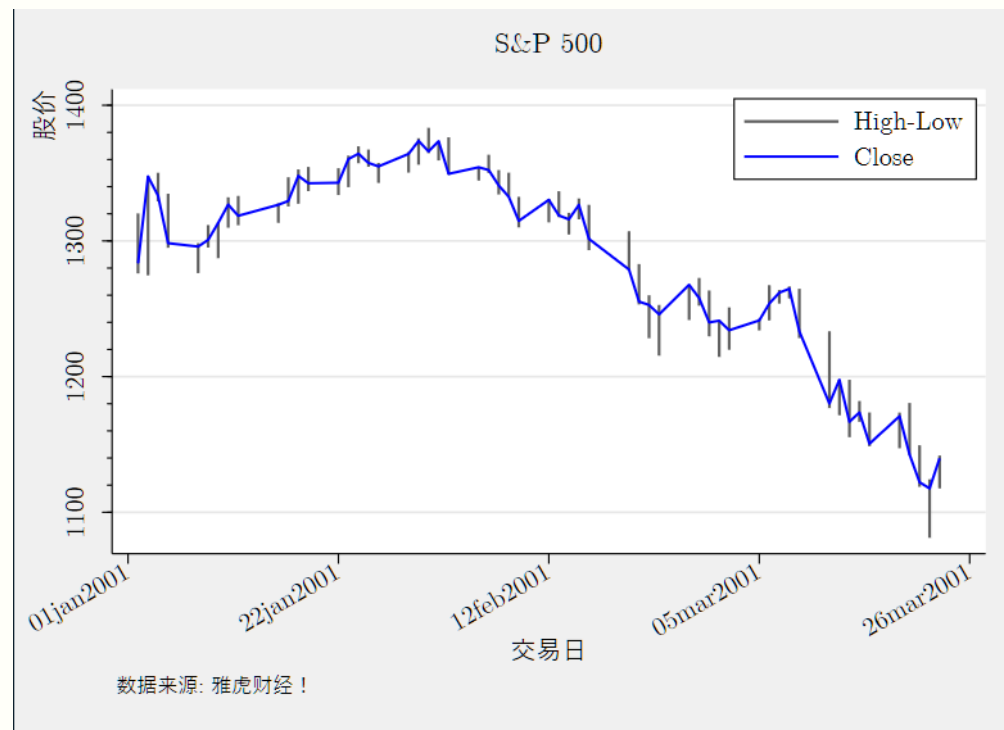
- 图形模版：图形的整体风格 `help scheme`
- 坐标轴、刻度等 `help axis_options`
- 标题类：主标题、副标题、注释等 `help title_options`
- 附加线类：横向或纵向附加线 `help addline_options`

```

sysuse "sp500", clear
replace volume = volume/1000
keep in 1/57
*-----Begin
#delimit ;
twoway                                //help twoway
    (rspike hi low date, lw(*1.3))
    (line close date,
        lpattern(solid) lwidth(*1.2) lcolor(blue))
,
yscale(range(1100 1400))              //help axis_options
ylabel(1100(100)1400, grid)           //help axis_options
ymtick(##5)                            //help axis_options
xlabel(, angle(30))                    //help axis_options
ytitle("股价", place(top))            //help title_options
xtitle("交易日")                      //help title_options
legend(order(1 "High-Low" 2 "Close")
        ring(0) position(2) row(2))    //help legend_options
subtitle("S&P 500", margin(b+2.5))    //help title_options
note("数据来源: 雅虎财经!")           //help title_options
scheme(s2mono);                        //help scheme
#delimit cr

graph export "sp500_rspike_01.png", replace
*-----Over

```

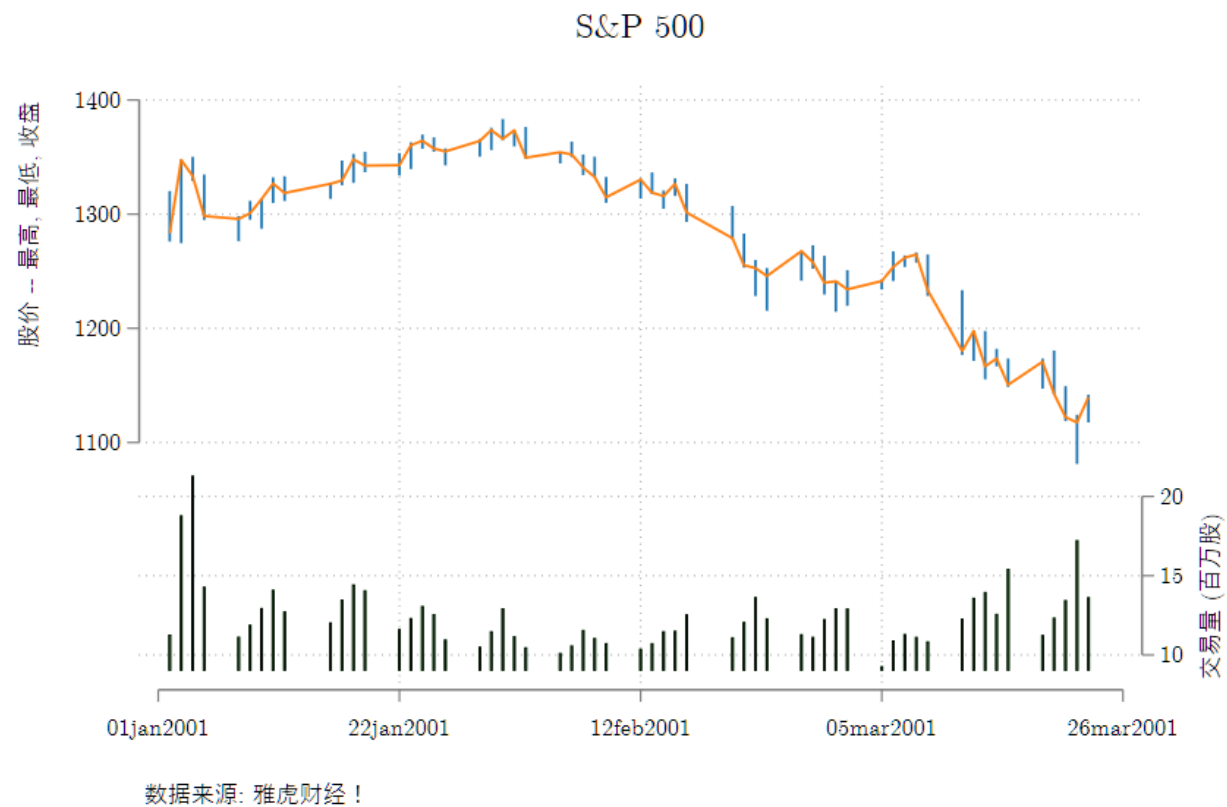


```

sysuse "sp500", clear
replace volume = volume/1000
keep in 1/57
*-----Begin
#delimit ;
twoway
    (rspike hi low date)
    (line close date)
    (bar volume date, barw(.25) yaxis(2))
,
yscale(axis(1) r(900 1400))
yscale(axis(2) r( 9 45))
ylabel(, axis(2) grid)
yttitle("股价：最高，最低，收盘",place(top))
yttitle("交易量 (百万股)", axis(2)
        bexpand just(left))
xttitle(" ")
legend(off)
subtitle("S&P 500", margin(b+2.5))
note("数据来源：雅虎财经！") ;
#delimit cr
*-----Over

graph export "sp500_rspike_02.png", replace

```



# 论文复现

- 专题：专题课程
  - 连享会公开课：论文复现的数据和代码-连玉君
- 专题：论文写作
  - 连享会：论文重现网站大全
  - Stata论文复现：社科领域Top5期刊复现资料



## 更多推文

.lianhx 复现 重现 可重复

- Stata论文复现：做一个优雅的码农
- Stata：论文重现代码模板
- 可重复性研究：如何保证你的研究结果可重现？
- Stata的版本兼容性问题：可重复研究
- 论文重现：外部命令的版本控制
- 会计期刊论文的结果可重现吗？
- 可重复研究：如何让你的研究明了易懂？

# Stata 中最重要的命令

## help cmd

- help xtreg
- help winsor2

## search keywords

- search dynamic panel data

## lianxh keywords -click-

- lianxh DID 倍分法
- songbl 合成控制