Lab1: Introduction to Stata

Introduction to Econometrics,Fall 2020

Yi Wang

Nanjing University

9/16/2020

Section 1

What is Stata?

What is Stata?

- Stata是经济学研究主流的数据分析软件，它功能强大，程序包丰富，可以说几乎涵盖了应用计量经济学领域所有的功能。
- Stata的**help**文件非常详细，完全可以自学。
- 想要完成规范的现代经济学实证研究，像Stata这样的计量软件是必不可少的工具。
- Stata最初由美国计算机资源中心（Computer Resource Center)研制，现在为Stata公司的产品，其最新版本为Stata 16。

What is Stata?

Stata is available in **four flavors**:

- Stata/MP–multiprocessor, the fastest version, up to 32,767 variables with either.
- Stata/SE–like Stata/MP, but for single CPUs, up to 32,767 variables with either.
- Stata/IC–standard Stata,up to 2,047 variables are allowed.
- Small Stata–intended for students and limited to 99 variables and 1200 observations.
- More Info:https://www.stata.com/manuals13/u5.pdf
- 商用版根据性能不同售价在$1000-$6500，学生版顶配为 Stata/MP 4 core，售价$995。

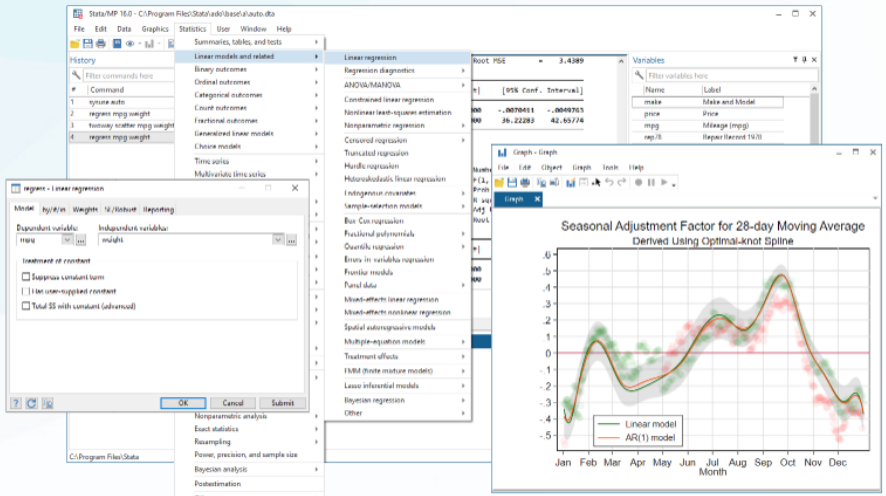
Section 2

Why Stata?

Why Stata?

Stata is statistical software for data science

- Master your data
- Broad suite of statistical features
- Publication-quality graphics
- Automated reporting
- Python integration
- Truly reproducible research
- Real documentation
- Trusted
- Easy to use
- Easy to grow with
- Easy to automate
- Easy to extend
- Advanced programming
- Automatic multicore support
- Community-contributed features
- World-class technical support
- Cross-platform compatible
- Widely used
- Comprehensive resources
- Vibrant community
- Affordable



New in STATA 16

https://www.stata.com/
http://www.uone-tech.cn/Stata.html

Why Stata?

工具名称	公司	是否免费	易用性	专业性	可编程	常用场景及领域
MATLAB	MathWorks	否	中	高	是	统计分析只是应用的一个方向，适合数据处理以及工程建模的各个领域
SPSS	IBM 公司	否	高	中	是	统计分析专业人士的入门级软件。可用于经济分析，市场调研等社会科学各个领域
Stata	Stata 公司	否	中	中	是	统计分析专业人士的进阶级软件。多用于医学，生物学研究领域
SAS	SAS 公司	否	低	高	是	统计分析专业人员的殿堂级软件。可应用于各个统计分析领域，是高级分析人员更青睐的统计分析利器
EViews	QMS 公司	否	高	中	是	能很好地处理时间序列分析等相关问题，主要应用于计量统计学领域
Excel	微软	否	高	低	是	非专业人士使用的简单统计分析软件。可以胜任日常工作中简单的数据统计、数据整理以及数据展示的工作
Python	-	是	中	高	是	完全的编程实现。可用于任何领域，并且与大数据组件结合可以方便地处理海量数据
R	AT&T	是	中	高	是	同 Python

- 总之，对于经济专业的学习者尤其是初学者而言，Stata和R是最佳选择，而Stata比R更易上手。

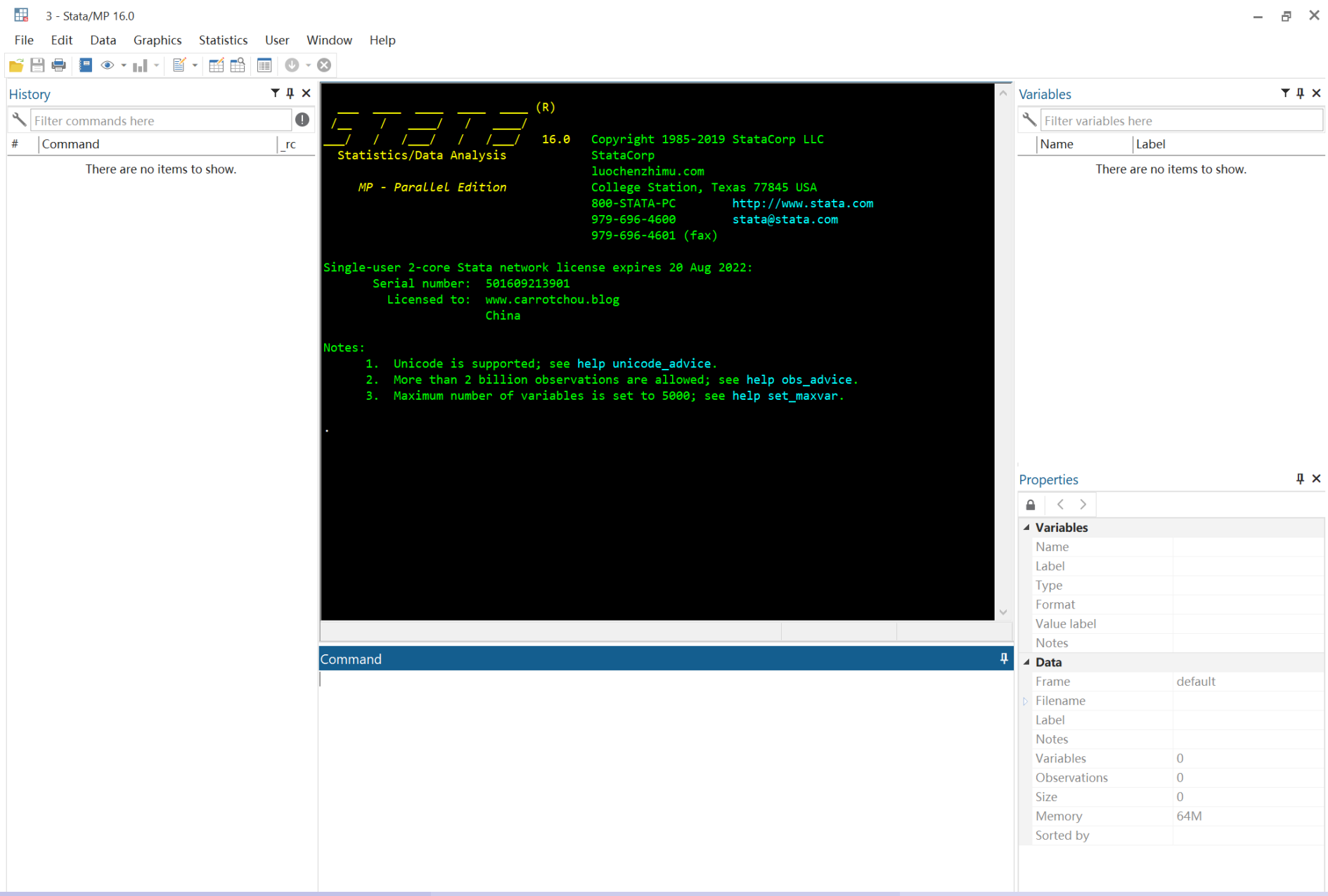
Section 3

How to install Stata?
(See **Stata Installation** for details.)

Section 4

Interface of Stata

Stata的使用界面


3 - Stata/MP 16.0
File Edit Data Graphics Statistics User Window Help
History
Filter commands here
Command _rc
There are no items to show.
(R)
16.0
Statistics/Data Analysis
MP - Parallel Edition
Copyright 1985-2019 StataCorp LLC
StataCorp
luochenzhimu.com
College Station, Texas 77845 USA
800-STATA-PC http://www.stata.com
979-696-4600 stata@stata.com
979-696-4601 (fax)
Single-user 2-core Stata network license expires 20 Aug 2022:
Serial number: 501609213901
Licensed to: www.carrotchou.blog
China
Notes:
1. Unicode is supported; see help unicode_advice.
2. More than 2 billion observations are allowed; see help obs_advice.
3. Maximum number of variables is set to 5000; see help set_maxvar.
.
Command
Variables
Filter variables here
Name Label
There are no items to show.
Properties
Variables
Name
Label
Type
Format
Value label
Notes
Data
Frame default
Filename
Label
Notes
Variables 0
Observations 0
Size 0
Memory 64M
Sorted by

Stata的使用界面

- 五个窗口，两组菜单条
  - Command–命令窗口
  - Results–结果窗口
  - History–历史窗口
  - Variables–变量窗口
  - Properties–属性窗口

  两种执行命令的方式

-
  - 菜单
  - 命令

Stata的使用界面

- An easy example

```
sysuse auto,clear
des
sum
twoway (scatter mpg weight)
reg price wei len mpg
```

Section 5

Basic Settings of Stata

Stata的基本设定

- 初次使用时Preference的设定
  - 设定方法 Edit–>preference–>General Preference
  - 保存设定 Edit–>Preference–>Save...–>New...任意名称

```
window manage prefs save kobold
```

  - 调入设定 Edit–>Preference–>Load...–>选择喜欢的设定

```
window manage prefs load kobold
```

  - 删除设定 Edit–>Preference–>Delete...–>选择要删除的设定

Stata的基本设定

- Stata帮助
  - -help-命令
  - -search-命令：Searches the [keywords] of the help files
  - -findit-命令：类似-search-命令，但可以进一步搜索网络上的信息

Examples

```
help regress
search panel data
findit panel unit root
```

Stata的基本设定

- stata系统目录的设定

```
. sysdir                                          //显示当前系统目录的设定

   STATA:  D:\Stata16\                            stata安装根目录
    BASE:  D:\Stata16\ado\base\                   【官方命令】存储地址
    SITE:  D:\Stata16\ado\site\                   【自编命令】存储地址
    PLUS:  C:\Users\huawei\ado\plus\              【外部命令】的储存地址
PERSONAL:  C:\Users\huawei\ado\personal\【自有文件夹】首次安装时需自建
OLDPLACE:  c:\ado\

. sysdir set PLUS "D:\stata16\ado\plus"       //设定外部命令的存放地址
. sysdir set PERSONAL "D:\stata16\ado\personal" //设定个人文件夹
. sysdir set OLDPLACE "D:\Stata16\ado\oldplace"

. mkdir "D:\stata16\ado\personal"          //在磁盘创建personal文件夹
. dir                                      //显示当前目录下的所有文件
. dir *.docx                               //显示后缀为 ".docx" 的所有文件
```

Stata的基本设定

- 文件目录或者工作路径设置

```
. pwd                                  //查看当前工作路径
  C:\Users\huawei\Documents
. cd                                   //查看当前工作路径
  C:\Users\huawei\Documents
. cd D:\Teaching\Stata\lab1            //进入（修改）指定文件夹
. cdout                                //打开当前文件夹
```

Stata的基本设定

- 程序运行过程中的目录管理

```
global root "D:\Teaching\Stata\lab1"
//设置你自己的根目录,并建立下面相应的文件夹

global do    $root/Dofiles          //可用于保存do文档
global log   $root/Logfiles         //可用于保存log文档
global work  $root/WorkData         //可用于保存临时数据
global raw   $root/RawData          //可用于保存原始数据
global save  $root/SaveData         //可用于保存修改后的最终数据
global fig   $root/Figures          //保存图片
global tab   $root/Tables           //保存表格

cd ${work}     //如在对数据进行处理时,可以直接指定该路径
```

Stata的基本设定

- 一些常用的设定

```
set more on              //开启分屏显示
sysuse auto, clear
list price               //列出变量

set more off             //禁止分屏显示
list price

set matsize 8000     //设置矩阵的最大维度
set maxvar 10000     //设置最大变量数
```

stata官方的范例数据

-
```
help dta_contents    //(File-->Example Datasets)
sysuse dir
```

Stata的基本设定

- Stata外部命令
  - 下载安装外部命令的三种方式

    help
    ssc install + cmd //不弹出界面
    findit

  - 外部命令查询

    ado //查看外部命令
    which+cmd

  - Example

```
help outreg2
findit outreg2
ssc install outreg2

ado
which outreg2
```

Section 6

Learning Materials of Stata

Learning Materials of Stata

- help文档（详尽,精确查找,最佳选择）
- 搜索引擎（Google,Bing,Yahoo,Baidu...）
- 线上线下交流学习
- 网络学习资源汇总
  - Stata website: http://www.Stata.com/
  - Stata journal: https://www.stata-journal.com/
  - Stata FAQs: https://www.stata.com/support/faqs/
  - Stata bookstore: https://www.stata.com/bookstore/books-on-stata/
  - Code and resources: https://geocenter.github.io/StataTraining/
  - 人大经济论坛【Stata专版】:https://bbs.pinggu.org/forum-67-1.html

Section 7

Import and Export Data

数据导入

- 实证分析的第一步即是数据处理。
- 收集数据、存储、修改、分析、输出结果。
- Stata所直接处理的是dta文件，类似txt文档，占用存储空间小，可以直接在菜单栏打开。
- 导入数据的三种方式：
  - 1.手动输入
  - 2.从txt或Excel文档中粘贴
  - 3.使用Stata命令

数据导入

- 手动输入 (极少使用)

```
clear
input x y z
1 2 3
4 5 6
end
save mydata, replace        //保存数据
use mydata, clear           //调入数据
rm mydata.dta  //删除数据(one at a time),与erase功能相同
```

数据导入

- 从txt或Excel文档中粘贴

```
shellout auto1.txt      //copy-paste(【!open】 in Mac)
shellout auto1.xls

edit    // 打开数据编辑器
```

数据导入

- 使用Stata命令
  - 【.dta文件】的导入

```
clear all
use "D:\Teaching\Stata\lab1\auto1.dta" //注意改路径

cd "D:\Teaching\Stata\lab1\"
use auto1.dta,clear

global root "D:\Teaching\Stata\lab1\" //定义全局宏
cd "$root"
use auto1.dta,clear

local root "D:\Teaching\Stata\lab1\"  //定义局部宏
cd `root'
use auto1.dta,clear
```

数据导入

- 使用Stata命令
  - 【.txt,.csv,.xlsx文件】的导入

```
insheet using auto1.txt,clear
insheet using auto1.csv,clear
import excel auto1.xls, firstrow clear
```

数据导出

- -export-导出

```
sysuse auto, clear
export excel auto2.xlsx
export excel make mpg weight using auto, replace
```

- -save-存储数据

```
sysuse auto, clear
keep in 1/10
save auto3.dta,replace
```

- 记得删除电脑上保存输出的数据，免得占内存

Section 8

Variables and Basic Statistics

Variables and Basic Statistics

- 变量名称基本规则
  - 由英文字母、数字或_组成，至多不超过32个;
  - 首字母不能为数字;
  - 英文字母大小写具有不同含义;
  - 尽量不要使用_作为第一个字母,因为许多stata的内部变量都是以“_”开头,如_n, _N,_cons,_b等等。

```
help _variables
```

Variables and Basic Statistics

- 查看数据结构

```
. sysuse auto, clear
(1978 Automobile Data)
. describe
Contains data from D:\Stata16\ado\base/a/auto.dta
  obs:            74                          1978 Automobile Data
 vars:            12                          13 Apr 2018 17:45
                                              (_dta has notes)
-------------------------------------------------------------------------------
              storage   display    value
variable name   type    format     label      variable label
-------------------------------------------------------------------------------
make            str18   %-18s                 Make and Model
price           int     %8.0gc                Price
mpg             int     %8.0g                 Mileage (mpg)
rep78           int     %8.0g                 Repair Record 1978
headroom        float   %6.1f                 Headroom (in.)
trunk           int     %8.0g                 Trunk space (cu. ft.)
weight          int     %8.0gc                Weight (lbs.)
length          int     %8.0g                 Length (in.)
turn            int     %8.0g                 Turn Circle (ft.)
displacement    int     %8.0g                 Displacement (cu. in.)
gear_ratio      float   %6.2f                 Gear Ratio
foreign         byte    %8.0g       origin    Car type
-------------------------------------------------------------------------------
Sorted by: foreign
```

Variables and Basic Statistics

- 变量的存储类型
  - 字符型数据：字母+特殊符号。
    - ★ 表示姓名、住址（文字信息）;性别（定性）;身份证号（数字）等。
    - ★ 一般用str#来表示字符。
    - ★ 每个汉字占两个字符。
    - ★ str18表示make变量最多容纳的字符个数是18。
  - 数值型数据：便于进行数字的算数运算。
    - ★ 整数的存储类型
      byte 字节型 (-100, +100)
      int 一般整数型 (-32000,+32000)
      long 长整数型 (−2.14*10^10, +2.14*10^10)，即，正负21亿
    - ★ 小数的存储类型
      float 浮点型 8 位有效数字
      double 双精度 16 位有效数字
  - 缺失数据："."被认为大于任何数。

Variables and Basic Statistics

- 更改变量的存储类型

```
sysuse auto, clear
list gear_ratio in 1/5
d gear_ratio

recast int gear_ratio, force  //更改变量的存储类型
d gear_ratio
list gear_ratio in 1/5

compress                      //自动精简资料的存储格式
```

Variables and Basic Statistics

- 定义变量的显示格式
  - 字符型变量%#s(提示符+字符数+显示格式)

    %-18s 靠左列印;
    %18s 靠右列印;
    %~18s 居中列印。
  - 数值变量%w.d+3种基本显示格式(c要求stata给出",")

    e.g.12345
    g一般格式：%9.0g(12345) %9.2gc(12,345)
    f固定格式：%9.4f(12345.0000) %9.0fc(12,345)
    e科学计数法格式：%9.2e(1.23e+04)

    %6.2f 总共占6个空格，小数位占两个空格。

Variables and Basic Statistics

- 定义变量的显示格式

```
list price gear in 1/5
format price %6.1f
format gear %6.4f
list price gear in 1/5
```

Variables and Basic Statistics

- 数据和变量的标签
  - 样本标签

```
. sysuse auto, clear
(1978 Automobile Data)
. label data "这是一份汽车价格资料"
. des
Contains data from D:\Stata16\ado\base/a/auto.dta
  obs:             74                          这是一份汽车价格资料
 vars:             12                          13 Apr 2018 17:45
                                               (_dta has notes)
------------------------------------------------------------------------------
              storage   display    value
variable name   type    format     label      variable label
------------------------------------------------------------------------------
make            str18   %-18s                 Make and Model
price           int     %8.0gc                Price
mpg             int     %8.0g                 Mileage (mpg)
rep78           int     %8.0g                 Repair Record 1978
headroom        float   %6.1f                 Headroom (in.)
trunk           int     %8.0g                 Trunk space (cu. ft.)
weight          int     %8.0gc                Weight (lbs.)
length          int     %8.0g                 Length (in.)
turn            int     %8.0g                 Turn Circle (ft.)
displacement    int     %8.0g                 Displacement (cu. in.)
gear_ratio      float   %6.2f                 Gear Ratio
foreign         byte    %8.0g      origin     Car type
------------------------------------------------------------------------------
Sorted by: foreign
```

Variables and Basic Statistics

- 数据和变量的标签
  - 变量标签

```
. label var price "汽车价格"
. label var foreign "汽车产地(1 国外; 2 国内)"
. des
Contains data from D:\Stata16\ado\base/a/auto.dta
  obs:            74                          这是一份汽车价格资料
 vars:            12                          13 Apr 2018 17:45
                                              (_dta has notes)
------------------------------------------------------------------------------
              storage   display    value
variable name   type    format     label      variable label
------------------------------------------------------------------------------
make            str18   %-18s                 Make and Model
price           int     %8.0gc                汽车价格
mpg             int     %8.0g                 Mileage (mpg)
rep78           int     %8.0g                 Repair Record 1978
headroom        float   %6.1f                 Headroom (in.)
trunk           int     %8.0g                 Trunk space (cu. ft.)
weight          int     %8.0gc                Weight (lbs.)
length          int     %8.0g                 Length (in.)
turn            int     %8.0g                 Turn Circle (ft.)
displacement    int     %8.0g                 Displacement (cu. in.)
gear_ratio      float   %6.2f                 Gear Ratio
foreign         byte    %8.0g       origin    汽车产地(1 国外; 2 国内)
------------------------------------------------------------------------------
Sorted by: foreign
```

Variables and Basic Statistics

- 数据和变量的标签
  - 值标签（数字和文字相对应）

```
browse

label define repair 1 "好" 2 "较      好" 3 "中" 4 "较
差" 5 "差"
//定义一个标签名repair

label values rep78 repair
//将变量值和标签联系起来

browse
```

Variables and Basic Statistics

- 数据和变量的标签
  - 管理值标签

```
label list            //列出值标签的名称和内容
label drop repair     //删除repair
label list
labelbook             // 推荐使用
```

```
. labelbook
-------------------------------------------------------------------------------

value label origin
-------------------------------------------------------------------------------

      values                                  labels
       range:  [0,1]                  string length:  [7,8]
           N:  2                unique at full length:  yes
        gaps:  no                 unique at length 12:  yes
  missing .*:  0                           null string:  no
                              leading/trailing blanks:  no
                                   numeric -> numeric:  no
  definition
           0  Domestic
           1  Foreign
   variables:  foreign
```

Variables and Basic Statistics

- 基本统计量
  - -summarize-命令

```
. sysuse auto, clear
(1978 Automobile Data)
. summarize mpg weight if foreign
```

Variable	Obs	Mean	Std. Dev.	Min	Max
mpg	22	24.77273	6.611187	14	41
weight	22	2315.909	433.0035	1760	3420

Variables and Basic Statistics

- 基本统计量
  - -codebook-命令

```
. codebook price
-------------------------------------------------------------------------------
price
-------------------------------------------------------------------------------

                  type:  numeric (int)
                 range:  [3291,15906]                 units:  1
         unique values:  74                       missing .:  0/74
                  mean:   6165.26
             std. dev:    2949.5
           percentiles:        10%       25%       50%       75%       90%
                              3895      4195    5006.5      6342     11385
```

Variables and Basic Statistics

- 基本统计量
  - -codebook-命令

```
. codebook rep78          //变量中的非重复值小于9，视为类别变量
```

```
rep78                                                                  Repair Rec
```

```
                  type:  numeric (int)
                 range:  [1,5]                        units:  1
         unique values:  5                    missing .:  5/74
            tabulation:  Freq.  Value
                            2  1
                            8  2
                           30  3
                           18  4
                           11  5
                            5  .
```

Variables and Basic Statistics

- 基本统计量
  - 列表统计-table-, -tabulate-

```
. sysuse auto,clear
(1978 Automobile Data)
. tabulate foreign
```

Car type	Freq.	Percent	Cum.
Domestic	52	70.27	70.27
Foreign	22	29.73	100.00
Total	74	100.00	

```
. table foreign
```

Car type	Freq.
Domestic	52
Foreign	22

Variables and Basic Statistics

- 基本统计量
  - 列表统计-table-, -tabulate-

```
. tabulate foreign rep78, summarize(mpg)
```

Means, Standard Deviations and Frequencies of Mileage (mpg)

	Repair Record 1978					
Car type	1	2	3	4	5	Total
Domestic	21	19.125	19	18.444444	32	19.541667
	4.2426407	3.7583241	4.0856221	4.5856055	2.8284271	4.7533116
	2	8	27	9	2	48
Foreign	.	.	23.333333	24.888889	26.333333	25.285714
	.	.	2.5166115	2.7131368	9.367497	6.3098562
	0	0	3	9	9	21
Total	21	19.125	19.433333	21.666667	27.363636	21.289855
	4.2426407	3.7583241	4.1413252	4.9348699	8.7323849	5.8664085
	2	8	30	18	11	69

Variables and Basic Statistics

- 基本统计量
  - 列表统计-table-, -tabulate-

```
. table foreign rep78, c(mean price) f(%9.2f) center row col
```

Car type	Repair Record 1978					
	1	2	3	4	5	Total
Domestic	4564.50	5967.63	6607.07	5881.56	4204.50	6179.25
Foreign			4828.67	6261.44	6292.67	6070.14
Total	4564.50	5967.63	6429.23	6071.50	5913.00	6146.04

Variables and Basic Statistics

- 基本统计量
  - 统计表格-tabstat-

```
. sysuse auto,clear
(1978 Automobile Data)
. tabstat price weight length
```

stats	price	weight	length
mean	6165.257	3019.459	187.9324

```
. tabstat price weight length, stats(mean med min max) col(s) format(%6.2f)
```

variable	mean	p50	min	max
price	6165.26	5006.50	3291.00	15906.00
weight	3019.46	3190.00	1760.00	4840.00
length	187.93	192.50	142.00	233.00

```
. tabstat price weight length, s(mean p25 med p75 min max) c(s) f(%6.2f)
```

variable	mean	p25	p50	p75	min	max
price	6165.26	4195.00	5006.50	6342.00	3291.00	15906.00
weight	3019.46	2240.00	3190.00	3600.00	1760.00	4840.00
length	187.93	170.00	192.50	204.00	142.00	233.00

Variables and Basic Statistics

- 基本统计量
  - 统计表格-tabstat-

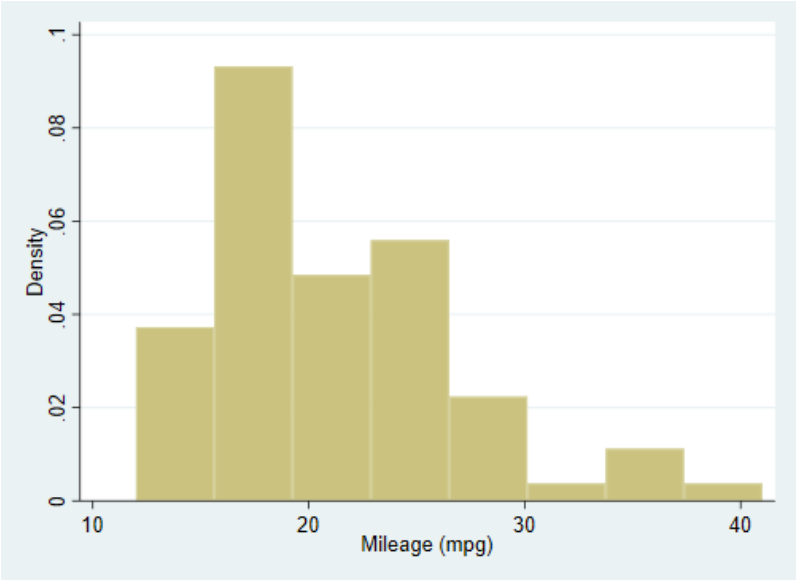
```
. tabstat price weight length, s(mean sd p25 med p75 min max) c(s) f(%6.2f) by(foreign)
Summary for variables: price weight length
     by categories of: foreign (Car type)
```

foreign	mean	sd	p25	p50	p75	min	max
Domestic	6072.42	3097.10	4184.00	4782.50	6234.00	3291.00	15906.00
	3317.12	695.36	2790.00	3360.00	3730.00	1800.00	4840.00
	196.13	20.05	179.50	200.00	209.50	147.00	233.00
Foreign	6384.68	2621.92	4499.00	5759.00	7140.00	3748.00	12990.00
	2315.91	433.00	2020.00	2180.00	2650.00	1760.00	3420.00
	168.55	13.68	156.00	170.00	175.00	142.00	193.00
Total	6165.26	2949.50	4195.00	5006.50	6342.00	3291.00	15906.00
	3019.46	777.19	2240.00	3190.00	3600.00	1760.00	4840.00
	187.93	22.27	170.00	192.50	204.00	142.00	233.00

Variables and Basic Statistics

- 基本图形分析
  - 直方图: 样本的总体分布情况

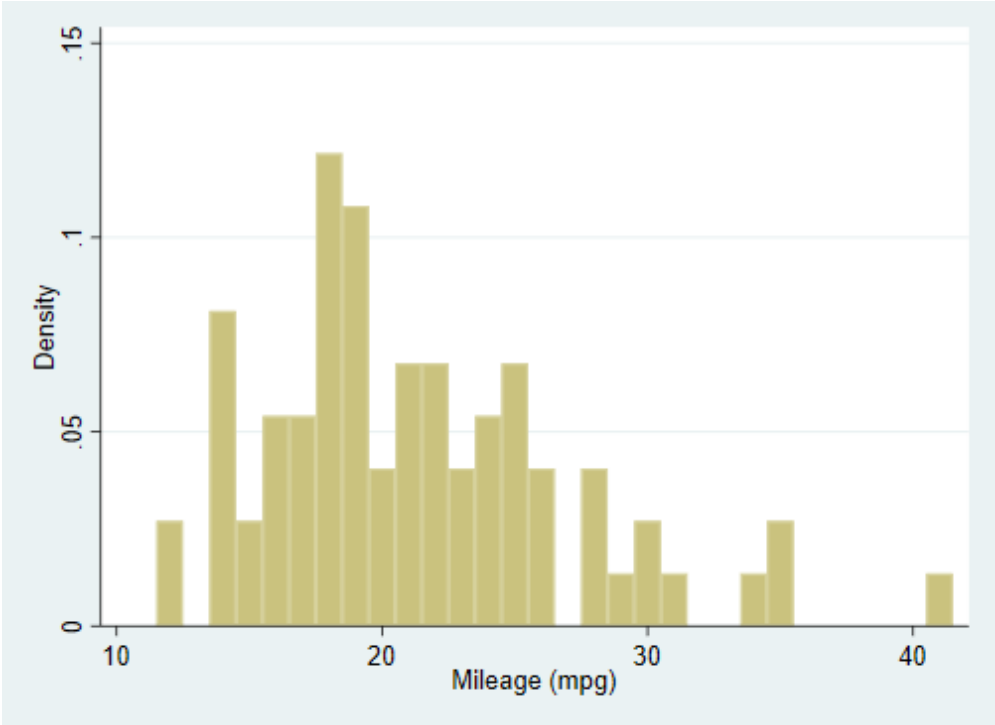
```
. sysuse auto,clear
(1978 Automobile Data)
. histogram mpg
(bin=8, start=12, width=3.625)
. graph export h1.png, width(500) replace
(file h1.png written in PNG format)
```



Variables and Basic Statistics

- 基本图形分析
  - 直方图: 样本的总体分布情况

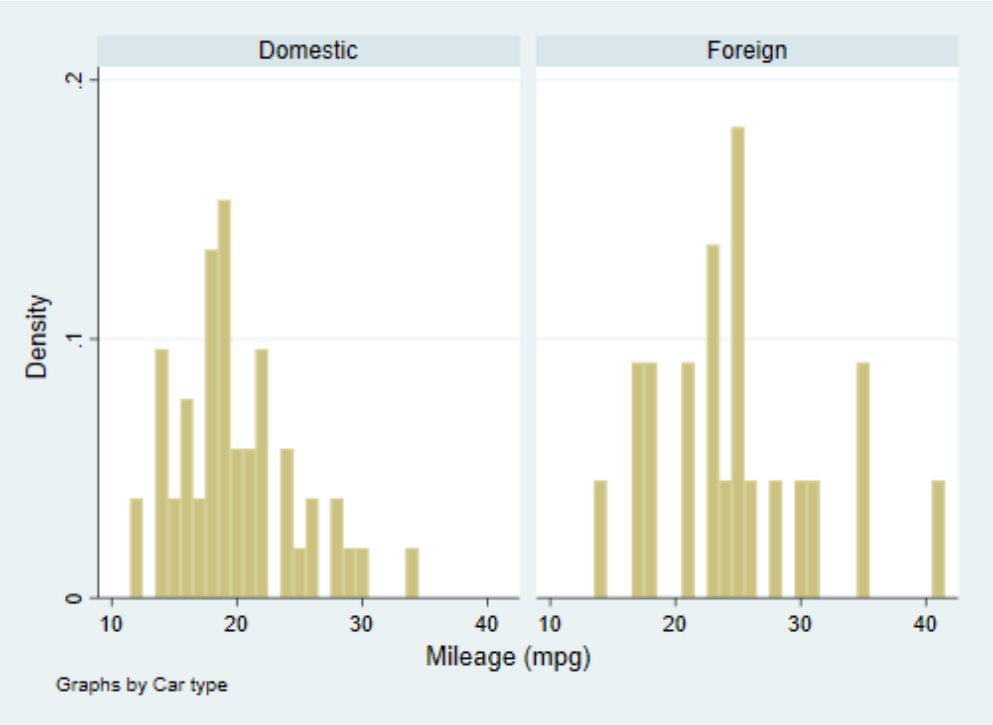
```
. histogram mpg, discrete    //discrete makes a histogram with a bin for each of the 21
> values.
(start=12, width=1)
. graph export h2.png, width(500) replace
(file h2.png written in PNG format)
```



Variables and Basic Statistics

- 基本图形分析
  - 直方图: 样本的总体分布情况

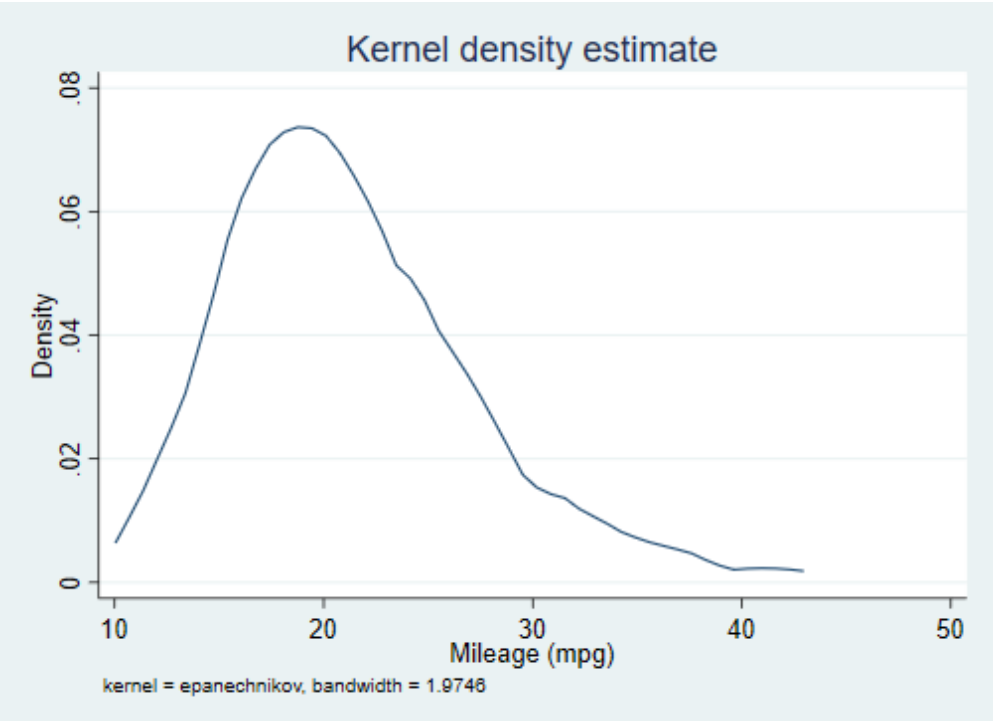
```
. histogram mpg, discrete by(foreign)
. graph export h3.png, width(500) replace
(file h3.png written in PNG format)
```



Variables and Basic Statistics

- 基本图形分析
  - 密度函数图

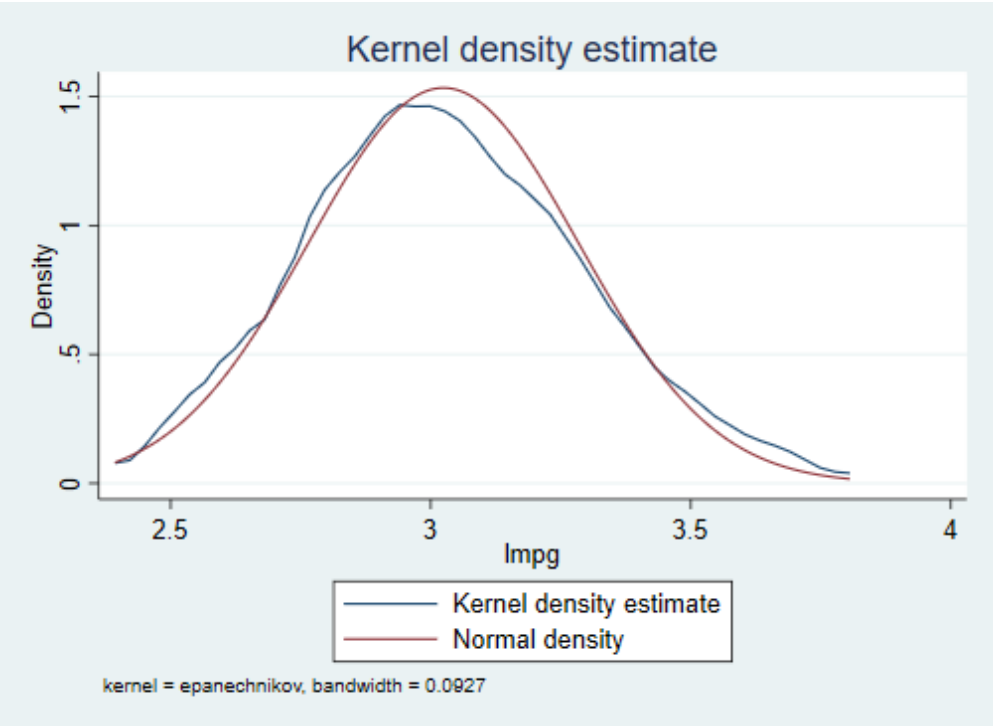
```
. sysuse auto,clear
(1978 Automobile Data)
. kdensity mpg
. graph export k1.png, width(500) replace
(file k1.png written in PNG format)
```



Variables and Basic Statistics

- 基本图形分析
  - 密度函数图

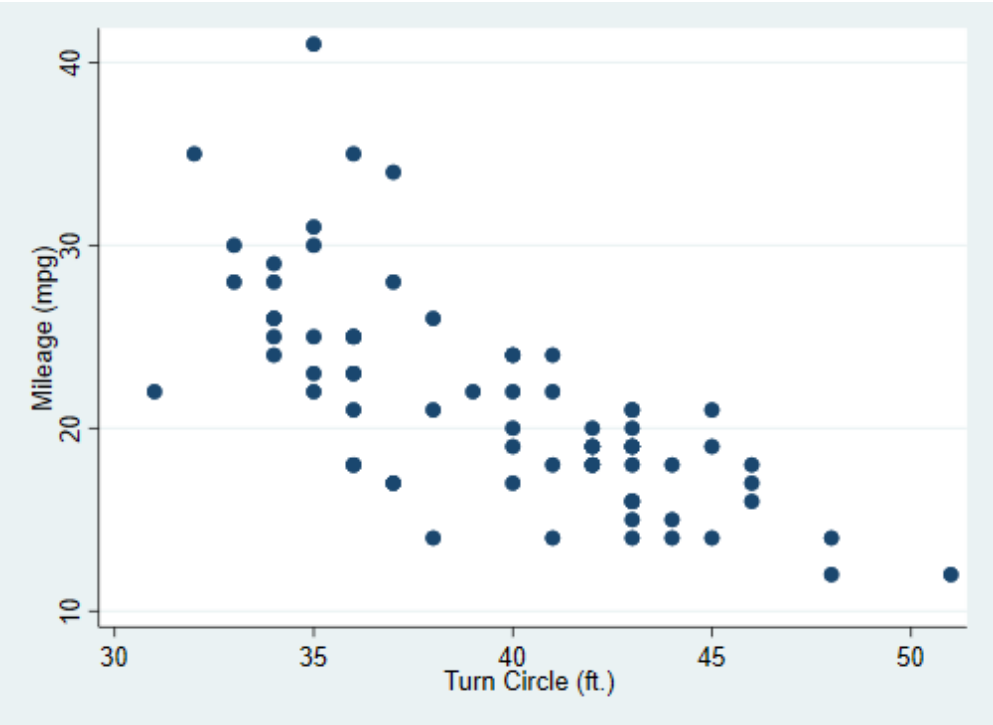
```
. gen lmpg=ln(mpg)
. kdensity lmpg,normal
. graph export k2.png, width(500) replace
(file k2.png written in PNG format)
```



Variables and Basic Statistics

- 散点图

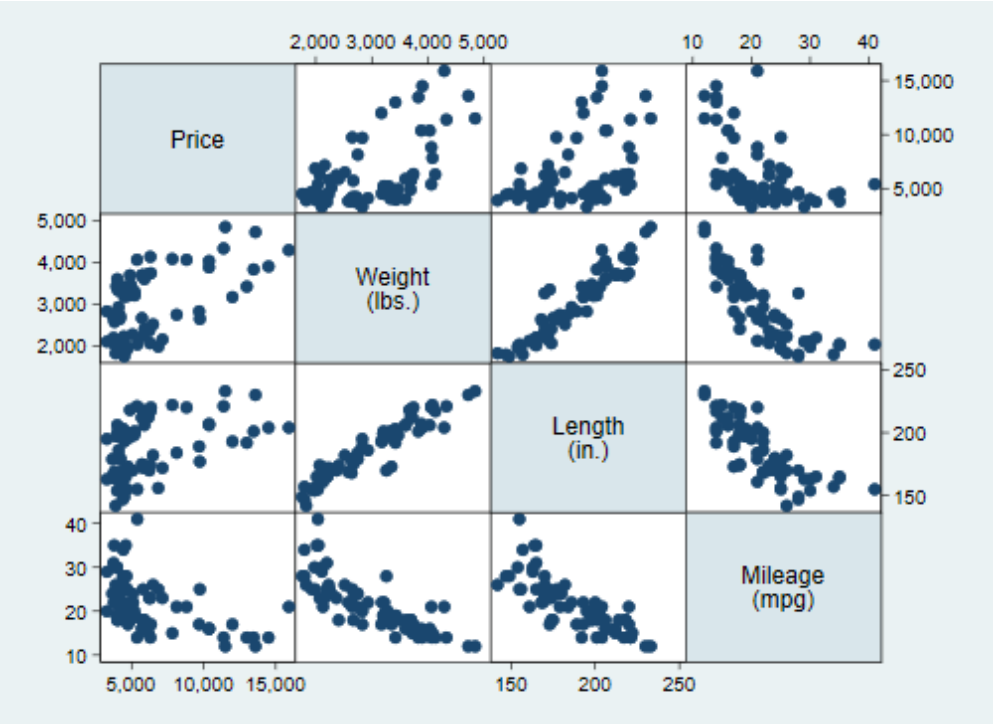
```
. sysuse auto, clear
(1978 Automobile Data)
. scatter mpg turn
. graph export s1.png, width(500) replace
(file s1.png written in PNG format)
```



Variables and Basic Statistics

- 相关系数矩阵

```
. sysuse auto, clear
(1978 Automobile Data)
. graph matrix price wei len mpg
. graph export m1.png, width(500) replace
(file m1.png written in PNG format)
```



Section 9

Do Files

Do Files

- 图形化界面的局限：
  - 命令不易保存、修改，软件关闭，命令即消失；
  - 操作繁琐,每次操作都要不断重复点击界面；
  - 功能组合有限，自由度低，不能进行软件开发。
- command&review窗口的局限：
  - 命令历史记录保存在Review窗口中，查找困难；
  - 零碎命令无条理，无法组织复杂的操作；
  - 与图形化界面类似，command窗口命令也无法长期保存。
- 所以我们需要一个记录、编辑命令的编辑器，Stata自带的编辑器即do文件编辑器，功能类似txt文档，所生成的文件扩展名为【.do】，也就是do文件。
- do文件实际上是Stata命令的集合，方便我们一次性执行多条命令，且使我们的分析工作具有可重复性。

Do Files

- 打开和新建do文档
  - 方法一:快捷键（常用）

Ctrl-key(Windows)	Ctrl-key(Mac)	Definition
Ctrl+D	Command+Shift+D	执行(Do)选中的命令(*)
Ctrl+R	Command+Shift+R	运行程序(Run)(*)
Ctrl+F	Command+F	在do-editor中搜索特定的关键词
Ctrl+O	Command+O	打开do文档
Ctrl+N	Command+N	新建do文档
Ctrl+S	Command+S	保存do文档(*)

(*) 表示仅适用于do-editor

Do Files

- 打开和新建do文档
  - 方法二

```
doedit           //打开do-editor
doedit auto.do //打开一个已存在的do文档，可指定完整路径
```

  - 方法三：Rusults窗口按钮
  - 设置界面属性

Do Files

- 执行do文档
  - 部分执行快捷键：选中需要执行的命令
    `Ctrl+D`（Windows）,`Command+shift+D`(Mac);
  - 整体执行：

```
do auto.do
```

Do Files

- 注释语句

```
help comments
clear all
sysuse auto

*示例:
*第一种注释方式
sum price weight   /*查看price与weight变量部分统计量*/
gen x = 5          //生成取值为5的变量x
```

Do Files

- 三种断行方式： “///” ,“/* */”,#delimit 命令

```
*-第一种断行方式:  /// 物理断行, 逻辑一行

sysuse auto, clear  //调用数据
sum price weight length gear turn
tabstat price weight length gear turn ,                ///
       stats(mean sd p5 p25 med p75 p95 min max)   ///
       format(%6.2f) c(s)

*-第二种断行方式:  /* */

sysuse auto, clear
sum price weight length gear turn
tabstat price weight length gear turn ,          /*
*/ stats(mean sd p5 p25 med p75 p95 min max)   /*
*/ format(%6.2f) c(s)
```

Do Files

- 三种断行方式： “///” ,“/* */”,#delimit 命令

```
*-第三种断行方式:  #delimit 命令      ///表示出现";"才结束

sysuse auto, clear
#delimit ;                                  //delimit声明
tabstat price weight length gear turn ,
stats(mean sd p5 p25 med p75 p95 min max)
      format(%6.2f) c(s) ;
      #delimit cr
```

Do Files

- 注意事项
  - Stata对大小写敏感
  - 注意中英文字符的切换，尤其是逗号，引号
  - 等于号==
  - 尽量避免使用系统预留字段作为变量名,如"_"
  - 各段代码采用一个或多个空行加以分隔
  - 每一行的语句不要过长，不用拖动下方引导条即可阅读

Section 10

Log File

5.录屏神器:log文件

```
    log using "$root\lab1_0916.log" //新建Log文件

*log using "$root\lab1_0916.log",append
                              //接着原来的日志记录
*log using "$root\lab1_0916.log",replace
                              //覆盖原来的日志文件重新记录

    matrix input a = (1,2\3,4)
    matrix list a
    matrix input b = (1,2\1,1)
    matrix list b
```

5.录屏神器:log文件

```
log off // 暂停录制

matrix c = a+b

log on // 继续录制

matrix list c

log close //结束录制

shellout "$root\lab1_0916.log"
```

Reference

- 连玉君Stata初级教程讲义
- Stata统计分析与应用(第3版).电子工业出版社
- https://data.princeton.edu/stata/markdown