

如何简洁地列出指定属性的变量？ds命令来了！

张春强2022 2021-01-16 | 584阅读 | 1转藏

转藏 全屏 朗读 分享



本文作者：寇晓璇，中南财经政法大学统计与数学学院

本文编辑：赵冰洁

技术总编：余术玲

爬虫俱乐部云端课程

爬虫俱乐部于2020年暑期在线上举办的**Stata与Python**编程技术训练营和**Stata数据分析法律与制度**专题训练营已经圆满结束啦~应广大学员需求，我们的课程现已在腾讯课堂全面上线，且**继续提供答疑服务**。现在关注公众号并在朋友圈转发推文《**来腾讯课堂学Stata和Python啦！**》或《**8月Stata数据分析法律与制度专场来啦！**》，即可获得**600元课程优惠券**，集赞50个再领**200元课程优惠券**！（截图发至本公众号后台领取）原价2400元的课程，现在只要**1600元**！

导读：

在处理数据时，若想一目了然地查看当前数据集中的所有变量名，或是想查看指定变量类型的变量名等等，这时ds命令就派上用场啦！简单来说，ds命令能够简洁地列出具有指定属性的变量，还可以列出数据中变量的其他特征并进行编辑。下面就来看一下它的具体用法吧！

一、语法——简单篇

ds [, alpha]

选项	描述
not	列出未在变量列表（ds后面指定的变量列表）中指定的变量
alpha	按字母顺序列出变量
detail	显示更多细节
skip()	变量之间的空格数(默认是skip(2))
varwidth()	显示变量名的宽度(默认是varwidth (12))

1. ds命令单独用可以将所有变量名列出

```
clear
use /data/r16/educ3 //导入数据集
ds //列出当前数据集所有变量名
ds,skip(4) //列出当前数据集所有变量名，变量间距为4
ds,varwidth(5) //列出当前数据集所有变量名，变量名宽度为5
```

```
. ds
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
fips    pop25pls  prcclpls  medhhinc  region  cliffem  emp      osigind  famnw    z         perserv
crimes  pophsppls  prcolhs   medfinc   dc       clfue    empmanuf osigindp fam2w    adjinc    perother
pcrimes popcol     medage    state     tlf      empgovt  emptrade cifbls   famwsamp perman
crimrate perhsppls  perwhite  division  clf      empself  empself  cifuebls pop18pls pertrade

. ds,skip(4)
fips    pophsppls  medage    division  cliffem  empmanuf  cifbls   pop18pls  perserv
crimes  popcol     perwhite  region    clfue    emptrade  cifuebls z         perother
pcrimes perhsppls  medhhinc  dc         empgovt  empself  famnw    adjinc    per~v
crimrate prcclpls  medfinc   tlf        empself  osigind  fam2w    perman
pop25pls prcolhs   state     clf        emp      osigindp  famwsamp pertrade

. ds,varwidth(5)
fips  cri~e  pop~l  prc~s  m~h~c  div~n  tlf  clfue  emp  emp~v  clfb~  fam2w  z  pe~de
cri~s  p~2~s  perh~  med~e  m~f~c  reg~n  clf  emp~t  em~uf  osi~d  ~eb1s  fam~p  adj~c  per~v
pcr~s  poph~  perc~  pe~te  state  dc    clf~m  em~lf  emp~e  osi~p  famnw  p~1~s  per~n  per~r
```

当设定varwidth为5时，变量名宽度大于5的不能显示完全，应当根据变量名长度选择合适的varwidth。可见，若是有一千万个变量以上的数据集出现在你面前时，用ds查看变量名是十分方便的。



张春强2022

★★★★★

+ 关注

对话

TA的最新馆藏

- 当把交互项加入后, 主项的系数符号竟变…
- 控制变量! 控制变量! Good-Controls-…
- 使用Shock-IV的文献汇总
- 文心一言的chatPDF
- 九个独特的GPT提示词
- 一个指令，让chatGPT给你润色英文段…

喜欢该文的人也喜欢

更多

- 一张图教你--如何成为一个驾驶高手 阅28934
- 苏联领导人一览表 74年的历史九位领导人，印象最深的有三位：列… 阅90
- 趁妻子去买菜，江苏语文老师锤杀19岁女儿，留下信件坚称不后悔… 阅2738
- 学会10个强者思维，你的人生会开挂！ 阅178
- 两性关系中的50条铁律 ??? 阅82

热门阅读

换一换

- 谜语大全及答案100个 阅117120
- 原 委托加工合同范本 阅11130
- 苏教版 三年级上册数学教学计划 阅23367
- 中国的传统节日及习俗 阅701319
- 领导班子和领导干部评价用语100例 阅18508

2. ds命令后加“*”或“-”可以有条件的列举变量

```
ds p*           //列出当前数据集所有p开头的变量名
ds pop*, not    //列出所有不以pop开头的变量名
ds popcol-clfue //按顺序列出当前数据集中popcol到clfue的原数据变量名

. ds p*
pcrimes  pophspls  perhspls  prcolhs  pop18pls  pertrade  perother
pop25pls  popcol   perclpls  perwhite  perman    perserv

. ds pop*, not
fips  perhspls  perwhite  division  clf  empself  emperv  clfuebls  z  perserv
crimes  perclpls  medhhinc  region  clffem  emp  osigind  famnw  adjinc  perother
pcrimes  prcolhs  medfinc  dc  clfue  empmanuf  osigindp  fam2w  perman
crimrate  medage  state  tlf  empgovt  emptrade  clfb1s  famwsamp  pertrade

. ds popcol-clfue
popcol  perclpls  medage  medhhinc  state  region  tlf  clffem
perhspls  prcolhs  perwhite  medfinc  division  dc  clf  clfue
```

3. ds命令后加“, alpha”可以按字母顺序显示变量名称

```
ds, alpha //按字母顺序纵向列出当前数据集中所有变量名

adjinc  clfue  dc  empmanuf  fam2w  medage  osigindp  perman  perwhite  pophspls  tlf
clf  clfuebls  division  empself  famnw  medfinc  pcrimes  perother  pop18pls  prcolhs  z
clfb1s  crimes  emp  emperv  famwsamp  medhhinc  perclpls  perserv  pop25pls  region
clffem  crimrate  empgovt  emptrade  fips  osigind  perhspls  pertrade  popcol  state
```

4. ds命令后加“,detail”可以详细的显示变量信息

```
ds, detail //显示当前数据集中变量的详细信息

variable name      storage  display  value  variable label
type              format  label
fips               long      %10.0g  state/place code
crimes             long      %10.0g  Crimes known to police 1981
pcrimes            long      %10.0g  Property crimes 1981
crimrate           int       %8.0g   Crimes per 100,000 pop. 1981
pop25pls           long      %10.0g  Persons aged 25+
pophspls           long      %10.0g  Persons 25+ H.S. or more
popcol             long      %10.0g  Persons 25+ 4+ yrs college
perhspls           float     %9.0g   % HS | 25+
```

二、语法——进阶篇

ds [varlist] [, options]

选项	描述
has(spec)	描述符合该规范的子集
not(spec)	描述不符合该规范的子集
insensitive	执行不区分大小写的模式匹配
indent	缩进输出(很少使用)

概括来说，可以分为has()和not()两大“派别”，我们从字面意思也很容易理解，“有”和“无”，即has是列举出括号中的指定类型，而not是列举出除括号指定类型外的类型。其中括号中的spec包含如下：

选项	描述
type typelist	指定类型
format patternlist	显示格式匹配模式列表
varlabel [patternlist]	变量标签/变量标签匹配模式列表
char [patternlist]	特征/特征匹配模式列表
vallabel [patternlist]	值标签/值标签匹配模式列表

1.ds命令后加“, has(type string)”可以显示所有字符型变量名

```
clear
sysuse auto
ds, has(type string) //列举所有字符型变量名
```

```
. clear

. sysuse auto
(1978 Automobile Data)

. ds, has(type string)
make
```

上例中的typelist为string，一般而言，typelist可以包含一种或多种类型，可以是数值型（numeric）、字符型（string）、整型（int）、浮点型（float）等，还可以是数组（numlist）。

2.ds命令后加”，has(format *f)”可以显示所有浮点型变量的名称

```
ds, has(format *f) //列举所有浮点型变量名
```

```
. ds, has(format *f)
headroom    gear_ratio
```

同理，

```
ds, has(format *c) //列出使用逗号格式的变量名
```

```
. ds, has(format *c)
price       weight
```

三、存储结果

使用ds命令还能够生成局部宏，可以调用局部宏来对变量进行运算、排序或描述性统计等。

```
clear
sysuse auto
ds, has(type string)
```

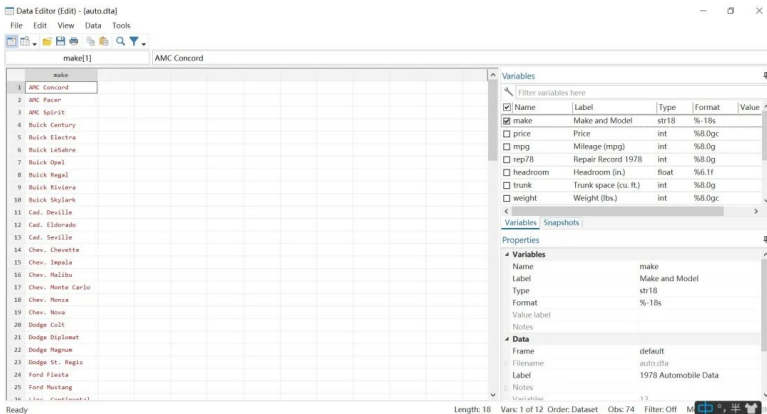
```
edit `r(varlist)' //列出所有字符串变量并弹出包含它们的数据编辑器
```

```
. clear

. sysuse auto
(1978 Automobile Data)

. ds, has(type string)
make

. edit `r(varlist)'
```



```
clear
sysuse auto
ds, has(type numeric)
```

```
order `r(varlist)' //列出所有数值型变量并将它们重新定位到数据集的开头
```

```
ds, has(type numeric)
```

```
summarize `r(varlist)' //列出所有数值型变量并对其进行描述性统计
```

Variable	Obs	Mean	Std. Dev.	Min	Max
price	74	6165.257	2949.496	3291	15906
mpg	74	21.2973	5.785503	12	41
rep78	69	3.405797	.9899323	1	5
headroom	74	2.993243	.8459948	1.5	5
trunk	74	13.75676	4.277404	5	23
weight	74	3019.459	777.1936	1760	4840
length	74	187.9324	22.26634	142	233
turn	74	39.64865	4.399354	31	51
displacement	74	197.2973	91.83722	79	425
gear_ratio	74	3.014865	.4562871	2.19	3.89
foreign	74	.2972973	.4601885	0	1

通过以上方法可以快速、准确的对变量进行排序和描述性统计。今天对ds命令的介绍就先到这，快去着手试试吧！



对我们的推文累计打赏超过1000元，我们即可给您开具发票，发票类别为“咨询费”。用心做事，不负您的支持！

往期推文推荐

- 如何在Python中进行描述性统计分析？
- 分析师和他们的雇主重视与管理层接触吗？——分析师参与盈余电话会议的研究
- 繁忙的董事与公司业绩：来自并购的证据
- 使用Python接口读取CSMAR数据
- 列表生成式让你的代码简洁又美观
- Stata中变量观测值的亲密伙伴——levelsof命令
- 爬虫俱乐部开发的命令更新及常见问题说明
- Seminar | 道德培训真的有用吗？
- 网络爬虫入门之requests 库的基本使用——以亚马逊图书界面为例
- reduce()函数和filter()函数闪亮登场
- “环环”入扣之foreach命令
- 统计年鉴数据整理小技巧
- Seminar | 作为飞行员，我比别的CEO多了什么？
- 利用TensorFlow构建前馈神经网络
- 推文合集（1）| Stata学习者必看的n篇推文！
- Seminar | 诚信的价值
- 利用tushare获取股票数据及实现可视化
- 从Excel到Stata的“摆渡车”——import excel命令
- 光阴十载，见证了《经济研究》中的“高被引”
- 利用tushare获取股票数据

在Python中实现Stata的stack功能

关于我们

微信公众号“Stata and Python数据分析”分享实用的stata、python等软件的数据处理知识，欢迎转载、打赏。我们是由李春涛教授领导下的研究生及本科生组成的大数据处理和分析团队。

此外，欢迎大家踊跃投稿，介绍一些关于stata和python的数据处理和分析技巧。
投稿邮箱：statatraining@163.com
投稿要求：
1) 必须原创，禁止抄袭；
2) 必须准确，详细，有例子，有截图；
注意事项：
1) 所有投稿都会经过本公众号运营团队成员的审核，审核通过才可录用，一经录用，会在该推文里作为作者署名，并有奖金分成。
2) 邮件请注明投稿，邮件名称为“投稿+推文名称”。
3) 应广大读者要求，现开通有偿问答服务，如果大家遇到有关数据处理、分析等问题，可以在公众号中提出，只需支付少量赏金，我们会在后期的推文里给予解答。



本站是提供个人知识管理的网络存储空间，所有内容均由用户发布，不代表本站观点。请注意甄别内容中的联系方式、诱导购买等信息，谨防诈骗。如发现有害或侵权内容，请点击一键举报。

☆ 转藏

📄 分享

🌸 献花 (0)

类似文章

更多



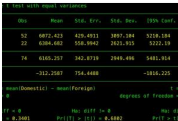
前沿: 删失数据分位数工具变量(CQIV)估计, 做删失数据异质性效应分析

前沿: 删失数据分位数工具变量(CQIV)估计, 做删失数据异质性效应分析。在本文中, 我们介绍了一个Stata命令——cqiv, 该命令简化了CQIV估...



Stata1

Stata1Excel格式数据的导入_变量名称修改_创建值标签。Stata入门——导入数据。Stata字符型与数值型变量的相互转换。公众...



均值T检验

均值T检验。相较于通常的描述性统计, 均值T检验则进一步提供了不同组别之间的差异信息, 并能够为接下来通过实证考察组别异质性提供一定...



互助问答第279期: 泊松分布工具变量回归

互助问答第279期: 泊松分布工具变量回归。尊敬的老师, 请教如何进行泊松分布工具变量回归 (stata)。请问有相关的命令或者论文介绍么? ...

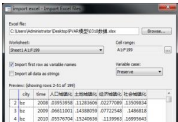
用stata做空间计量的经验贴

用stata做空间计量的经验贴本人准大四生, 之前为了完成发展经济学的论文接触了一下空间计量的内容, 使用的软件是stata, 折腾了好久算是有点心得, 想和和我一样的纯新手分享一下。于是我们有了矩阵 (说...

表 2.2 数值型数据	
	最大值
	100
	32740
	2147483620
19*10 ¹⁸	1.70141173319*10 ¹⁶
3*10 ¹⁰⁷	8.9884656743*10 ¹⁰⁸

STATA数据的类型与压缩[转载]

STATA数据的类型与压缩[转载]1. 数据的类型Stata常用的数据类型主要有三类: 数值型、字符型和日期型, 下面逐一进行介绍。字符型数据一般...



Stata系列02: Stata数据导入

Stata系列02: Stata数据导入导读 大家好, 在上一期介绍了Stata软件版本, 以及Stata做实证分析的优势。本期将介绍Stata最常用的数据导...

Stata中Mata的st_view函数-小心使用!

Stata中Mata的st_view函数-小心使用! 这意味着, 创建了Mata对象x, 它是Stata数据集中变量x的视图, 如果Stata数据集中的变量x改变位置, 则Mata st_view对象x将不再指向正确的变量。mata: mata clear vo...

Table 1: Summary statistics		
mean	sd	min
7.767	5.756	1.005
39.153	3.060	34.000
1.283	0.475	1.000
37.218	10.509	1.000

Stata: 毕业论文大礼包 C——新版 esttab

LaTeXlhttest—LaTeX连享会计量方法专题……, https://gitee.com/arlionn/Course6. 相关系数矩阵输出 (lxcorr) 输出至单个 Word 文件...