# A fuel consumption study of Stata's auto dataset

We conduct a study of the fuel consumption of cars in Stata's auto dataset.

## 生成数据

从变量里程生成新变量**油耗**(公升每一百公里)。

. generate 油耗 = 100/里程  
. label variable 油耗 "油量消耗(公升每一百公里)"

## 检验数据

## 描述变量

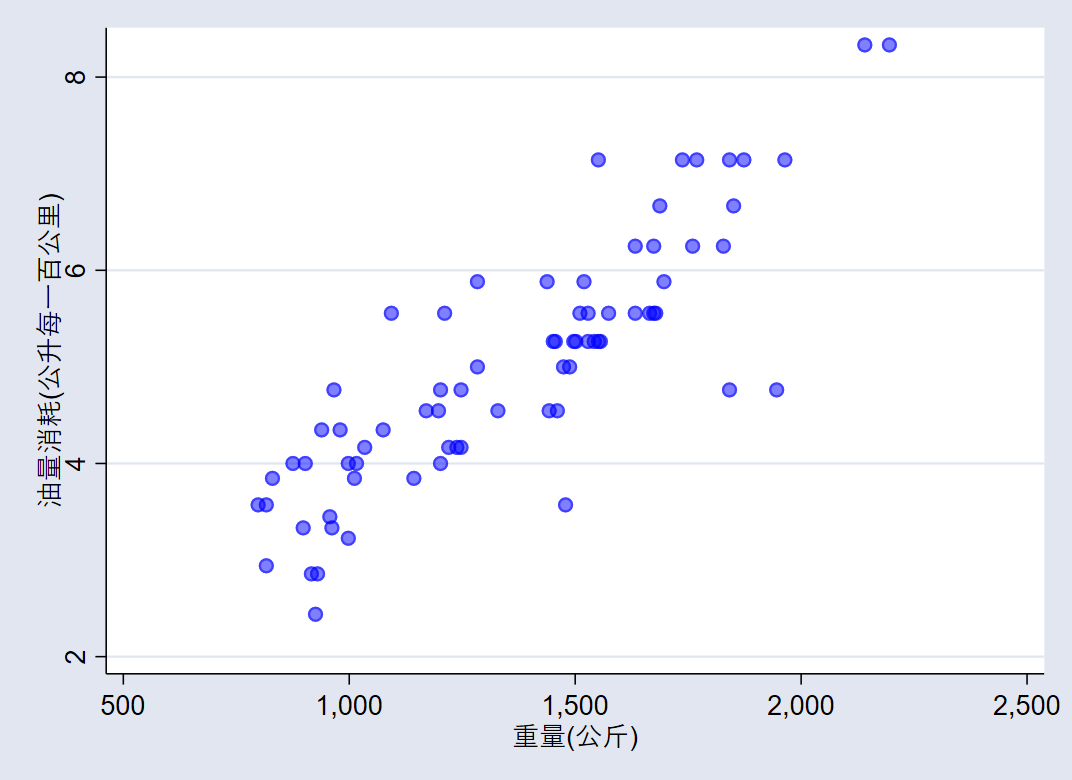
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 变量名 | 类型 | 格式 | 值标签 | 标签 |
| 油耗 | float | %9.0g |  | 油量消耗(公升每一百公里) |
| 重量 | float | %8.0gc |  | 重量(公斤) |

## 摘要统计

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 变量 | 观测 | 均值 | 标准差 | 最小值 | 最大值 |
| 油耗 | 74 | 5.019 | 1.28 | 2.439 | 8.333 |
| 重量 | 74 | 1370 | 352.5 | 798.3 | 2195 |

变量**重量**的最小值798.32,最大值2195.39,极差1397.06.

## 图: 油耗与重量

  
**图1: 油耗与重量**

## 研究油耗与重量关系 - 线性回归

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 油耗 | Coef. | Std. Err. | t | P>|t| | [95% Conf. Interval] | |
| 重量 | .003102 | .0002223 | 13.95 | 0.000 | .0026589 | .0035452 |
| \_cons | .7707669 | .3142571 | 2.45 | 0.017 | .1443069 | 1.397227 |

The regression shows that for every unit increase in weight, a *0.0031* unit increase in fuel consumption is predicted.

## Produce a table from -estimates table-

We list the results from two regressions.

|  |  |  |
| --- | --- | --- |
| Variable | model1 | model2 |
| 重量(公斤) | 0.0030\*\*\* | 0.0028\*\*\* |
| 变速比 | 0.1706 | -0.3367 |
| 转弯半径(米) | 0.0798 | 0.2010 |
| 国籍 |  | 0.8650\*\*\* |
| Constant | -0.5814 | -0.4661 |
| N | 74 | 74 |
| r2 | 0.7332 | 0.7767 |
| r2\_a | 0.7218 | 0.7637 |

## Produce a table from community-contributed -esttab-

We list the results from the same two regressions as above using -esttab- and -putdocx-. -esttab- is a popular community-contributed command which generates tables for report.

|  |  |  |
| --- | --- | --- |
| 线性回归表使用esttab | | |
|  | 油耗 | 油耗 |
|  | b/t | b/t |
| 重量(公斤) | 0.003\*\*\* | 0.003\*\*\* |
|  | (6.09) | (6.06) |
| 变速比 | 0.171 | -0.337 |
|  | (0.64) | (-1.19) |
| 转弯半径(米) | 0.080 | 0.201 |
|  | (0.70) | (1.81) |
| 国籍 |  | 0.865\*\*\* |
|  |  | (3.66) |
| Constant | -0.581 | -0.466 |
|  | (-0.38) | (-0.33) |
| Observations | 74 | 74 |
| R-squared | 0.73 | 0.78 |
| Adjusted R-squared | 0.72 | 0.76 |
| *t statistics in parentheses* | | |
| *\* p<0.05, \*\* p<0.01, \*\*\* p<0.001* | | |

## 不同国籍车辆对比

|  |  |
| --- | --- |
| 国外 | 国内 |
| |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  | 观测 | 均值 | 标准差 | 最小值 | 最大值 | | 油耗 | 22 | 4.313 | 1.144 | 2.439 | 7.143 | | 重量 | 22 | 1050 | 196.4 | 798.3 | 1551 | | |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  | 观测 | 均值 | 标准差 | 最小值 | 最大值 | |  | 52 | 5.318 | 1.224 | 2.941 | 8.333 | |  | 52 | 1505 | 315.4 | 816.5 | 2195 | |
|  |  |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | b | t | b | t | | 重量(公斤) | 0.003\*\* | (3.49) | 0.003\*\*\* | (5.07) | | 变速比 | -0.577 | (-1.12) | -0.234 | (-0.72) | | 转弯半径(米) | 0.672 | (1.67) | 0.217 | (1.92) | | Constant | -4.478 | (-1.00) | -0.640 | (-0.43) | | Observations | 22 |  | 52 |  | | R-squared | 0.73 |  | 0.78 |  | | Adjusted R-squared | 0.69 |  | 0.77 |  | | |

## Stata命令输出

. use auto\_zh, clear  
(1978年汽车数据)  
  
. generate 油耗 = 100/里程  
  
. label variable 油耗 "油量消耗(公升每一百公里)"  
  
. regress 油耗 重量  
  
 Source | SS df MS Number of obs = 74  
-------------+---------------------------------- F(1, 72) = 194.71  
 Model | 87.2964971 1 87.2964971 Prob > F = 0.0000  
 Residual | 32.2797637 72 .448330051 R-squared = 0.7300  
-------------+---------------------------------- Adj R-squared = 0.7263  
 Total | 119.576261 73 1.63803097 Root MSE = .66957  
  
------------------------------------------------------------------------------  
 油耗 | Coef. Std. Err. t P>|t| [95% Conf. Interval]  
-------------+----------------------------------------------------------------  
 重量 | .003102 .0002223 13.95 0.000 .0026589 .0035452  
 \_cons | .7707669 .3142571 2.45 0.017 .1443069 1.397227  
------------------------------------------------------------------------------  
  
. mata:  
------------------------------------------------- mata (type end to exit) -----  
: st\_view(Y=.,.,("油耗"), .)  
  
: st\_view(X=.,.,("重量"), .)  
  
: X=X,J(rows(X),1,1)  
  
: b=invsym(X'\*X)\*X'\*Y  
  
: v=((Y- X\*b)'\*(Y- X\*b))/(rows(X)-cols(X))\*invsym(X'\*X)   
  
: se=sqrt(diagonal(v))  
  
: t=b:/se  
  
: p=2\*ttail(rows(X)-cols(X),abs(t))  
  
: b,se,t,p  
 1 2 3 4  
 +---------------------------------------------------------+  
 1 | .0031020043 .0002223018 13.95402458 3.66622e-22 |  
 2 | .7707668702 .3142570686 2.45266359 .0166043065 |  
 +---------------------------------------------------------+  
  
: end  
-------------------------------------------------------------------------------