# A fuel consumption study of Stata's auto dataset

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

## Perform data transformation

We generate a variable, **fuel**, in the unit of Gallons per 100 Miles based on the variable **mpg**.

. generate fuel = 100/mpg
. label variable fuel "Fuel consumption (Gallons per 100 Miles)"

## Examine the variables

We examine variables for possible errors in the data.

### Describe the variables

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Name | Storage type | Display format | Value label | Variable label |
| fuel | float | %9.0g |  | Fuel consumption (Gallons per 100 Miles) |
| weight | int | %8.0gc |  | Weight (lbs.) |

### Summarize the variables

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Observations | Mean | Standard Deviation | Min | Max |
| fuel | 74 | 5.019 | 1.28 | 2.439 | 8.333 |
| weight | 74 | 3019 | 777.2 | 1760 | 4840 |

The variable **weight** has minimum value 1760.00, maximum value 4840.00, and range 3080.00.

## Plot fuel consumption and vehicle weight


**Figure 1: scatter plot fuel consumption and weight**

## Explore relationship between fuel consumption and vehicle weight - linear regression

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| fuel | Coef. | Std. Err. | t | P>|t| | [95% Conf. Interval] |
| weight | .001407 | .0001008 | 13.95 | 0.000 | .001206 | .0016081 |
| \_cons | .7707669 | .3142571 | 2.45 | 0.017 | .1443069 | 1.397227 |

The regression shows that for every unit increase in weight, a *0.0014* unit increase in fuel consumption is predicted.

## Produce a table from -estimates table-

We list the results from two regressions.

|  |  |  |
| --- | --- | --- |
| Variable | model1    | model2    |
| Weight (lbs.) | 0.0014\*\*\* | 0.0013\*\*\* |
| Gear Ratio | 0.1706    | -0.3367    |
| Turn Circle (ft.)  | 0.0243    | 0.0613    |
| Car type |     | 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.

|  |
| --- |
| Regression table using -esttab- |
|  | Fuel consumption | Fuel consumption |
|  | b/t | b/t |
| Weight (lbs.) | 0.001\*\*\* | 0.001\*\*\* |
|  | (6.09) | (6.06) |
| Gear Ratio | 0.171 | -0.337 |
|  | (0.64) | (-1.19) |
| Turn Circle (ft.)  | 0.024 | 0.061 |
|  | (0.70) | (1.81) |
| Car type |  | 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* |

## Relationship based on car type - a nested table

|  |  |
| --- | --- |
| Foreign | Domestic |
|

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Obs | Mean | SD | Min | Max |
| fuel | 22 | 4.313 | 1.144 | 2.439 | 7.143 |
| weight | 22 | 2316 | 433 | 1760 | 3420 |

 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Obs | Mean | SD | Min | Max |
|  | 52 | 5.318 | 1.224 | 2.941 | 8.333 |
|  | 52 | 3317 | 695.4 | 1800 | 4840 |

 |
|  |  |
|

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | b | t | b | t |
| Weight (lbs.) | 0.002\*\* | (3.49) | 0.001\*\*\* | (5.07) |
| Gear Ratio | -0.577 | (-1.12) | -0.234 | (-0.72) |
| Turn Circle (ft.)  | 0.205 | (1.67) | 0.066 | (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 |  |

 |

## Output from Stata commands

. sysuse auto, clear
(1978 Automobile Data)

. generate fuel = 100/mpg

. label variable fuel "Fuel consumption (Gallons per 100 Miles)"

. regress fuel weight

 Source | SS df MS Number of obs = 74
-------------+---------------------------------- F(1, 72) = 194.71
 Model | 87.2964969 1 87.2964969 Prob > F = 0.0000
 Residual | 32.2797639 72 .448330054 R-squared = 0.7300
-------------+---------------------------------- Adj R-squared = 0.7263
 Total | 119.576261 73 1.63803097 Root MSE = .66957

------------------------------------------------------------------------------
 fuel | Coef. Std. Err. t P>|t| [95% Conf. Interval]
-------------+----------------------------------------------------------------
 weight | .001407 .0001008 13.95 0.000 .001206 .0016081
 \_cons | .7707669 .3142571 2.45 0.017 .1443069 1.397227
------------------------------------------------------------------------------

. mata:
------------------------------------------------- mata (type end to exit) -----
: st\_view(Y=.,.,("fuel"), .)

: st\_view(X=.,.,("weight"), .)

: 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 | .0014070443 .0001008343 13.95402451 3.66623e-22 |
 2 | .7707669021 .3142570678 2.452663698 .0166043019 |
 +---------------------------------------------------------+

: end
-------------------------------------------------------------------------------