How can I create Added Variable Plots in R?

Authored by

stats writer

May 1, 2024

Added Variable Plots, also known as partial regression plots, are a graphical tool used to visualize the relationship between a response variable and a predictor variable while controlling for the influence of other predictor variables. In order to create Added Variable Plots in R, you can use the "car" package and the "avPlots()" function. This function takes in a model object and produces a plot with the response variable on the y-axis and the partial residuals of the predictor variable on the x-axis. It also includes a fitted line to show the relationship between the two variables. By adding multiple predictor variables to the model, you can create multiple Added Variable Plots to compare the effects of different predictors on the response variable. These plots are useful for identifying influential predictors and detecting nonlinear relationships.

Create Added Variable Plots in R

In statistics, added variable plots are individual plots that display the relationship between a and one predictor variable in a multiple linear regression model, while controlling for the presence of other predictor variables in the model.

Note: Sometimes these plots are also called "partial regression plots."

These type of plots allow us to observe the relationship between each individual predictor variable and the response variable in a model while holding other predictor variables constant.

To create added variable plots in R, we can use the avPlots() function from the car package:

```
#load car package
library(car)

#fit multiple linear regression model
model <- lm(y ~ x1 + x2 + ..., data = df)

#create added variable plots
avPlots(model)
```

The following example shows how to use this syntax in practice.

Example: Added Variable Plots in R

Suppose we fit the following multiple linear regression model in R, using data from the mtcars dataset:

```
#fit multiple linear regression model
model <- lm(mpg ~ disp + hp + drat, data = mtcars)

#view summary of model
summary(model)

Call:
lm(formula = mpg ~ disp + hp + drat, data = mtcars)

Residuals:
```

Min 1Q Median 3Q Max
-5.1225 -1.8454 -0.4456 1.1342 6.4958

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 19.344293 6.370882 3.036 0.00513 **
disp -0.019232 0.009371 -2.052 0.04960 *
hp -0.031229 0.013345 -2.340 0.02663 *
drat 2.714975 1.487366 1.825 0.07863 .

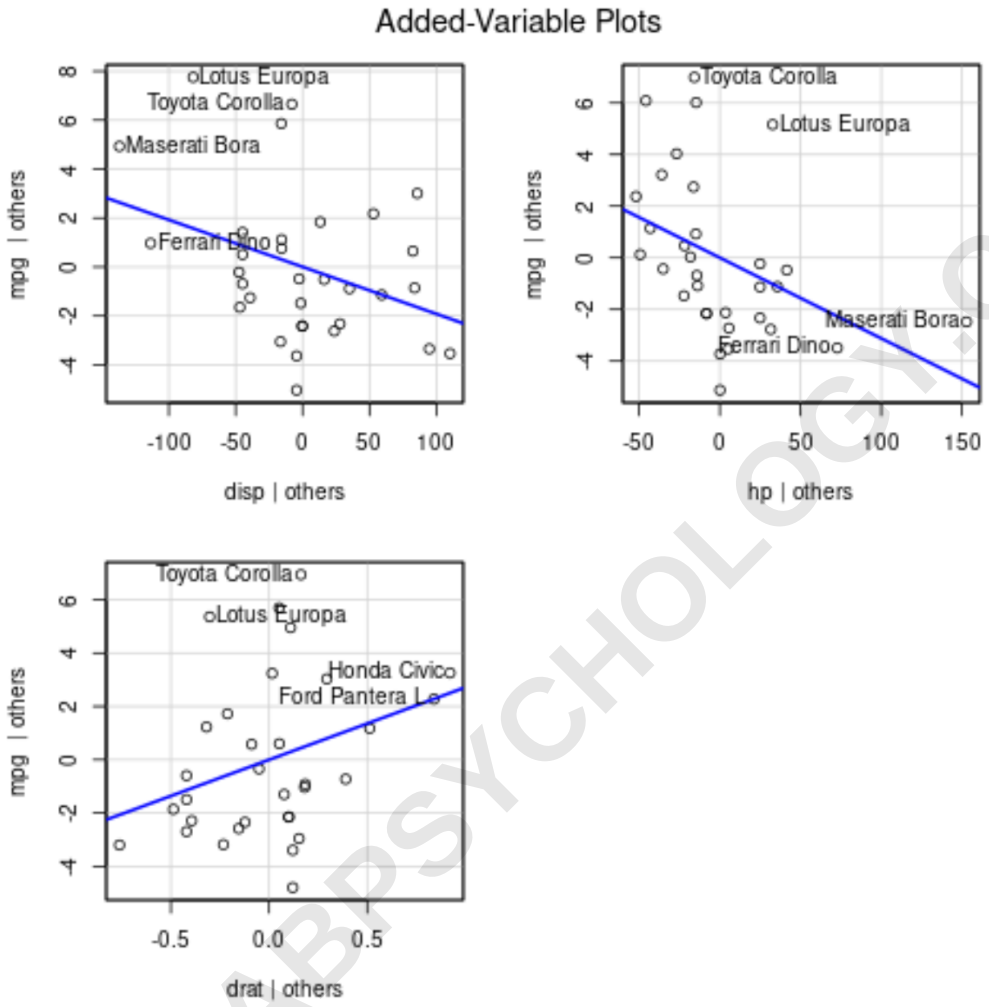
Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**

Residual standard error: 3.008 on 28 degrees of freedom
Multiple R-squared: 0.775, Adjusted R-squared: 0.7509
F-statistic: 32.15 on 3 and 28 DF, p-value: 3.28e-09

To visualize the relationship between the response variable "mpg" and each individual predictor variable in the model, we can produce added variable plots using the avPlots() function:

#load car package
library(car)

#produce added variable plots
avPlots(model)



Here is how to interpret each plot:

The x-axis displays a single predictor variable and the y-axis displays the response variable.The blue line shows the association between the predictor variable and the response variable, *while holding the value of all*

***other predictor variables constant*.The points that are labelled in each plot represent the two observations with the largest and the two observations with the largest partial leverage.**

Note that the angle of the line in each plot matches the sign of the coefficient from the estimated regression equation.

For example, here are the estimated coefficients for each predictor variable from the model:

disp: -0.019232hp: -0.031229drat: 2.714975