

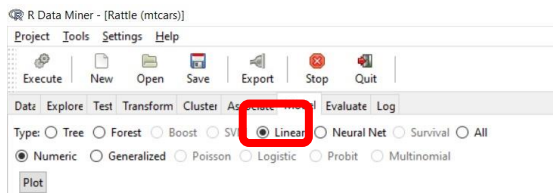
This document was prepared using R version 3.6.3
Depending on R version, the output may be slightly different, and that's acceptable.

Action 1 – Read the data and press execute.

First Model - Weight – Miles_per_Gallon(mpg)

Action2: Make mpg as target and wt as an input variable and ignore all other variables.

Action3: Go to the “Model” tab.



The screenshot shows the R Data Miner - Rattle (mtcars) interface. The 'Model' tab is selected, and the 'Linear' radio button is chosen. The 'Numeric' radio button is also selected. The 'Plot' button is visible. Below the model selection, the summary of the Linear Regression model (built using lm) is displayed.

```
Summary of the Linear Regression model (built using lm):  
Call:  
lm(formula = mpg ~ ., data = crs$dataset[, c(crs$input, crs$target)])  
Residuals:  
    Min       1Q   Median       3Q      Max   
-4.5432 -2.3647 -0.1252  1.4096  6.8727   
Coefficients:  
            Estimate Std. Error t value Pr(>|t|)      
(Intercept)  37.2851    1.8776   19.858 < 2e-16 ***  
wt           -5.3445    0.5591   -9.559 1.29e-10 ***  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
Residual standard error: 3.046 on 30 degrees of freedom  
Multiple R-squared:  0.7528,    Adjusted R-squared:  0.7446   
F-statistic: 91.38 on 1 and 30 DF,  p-value: 1.294e-10  
==== ANOVA ====  
Analysis of Variance Table  
Response: mpg  
          Df Sum Sq Mean Sq F value    Pr(>F)      
wt          1  847.73   847.73   91.375 1.294e-10 ***  
Residuals  30  278.32    9.28                   
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
[1] "\n"  
Time taken: 0.01 secs  
Rattle timestamp: 2020-05-09 10:22:04 ruchl
```

Action4: Select type as Linear and Numeric and then “Execute”.

Interpretation: The **Weight – Miles_per_Gallon(mpg)** model estimates a decrease in mileage of 5.34 miles per gallon with 1 unit increase in weight, and about 75 % of the variation in percentage of Miles_per_Gallon is associated with variation in weight.

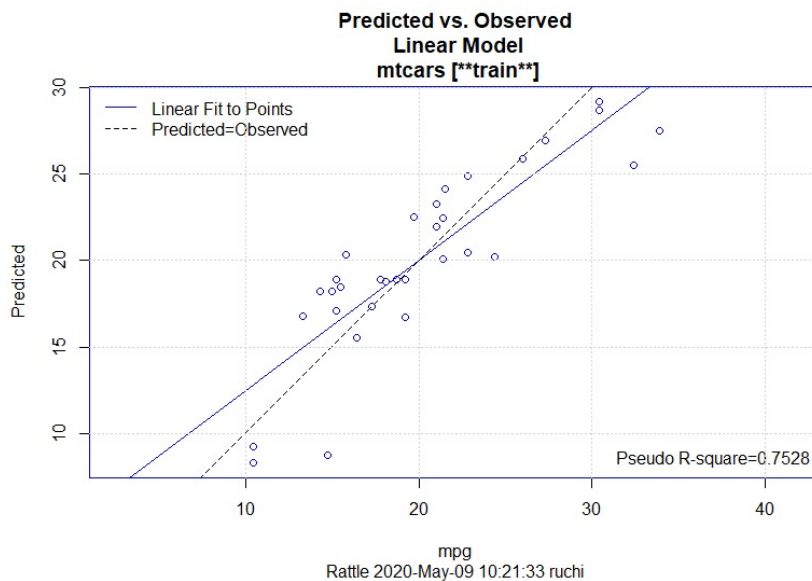
Assessing the Visual Fit

Action5: Go to “Evaluate” tab.

This document was prepared using R version 3.6.3
Depending on R version, the output may be slightly different, and that's acceptable.

Action6: Select type “Pr v Ob” and then “Execute”.

Output : Available in plot/ graph section of RStudio.



Interpretation: Pseudo R squared is 0.7528 . The “predicted vs observed” line is the line when our predictions are perfect. As we see the “linear fit to points” has almost similar slope, we would conclude we have a very good fit.

Horse_Power – Miles_per_Gallon (mpg):

Action7: Make mpg as target and hp as an input variable.

Action8: Go to the “Model” tab.

Action9: Select type as Linear and Numeric and then “Execute”.

This document was prepared using R version 3.6.3

Depending on R version, the output may be slightly different, and that's acceptable.

Interpretation: The **Horse_Power – Miles_per_Gallon (mpg)** model estimates a decrease in mileage of 0.06 miles per gallon with 1 unit increase in horse power, and about 60% of the variation in percentage of Miles_per_Gallon is associated with variation in horse power.

Why do you think the joint estimate produces different estimates for the effect of Horse Power and Weight when compared to individual regressions?

Answer – The variation of weight explains around 75% variance in mpg. The rest 25 % was unexplained. When we added the horse_power to the model, it could explain some part of the 25% (unexplained by the weight) also. Thus, overall variance explained increases to 82% when we add both horse_power and weight.

Assessing the Visual Fit

Action10: Go to “Evaluate” tab.

Action11: Select type “Pr v Ob” and then “Execute”.

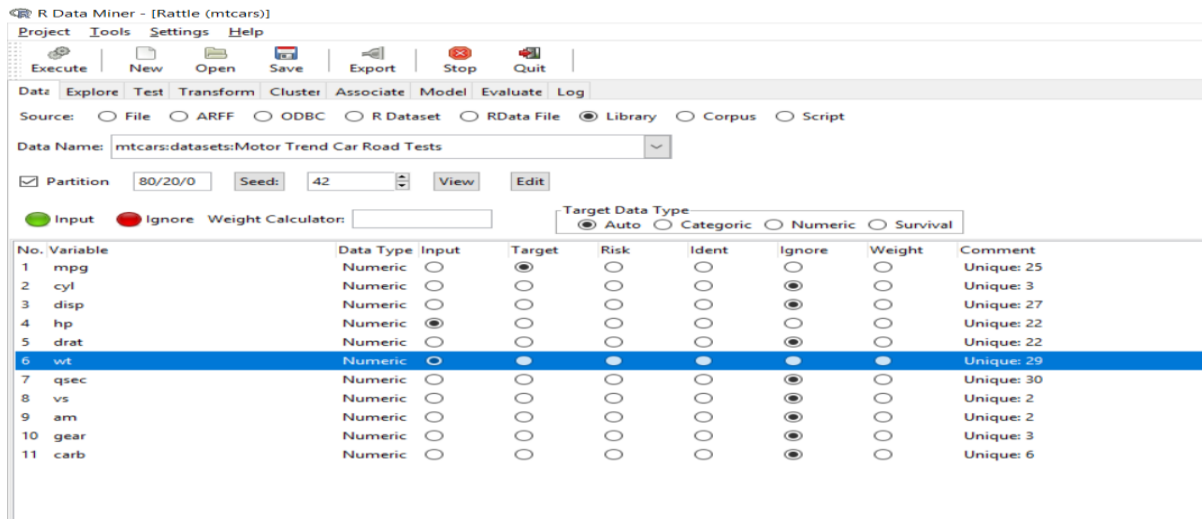
Output : Available plot in graph section of RStudio.

Interpretation: Pseudo R squared is 0.6616. The “predicted vs observed” line is the line when our predictions are perfect. As we see the “linear fit to points” has slope somewhat similar, we would conclude we have a good fit.

Action12: Make mpg as target and both wt & hp should be taken as input variables.

This document was prepared using R version 3.6.3

Depending on R version, the output may be slightly different, and that's acceptable.



Action13: Go to the “Model” tab.

Action14: Select type as Linear and Numeric and then “Execute”.

Interpretation: About 82% of the variation in percentage of Miles_per_Gallon is associated with variation in horse power and weight.

Assessing the Visual Fit

Action15: Go to Evaluate tab.

Action16: Select type Pr v Ob and then Execute.

Output : On the right side of RStudio.

Interpretation: Pseudo R squared has increased to 0.82. The “predicted vs observed” line is the line when our predictions are perfect. As we see the “linear fit to points” has a slope very close to the slope of “predicted vs observed” line , we conclude we have excellent fit. This means that adding two predictors improves fit a lot.

This document was prepared using R version 3.6.3

Depending on R version, the output may be slightly different, and that's acceptable.