Fuel Economy Project

Kedar

July 12, 2018

library(readr)  
Fuel2010<- read.csv("E:\\Kedar Business\\Acadgild\\KYD\\Project\\FE2010.csv")  
Fuel2011<- read.csv("E:\\Kedar Business\\Acadgild\\KYD\\Project\\FE2011.csv")

dim(Fuel2010); dim(Fuel2011)

## [1] 1107 10

## [1] 245 10

View(Fuel2010)   
str(Fuel2010)

## 'data.frame': 1107 obs. of 10 variables:  
## $ EngDispl : num 4.7 4.7 4.2 4.2 5.2 5.2 2 6 3 3 ...  
## $ NumCyl : int 8 8 8 8 10 10 4 12 6 6 ...  
## $ FE : num 28 25.6 26.8 25 24.8 ...  
## $ NumGears : int 6 6 6 6 6 6 6 6 6 6 ...  
## $ TransLockup : int 1 1 1 1 0 0 0 0 1 0 ...  
## $ TransCreeperGear : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ IntakeValvePerCyl : int 2 2 2 2 2 2 2 2 2 2 ...  
## $ ExhaustValvesPerCyl: int 2 2 2 2 2 2 2 2 2 2 ...  
## $ VarValveTiming : int 1 1 1 1 1 1 1 1 1 1 ...  
## $ VarValveLift : int 0 0 0 0 0 0 0 0 1 1 ...

str(Fuel2011)

## 'data.frame': 245 obs. of 10 variables:  
## $ EngDispl : num 5.9 4.2 4.2 5.2 5.2 3 1.5 1.5 6.3 6 ...  
## $ NumCyl : int 12 8 8 10 10 6 4 4 8 12 ...  
## $ FE : num 22.9 26.8 24.3 24.3 23.1 ...  
## $ NumGears : int 6 6 6 6 6 7 6 1 7 5 ...  
## $ TransLockup : int 0 0 0 0 0 1 0 1 1 1 ...  
## $ TransCreeperGear : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ IntakeValvePerCyl : int 2 2 2 2 2 2 2 2 2 2 ...  
## $ ExhaustValvesPerCyl: int 2 2 2 2 2 2 2 2 2 1 ...  
## $ VarValveTiming : int 0 1 1 1 1 1 1 1 1 1 ...  
## $ VarValveLift : int 0 0 0 0 0 0 1 1 0 0 ...

names(Fuel2010)

## [1] "EngDispl" "NumCyl" "FE"   
## [4] "NumGears" "TransLockup" "TransCreeperGear"   
## [7] "IntakeValvePerCyl" "ExhaustValvesPerCyl" "VarValveTiming"   
## [10] "VarValveLift"

names(Fuel2011)

## [1] "EngDispl" "NumCyl" "FE"   
## [4] "NumGears" "TransLockup" "TransCreeperGear"   
## [7] "IntakeValvePerCyl" "ExhaustValvesPerCyl" "VarValveTiming"   
## [10] "VarValveLift"

class(Fuel2010)

## [1] "data.frame"

summary(Fuel2010)

## EngDispl NumCyl FE NumGears   
## Min. :1.000 Min. : 2.000 Min. :17.50 Min. :1.000   
## 1st Qu.:2.400 1st Qu.: 4.000 1st Qu.:29.09 1st Qu.:5.000   
## Median :3.500 Median : 6.000 Median :34.51 Median :6.000   
## Mean :3.507 Mean : 5.971 Mean :34.71 Mean :5.268   
## 3rd Qu.:4.300 3rd Qu.: 8.000 3rd Qu.:39.20 3rd Qu.:6.000   
## Max. :8.400 Max. :16.000 Max. :69.64 Max. :8.000   
## TransLockup TransCreeperGear IntakeValvePerCyl ExhaustValvesPerCyl  
## Min. :0.0000 Min. :0.00000 Min. :0.000 Min. :0.000   
## 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:2.000 1st Qu.:2.000   
## Median :1.0000 Median :0.00000 Median :2.000 Median :2.000   
## Mean :0.6802 Mean :0.04878 Mean :1.862 Mean :1.837   
## 3rd Qu.:1.0000 3rd Qu.:0.00000 3rd Qu.:2.000 3rd Qu.:2.000   
## Max. :1.0000 Max. :1.00000 Max. :3.000 Max. :2.000   
## VarValveTiming VarValveLift   
## Min. :0.0000 Min. :0.0000   
## 1st Qu.:1.0000 1st Qu.:0.0000   
## Median :1.0000 Median :0.0000   
## Mean :0.8229 Mean :0.1671   
## 3rd Qu.:1.0000 3rd Qu.:0.0000   
## Max. :1.0000 Max. :1.0000

cor(Fuel2010$FE, Fuel2010$EngDispl)

## [1] -0.7873938

cor(Fuel2010$FE, Fuel2010$NumCyl)

## [1] -0.740218

cor(Fuel2010$FE, Fuel2010$NumGears)

## [1] -0.2112849

cor(Fuel2010$FE, Fuel2010$TransLockup)

## [1] -0.2719389

cor(Fuel2010$FE, Fuel2010$TransCreeperGear)

## [1] -0.06962168

cor(Fuel2010$FE, Fuel2010$IntakeValvePerCyl)

## [1] 0.280344

cor(Fuel2010$FE, Fuel2010$ExhaustValvesPerCyl)

## [1] 0.3356529

cor(Fuel2010$FE, Fuel2010$VarValveTiming)

## [1] 0.1249528

cor(Fuel2010$FE, Fuel2010$VarValveLift)

## [1] 0.09621127

#################### linear Regression #################  
  
model <- lm(FE ~., data = Fuel2010)  
summary(model)

##   
## Call:  
## lm(formula = FE ~ ., data = Fuel2010)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -17.1153 -2.7142 -0.3535 2.4191 25.6521   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 54.3472 1.0973 49.530 < 2e-16 \*\*\*  
## EngDispl -3.8610 0.2805 -13.765 < 2e-16 \*\*\*  
## NumCyl -0.4888 0.1845 -2.649 0.00819 \*\*   
## NumGears -0.1725 0.1065 -1.620 0.10555   
## TransLockup -1.4450 0.3000 -4.817 1.66e-06 \*\*\*  
## TransCreeperGear -0.9138 0.6681 -1.368 0.17167   
## IntakeValvePerCyl -0.3737 0.9892 -0.378 0.70566   
## ExhaustValvesPerCyl -1.1105 0.9598 -1.157 0.24752   
## VarValveTiming 1.6870 0.3796 4.444 9.71e-06 \*\*\*  
## VarValveLift 0.6235 0.3719 1.676 0.09393 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.489 on 1097 degrees of freedom  
## Multiple R-squared: 0.6445, Adjusted R-squared: 0.6415   
## F-statistic: 220.9 on 9 and 1097 DF, p-value: < 2.2e-16

LmModel<- lm(FE~ EngDispl+NumCyl+TransLockup+VarValveTiming, data = Fuel2010 )  
summary(LmModel)

##   
## Call:  
## lm(formula = FE ~ EngDispl + NumCyl + TransLockup + VarValveTiming,   
## data = Fuel2010)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -14.6536 -2.8052 -0.5011 2.3228 26.4104   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 50.7368 0.5520 91.915 < 2e-16 \*\*\*  
## EngDispl -3.4521 0.2506 -13.777 < 2e-16 \*\*\*  
## NumCyl -0.6978 0.1713 -4.074 4.95e-05 \*\*\*  
## TransLockup -1.4537 0.3000 -4.846 1.44e-06 \*\*\*  
## VarValveTiming 1.4980 0.3623 4.134 3.83e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.517 on 1102 degrees of freedom  
## Multiple R-squared: 0.6385, Adjusted R-squared: 0.6372   
## F-statistic: 486.5 on 4 and 1102 DF, p-value: < 2.2e-16

lmPred <- predict(LmModel, newdata = Fuel2011)  
lmError <- Fuel2011$FE - lmPred  
  
MAPE<-sum(abs(Fuel2010$FE-lmPred)/Fuel2010$FE)\*100

## Warning in Fuel2010$FE - lmPred: longer object length is not a multiple of  
## shorter object length

MAPE

## [1] 27966.37

linear\_RSME <- sqrt(mean((lmError^2))) # RSME  
linear\_RSME

## [1] 5.001716

## R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

summary(cars)

## speed dist   
## Min. : 4.0 Min. : 2.00   
## 1st Qu.:12.0 1st Qu.: 26.00   
## Median :15.0 Median : 36.00   
## Mean :15.4 Mean : 42.98   
## 3rd Qu.:19.0 3rd Qu.: 56.00   
## Max. :25.0 Max. :120.00

## Including Plots

You can also embed plots, for example:



Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.