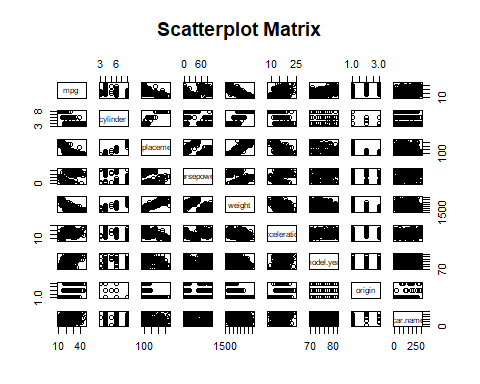
Final\_project.R

2019-05-12

## 1) Load data #######  
  
##open file   
  
autompg<-read.csv(file.choose())  
  
#view data  
  
#View(autompg)  
  
#structure of data  
  
#str(autompg)  
  
#five point summary of data  
  
summary(autompg)

## mpg cylinder displacement horsepower   
## Min. : 9.00 Min. :3.000 Min. : 68.0 150 : 22   
## 1st Qu.:17.50 1st Qu.:4.000 1st Qu.:104.2 90 : 20   
## Median :23.00 Median :4.000 Median :148.5 88 : 19   
## Mean :23.51 Mean :5.455 Mean :193.4 110 : 18   
## 3rd Qu.:29.00 3rd Qu.:8.000 3rd Qu.:262.0 100 : 17   
## Max. :46.60 Max. :8.000 Max. :455.0 75 : 14   
## (Other):288   
## weight acceleration model.year origin   
## Min. :1613 Min. : 8.00 Min. :70.00 Min. :1.000   
## 1st Qu.:2224 1st Qu.:13.82 1st Qu.:73.00 1st Qu.:1.000   
## Median :2804 Median :15.50 Median :76.00 Median :1.000   
## Mean :2970 Mean :15.57 Mean :76.01 Mean :1.573   
## 3rd Qu.:3608 3rd Qu.:17.18 3rd Qu.:79.00 3rd Qu.:2.000   
## Max. :5140 Max. :24.80 Max. :82.00 Max. :3.000   
##   
## car.name   
## ford pinto : 6   
## amc matador : 5   
## ford maverick : 5   
## toyota corolla: 5   
## amc gremlin : 4   
## amc hornet : 4   
## (Other) :369

###### 2 ) Explore data #######  
  
#When we have more than two variables and we want to find the correlation between one variable versus the remaining ones we use scatterplot matrix  
  
pairs(~mpg + cylinder + displacement + horsepower + weight + acceleration + model.year+origin + car.name,data=autompg,main="Scatterplot Matrix")



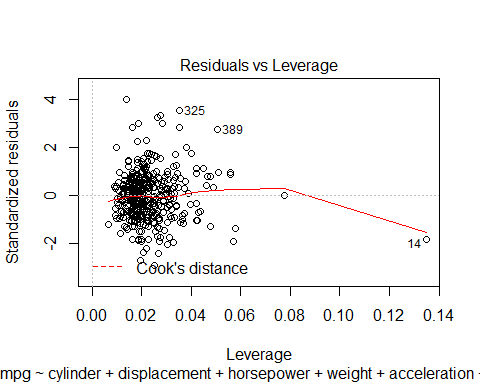
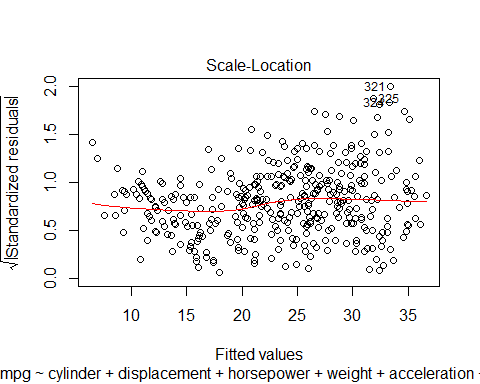
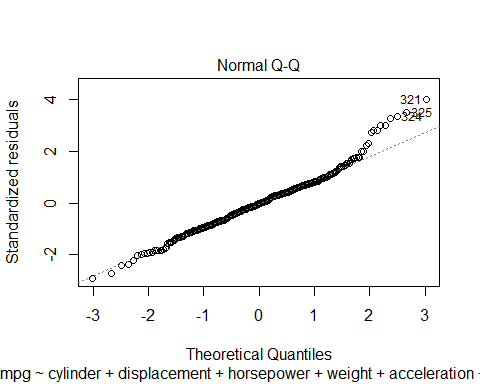
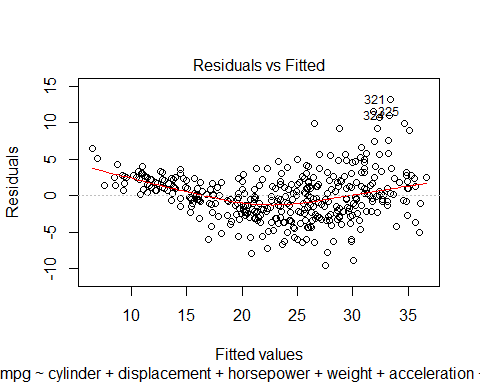
#correlation Matrix of data  
  
  
#cor(correlations)  
  
  
##### 3) Cleaning the data ############  
  
  
  
#Removing Null Values from horsepower Column   
  
autompg<-autompg[-c(33,127,331,337,355,375),]  
  
#view of dataset after removing column  
View(autompg)  
  
#Resquence Horsepower column   
row.names(autompg)<-seq(length=nrow(autompg))  
  
#Convert horsepower into numeric  
  
autompg$horsepower<- as.numeric(autompg$horsepower)  
  
#Convert carname into numeric  
  
autompg$car.name<- as.numeric(autompg$car.name)  
  
#structure of data  
  
str(autompg)

## 'data.frame': 392 obs. of 9 variables:  
## $ mpg : num 18 15 18 16 17 15 14 14 14 15 ...  
## $ cylinder : int 8 8 8 8 8 8 8 8 8 8 ...  
## $ displacement: num 307 350 318 304 302 429 454 440 455 390 ...  
## $ horsepower : num 17 35 29 29 24 42 47 46 48 40 ...  
## $ weight : int 3504 3693 3436 3433 3449 4341 4354 4312 4425 3850 ...  
## $ acceleration: num 12 11.5 11 12 10.5 10 9 8.5 10 8.5 ...  
## $ model.year : int 70 70 70 70 70 70 70 70 70 70 ...  
## $ origin : int 1 1 1 1 1 1 1 1 1 1 ...  
## $ car.name : num 50 37 232 15 162 142 55 224 242 2 ...

###### 4) splitting the dataset into training and test ###########  
  
training<-autompg[1:318,]  
  
test<-autompg[319:392,] ##### 6 was missing values in horsepower after removing that total 392 rows  
  
  
  
######### 4) Multiple Regression #########  
  
#i have used all variables in first model   
  
autoModel1 <- lm(mpg ~ cylinder + displacement + horsepower + weight + acceleration + model.year + origin + car.name, data = autompg)  
summary(autoModel1)

##   
## Call:  
## lm(formula = mpg ~ cylinder + displacement + horsepower + weight +   
## acceleration + model.year + origin + car.name, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -9.5122 -2.2025 -0.0471 1.8939 13.1844   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -2.180e+01 4.289e+00 -5.082 5.86e-07 \*\*\*  
## cylinder -2.685e-01 3.389e-01 -0.792 0.4287   
## displacement 1.619e-02 7.271e-03 2.227 0.0265 \*   
## horsepower 1.023e-02 7.063e-03 1.448 0.1484   
## weight -6.832e-03 5.829e-04 -11.720 < 2e-16 \*\*\*  
## acceleration 1.496e-01 7.757e-02 1.928 0.0546 .   
## model.year 7.711e-01 4.961e-02 15.544 < 2e-16 \*\*\*  
## origin 1.231e+00 2.772e-01 4.440 1.18e-05 \*\*\*  
## car.name 3.498e-03 2.023e-03 1.729 0.0846 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.316 on 383 degrees of freedom  
## Multiple R-squared: 0.8232, Adjusted R-squared: 0.8195   
## F-statistic: 222.9 on 8 and 383 DF, p-value: < 2.2e-16

plot(autoModel1)



coefficients(autoModel1)

## (Intercept) cylinder displacement horsepower weight   
## -21.796585594 -0.268534045 0.016192212 0.010228569 -0.006832202   
## acceleration model.year origin car.name   
## 0.149554354 0.771085956 1.230845717 0.003497711

#### Anova is a technique for analysis of variance #####  
  
anova(autoModel1)

## Analysis of Variance Table  
##   
## Response: mpg  
## Df Sum Sq Mean Sq F value Pr(>F)   
## cylinder 1 14403.1 14403.1 1310.0826 < 2.2e-16 \*\*\*  
## displacement 1 1073.3 1073.3 97.6297 < 2.2e-16 \*\*\*  
## horsepower 1 33.0 33.0 3.0037 0.08388 .   
## weight 1 1154.8 1154.8 105.0390 < 2.2e-16 \*\*\*  
## acceleration 1 64.8 64.8 5.8932 0.01566 \*   
## model.year 1 2572.9 2572.9 234.0262 < 2.2e-16 \*\*\*  
## origin 1 273.5 273.5 24.8743 9.296e-07 \*\*\*  
## car.name 1 32.9 32.9 2.9907 0.08455 .   
## Residuals 383 4210.7 11.0   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

######## Regression model ######  
  
autoModel1\_1 <- lm(mpg ~ cylinder + displacement + horsepower + weight + acceleration + model.year + origin + car.name, data = training)  
summary(autoModel1\_1)

##   
## Call:  
## lm(formula = mpg ~ cylinder + displacement + horsepower + weight +   
## acceleration + model.year + origin + car.name, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -9.6984 -1.8522 0.0777 1.6758 12.9324   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -7.0012229 4.6448796 -1.507 0.1328   
## cylinder -0.3250702 0.3155775 -1.030 0.3038   
## displacement 0.0062862 0.0067712 0.928 0.3539   
## horsepower 0.0108699 0.0066227 1.641 0.1018   
## weight -0.0059156 0.0005462 -10.831 <2e-16 \*\*\*  
## acceleration -0.0436253 0.0772963 -0.564 0.5729   
## model.year 0.6169766 0.0576356 10.705 <2e-16 \*\*\*  
## origin 0.7052943 0.2948634 2.392 0.0174 \*   
## car.name 0.0017061 0.0019288 0.885 0.3771   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.859 on 309 degrees of freedom  
## Multiple R-squared: 0.8241, Adjusted R-squared: 0.8195   
## F-statistic: 180.9 on 8 and 309 DF, p-value: < 2.2e-16

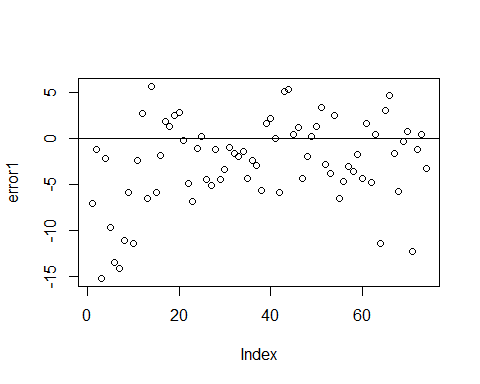
### use model on test dataset ###  
  
B0<-autoModel1\_1$coefficients[1]  
  
B1<-autoModel1\_1$coefficients[2]  
  
B2<-autoModel1\_1$coefficients[3]  
  
B3<-autoModel1\_1$coefficients[4]  
  
B4<-autoModel1\_1$coefficients[5]  
  
B5<-autoModel1\_1$coefficients[6]  
  
B6<-autoModel1\_1$coefficients[7]  
  
B7<-autoModel1\_1$coefficients[8]  
  
B8<-autoModel1\_1$coefficients[9]  
  
  
  
##### calculate predicted values #######  
  
y\_predicted1<- B0+ B1 \*test$cylinder + B2 \* test$displacement + B3 \* test$horsepower + B4 \* test$weight +  
 B5 \* test$acceleration + B6 \* test$ model.year + B7 \* test$origin + B8 \* test$car.name  
  
  
### shows all predicted values ####  
fitted(autoModel1\_1)

## 1 2 3 4 5 6 7   
## 15.239887 14.587427 16.195865 15.711766 15.866423 11.571358 11.601156   
## 8 9 10 11 12 13 14   
## 12.060874 11.473709 14.035630 15.761975 15.468682 14.557329 19.034781   
## 15 16 17 18 19 20 21   
## 24.485111 20.117653 20.133018 21.335539 25.456458 26.183984 21.682760   
## 22 23 24 25 26 27 28   
## 22.926292 23.569134 23.445142 20.822406 9.372055 10.265599 10.461498   
## 29 30 31 32 33 34 35   
## 8.128418 26.073434 24.067387 25.992448 20.903826 16.405137 16.868194   
## 36 37 38 39 40 41 42   
## 18.112031 16.931049 12.160826 11.341806 12.531982 12.807616 8.126265   
## 43 44 45 46 47 48 49   
## 9.448126 7.334665 19.173753 22.876351 17.499018 19.101140 24.444622   
## 50 51 52 53 54 55 56   
## 25.765324 25.379548 25.613413 27.879871 28.593908 26.955228 25.423672   
## 57 58 59 60 61 62 63   
## 26.249910 25.127666 24.881522 23.654597 24.867566 12.393286 12.402600   
## 64 65 66 67 68 69 70   
## 13.172070 13.318663 15.533220 11.143066 10.907231 11.494533 11.900442   
## 71 72 73 74 75 76 77   
## 25.954213 14.203507 12.870990 11.831006 13.510425 20.911692 23.919735   
## 78 79 80 81 82 83 84   
## 21.077927 25.551018 23.887928 25.673095 25.022453 24.991560 27.075778   
## 85 86 87 88 89 90 91   
## 13.979093 16.163845 14.562389 13.969848 15.755983 9.831105 12.118318   
## 92 93 94 95 96 97 98   
## 12.585060 13.140319 11.046219 9.807699 15.737019 19.481759 18.316961   
## 99 100 101 102 103 104 105   
## 20.168844 20.944183 21.526759 27.349863 8.895088 9.701187 11.084792   
## 106 107 108 109 110 111 112   
## 11.961936 21.133604 26.516260 24.128195 25.867860 27.698369 24.886095   
## 113 114 115 116 117 118 119   
## 22.920731 25.832239 14.013142 13.886300 27.459123 26.565449 23.956436   
## 120 121 122 123 124 125 126   
## 21.867854 18.064359 23.094870 22.614769 16.961820 20.950625 21.046109   
## 127 128 129 130 131 132 133   
## 18.632749 28.457200 24.701867 29.270864 24.041869 15.981520 16.866592   
## 134 135 136 137 138 139 140   
## 17.186532 14.055598 10.892743 12.308401 11.029977 13.147373 26.431015   
## 141 142 143 144 145 146 147   
## 28.155254 26.299202 30.513830 28.096772 26.371691 27.169753 26.578860   
## 148 149 150 151 152 153 154   
## 26.144081 26.743458 27.646712 20.828103 18.598336 19.462510 21.073769   
## 155 156 157 158 159 160 161   
## 12.590344 13.056666 12.833701 11.992782 15.580049 15.864099 16.860210   
## 162 163 164 165 166 167 168   
## 17.589585 20.993642 19.571521 20.443347 28.383789 24.330650 21.584476   
## 169 170 171 172 173 174 175   
## 24.632207 25.821030 27.303165 26.201466 22.095724 29.098351 20.705504   
## 176 177 178 179 180 181 182   
## 24.533977 23.177228 23.617279 24.318174 30.158054 26.525876 28.052047   
## 183 184 185 186 187 188 189   
## 25.335514 26.790778 28.234255 14.757320 15.140232 16.020618 15.277385   
## 190 191 192 193 194 195 196   
## 20.685515 19.907809 22.409969 22.034852 27.486334 26.878368 29.732194   
## 197 198 199 200 201 202 203   
## 30.772569 17.882118 19.210217 18.464745 21.613557 30.511279 29.610399   
## 204 205 206 207 208 209 210   
## 29.077966 25.425012 22.038321 16.833912 21.728844 23.574734 17.537566   
## 211 212 213 214 215 216 217   
## 14.247660 16.081069 16.755916 17.675031 30.022805 28.097292 30.552576   
## 218 219 220 221 222 223 224   
## 27.765967 30.507423 17.389354 15.831995 16.005198 14.960351 19.472946   
## 225 226 227 228 229 230 231   
## 19.830364 18.858760 20.420003 16.520521 16.177713 15.563700 15.119290   
## 232 233 234 235 236 237 238   
## 30.414543 25.296321 28.970814 24.994252 28.393321 28.601400 30.617689   
## 239 240 241 242 243 244 245   
## 28.946258 25.468551 25.565730 25.867504 30.167634 30.791293 30.724887   
## 246 247 248 249 250 251 252   
## 30.304712 32.060003 20.711329 19.012966 20.033240 20.061378 22.761912   
## 253 254 255 256 257 258 259   
## 24.018372 25.792441 20.678972 22.538270 20.749590 23.462035 19.317639   
## 260 261 262 263 264 265 266   
## 20.818399 20.691297 20.804401 22.168418 16.951962 28.471232 28.470574   
## 267 268 269 270 271 272 273   
## 29.595244 28.338338 28.700366 24.990065 25.053816 28.953433 24.374357   
## 274 275 276 277 278 279 280   
## 23.112491 25.337882 21.319461 30.607897 30.191968 22.621475 24.488532   
## 281 282 283 284 285 286 287   
## 25.322576 22.791378 21.566002 18.635591 19.552549 18.643865 18.947462   
## 288 289 290 291 292 293 294   
## 16.003308 17.938768 19.803077 18.674380 31.670832 31.720422 30.812979   
## 295 296 297 298 299 300 301   
## 26.309821 22.028969 18.444302 23.884603 21.558506 29.328566 29.555303   
## 302 303 304 305 306 307 308   
## 31.113061 30.114741 26.601013 25.948231 25.518314 27.757526 31.058387   
## 309 310 311 312 313 314 315   
## 32.351754 29.977598 31.869509 27.174443 26.017847 25.011210 22.813282   
## 316 317 318   
## 30.281044 28.703614 29.197860

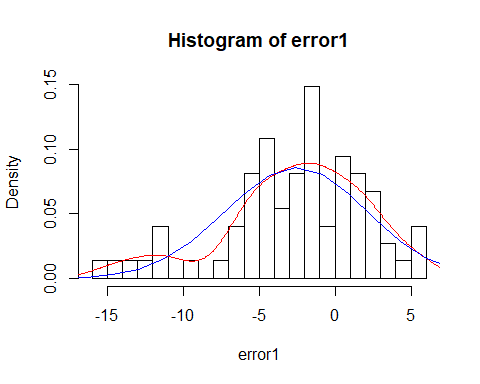
# #calculating error ##  
  
error1<- y\_predicted1 - test$mpg  
  
summary(error1)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -15.1739 -4.8103 -1.9053 -2.5683 0.4937 5.7104

plot(error1)  
  
abline(0,0)



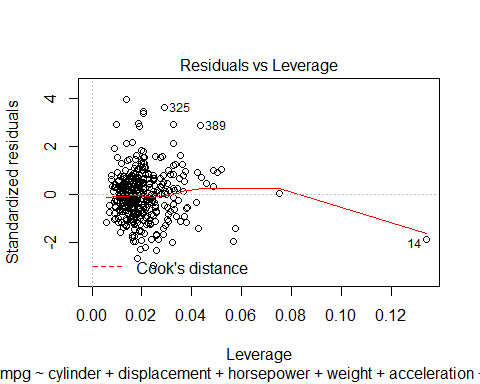
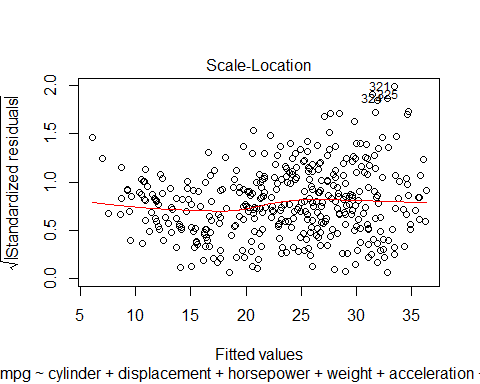
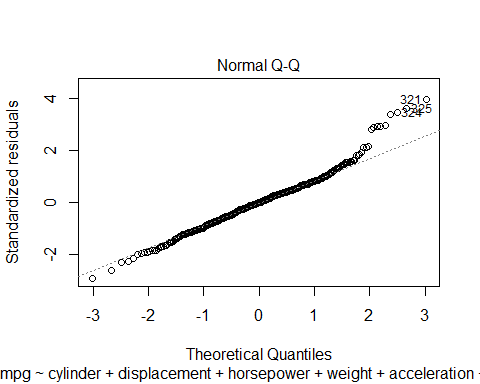
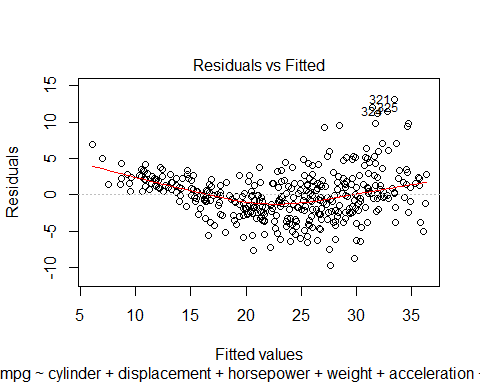
hist(error1,prob=T,breaks = 20)  
  
lines(density(error1),col='red')  
  
  
  
#Normal Curve for automodel1  
mean\_error1<-mean(error1)  
variance\_error1<-var(error1)  
sd\_error1<-sqrt(variance\_error1)  
x\_error1<-seq(-25,25,length=30)  
y\_error1<-dnorm(x\_error1,mean\_error1,sd\_error1)  
lines(x\_error1,y\_error1,col='blue')



############ Model 2#############  
  
autoModel2 = lm(mpg ~ cylinder + displacement + horsepower + weight + acceleration + model.year + origin , data = autompg)  
summary(autoModel2)

##   
## Call:  
## lm(formula = mpg ~ cylinder + displacement + horsepower + weight +   
## acceleration + model.year + origin, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -9.656 -2.069 -0.043 1.775 13.098   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -2.134e+01 4.292e+00 -4.971 1.00e-06 \*\*\*  
## cylinder -2.739e-01 3.398e-01 -0.806 0.4208   
## displacement 1.539e-02 7.275e-03 2.116 0.0350 \*   
## horsepower 1.072e-02 7.076e-03 1.515 0.1306   
## weight -6.756e-03 5.828e-04 -11.592 < 2e-16 \*\*\*  
## acceleration 1.489e-01 7.777e-02 1.914 0.0563 .   
## model.year 7.688e-01 4.972e-02 15.463 < 2e-16 \*\*\*  
## origin 1.344e+00 2.701e-01 4.975 9.89e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.324 on 384 degrees of freedom  
## Multiple R-squared: 0.8218, Adjusted R-squared: 0.8186   
## F-statistic: 253.1 on 7 and 384 DF, p-value: < 2.2e-16

plot(autoModel2)



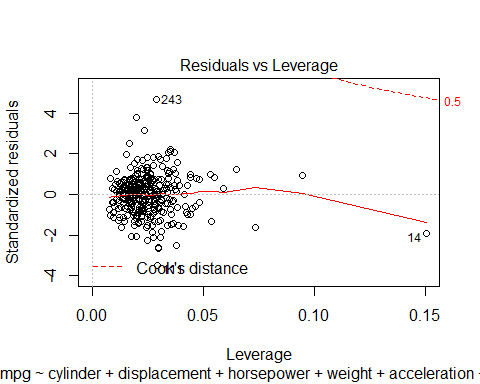
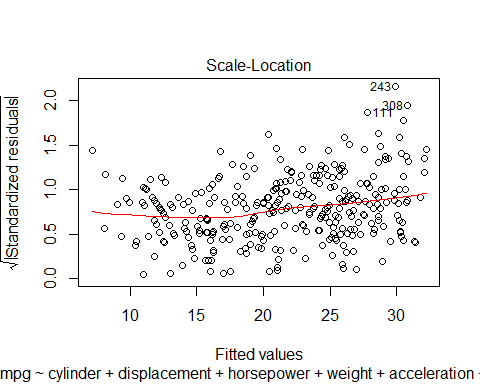
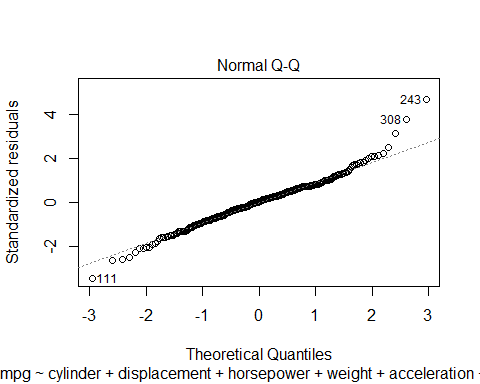
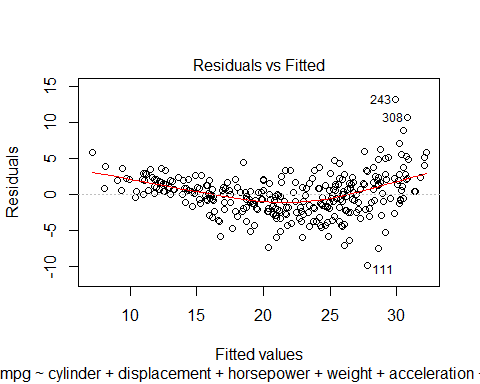
coefficients(autoModel2)

## (Intercept) cylinder displacement horsepower weight   
## -21.336705997 -0.273860004 0.015389602 0.010718491 -0.006755643   
## acceleration model.year origin   
## 0.148888529 0.768801925 1.343758663

######## Regression model ######  
  
autoModel2\_2 <- lm(mpg ~ cylinder + displacement + horsepower + weight + acceleration + model.year + origin , data = training)  
summary(autoModel2\_2)

##   
## Call:  
## lm(formula = mpg ~ cylinder + displacement + horsepower + weight +   
## acceleration + model.year + origin, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -9.7838 -1.7970 0.0621 1.7175 13.1697   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -6.6829720 4.6292978 -1.444 0.14985   
## cylinder -0.3270269 0.3154590 -1.037 0.30070   
## displacement 0.0056390 0.0067292 0.838 0.40268   
## horsepower 0.0110226 0.0066181 1.666 0.09682 .   
## weight -0.0058500 0.0005409 -10.815 < 2e-16 \*\*\*  
## acceleration -0.0469405 0.0771783 -0.608 0.54349   
## model.year 0.6146689 0.0575564 10.679 < 2e-16 \*\*\*  
## origin 0.7644817 0.2870698 2.663 0.00815 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.858 on 310 degrees of freedom  
## Multiple R-squared: 0.8236, Adjusted R-squared: 0.8196   
## F-statistic: 206.8 on 7 and 310 DF, p-value: < 2.2e-16

plot(autoModel2\_2)



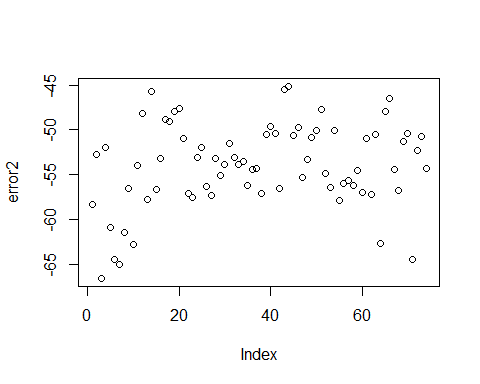
### use model on test dataset ###  
  
B0<-autoModel2\_2$coefficients[1]  
  
B1<-autoModel2\_2$coefficients[2]  
  
B2<-autoModel2\_2$coefficients[3]  
  
B3<-autoModel2\_2$coefficients[4]  
  
B4<-autoModel2\_2$coefficients[5]  
  
B5<-autoModel2\_2$coefficients[6]  
  
B6<-autoModel2\_2$coefficients[7]  
  
B7<-autoModel2\_2$coefficients[8]  
  
  
  
##### calculate predicted values #######  
  
y\_predicted2<- B0+ B1 \*test$cylinder + B2 \* test$displacement + B3 \* test$horsepower + B4 \* test$weight +B5 \* test$acceleration   
 + B6 \* test$ model.year + B7 \* test$origin

## [1] 51.46696 51.46696 51.46696 49.93800 51.46696 50.70248 50.70248  
## [8] 50.70248 50.70248 51.46696 51.46696 50.70248 51.46696 51.46696  
## [15] 50.70248 51.46696 50.55266 50.55266 50.55266 50.55266 50.55266  
## [22] 52.08163 50.55266 52.08163 52.08163 52.08163 52.08163 52.08163  
## [29] 50.55266 50.55266 50.55266 51.31715 52.08163 52.08163 52.08163  
## [36] 52.08163 51.31715 51.31715 52.08163 52.08163 50.55266 50.55266  
## [43] 50.55266 50.55266 51.16733 51.16733 51.16733 51.16733 51.16733  
## [50] 51.16733 51.16733 51.93182 52.69630 52.69630 51.16733 51.16733  
## [57] 52.69630 52.69630 52.69630 52.69630 52.69630 52.69630 51.16733  
## [64] 51.16733 51.16733 51.16733 52.69630 51.16733 51.16733 51.16733  
## [71] 51.93182 51.16733 51.16733 51.16733

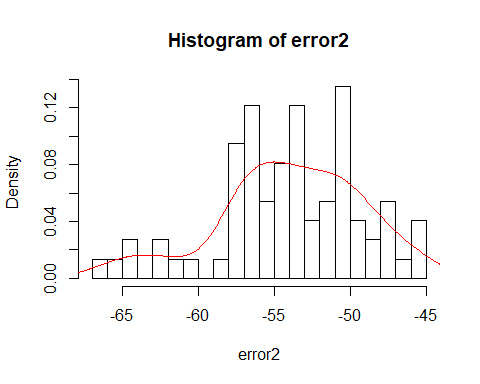
# #calculating error ##  
  
error2<- y\_predicted2 - test$mpg  
  
summary(error2)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -66.61 -56.67 -53.89 -53.99 -50.55 -45.13

plot(error2)  
  
abline(0,0)



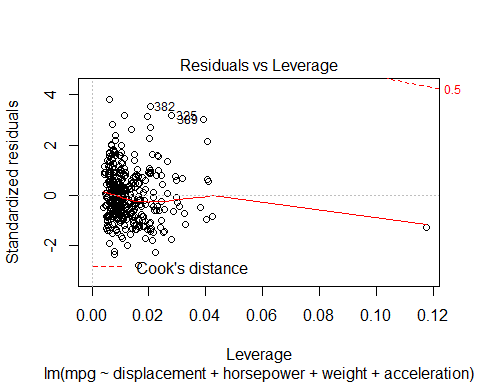
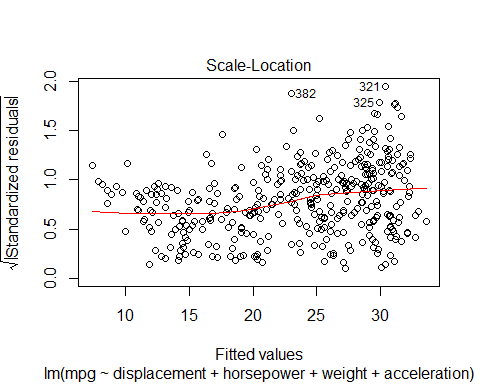
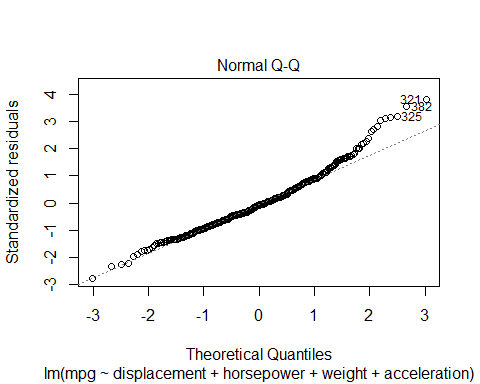
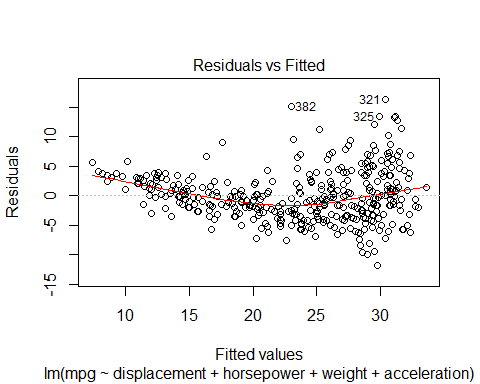
hist(error2,prob=T,breaks = 20)  
  
lines(density(error2),col='red')  
  
  
  
#Normal Curve for automodel2  
mean\_error2<-mean(error2)  
variance\_error2<-var(error2)  
sd\_error2<-sqrt(variance\_error2)  
x\_error2<-seq(-25,25,length=30)  
y\_error2<-dnorm(x\_error2,mean\_error2,sd\_error2)  
lines(x\_error2,y\_error2,col='blue')



############ Model 3 ###############  
  
autoModel3 = lm(mpg ~ displacement + horsepower + weight + acceleration , data = autompg)  
summary(autoModel3)

##   
## Call:  
## lm(formula = mpg ~ displacement + horsepower + weight + acceleration,   
## data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -11.7797 -2.7837 -0.3145 2.3950 16.2489   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 40.4449493 2.0088844 20.133 < 2e-16 \*\*\*  
## displacement -0.0102652 0.0065455 -1.568 0.1176   
## horsepower 0.0064381 0.0085818 0.750 0.4536   
## weight -0.0061061 0.0007478 -8.165 4.58e-15 \*\*\*  
## acceleration 0.1828906 0.0981158 1.864 0.0631 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.281 on 387 degrees of freedom  
## Multiple R-squared: 0.7022, Adjusted R-squared: 0.6991   
## F-statistic: 228.1 on 4 and 387 DF, p-value: < 2.2e-16

plot(autoModel3)



coefficients(autoModel3)

## (Intercept) displacement horsepower weight acceleration   
## 40.444949303 -0.010265205 0.006438145 -0.006106136 0.182890644

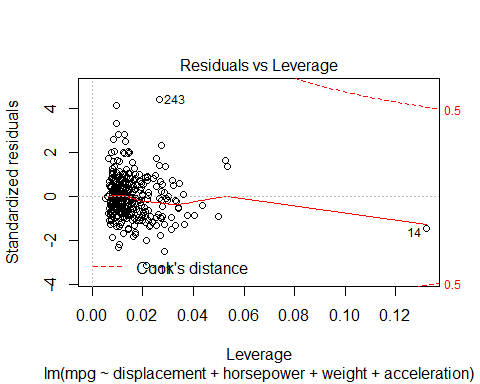
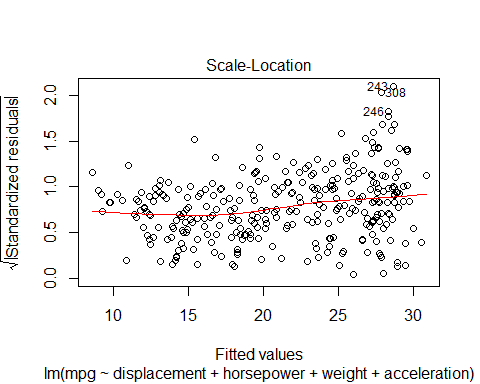
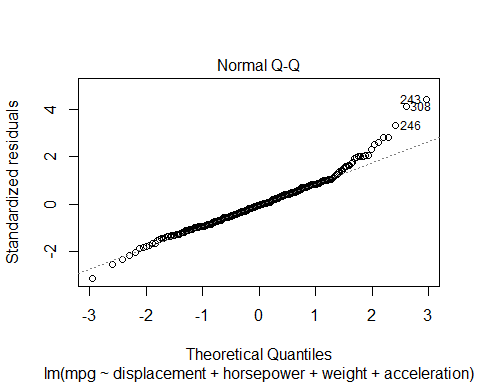
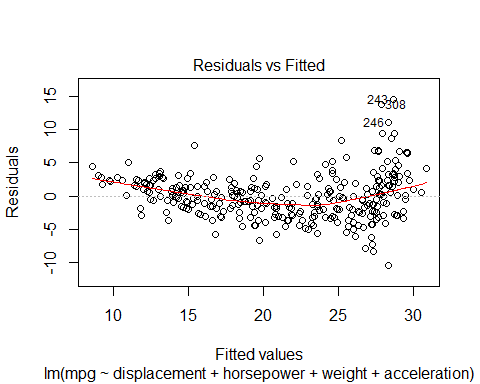
anova(autoModel3)

## Analysis of Variance Table  
##   
## Response: mpg  
## Df Sum Sq Mean Sq F value Pr(>F)   
## displacement 1 15440.2 15440.2 842.3156 < 2.2e-16 \*\*\*  
## horsepower 1 55.4 55.4 3.0227 0.08290 .   
## weight 1 1165.8 1165.8 63.5968 1.74e-14 \*\*\*  
## acceleration 1 63.7 63.7 3.4746 0.06307 .   
## Residuals 387 7094.0 18.3   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

######## Regression model ######  
  
autoModel3\_3 <- lm(mpg ~ displacement + horsepower + weight + acceleration , data = training)  
summary(autoModel3\_3)

##   
## Call:  
## lm(formula = mpg ~ displacement + horsepower + weight + acceleration,   
## data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -10.3555 -2.2411 -0.1792 1.8035 14.4747   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 39.3652498 1.7873671 22.024 < 2e-16 \*\*\*  
## displacement -0.0117150 0.0055153 -2.124 0.0344 \*   
## horsepower 0.0076990 0.0073412 1.049 0.2951   
## weight -0.0051545 0.0006255 -8.240 4.75e-15 \*\*\*  
## acceleration 0.0071379 0.0878216 0.081 0.9353   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.338 on 313 degrees of freedom  
## Multiple R-squared: 0.757, Adjusted R-squared: 0.7539   
## F-statistic: 243.7 on 4 and 313 DF, p-value: < 2.2e-16

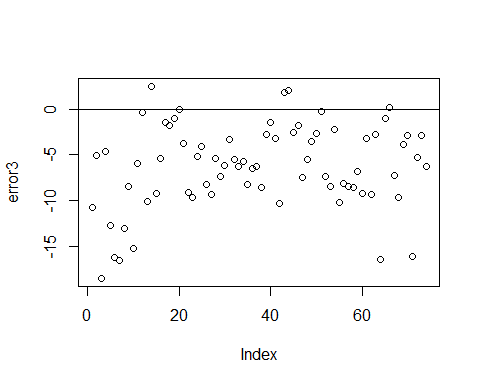
plot(autoModel3\_3)



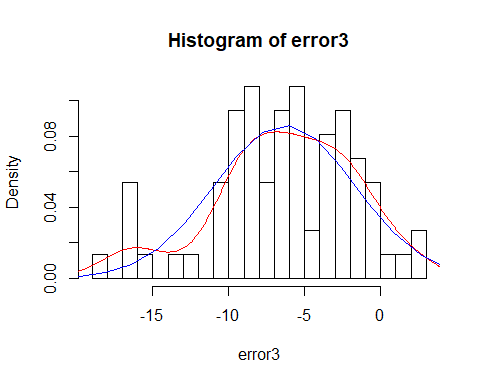
### use model on test dataset ###  
  
B0<-autoModel3\_3$coefficients[1]  
  
B1<-autoModel3\_3$coefficients[2]  
  
B2<-autoModel3\_3$coefficients[3]  
  
B3<-autoModel3\_3$coefficients[4]  
  
B4<-autoModel3\_3$coefficients[5]  
  
  
  
##### calculate predicted values #######  
  
y\_predicted3<- B0+ B1\* test$displacement + B2 \* test$horsepower + B3 \* test$weight +B4 \* test$acceleration  
  
  
  
  
  
# #calculating error ##  
  
error3<- y\_predicted3 - test$mpg  
  
summary(error3)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -18.513 -9.023 -6.069 -6.323 -2.874 2.476

plot(error3)  
  
abline(0,0)



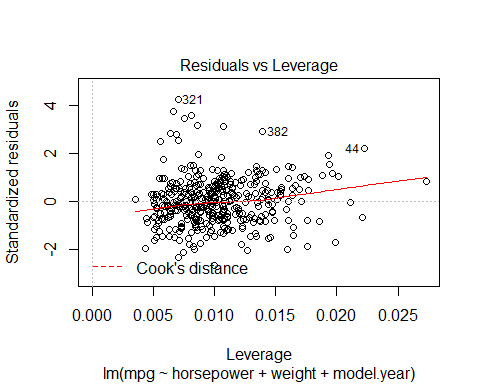
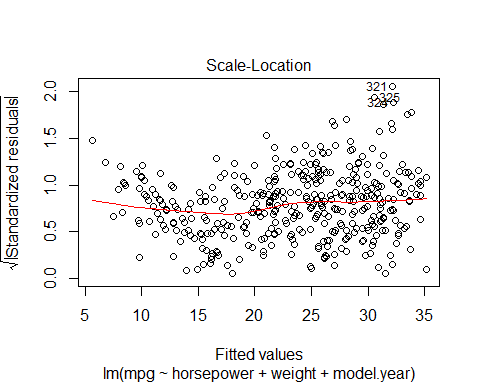
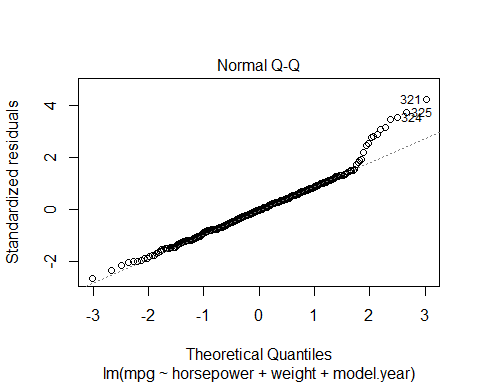
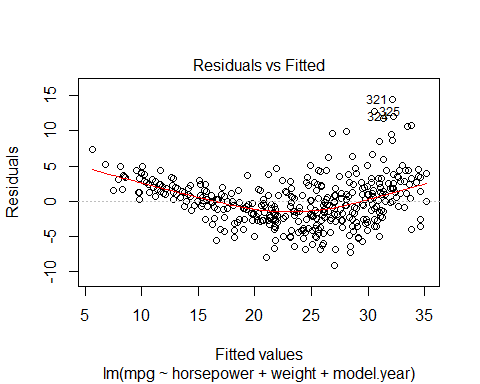
hist(error3,prob=T,breaks = 20)  
  
lines(density(error3),col='red')  
  
  
  
#Normal Curve for automodel3  
mean\_error3<-mean(error3)  
variance\_error3<-var(error3)  
sd\_error3<-sqrt(variance\_error3)  
x\_error3<-seq(-25,25,length=30)  
y\_error3<-dnorm(x\_error3,mean\_error3,sd\_error3)  
lines(x\_error3,y\_error3,col='blue')



######### Model 4 ######  
  
autoModel4 = lm(mpg ~ horsepower + weight + model.year, data = autompg)   
summary(autoModel4)

##   
## Call:  
## lm(formula = mpg ~ horsepower + weight + model.year, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -9.0692 -2.3361 -0.0855 1.9037 14.4160   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.578e+01 4.076e+00 -3.871 0.000127 \*\*\*  
## horsepower 1.211e-02 6.807e-03 1.778 0.076108 .   
## weight -6.416e-03 2.462e-04 -26.062 < 2e-16 \*\*\*  
## model.year 7.594e-01 4.935e-02 15.388 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.418 on 388 degrees of freedom  
## Multiple R-squared: 0.8097, Adjusted R-squared: 0.8083   
## F-statistic: 550.4 on 3 and 388 DF, p-value: < 2.2e-16

plot(autoModel4)



anova(autoModel4)

## Analysis of Variance Table  
##   
## Response: mpg  
## Df Sum Sq Mean Sq F value Pr(>F)   
## horsepower 1 4829.0 4829.0 413.42 < 2.2e-16 \*\*\*  
## weight 1 11692.4 11692.4 1001.02 < 2.2e-16 \*\*\*  
## model.year 1 2765.7 2765.7 236.78 < 2.2e-16 \*\*\*  
## Residuals 388 4532.0 11.7   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

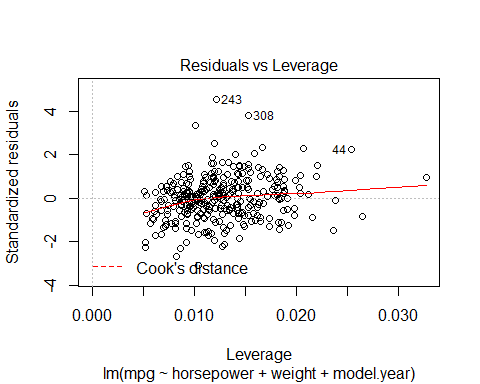
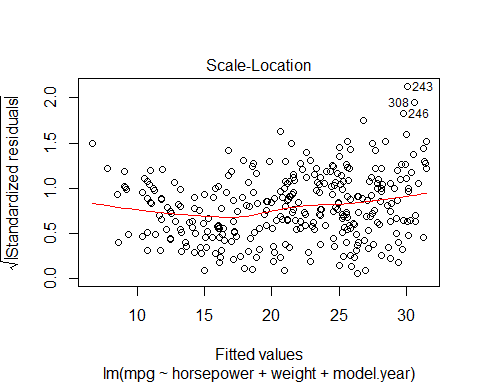
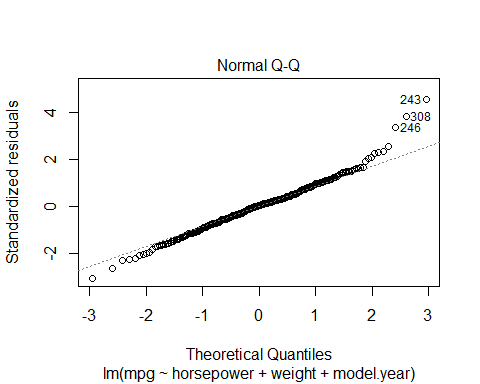
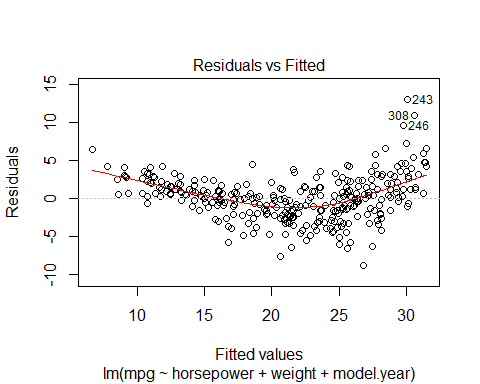
coefficients(autoModel4)

## (Intercept) horsepower weight model.year   
## -15.778499853 0.012105190 -0.006415582 0.759360980

######## Regression model ######  
  
autoModel4\_4 <- lm(mpg ~ horsepower + weight + model.year , data = training)  
summary(autoModel4\_4)

##   
## Call:  
## lm(formula = mpg ~ horsepower + weight + model.year, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -8.8046 -1.7384 0.0964 1.5661 12.9786   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -4.7449777 4.3782160 -1.084 0.2793   
## horsepower 0.0139959 0.0063277 2.212 0.0277 \*   
## weight -0.0060750 0.0002196 -27.663 <2e-16 \*\*\*  
## model.year 0.5924535 0.0554643 10.682 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.883 on 314 degrees of freedom  
## Multiple R-squared: 0.8182, Adjusted R-squared: 0.8165   
## F-statistic: 471.1 on 3 and 314 DF, p-value: < 2.2e-16

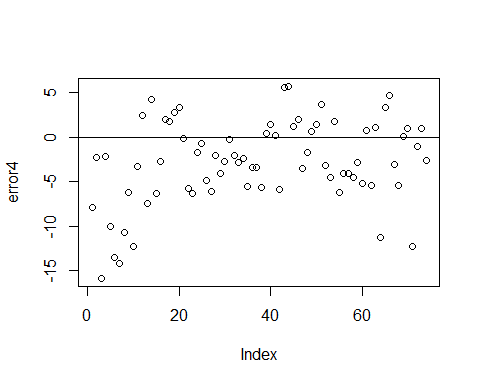
plot(autoModel4\_4)



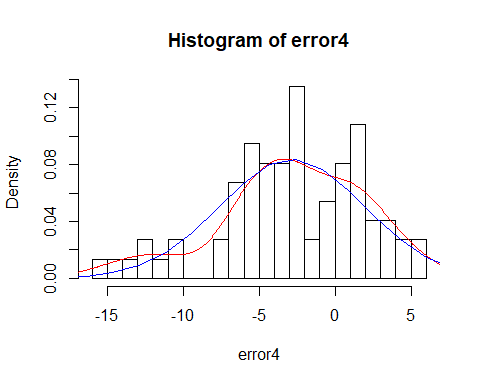
### use model on test dataset ###  
  
B0<-autoModel4\_4$coefficients[1]  
  
B1<-autoModel4\_4$coefficients[2]  
  
B2<-autoModel4\_4$coefficients[3]  
  
B3<-autoModel4\_4$coefficients[4]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted4<- B0+ B1 \* test$horsepower + B2 \* test$weight +B3 \*test$ model.year   
  
  
  
  
  
# #calculating error ##  
  
error4<- y\_predicted4- test$mpg  
  
summary(error4)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -15.899 -5.496 -2.703 -2.835 0.903 5.728

plot(error4)  
  
abline(0,0)



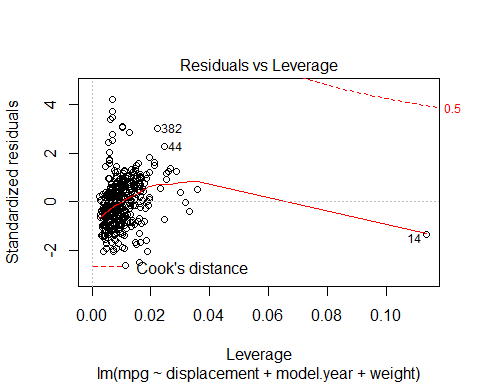
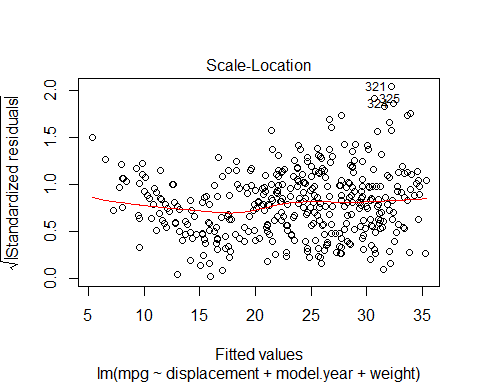
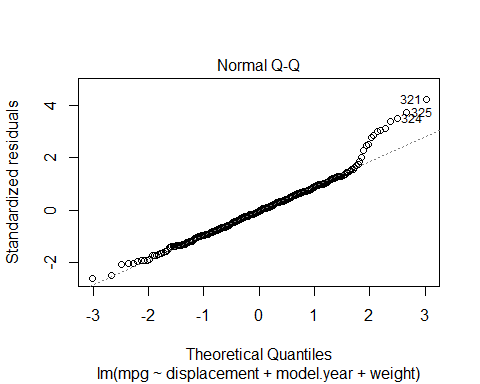
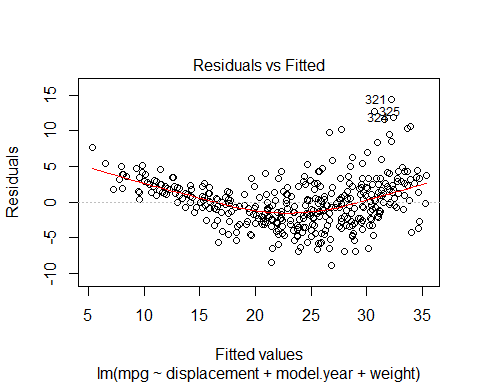
hist(error4,prob=T,breaks = 20)  
  
lines(density(error4),col='red')  
  
  
  
#Normal Curve for automodel4  
mean\_error4<-mean(error4)  
variance\_error4<-var(error4)  
sd\_error4<-sqrt(variance\_error4)  
x\_error4<-seq(-25,25,length=30)  
y\_error4<-dnorm(x\_error4,mean\_error4,sd\_error4)  
lines(x\_error4,y\_error4,col='blue')



########### Model 5 ###############  
  
  
autoModel5 = lm(mpg ~ displacement + model.year + weight , data = autompg)  
summary(autoModel5)

##   
## Call:  
## lm(formula = mpg ~ displacement + model.year + weight, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -8.8400 -2.2917 -0.1177 2.0420 14.3559   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.436e+01 4.021e+00 -3.572 0.000398 \*\*\*  
## displacement 2.835e-04 4.744e-03 0.060 0.952382   
## model.year 7.580e-01 5.100e-02 14.863 < 2e-16 \*\*\*  
## weight -6.664e-03 5.710e-04 -11.670 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.432 on 388 degrees of freedom  
## Multiple R-squared: 0.8082, Adjusted R-squared: 0.8067   
## F-statistic: 544.9 on 3 and 388 DF, p-value: < 2.2e-16

plot(autoModel5)



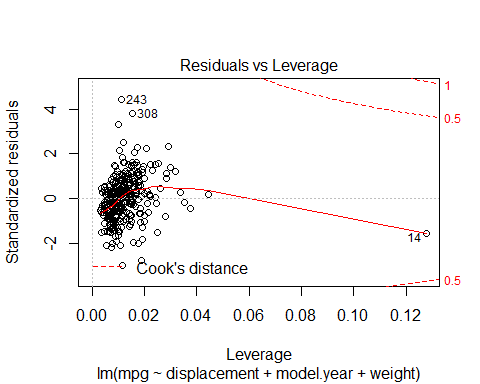
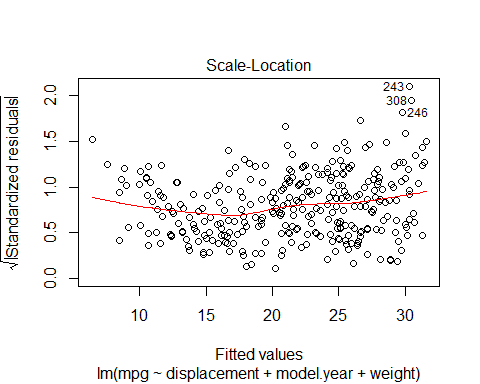
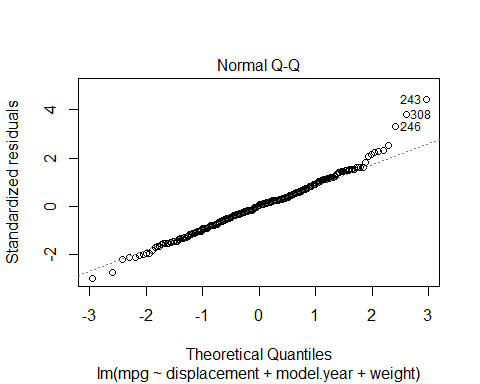
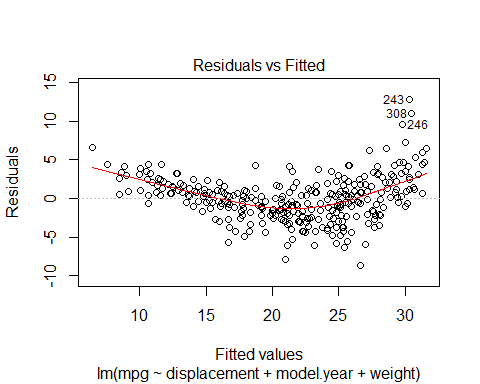
coefficients(autoModel5)

## (Intercept) displacement model.year weight   
## -14.363338390 0.000283494 0.758043499 -0.006663689

######## Regression model ######  
  
autoModel5\_5 <- lm(mpg ~ displacement + model.year + weight , data = training)  
summary(autoModel5\_5)

##   
## Call:  
## lm(formula = mpg ~ displacement + model.year + weight, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -8.6308 -1.8200 0.1165 1.6250 12.7711   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -2.4285828 4.3261272 -0.561 0.575   
## displacement -0.0019590 0.0042871 -0.457 0.648   
## model.year 0.5776598 0.0574503 10.055 <2e-16 \*\*\*  
## weight -0.0061076 0.0005206 -11.732 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.904 on 314 degrees of freedom  
## Multiple R-squared: 0.8155, Adjusted R-squared: 0.8138   
## F-statistic: 462.7 on 3 and 314 DF, p-value: < 2.2e-16

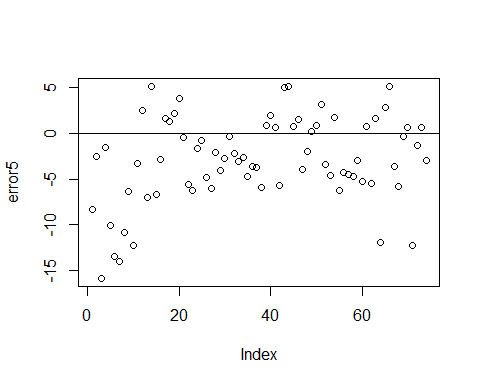
plot(autoModel5\_5)



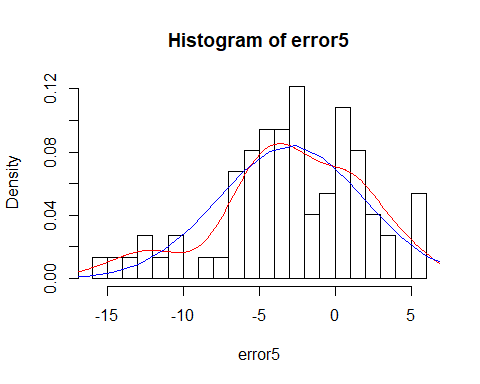
### use model on test dataset ###  
  
B0<-autoModel5\_5$coefficients[1]  
  
B1<-autoModel5\_5$coefficients[2]  
  
B2<-autoModel5\_5$coefficients[3]  
  
B3<-autoModel5\_5$coefficients[4]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted5 <- B0 + B1 \* test$displacement + B2 \* test$model.year +B3 \*test$ weight  
  
  
  
  
  
# #calculating error ##  
  
error5<- y\_predicted5- test$mpg  
  
summary(error5)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -15.8714 -5.5647 -2.8591 -2.9189 0.7378 5.1699

plot(error5)  
  
abline(0,0)



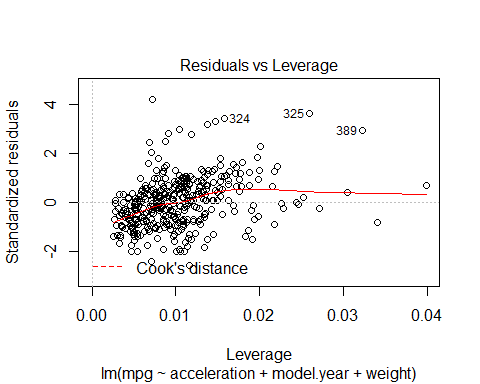
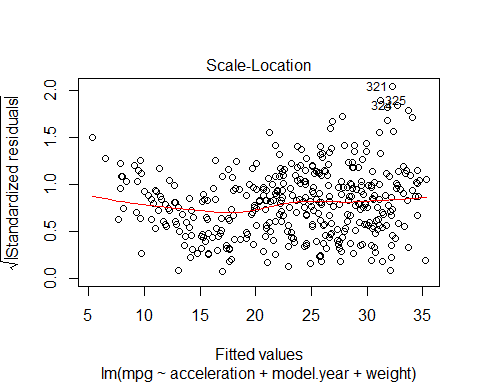
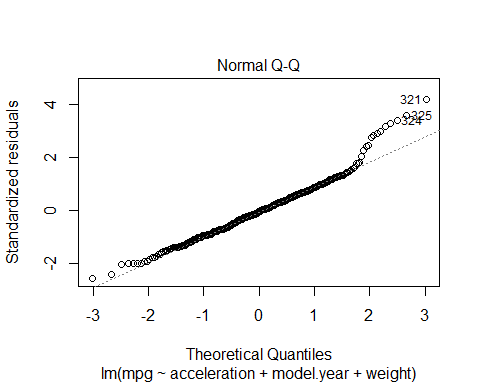
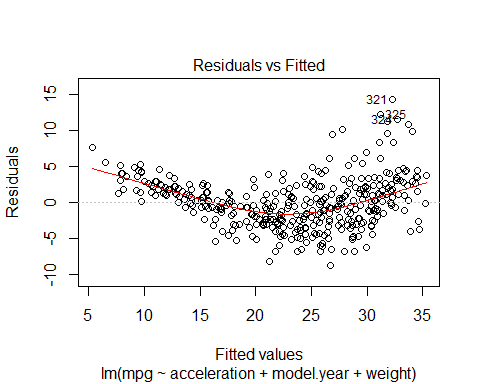
hist(error5,prob=T,breaks = 20)  
  
lines(density(error5),col='red')  
  
  
  
#Normal Curve for automodel4  
mean\_error5<-mean(error5)  
variance\_error5<-var(error5)  
sd\_error5<-sqrt(variance\_error5)  
x\_error5<-seq(-25,25,length=30)  
y\_error5<-dnorm(x\_error5,mean\_error5,sd\_error5)  
lines(x\_error5,y\_error5,col='blue')



############ Model 6 #############  
  
  
autoModel6 = lm(mpg ~ acceleration + model.year + weight , data = autompg)  
summary(autoModel6)

##   
## Call:  
## lm(formula = mpg ~ acceleration + model.year + weight, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -8.6749 -2.3528 -0.1082 2.0168 14.3022   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -14.936555 4.055512 -3.683 0.000263 \*\*\*  
## acceleration 0.066359 0.070361 0.943 0.346204   
## model.year 0.748446 0.050366 14.860 < 2e-16 \*\*\*  
## weight -0.006554 0.000230 -28.502 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.428 on 388 degrees of freedom  
## Multiple R-squared: 0.8086, Adjusted R-squared: 0.8071   
## F-statistic: 546.5 on 3 and 388 DF, p-value: < 2.2e-16

plot(autoModel6)



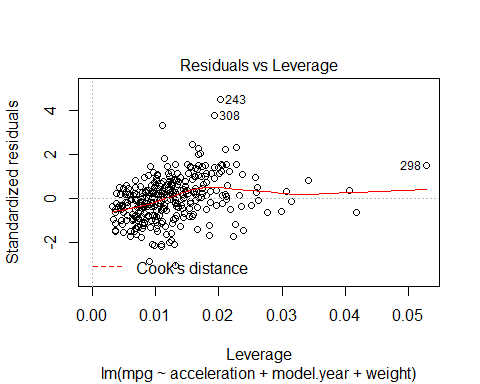
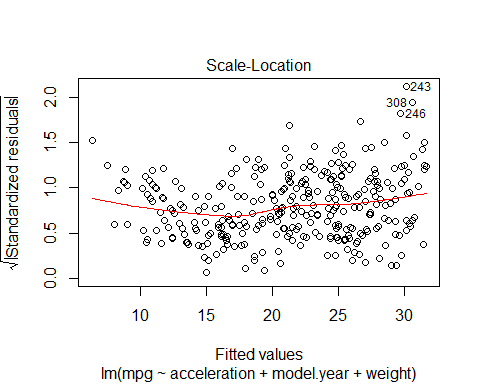
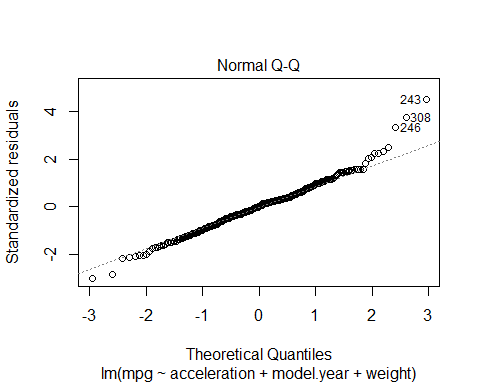
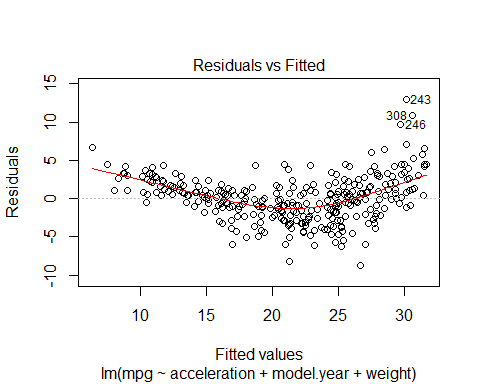
coefficients(autoModel6)

## (Intercept) acceleration model.year weight   
## -14.936555253 0.066359316 0.748446024 -0.006554126

#######regression Model ######  
  
  
autoModel6\_6 <- lm(mpg ~ acceleration + model.year + weight , data = training)  
summary(autoModel6\_6)

##   
## Call:  
## lm(formula = mpg ~ acceleration + model.year + weight, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -8.700 -1.767 0.026 1.613 12.949   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -2.3465954 4.3183561 -0.543 0.587   
## acceleration -0.0505180 0.0680079 -0.743 0.458   
## model.year 0.5933003 0.0570995 10.391 <2e-16 \*\*\*  
## weight -0.0063948 0.0002079 -30.755 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.902 on 314 degrees of freedom  
## Multiple R-squared: 0.8157, Adjusted R-squared: 0.814   
## F-statistic: 463.3 on 3 and 314 DF, p-value: < 2.2e-16

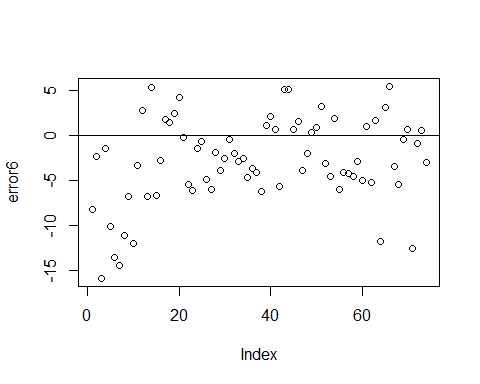
plot(autoModel6\_6)



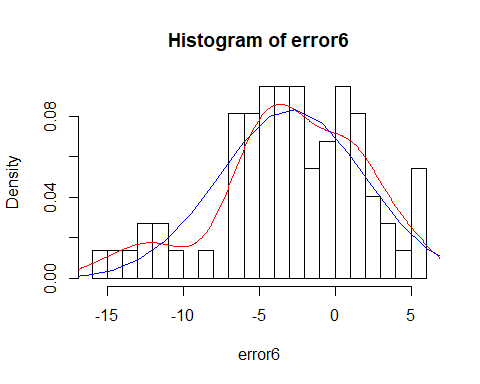
### use model on test dataset ###  
  
B0<-autoModel6\_6$coefficients[1]  
  
B1<-autoModel6\_6$coefficients[2]  
  
B2<-autoModel6\_6$coefficients[3]  
  
B3<-autoModel6\_6$coefficients[4]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted6 <- B0 + B1 \* test$acceleration + B2 \* test$model.year +B3 \*test$ weight  
  
  
  
  
  
# #calculating error ##  
  
error6<- y\_predicted6- test$mpg  
  
summary(error6)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -15.8799 -5.4174 -2.8174 -2.8606 0.6744 5.4321

plot(error6)  
  
abline(0,0)



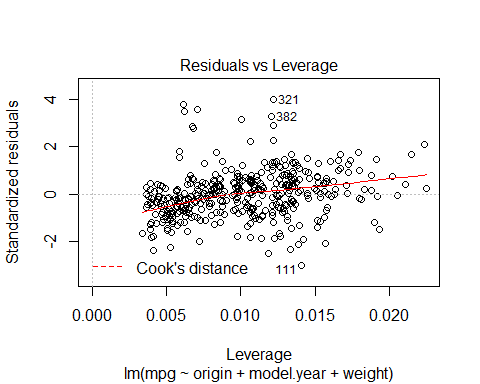
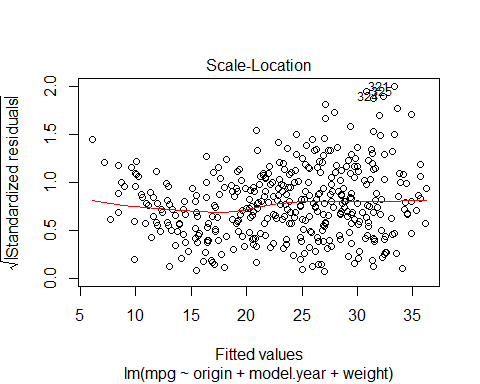
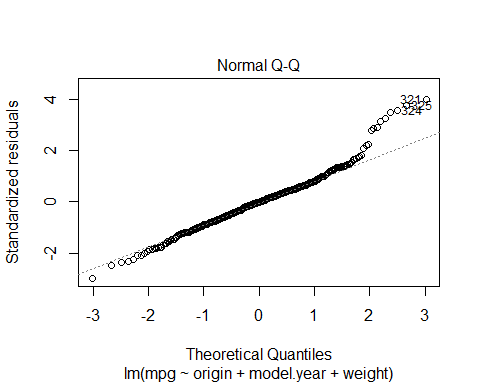
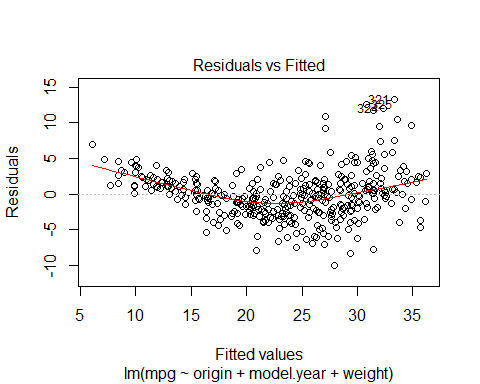
hist(error6,prob=T,breaks = 20)  
  
lines(density(error6),col='red')  
  
  
  
#Normal Curve for automodel4  
mean\_error6<-mean(error6)  
variance\_error6<-var(error6)  
sd\_error6<-sqrt(variance\_error6)  
x\_error6<-seq(-25,25,length=30)  
y\_error6<-dnorm(x\_error6,mean\_error6,sd\_error6)  
lines(x\_error6,y\_error6,col='blue')



############ Model 7 ##############  
  
  
autoModel7 = lm(mpg ~ origin + model.year + weight , data = autompg)  
summary(autoModel7)

##   
## Call:  
## lm(formula = mpg ~ origin + model.year + weight, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -9.9440 -2.0948 -0.0389 1.7255 13.2722   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.805e+01 4.001e+00 -4.510 8.60e-06 \*\*\*  
## origin 1.150e+00 2.591e-01 4.439 1.18e-05 \*\*\*  
## model.year 7.571e-01 4.832e-02 15.668 < 2e-16 \*\*\*  
## weight -5.994e-03 2.541e-04 -23.588 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.348 on 388 degrees of freedom  
## Multiple R-squared: 0.8175, Adjusted R-squared: 0.816   
## F-statistic: 579.2 on 3 and 388 DF, p-value: < 2.2e-16

plot(autoModel7)



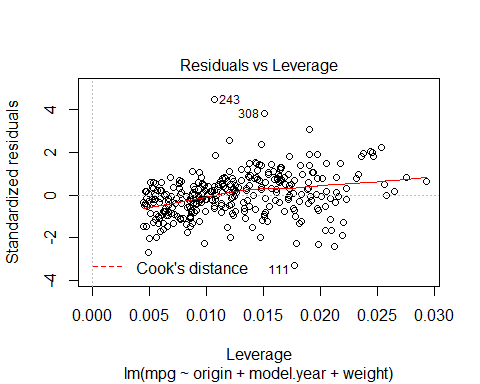
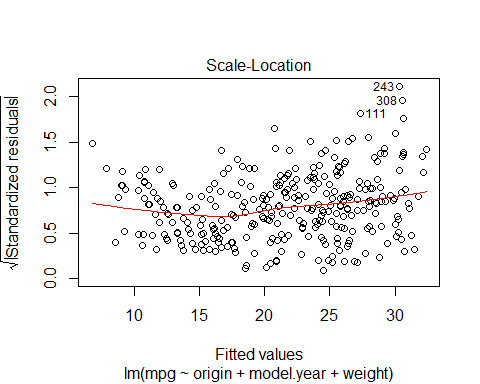
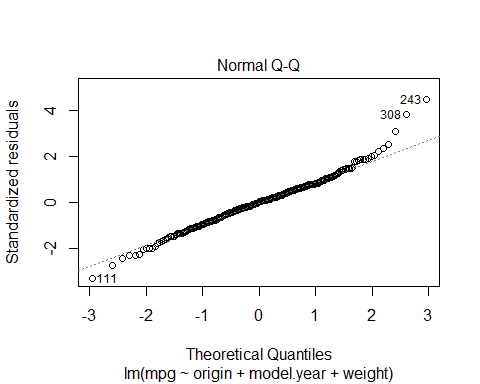
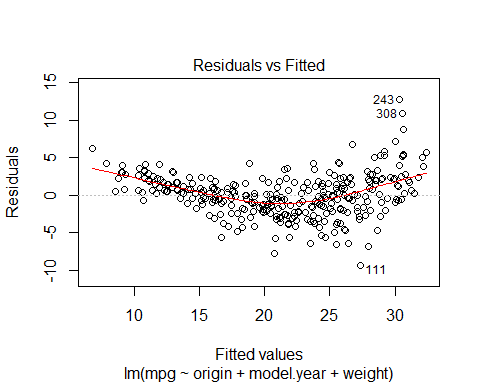
coefficients(autoModel7)

## (Intercept) origin model.year weight   
## -18.045850149 1.150390789 0.757126111 -0.005994118

######### regression model ###########  
  
  
autoModel7\_7 <- lm(mpg ~ origin + model.year + weight , data = training)  
summary(autoModel7\_7)

##   
## Call:  
## lm(formula = mpg ~ origin + model.year + weight, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -9.3469 -1.7854 0.0849 1.7272 12.7686   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -5.5634279 4.3640262 -1.275 0.20331   
## origin 0.7910404 0.2697267 2.933 0.00361 \*\*   
## model.year 0.5905475 0.0550589 10.726 < 2e-16 \*\*\*  
## weight -0.0059194 0.0002327 -25.441 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.866 on 314 degrees of freedom  
## Multiple R-squared: 0.8203, Adjusted R-squared: 0.8186   
## F-statistic: 477.8 on 3 and 314 DF, p-value: < 2.2e-16

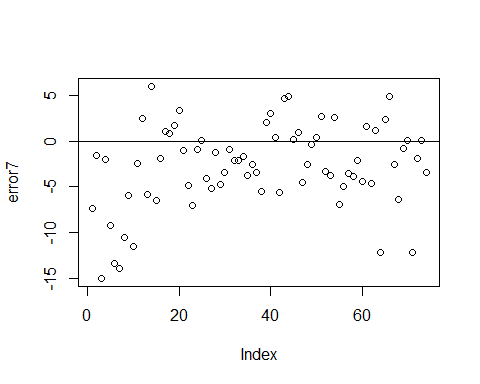
plot(autoModel7\_7)



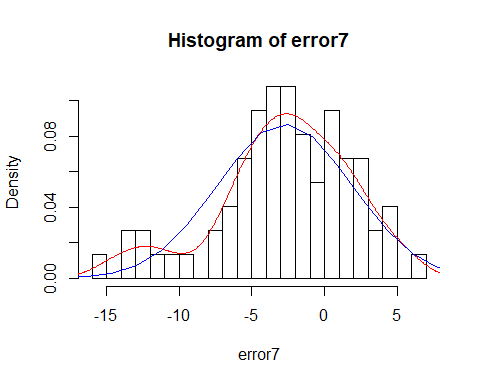
### use model on test dataset ###  
  
B0<-autoModel7\_7$coefficients[1]  
  
B1<-autoModel7\_7$coefficients[2]  
  
B2<-autoModel7\_7$coefficients[3]  
  
B3<-autoModel7\_7$coefficients[4]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted7 <- B0 + B1 \* test$origin + B2 \* test$model.year +B3 \*test$ weight  
  
  
  
  
  
# #calculating error ##  
  
error7<- y\_predicted7- test$mpg  
  
summary(error7)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -15.0364 -4.9057 -2.2907 -2.7447 0.3936 6.0286

plot(error7)  
  
abline(0,0)



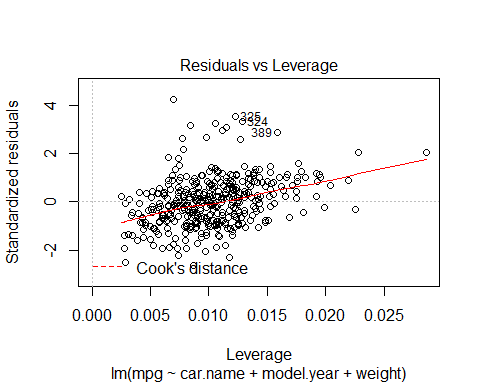
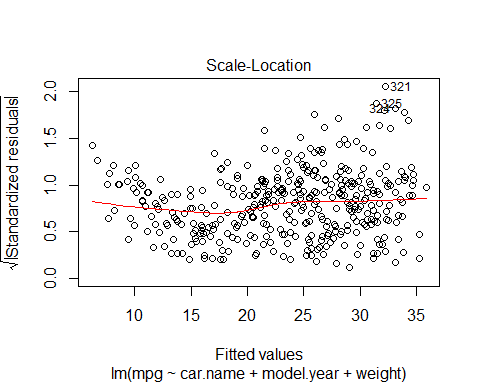
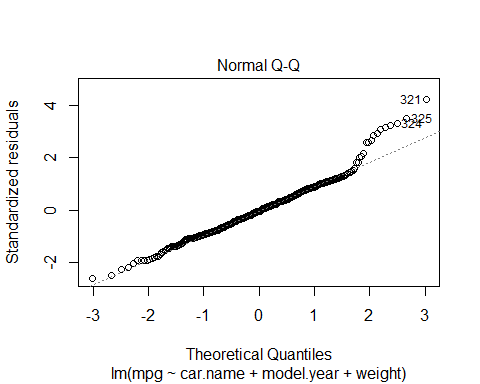
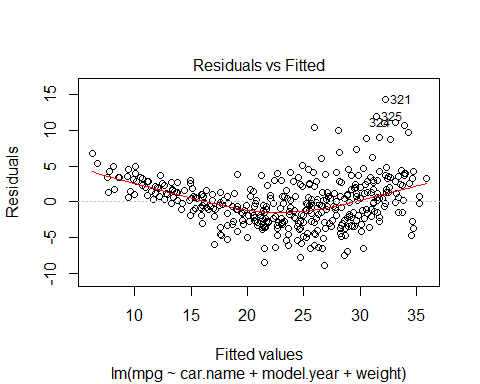
hist(error7,prob=T,breaks = 20)  
  
lines(density(error7),col='red')  
  
  
  
#Normal Curve for automodel4  
mean\_error7<-mean(error7)  
variance\_error7<-var(error7)  
sd\_error7<-sqrt(variance\_error7)  
x\_error7<-seq(-25,25,length=30)  
y\_error7<-dnorm(x\_error7,mean\_error7,sd\_error7)  
lines(x\_error7,y\_error7,col='blue')



########### Model 8 ###################  
  
  
  
autoModel8 = lm(mpg ~ car.name + model.year + weight , data = autompg)  
summary(autoModel8)

##   
## Call:  
## lm(formula = mpg ~ car.name + model.year + weight, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -8.8659 -2.2954 -0.1699 2.0068 14.3290   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.562e+01 4.000e+00 -3.904 0.000112 \*\*\*  
## car.name 5.449e-03 1.983e-03 2.748 0.006269 \*\*   
## model.year 7.576e-01 4.906e-02 15.442 < 2e-16 \*\*\*  
## weight -6.487e-03 2.192e-04 -29.597 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.399 on 388 degrees of freedom  
## Multiple R-squared: 0.8118, Adjusted R-squared: 0.8104   
## F-statistic: 558 on 3 and 388 DF, p-value: < 2.2e-16

plot(autoModel8)



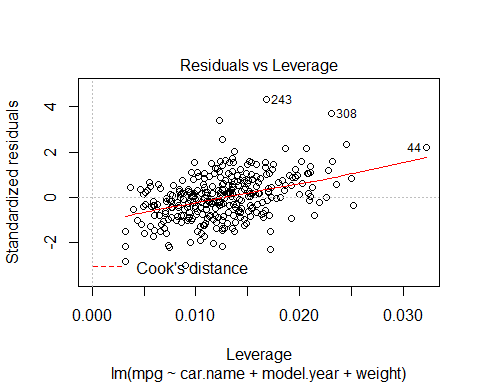
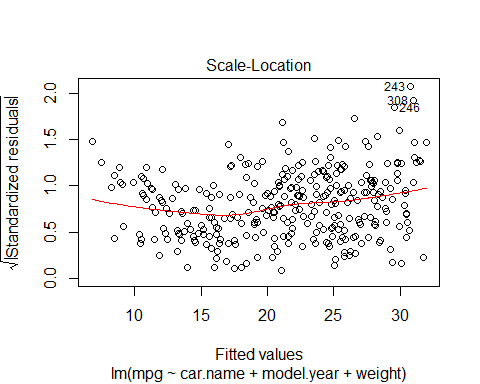
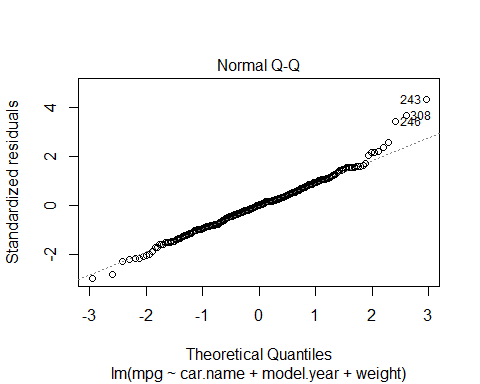
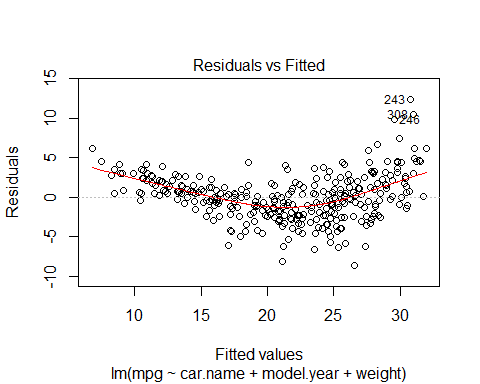
coefficients(autoModel8)

## (Intercept) car.name model.year weight   
## -15.615348708 0.005449309 0.757626282 -0.006487357

########## Regression Model ##########   
  
autoModel8\_8 <- lm(mpg ~ car.name + model.year + weight , data = training)  
summary(autoModel8\_8)

##   
## Call:  
## lm(formula = mpg ~ car.name + model.year + weight, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -8.5797 -1.9285 0.0098 1.6797 12.3583   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -3.4943025 4.3161696 -0.810 0.4188   
## car.name 0.0031308 0.0018450 1.697 0.0907 .   
## model.year 0.5864794 0.0555273 10.562 <2e-16 \*\*\*  
## weight -0.0062556 0.0001926 -32.473 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.892 on 314 degrees of freedom  
## Multiple R-squared: 0.8171, Adjusted R-squared: 0.8153   
## F-statistic: 467.5 on 3 and 314 DF, p-value: < 2.2e-16

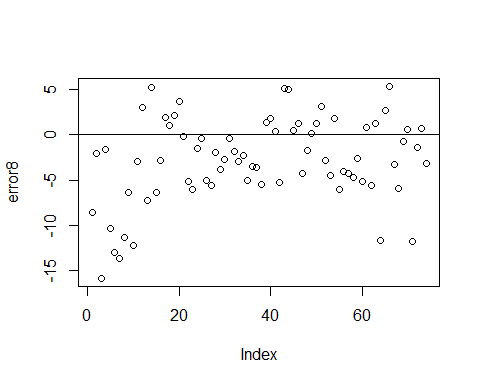
plot(autoModel8\_8)



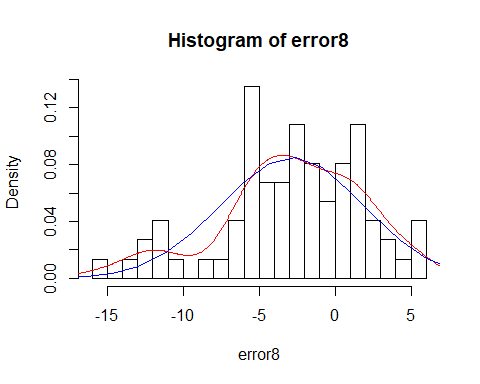
### use model on test dataset ###  
  
B0<-autoModel8\_8$coefficients[1]  
  
B1<-autoModel8\_8$coefficients[2]  
  
B2<-autoModel8\_8$coefficients[3]  
  
B3<-autoModel8\_8$coefficients[4]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted8 <- B0 + B1 \* test$car.name + B2 \* test$model.year +B3 \*test$ weight  
  
  
  
  
  
# #calculating error ##  
  
error8<- y\_predicted8- test$mpg  
  
summary(error8)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -15.8211 -5.2460 -2.7767 -2.8537 0.6643 5.3258

plot(error8)  
  
abline(0,0)



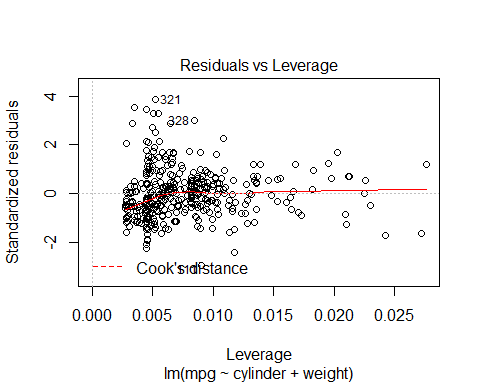
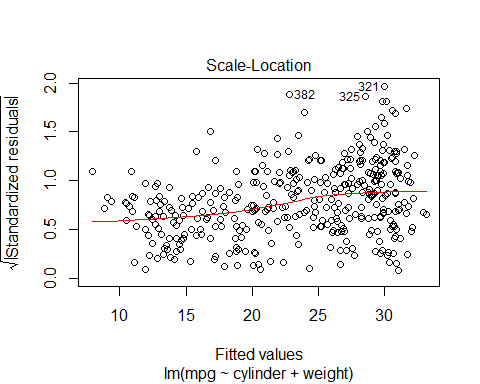
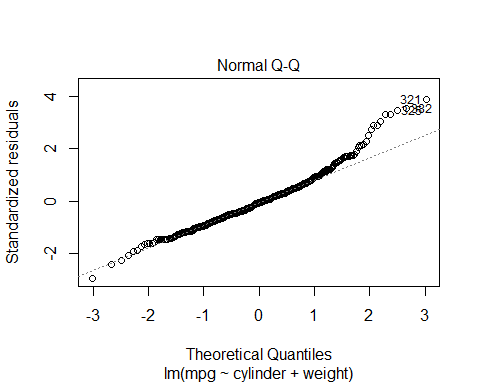
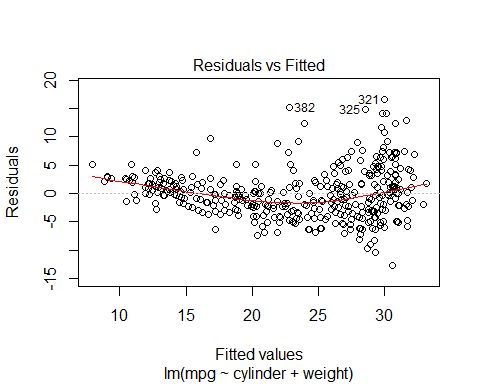
hist(error8,prob=T,breaks = 20)  
  
lines(density(error8),col='red')  
  
  
  
#Normal Curve for automodel4  
mean\_error8<-mean(error8)  
variance\_error8<-var(error8)  
sd\_error8<-sqrt(variance\_error8)  
x\_error8<-seq(-25,25,length=30)  
y\_error8<-dnorm(x\_error8,mean\_error8,sd\_error8)  
lines(x\_error8,y\_error8,col='blue')



############# Model 9 ##############  
  
  
  
autoModel9= lm(mpg ~ cylinder + weight, data = autompg)  
summary(autoModel9)

##   
## Call:  
## lm(formula = mpg ~ cylinder + weight, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -12.6469 -2.8282 -0.2905 2.1606 16.5856   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 46.2923105 0.7939685 58.305 <2e-16 \*\*\*  
## cylinder -0.7213779 0.2893780 -2.493 0.0131 \*   
## weight -0.0063471 0.0005811 -10.922 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.304 on 389 degrees of freedom  
## Multiple R-squared: 0.6975, Adjusted R-squared: 0.6959   
## F-statistic: 448.4 on 2 and 389 DF, p-value: < 2.2e-16

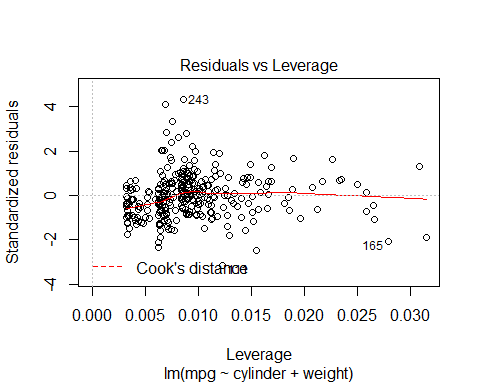
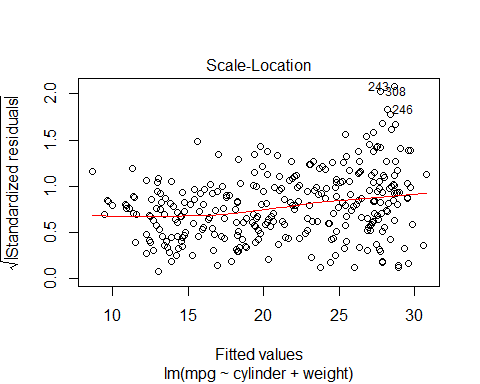
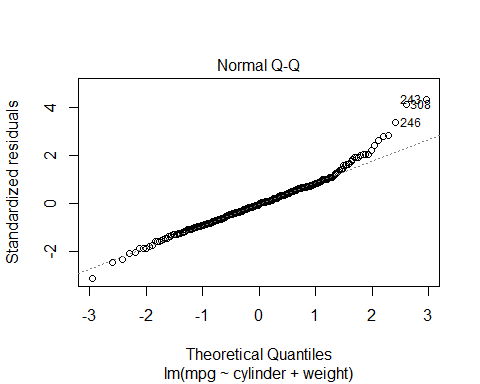
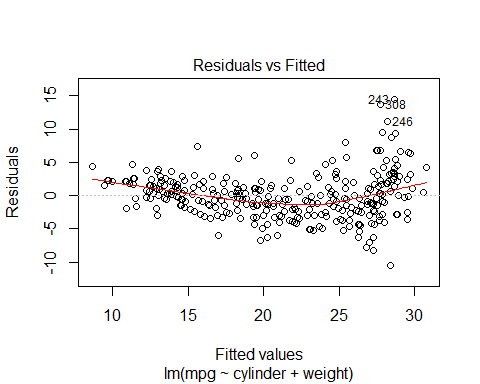
plot(autoModel9)



########## Regression Model ##########   
  
autoModel9\_9 <- lm(mpg ~ cylinder + weight , data = training)  
summary(autoModel9\_9)

##   
## Call:  
## lm(formula = mpg ~ cylinder + weight, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -10.4432 -2.1948 -0.1558 1.8771 14.4162   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 42.1207884 0.6904495 61.005 <2e-16 \*\*\*  
## cylinder -0.5470653 0.2458245 -2.225 0.0268 \*   
## weight -0.0056669 0.0004904 -11.556 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.343 on 315 degrees of freedom  
## Multiple R-squared: 0.7548, Adjusted R-squared: 0.7532   
## F-statistic: 484.7 on 2 and 315 DF, p-value: < 2.2e-16

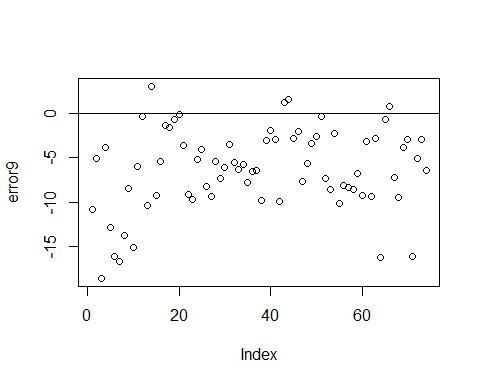
plot(autoModel9\_9)



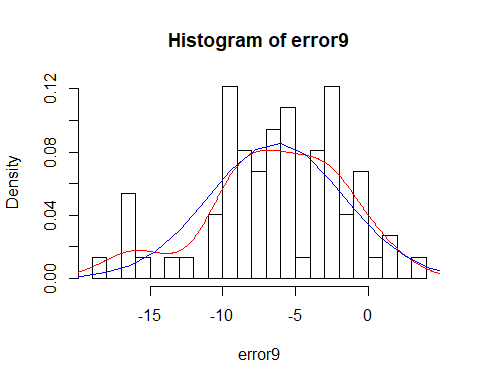
### use model on test dataset ###  
  
B0<-autoModel9\_9$coefficients[1]  
  
B1<-autoModel9\_9$coefficients[2]  
  
B2<-autoModel9\_9$coefficients[3]  
  
  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted9 <- B0 + B1 \* test$cylinder + B2 \*test$ weight  
  
  
  
  
  
# #calculating error ##  
  
error9<- y\_predicted9- test$mpg  
  
summary(error9)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -18.625 -9.180 -6.040 -6.315 -2.921 3.066

plot(error9)  
  
abline(0,0)



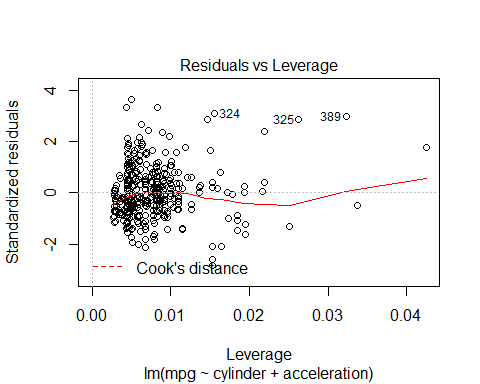
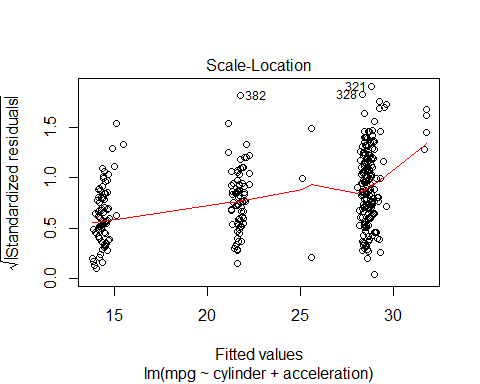
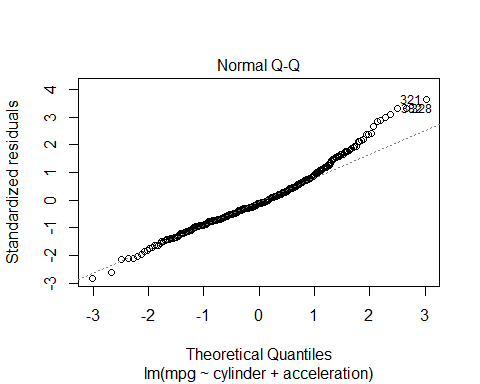
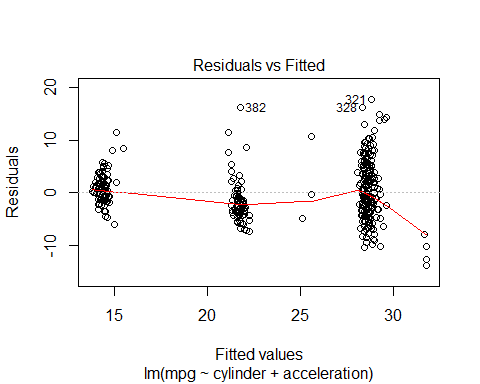
hist(error9,prob=T,breaks = 20)  
  
lines(density(error9),col='red')  
  
  
  
#Normal Curve for automodel4  
mean\_error9<-mean(error9)  
variance\_error9<-var(error9)  
sd\_error9<-sqrt(variance\_error9)  
x\_error9<-seq(-25,25,length=30)  
y\_error9<-dnorm(x\_error9,mean\_error9,sd\_error9)  
lines(x\_error9,y\_error9,col='blue')



########### Model 10 ###############  
  
  
  
autoModel10 = lm(mpg ~ cylinder + acceleration, data = autompg)  
summary(autoModel10)

##   
## Call:  
## lm(formula = mpg ~ cylinder + acceleration, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -13.7655 -3.1853 -0.5624 2.4918 17.7812   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 40.5704 2.2475 18.052 <2e-16 \*\*\*  
## cylinder -3.4624 0.1687 -20.526 <2e-16 \*\*\*  
## acceleration 0.1172 0.1043 1.124 0.262   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.912 on 389 degrees of freedom  
## Multiple R-squared: 0.606, Adjusted R-squared: 0.6039   
## F-statistic: 299.1 on 2 and 389 DF, p-value: < 2.2e-16

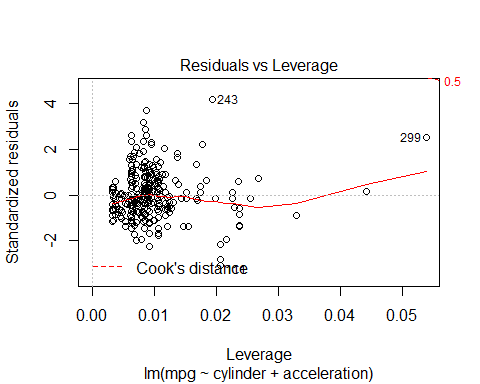
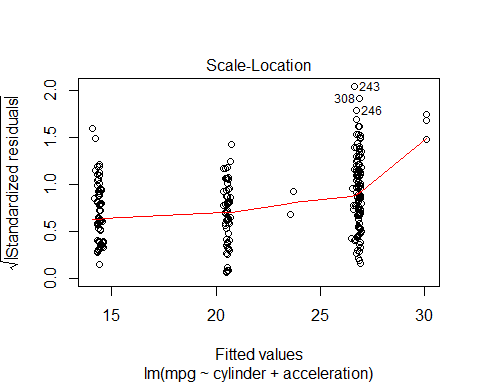
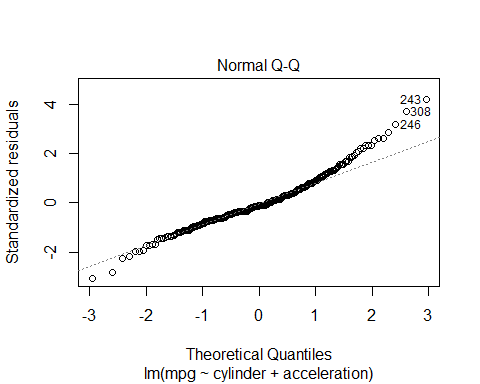
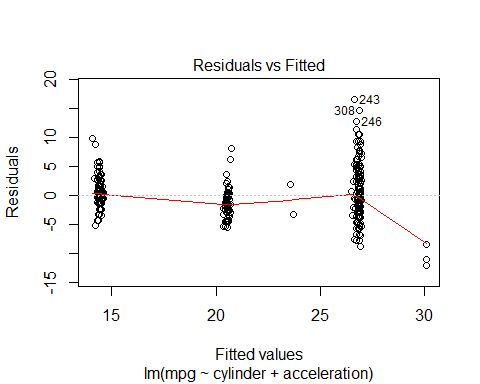
plot(autoModel10)



########## Regression Model ##########   
  
autoModel10\_10 <- lm(mpg ~ cylinder + acceleration , data = training)  
summary(autoModel10\_10)

##   
## Call:  
## lm(formula = mpg ~ cylinder + acceleration, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -12.067 -2.483 -0.485 2.039 16.475   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 39.99110 2.08964 19.138 <2e-16 \*\*\*  
## cylinder -3.13504 0.15155 -20.686 <2e-16 \*\*\*  
## acceleration -0.03841 0.09660 -0.398 0.691   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.988 on 315 degrees of freedom  
## Multiple R-squared: 0.651, Adjusted R-squared: 0.6488   
## F-statistic: 293.8 on 2 and 315 DF, p-value: < 2.2e-16

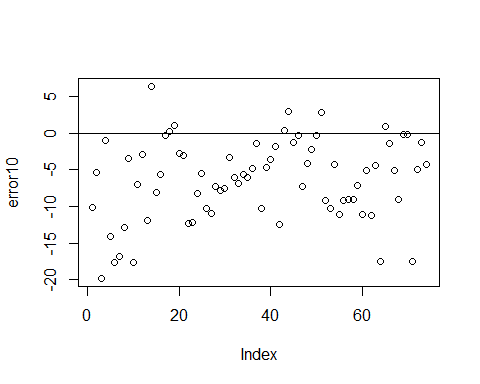
plot(autoModel10\_10)



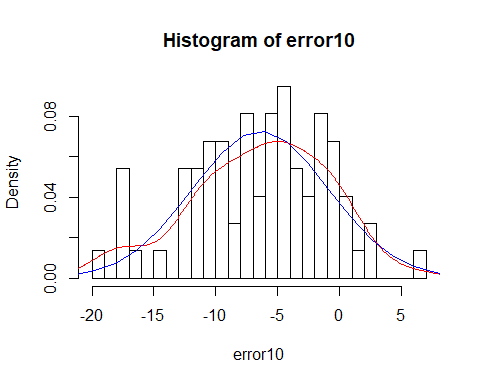
### use model on test dataset ###  
  
B0<-autoModel10\_10$coefficients[1]  
  
B1<-autoModel10\_10$coefficients[2]  
  
B2<-autoModel10\_10$coefficients[3]  
  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted10 <- B0 + B1 \* test$cylinder + B2 \*test$ acceleration  
  
  
  
  
  
# #calculating error ##  
  
error10<- y\_predicted10- test$mpg  
  
summary(error10)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -19.837 -10.217 -5.598 -6.391 -2.324 6.406

plot(error10)  
  
abline(0,0)



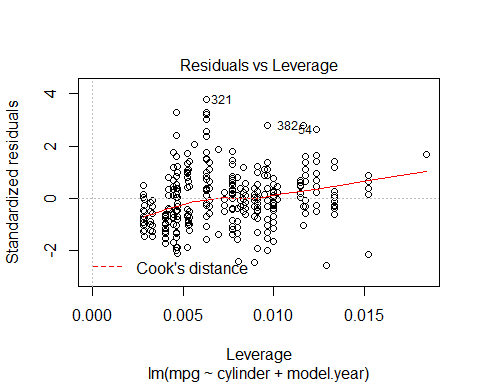
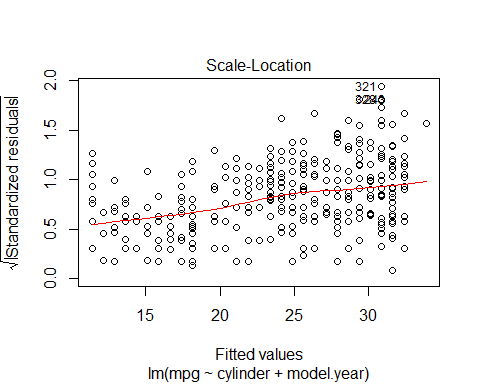
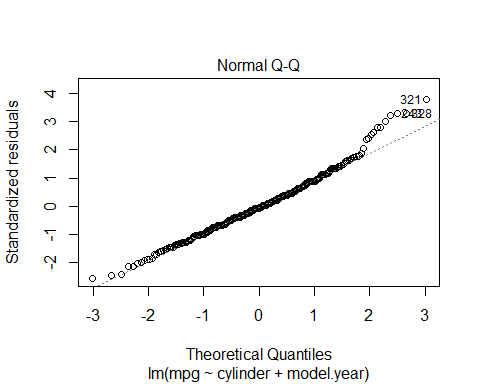
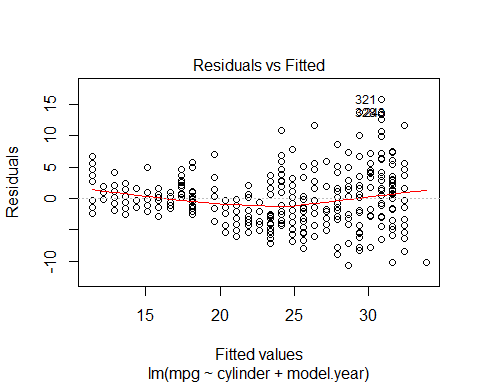
hist(error10,prob=T,breaks = 20)  
  
lines(density(error10),col='red')  
  
  
  
#Normal Curve for automodel4  
mean\_error10<-mean(error10)  
variance\_error10<-var(error10)  
sd\_error10<-sqrt(variance\_error10)  
x\_error10<-seq(-25,25,length=30)  
y\_error10<-dnorm(x\_error10,mean\_error10,sd\_error10)  
lines(x\_error10,y\_error10,col='blue')



############ Model 11 #################  
  
  
  
autoModel11= lm(mpg ~ cylinder + model.year, data = autompg)  
summary(autoModel11)

##   
## Call:  
## lm(formula = mpg ~ cylinder + model.year, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -10.6219 -2.8738 -0.3014 2.4856 15.7250   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -17.14640 4.94386 -3.468 0.000582 \*\*\*  
## cylinder -2.99812 0.13197 -22.718 < 2e-16 \*\*\*  
## model.year 0.75017 0.06111 12.276 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.177 on 389 degrees of freedom  
## Multiple R-squared: 0.7151, Adjusted R-squared: 0.7136   
## F-statistic: 488.1 on 2 and 389 DF, p-value: < 2.2e-16

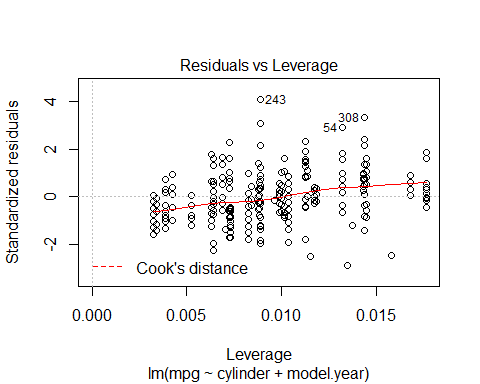
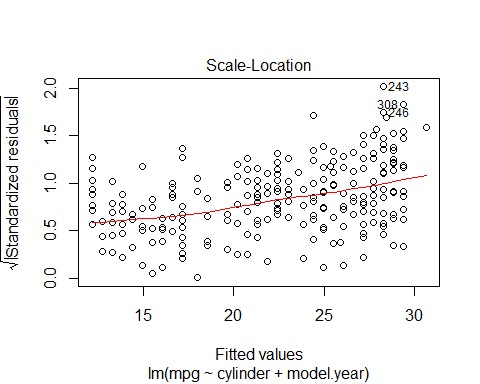
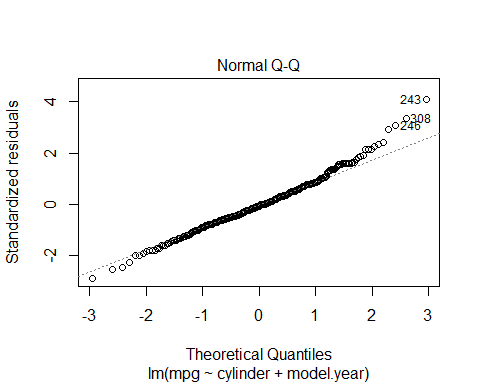
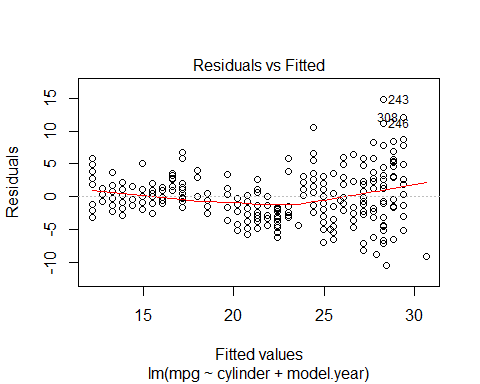
plot(autoModel11)



########## Regression Model ##########   
  
autoModel11\_11 <- lm(mpg ~ cylinder + model.year , data = training)  
summary(autoModel11\_11)

##   
## Call:  
## lm(formula = mpg ~ cylinder + model.year, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -10.4336 -2.2344 -0.2407 2.0436 14.8099   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -3.36606 5.43600 -0.619 0.536   
## cylinder -2.92194 0.11921 -24.510 < 2e-16 \*\*\*  
## model.year 0.55569 0.07034 7.900 4.71e-14 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.644 on 315 degrees of freedom  
## Multiple R-squared: 0.7085, Adjusted R-squared: 0.7067   
## F-statistic: 382.9 on 2 and 315 DF, p-value: < 2.2e-16

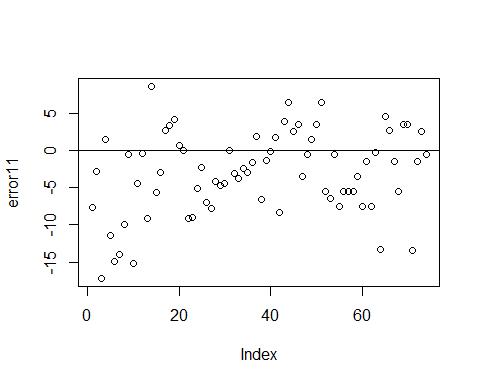
plot(autoModel11\_11)



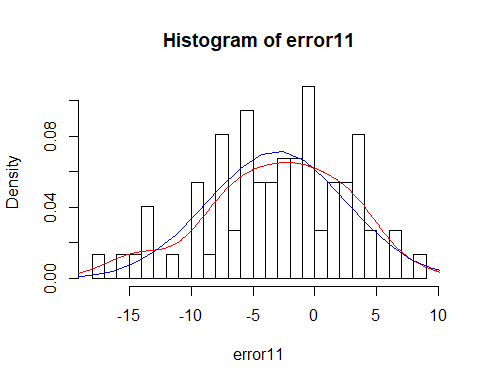
### use model on test dataset ###  
  
B0<-autoModel11\_11$coefficients[1]  
  
B1<-autoModel11\_11$coefficients[2]  
  
B2<-autoModel11\_11$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted11 <- B0 + B1 \* test$cylinder + B2 \*test$ model.year  
  
  
  
  
# #calculating error ##  
  
error11<- y\_predicted11- test$mpg  
  
summary(error11)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -17.199 -6.562 -2.871 -3.070 1.279 8.623

plot(error11)  
  
abline(0,0)



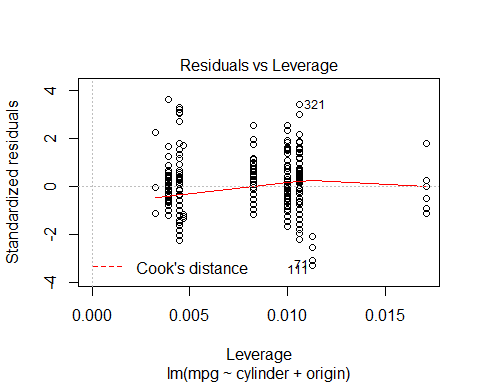
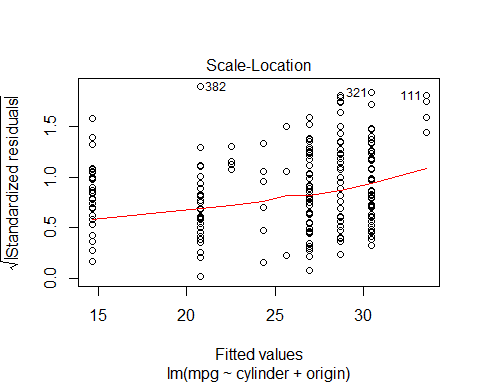
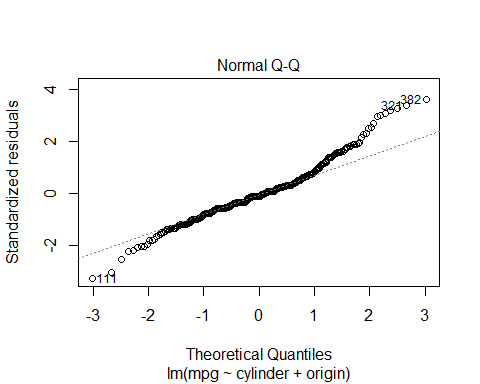
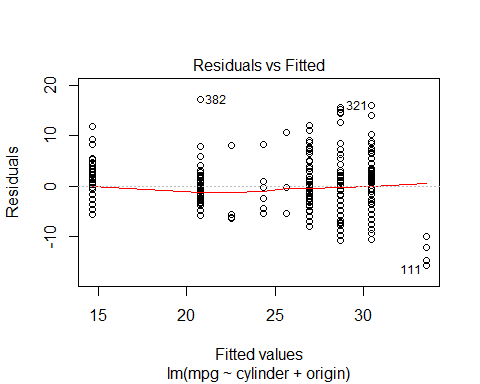
hist(error11,prob=T,breaks = 20)  
  
lines(density(error11),col='red')  
  
  
  
#Normal Curve for automodel4  
mean\_error11<-mean(error11)  
variance\_error11<-var(error11)  
sd\_error11<-sqrt(variance\_error11)  
x\_error11<-seq(-25,25,length=30)  
y\_error11<-dnorm(x\_error11,mean\_error11,sd\_error11)  
lines(x\_error11,y\_error11,col='blue')



############## Model 12 ###################  
  
  
autoModel12= lm(mpg ~ cylinder + origin, data = autompg)  
summary(autoModel12)

##   
## Call:  
## lm(formula = mpg ~ cylinder + origin, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -15.5772 -2.7499 -0.6023 2.0695 17.1977   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 37.5556 1.3766 27.282 < 2e-16 \*\*\*  
## cylinder -3.0854 0.1723 -17.908 < 2e-16 \*\*\*  
## origin 1.7593 0.3648 4.822 2.04e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.779 on 389 degrees of freedom  
## Multiple R-squared: 0.627, Adjusted R-squared: 0.6251   
## F-statistic: 326.9 on 2 and 389 DF, p-value: < 2.2e-16

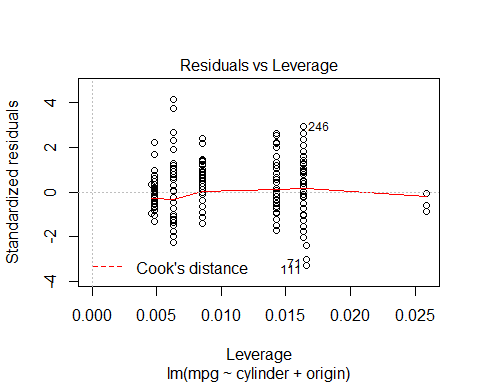
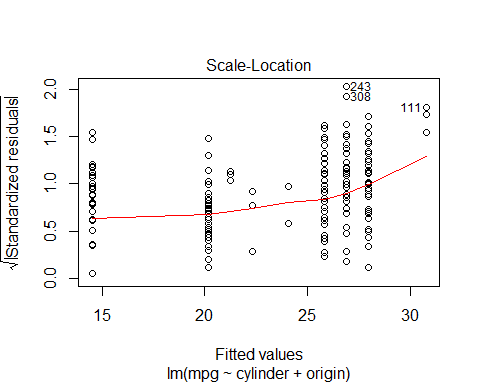
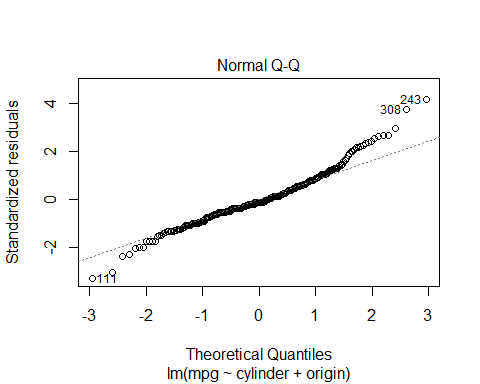
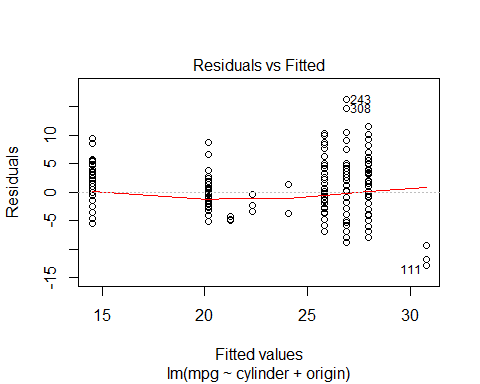
plot(autoModel12)



########## Regression Model ##########   
  
autoModel12\_12 <- lm(mpg ~ cylinder + origin , data = training)  
summary(autoModel12\_12)

##   
## Call:  
## lm(formula = mpg ~ cylinder + origin, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -12.7672 -2.1518 -0.5116 2.1111 16.2305   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 35.9946 1.3519 26.625 < 2e-16 \*\*\*  
## cylinder -2.8201 0.1601 -17.617 < 2e-16 \*\*\*  
## origin 1.0776 0.3743 2.879 0.00426 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.938 on 315 degrees of freedom  
## Multiple R-squared: 0.6598, Adjusted R-squared: 0.6576   
## F-statistic: 305.4 on 2 and 315 DF, p-value: < 2.2e-16

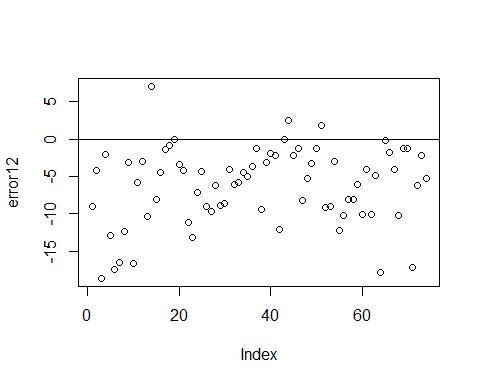
plot(autoModel12\_12)



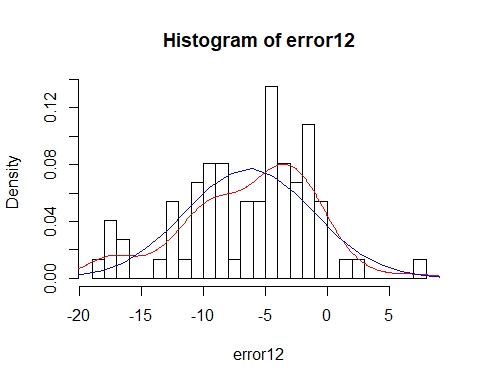
### use model on test dataset ###  
  
B0<-autoModel12\_12$coefficients[1]  
  
B1<-autoModel12\_12$coefficients[2]  
  
B2<-autoModel12\_12$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted12 <- B0 + B1 \* test$cylinder + B2 \*test$origin  
  
  
  
  
# #calculating error ##  
  
error12<- y\_predicted12- test$mpg  
  
summary(error12)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -18.653 -9.386 -5.208 -6.270 -2.419 7.067

plot(error12)  
  
abline(0,0)



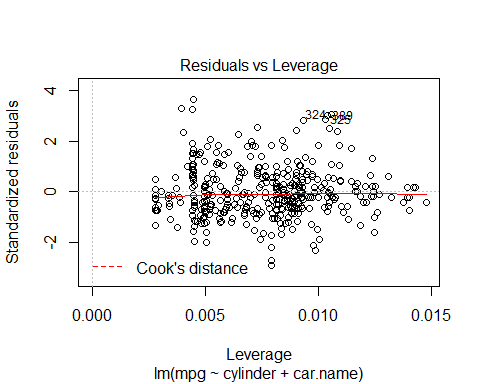
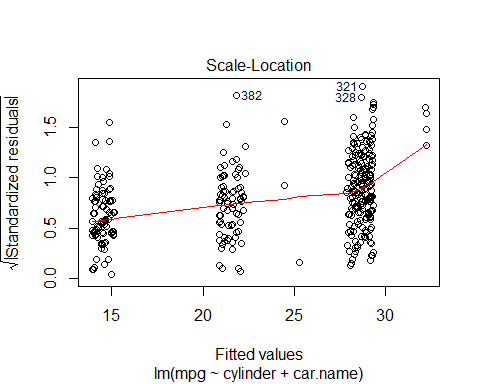
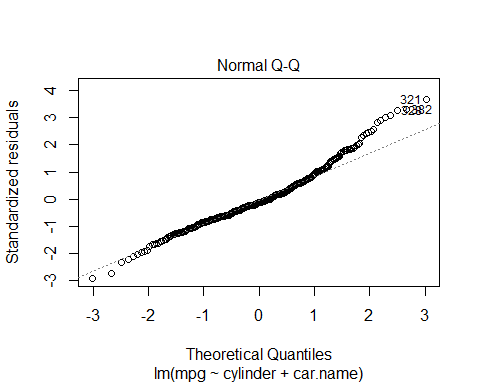
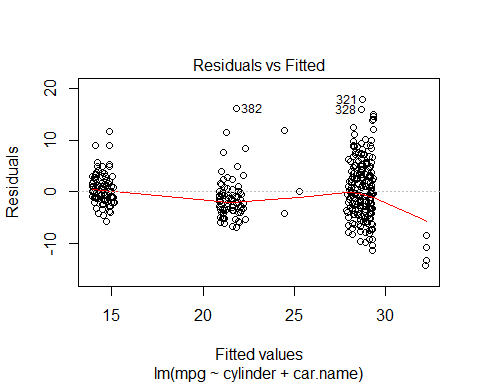
hist(error12,prob=T,breaks = 20)  
  
lines(density(error12),col='red')  
  
  
  
#Normal Curve for automodel4  
mean\_error12<-mean(error12)  
variance\_error12<-var(error12)  
sd\_error12<-sqrt(variance\_error12)  
x\_error12<-seq(-25,25,length=30)  
y\_error12<-dnorm(x\_error12,mean\_error12,sd\_error12)  
lines(x\_error12,y\_error12,col='blue')



############## Model 13 ###############  
  
  
  
autoModel13= lm(mpg ~ cylinder + car.name, data = autompg)  
summary(autoModel13)

##   
## Call:  
## lm(formula = mpg ~ cylinder + car.name, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -14.1913 -3.1153 -0.6201 2.6171 17.8864   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 41.803498 1.062368 39.349 <2e-16 \*\*\*  
## cylinder -3.487433 0.151251 -23.057 <2e-16 \*\*\*  
## car.name 0.004858 0.002881 1.686 0.0926 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.902 on 389 degrees of freedom  
## Multiple R-squared: 0.6076, Adjusted R-squared: 0.6055   
## F-statistic: 301.1 on 2 and 389 DF, p-value: < 2.2e-16

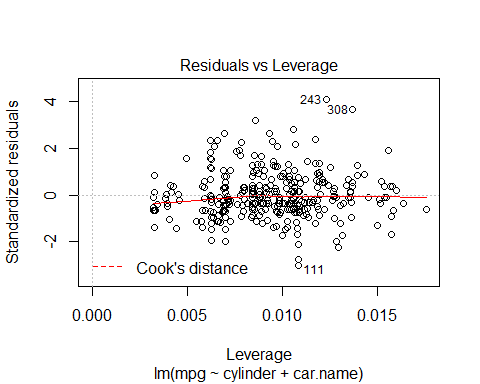
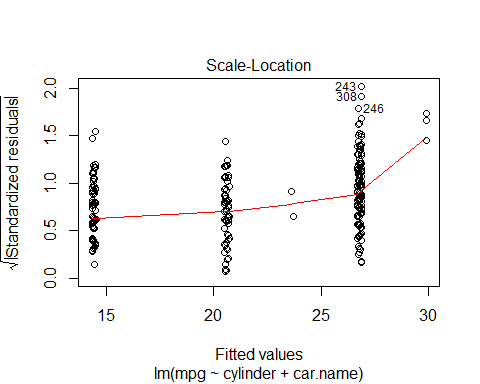
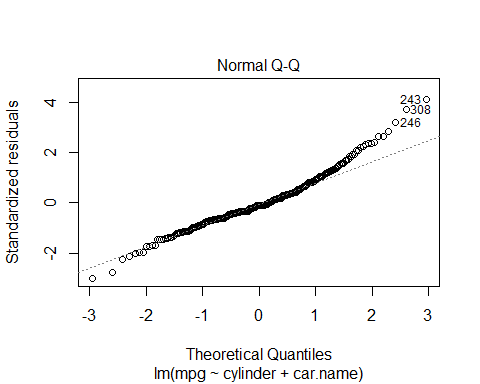
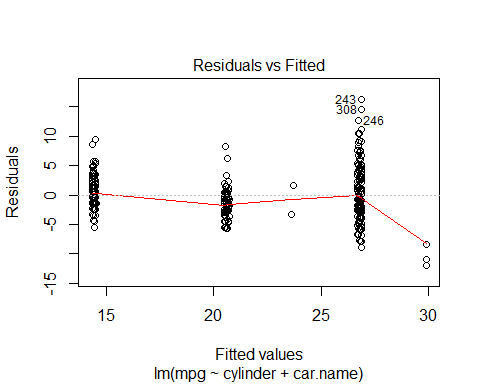
plot(autoModel13)



########## Regression Model ##########   
  
autoModel13\_13 <- lm(mpg ~ cylinder + car.name, data = training)  
summary(autoModel13\_13)

##   
## Call:  
## lm(formula = mpg ~ cylinder + car.name, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -11.9039 -2.4937 -0.4783 2.0338 16.2032   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 39.0512397 0.9616253 40.610 <2e-16 \*\*\*  
## cylinder -3.0923977 0.1330316 -23.246 <2e-16 \*\*\*  
## car.name 0.0007421 0.0025754 0.288 0.773   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.988 on 315 degrees of freedom  
## Multiple R-squared: 0.6509, Adjusted R-squared: 0.6487   
## F-statistic: 293.6 on 2 and 315 DF, p-value: < 2.2e-16

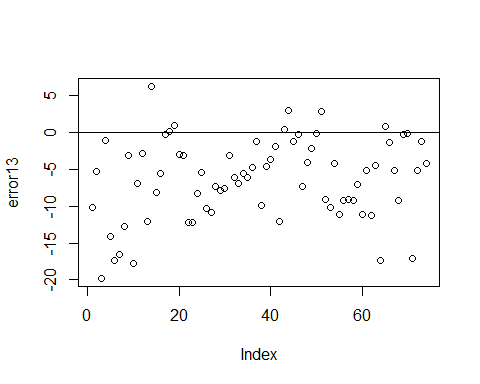
plot(autoModel13\_13)



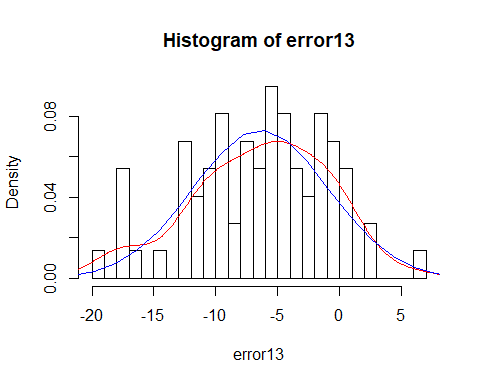
### use model on test dataset ###  
  
B0<-autoModel13\_13$coefficients[1]  
  
B1<-autoModel13\_13$coefficients[2]  
  
B2<-autoModel13\_13$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted13 <- B0 + B1 \* test$cylinder + B2 \*test$car.name  
  
  
  
  
# #calculating error ##  
  
error13<- y\_predicted13- test$mpg  
  
summary(error13)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -19.787 -10.134 -5.558 -6.385 -2.412 6.210

plot(error13)  
  
abline(0,0)



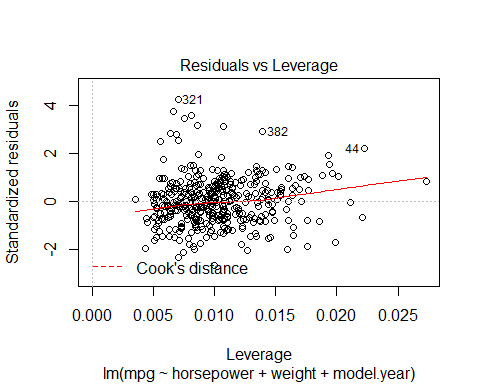
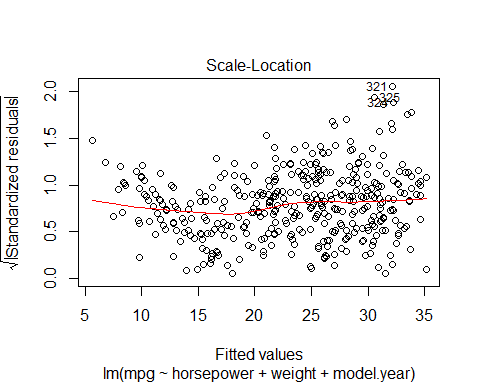
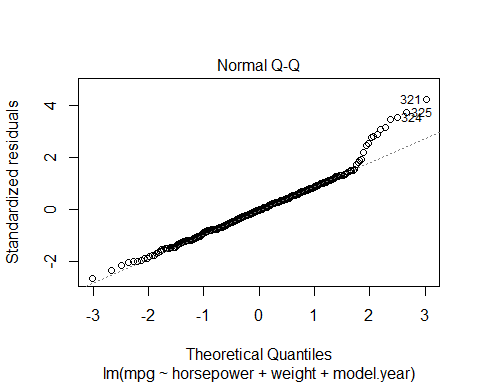
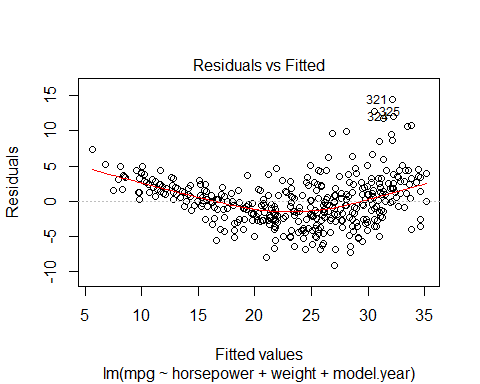
hist(error13,prob=T,breaks = 20)  
  
lines(density(error13),col='red')  
  
  
  
#Normal Curve for automodel4  
mean\_error13<-mean(error13)  
variance\_error13<-var(error13)  
sd\_error13<-sqrt(variance\_error13)  
x\_error13<-seq(-25,25,length=30)  
y\_error13<-dnorm(x\_error13,mean\_error13,sd\_error13)  
lines(x\_error13,y\_error13,col='blue')



######## Model 14 #############  
  
autoModel14= lm(mpg ~ displacement + horsepower , data = autompg)  
summary(autoModel4)

##   
## Call:  
## lm(formula = mpg ~ horsepower + weight + model.year, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -9.0692 -2.3361 -0.0855 1.9037 14.4160   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.578e+01 4.076e+00 -3.871 0.000127 \*\*\*  
## horsepower 1.211e-02 6.807e-03 1.778 0.076108 .   
## weight -6.416e-03 2.462e-04 -26.062 < 2e-16 \*\*\*  
## model.year 7.594e-01 4.935e-02 15.388 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.418 on 388 degrees of freedom  
## Multiple R-squared: 0.8097, Adjusted R-squared: 0.8083   
## F-statistic: 550.4 on 3 and 388 DF, p-value: < 2.2e-16

plot(autoModel4)



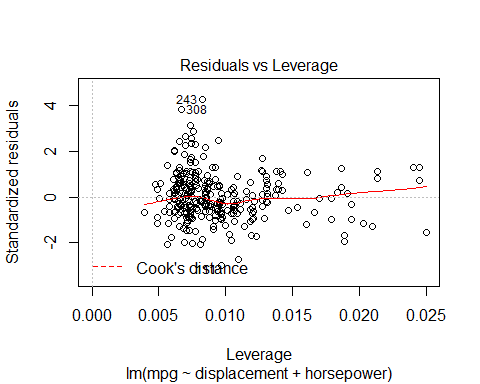
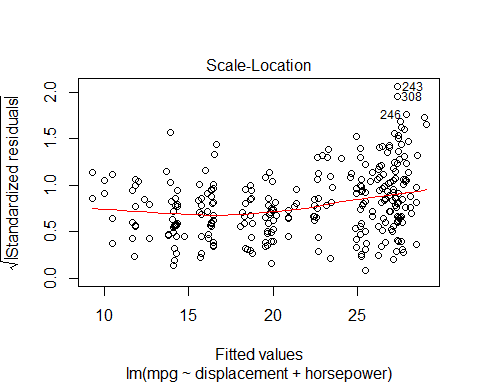
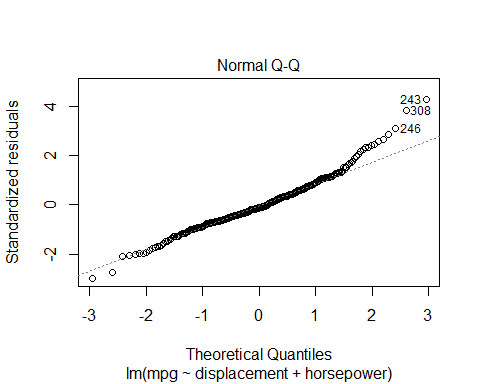
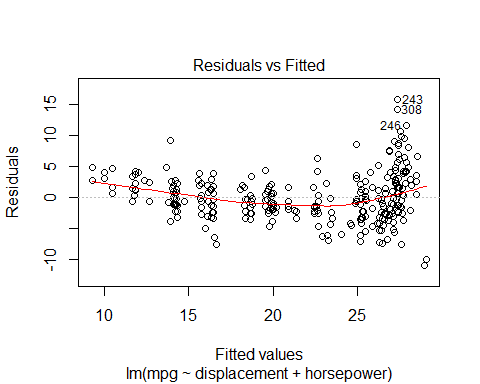
coefficients(autoModel14)

## (Intercept) displacement horsepower   
## 33.93597537 -0.05793190 0.01481191

########## Regression Model ##########   
  
autoModel14\_14 <- lm(mpg ~ displacement + horsepower , data = training)  
summary(autoModel14\_14)

##   
## Call:  
## lm(formula = mpg ~ displacement + horsepower, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -10.959 -2.363 -0.445 2.016 15.704   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 31.017561 0.787514 39.387 <2e-16 \*\*\*  
## displacement -0.049527 0.002240 -22.112 <2e-16 \*\*\*  
## horsepower 0.016381 0.008065 2.031 0.0431 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.703 on 315 degrees of freedom  
## Multiple R-squared: 0.6991, Adjusted R-squared: 0.6972   
## F-statistic: 366 on 2 and 315 DF, p-value: < 2.2e-16

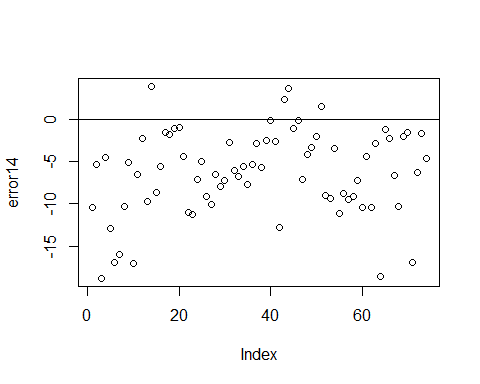
plot(autoModel14\_14)



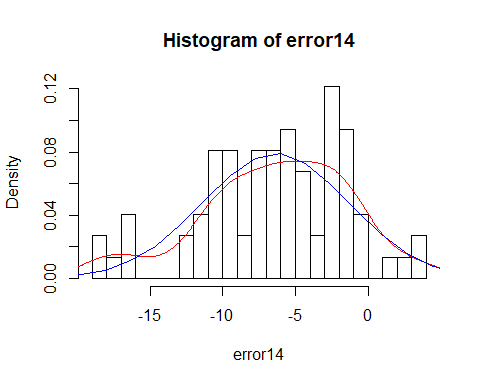
### use model on test dataset ###  
  
B0<-autoModel14\_14$coefficients[1]  
  
B1<-autoModel14\_14$coefficients[2]  
  
B2<-autoModel14\_14$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted14 <- B0 + B1 \* test$displacement + B2 \*test$horsepower  
  
  
  
  
# #calculating error ##  
  
error14<- y\_predicted14- test$mpg  
  
summary(error14)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -18.826 -9.518 -5.860 -6.370 -2.553 3.883

plot(error14)  
  
abline(0,0)



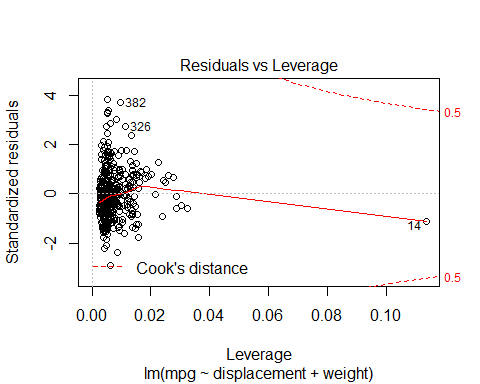
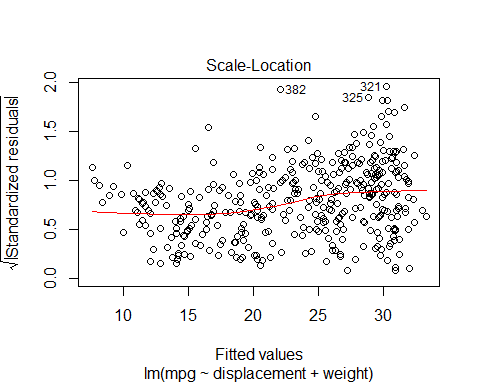
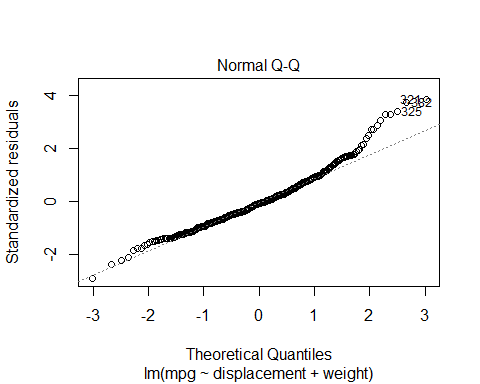
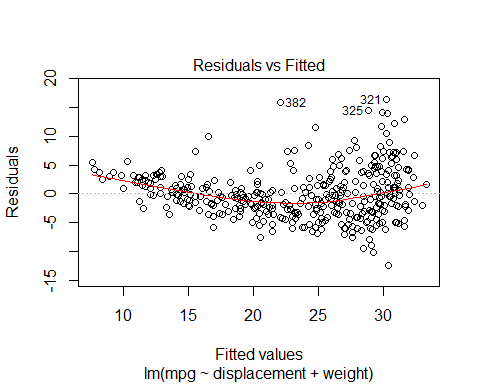
hist(error14,prob=T,breaks = 20)  
  
lines(density(error14),col='red')  
  
  
  
#Normal Curve for automodel14  
mean\_error14<-mean(error14)  
variance\_error14<-var(error14)  
sd\_error14<-sqrt(variance\_error14)  
x\_error14<-seq(-25,25,length=30)  
y\_error14<-dnorm(x\_error14,mean\_error14,sd\_error14)  
lines(x\_error14,y\_error14,col='blue')



########## Model 15 ##############  
  
  
autoModel15= lm(mpg ~ displacement + weight, data = autompg)  
summary(autoModel15)

##   
## Call:  
## lm(formula = mpg ~ displacement + weight, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -12.407 -2.928 -0.357 2.320 16.376   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 43.7776194 1.1630993 37.639 < 2e-16 \*\*\*  
## displacement -0.0164971 0.0057653 -2.861 0.00444 \*\*   
## weight -0.0057511 0.0007103 -8.097 7.31e-15 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.293 on 389 degrees of freedom  
## Multiple R-squared: 0.699, Adjusted R-squared: 0.6974   
## F-statistic: 451.6 on 2 and 389 DF, p-value: < 2.2e-16

plot(autoModel15)



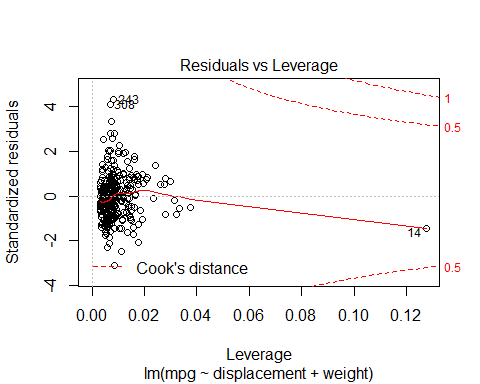
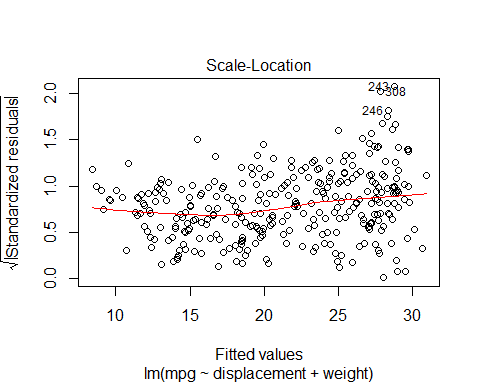
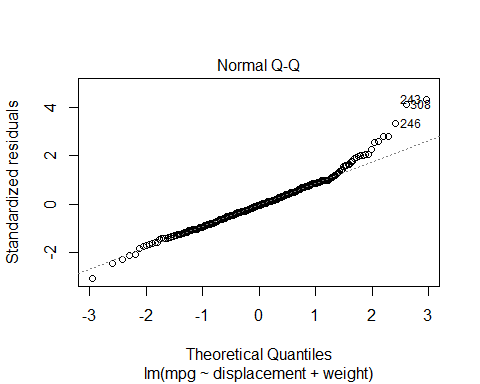
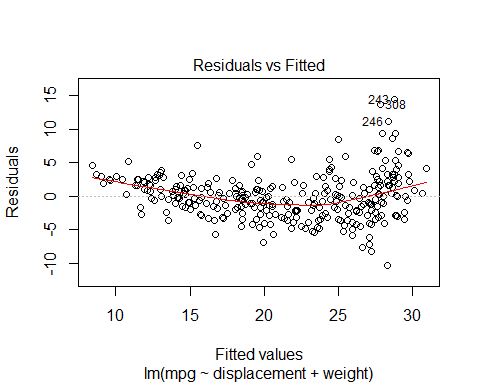
coefficients(autoModel15)

## (Intercept) displacement weight   
## 43.777619395 -0.016497109 -0.005751127

########## Regression Model ##########   
  
autoModel15\_15 <- lm(mpg ~ displacement + weight , data = training)  
summary(autoModel15\_15)

##   
## Call:  
## lm(formula = mpg ~ displacement + weight, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -10.2525 -2.1656 -0.1836 1.8247 14.3691   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 40.216518 0.979180 41.072 < 2e-16 \*\*\*  
## displacement -0.012388 0.004775 -2.594 0.00992 \*\*   
## weight -0.005224 0.000589 -8.870 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.334 on 315 degrees of freedom  
## Multiple R-squared: 0.7561, Adjusted R-squared: 0.7546   
## F-statistic: 488.3 on 2 and 315 DF, p-value: < 2.2e-16

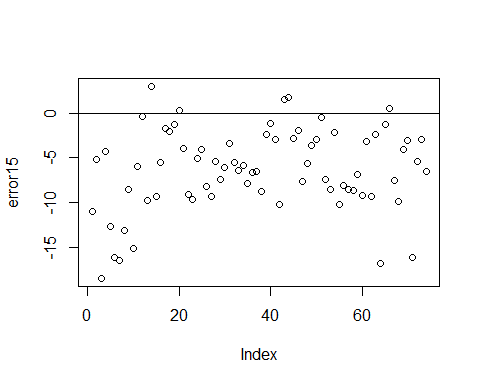
plot(autoModel15\_15)



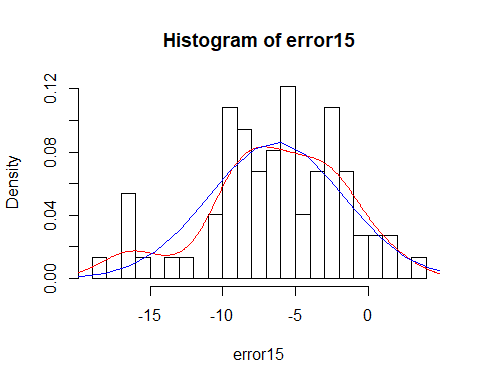
### use model on test dataset ###  
  
B0<-autoModel15\_15$coefficients[1]  
  
B1<-autoModel15\_15$coefficients[2]  
  
B2<-autoModel15\_15$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted15 <- B0 + B1 \* test$displacement + B2 \*test$weight  
  
  
  
  
# #calculating error ##  
  
error15<- y\_predicted15- test$mpg  
  
summary(error15)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -18.473 -8.971 -6.037 -6.326 -2.953 3.006

plot(error15)  
  
abline(0,0)



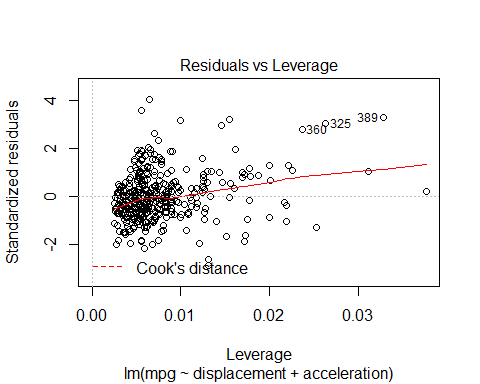
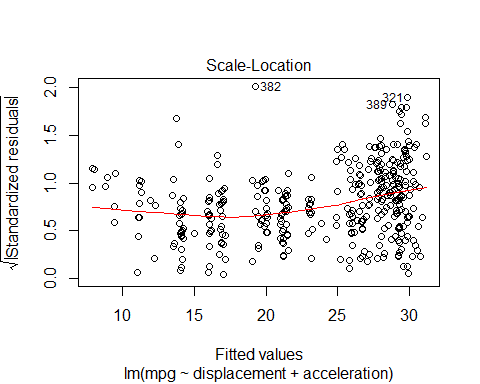
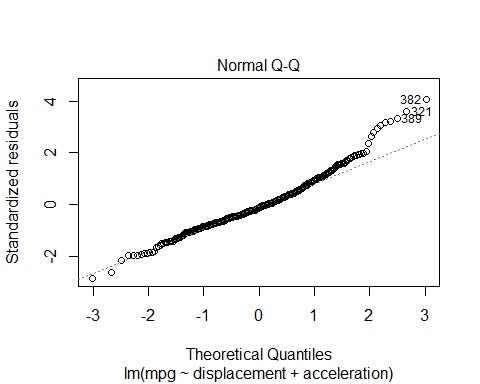
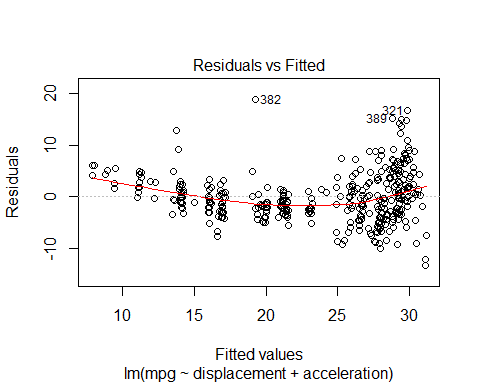
hist(error15,prob=T,breaks = 20)  
  
lines(density(error15),col='red')  
  
  
  
#Normal Curve for automodel15  
  
  
mean\_error15<-mean(error15)  
variance\_error15<-var(error15)  
sd\_error15<-sqrt(variance\_error15)  
x\_error15<-seq(-25,25,length=30)  
y\_error15<-dnorm(x\_error15,mean\_error15,sd\_error15)  
lines(x\_error15,y\_error15,col='blue')



########### Model 16 ##############  
  
autoModel16= lm(mpg ~ displacement + acceleration, data = autompg)  
summary(autoModel16)

##   
## Call:  
## lm(formula = mpg ~ displacement + acceleration, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -13.1398 -3.0403 -0.5378 2.3410 18.7542   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 36.188193 1.922043 18.828 <2e-16 \*\*\*  
## displacement -0.060886 0.002672 -22.791 <2e-16 \*\*\*  
## acceleration -0.058245 0.101331 -0.575 0.566   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.639 on 389 degrees of freedom  
## Multiple R-squared: 0.6485, Adjusted R-squared: 0.6467   
## F-statistic: 358.9 on 2 and 389 DF, p-value: < 2.2e-16

plot(autoModel16)



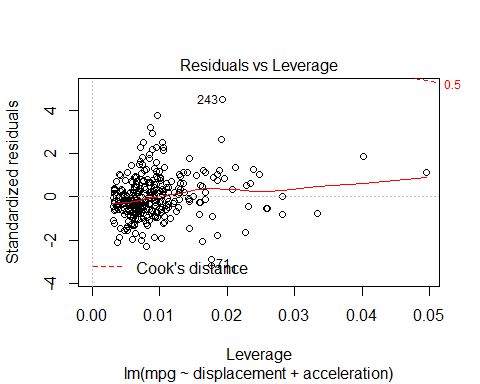
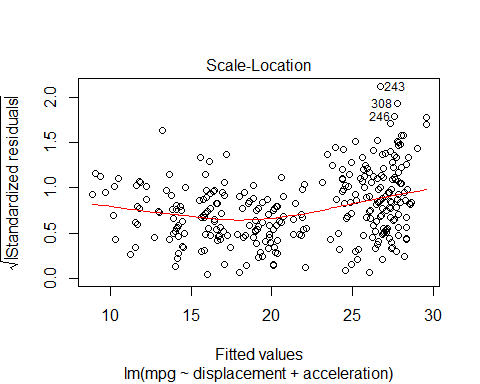
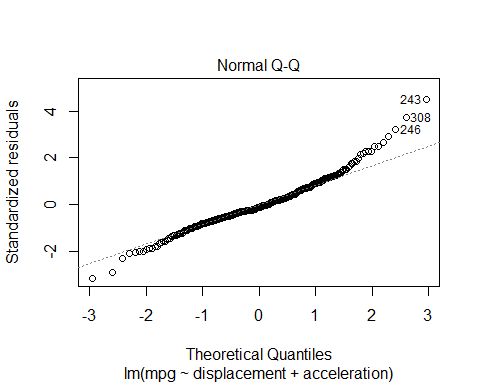
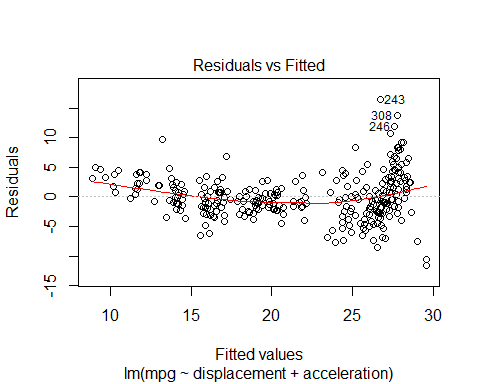
coefficients(autoModel16)

## (Intercept) displacement acceleration   
## 36.18819296 -0.06088649 -0.05824544

###### Regression Model #####  
  
  
autoModel16\_16 <- lm(mpg ~ displacement + acceleration , data = training)  
summary(autoModel16\_16)

##   
## Call:  
## lm(formula = mpg ~ displacement + acceleration, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -11.5400 -2.1652 -0.4009 1.9708 16.4167   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 36.356952 1.758778 20.672 <2e-16 \*\*\*  
## displacement -0.055069 0.002346 -23.472 <2e-16 \*\*\*  
## acceleration -0.219419 0.092473 -2.373 0.0183 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.694 on 315 degrees of freedom  
## Multiple R-squared: 0.7006, Adjusted R-squared: 0.6987   
## F-statistic: 368.5 on 2 and 315 DF, p-value: < 2.2e-16

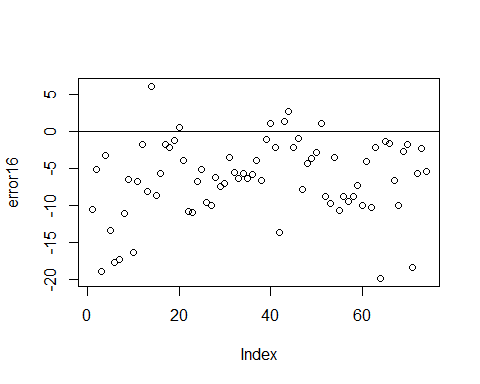
plot(autoModel16\_16)



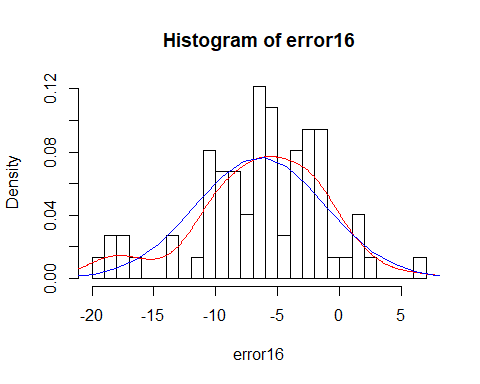
### use model on test dataset ###  
  
B0<-autoModel16\_16$coefficients[1]  
  
B1<-autoModel16\_16$coefficients[2]  
  
B2<-autoModel16\_16$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted16 <- B0 + B1 \* test$displacement + B2 \*test$acceleration  
  
  
  
  
# #calculating error ##  
  
error16<- y\_predicted16- test$mpg  
  
summary(error16)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -19.801 -9.544 -6.066 -6.384 -2.438 6.059

plot(error16)  
  
abline(0,0)



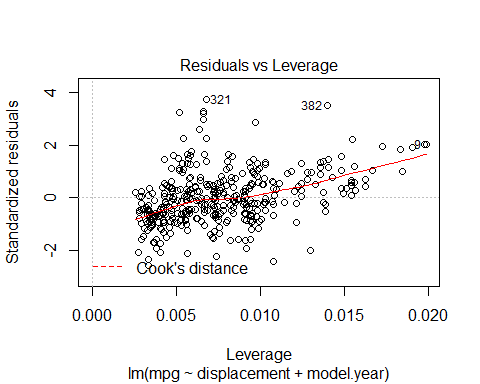
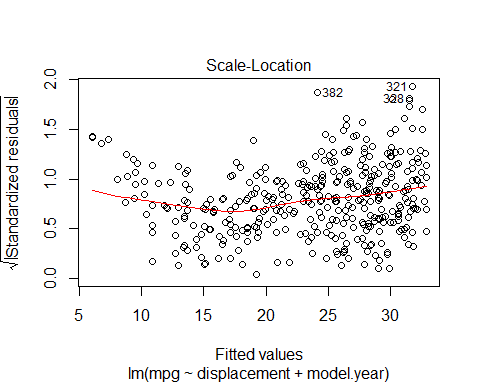
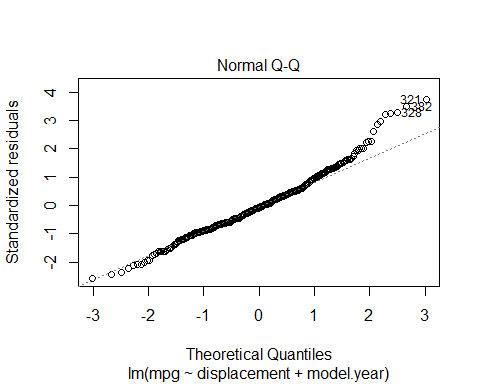
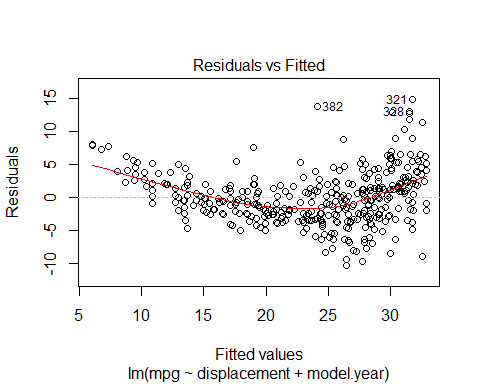
hist(error16,prob=T,breaks = 20)  
  
lines(density(error16),col='red')  
  
  
  
#Normal Curve for automodel16  
  
  
mean\_error16<-mean(error16)  
variance\_error16<-var(error16)  
sd\_error16<-sqrt(variance\_error16)  
x\_error16<-seq(-25,25,length=30)  
y\_error16<-dnorm(x\_error16,mean\_error16,sd\_error16)  
lines(x\_error16,y\_error16,col='blue')



######### Model 17 ##########  
  
  
autoModel17= lm(mpg ~ displacement + model.year, data = autompg)  
summary(autoModel17)

##   
## Call:  
## lm(formula = mpg ~ displacement + model.year, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -10.2507 -2.5546 -0.3074 2.0580 14.8331   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -19.369518 4.640695 -4.174 3.7e-05 \*\*\*  
## displacement -0.051015 0.002072 -24.621 < 2e-16 \*\*\*  
## model.year 0.694047 0.058861 11.791 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.983 on 389 degrees of freedom  
## Multiple R-squared: 0.7409, Adjusted R-squared: 0.7395   
## F-statistic: 556 on 2 and 389 DF, p-value: < 2.2e-16

plot(autoModel17)



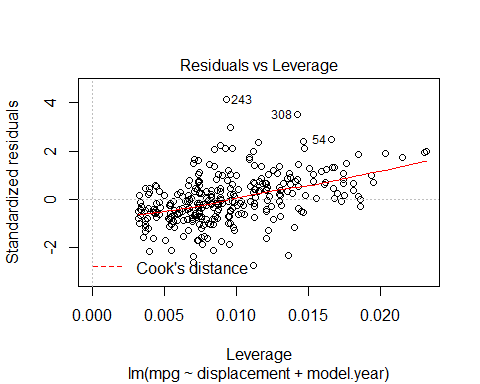
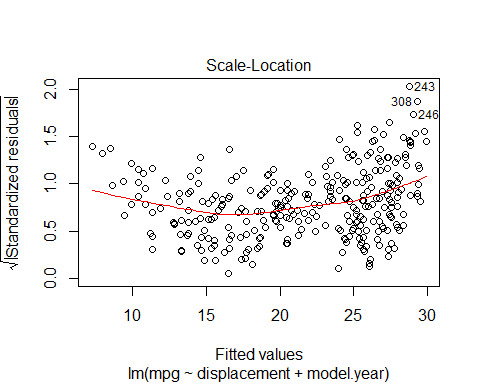
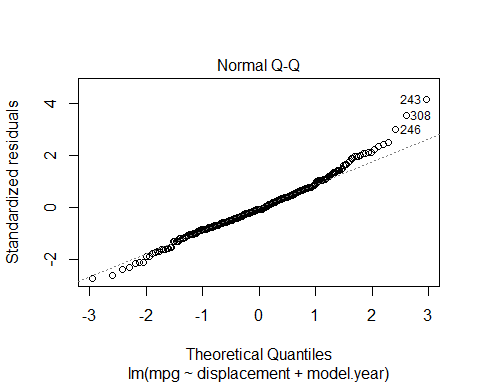
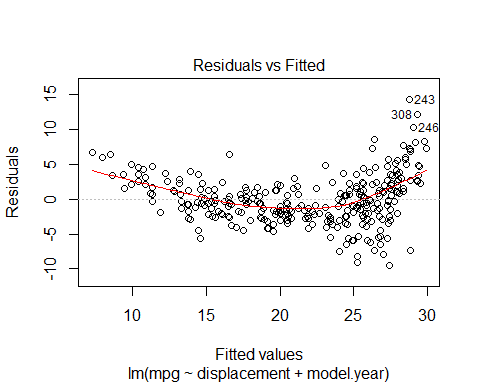
coefficients(autoModel17)

## (Intercept) displacement model.year   
## -19.36951823 -0.05101505 0.69404655

###### Regression Model #####  
  
  
  
autoModel17\_17<- lm(mpg ~ displacement + model.year , data = training)  
summary(autoModel17\_17)

##   
## Call:  
## lm(formula = mpg ~ displacement + model.year, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -9.4518 -2.1511 -0.3291 1.9686 14.3054   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -2.996930 5.179854 -0.579 0.563   
## displacement -0.048845 0.001858 -26.284 < 2e-16 \*\*\*  
## model.year 0.463943 0.067806 6.842 4.07e-11 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.477 on 315 degrees of freedom  
## Multiple R-squared: 0.7346, Adjusted R-squared: 0.733   
## F-statistic: 436 on 2 and 315 DF, p-value: < 2.2e-16

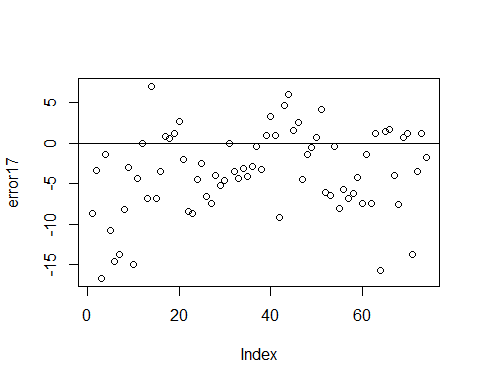
plot(autoModel17\_17)



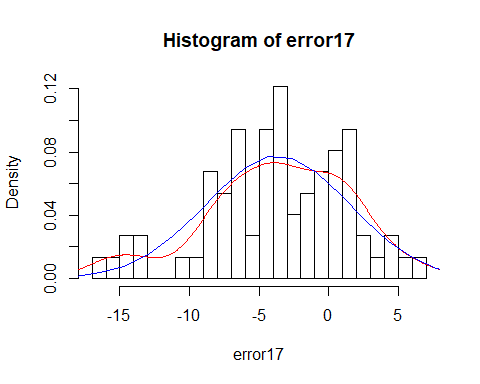
### use model on test dataset ###  
  
B0<-autoModel17\_17$coefficients[1]  
  
B1<-autoModel17\_17$coefficients[2]  
  
B2<-autoModel17\_17$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted17 <- B0 + B1 \* test$displacement + B2 \*test$model.year  
  
  
  
  
# #calculating error ##  
  
error17<- y\_predicted17- test$mpg  
  
summary(error17)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -16.6822 -6.8081 -3.5271 -3.6484 0.6548 6.9994

plot(error17)  
  
abline(0,0)



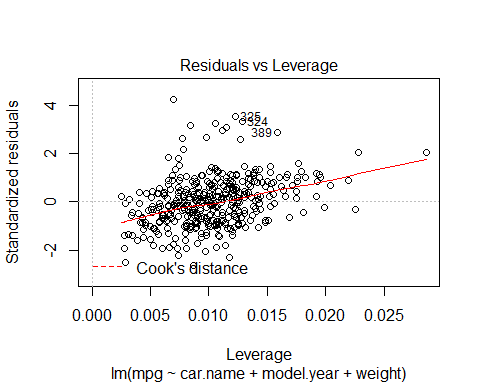
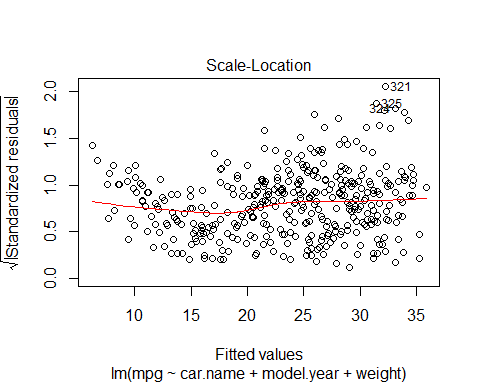
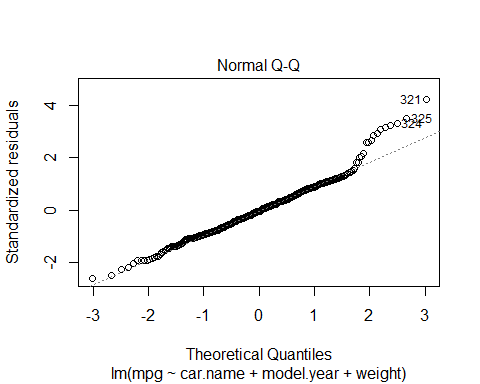
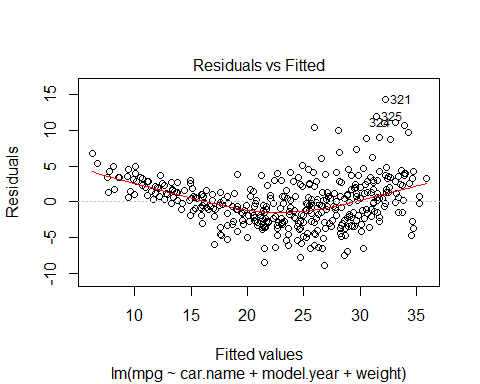
hist(error17,prob=T,breaks = 20)  
  
lines(density(error17),col='red')  
  
  
  
#Normal Curve for automodel17  
  
  
mean\_error17<-mean(error17)  
variance\_error17<-var(error17)  
sd\_error17<-sqrt(variance\_error17)  
x\_error17<-seq(-25,25,length=30)  
y\_error17<-dnorm(x\_error17,mean\_error17,sd\_error17)  
lines(x\_error17,y\_error17,col='blue')



######## Model 18 #############  
  
autoModel18= lm(mpg ~ displacement + origin, data = autompg)  
summary(autoModel8)

##   
## Call:  
## lm(formula = mpg ~ car.name + model.year + weight, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -8.8659 -2.2954 -0.1699 2.0068 14.3290   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.562e+01 4.000e+00 -3.904 0.000112 \*\*\*  
## car.name 5.449e-03 1.983e-03 2.748 0.006269 \*\*   
## model.year 7.576e-01 4.906e-02 15.442 < 2e-16 \*\*\*  
## weight -6.487e-03 2.192e-04 -29.597 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.399 on 388 degrees of freedom  
## Multiple R-squared: 0.8118, Adjusted R-squared: 0.8104   
## F-statistic: 558 on 3 and 388 DF, p-value: < 2.2e-16

plot(autoModel8)



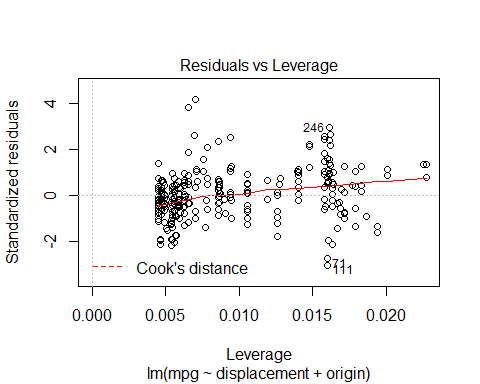
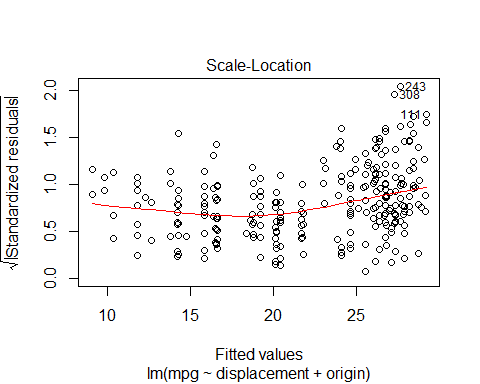
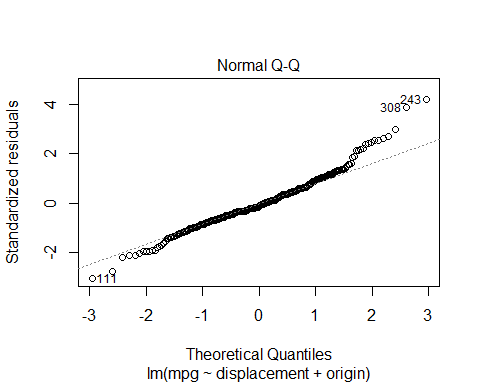
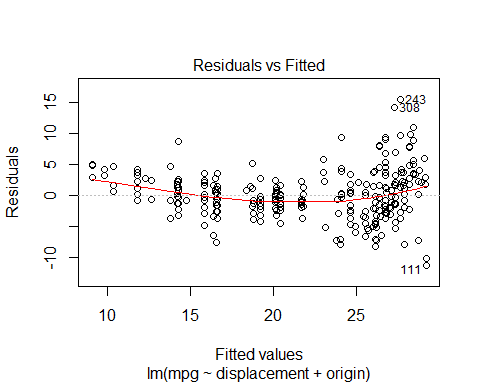
coefficients(autoModel18)

## (Intercept) displacement origin   
## 32.38346312 -0.05486425 1.09653677

###### Regression Model #####  
  
  
autoModel18\_18 <- lm(mpg ~ displacement + origin , data = training)  
summary(autoModel18\_18)

##   
## Call:  
## lm(formula = mpg ~ displacement + origin, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -11.1778 -2.1907 -0.3939 1.8709 15.4535   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 31.013865 0.988924 31.361 <2e-16 \*\*\*  
## displacement -0.049461 0.002511 -19.699 <2e-16 \*\*\*  
## origin 0.542047 0.363042 1.493 0.136   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.714 on 315 degrees of freedom  
## Multiple R-squared: 0.6973, Adjusted R-squared: 0.6954   
## F-statistic: 362.9 on 2 and 315 DF, p-value: < 2.2e-16

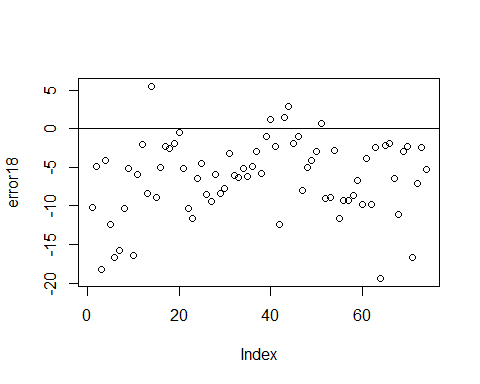
plot(autoModel18\_18)



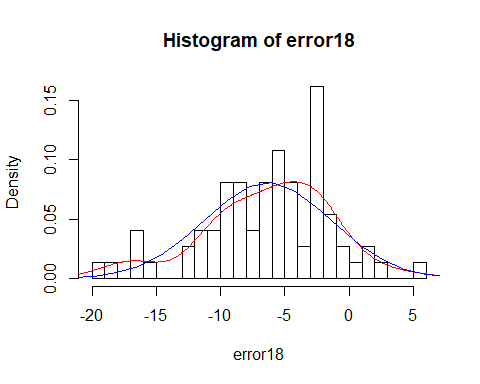
### use model on test dataset ###  
  
B0<-autoModel18\_18$coefficients[1]  
  
B1<-autoModel18\_18$coefficients[2]  
  
B2<-autoModel18\_18$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted18 <- B0 + B1 \* test$displacement + B2 \*test$origin  
  
  
  
  
# #calculating error ##  
  
error18<- y\_predicted18- test$mpg  
  
summary(error18)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -19.403 -9.242 -5.866 -6.321 -2.600 5.478

plot(error18)  
  
abline(0,0)



hist(error18,prob=T,breaks = 20)  
  
lines(density(error18),col='red')  
  
  
  
#Normal Curve for automodel15  
  
  
mean\_error18<-mean(error18)  
variance\_error18<-var(error18)  
sd\_error18<-sqrt(variance\_error18)  
x\_error18<-seq(-25,25,length=30)  
y\_error18<-dnorm(x\_error18,mean\_error18,sd\_error18)  
lines(x\_error18,y\_error18,col='blue')



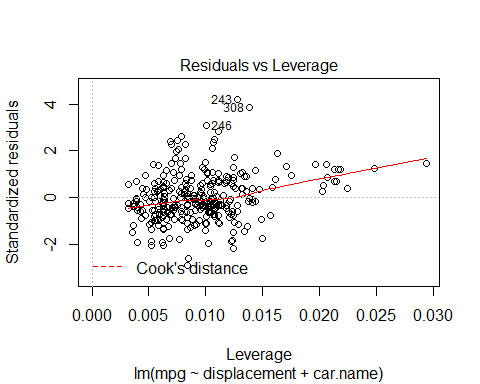
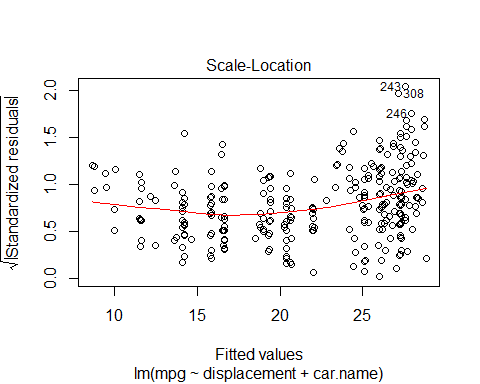
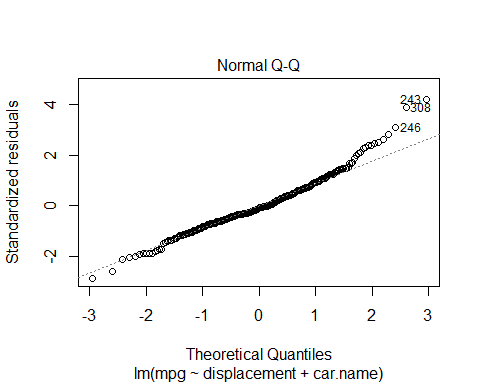
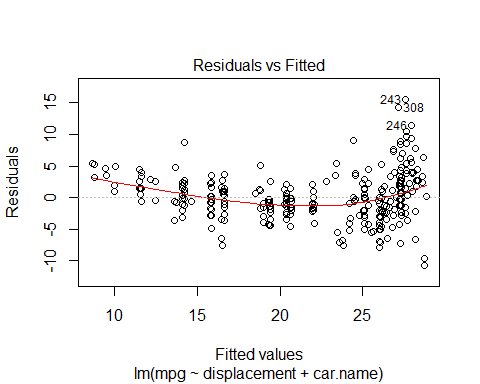
########### Model 19 ################  
  
  
autoModel19= lm(mpg ~ displacement + car.name, data = autompg)  
summary(autoModel19)

##   
## Call:  
## lm(formula = mpg ~ displacement + car.name, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -12.9005 -2.9406 -0.4973 2.3287 18.4023   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 34.503778 0.732510 47.104 <2e-16 \*\*\*  
## displacement -0.059275 0.002340 -25.328 <2e-16 \*\*\*  
## car.name 0.003120 0.002735 1.141 0.255   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.633 on 389 degrees of freedom  
## Multiple R-squared: 0.6494, Adjusted R-squared: 0.6476   
## F-statistic: 360.3 on 2 and 389 DF, p-value: < 2.2e-16

autoModel19\_19 <- lm(mpg ~ displacement + car.name , data = training)  
summary(autoModel19\_19)

##   
## Call:  
## lm(formula = mpg ~ displacement + car.name, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -10.6957 -2.2623 -0.3548 2.1732 15.5083   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 32.4328477 0.6545582 49.549 <2e-16 \*\*\*  
## displacement -0.0519934 0.0020164 -25.786 <2e-16 \*\*\*  
## car.name -0.0005579 0.0024138 -0.231 0.817   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.726 on 315 degrees of freedom  
## Multiple R-squared: 0.6953, Adjusted R-squared: 0.6933   
## F-statistic: 359.3 on 2 and 315 DF, p-value: < 2.2e-16

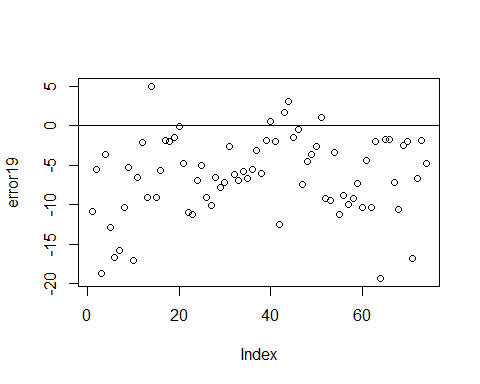
plot(autoModel19\_19)



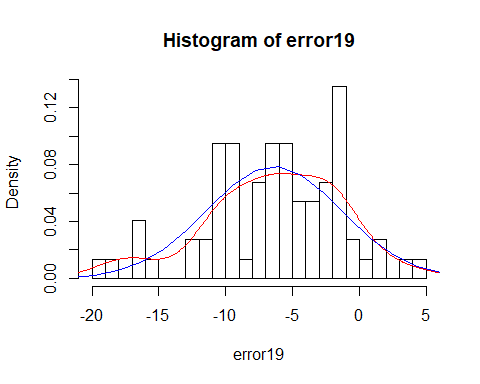
### use model on test dataset ###  
  
B0<-autoModel19\_19$coefficients[1]  
  
B1<-autoModel19\_19$coefficients[2]  
  
B2<-autoModel19\_19$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted19 <- B0 + B1 \* test$displacement + B2 \*test$car.name  
  
  
  
  
# #calculating error ##  
  
error19<- y\_predicted19- test$mpg  
  
summary(error19)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -19.301 -9.355 -6.080 -6.381 -2.225 4.991

plot(error19)  
  
abline(0,0)



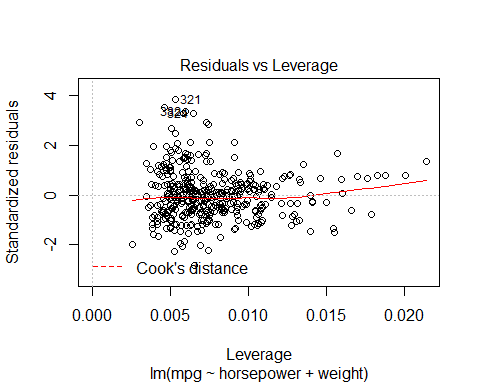
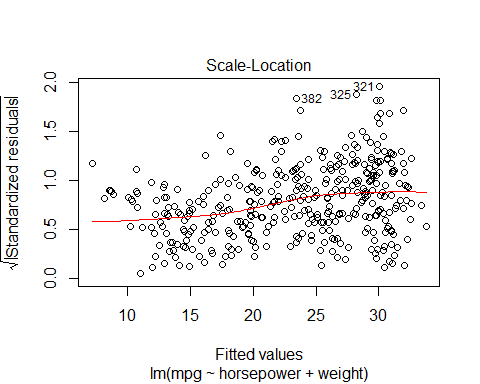
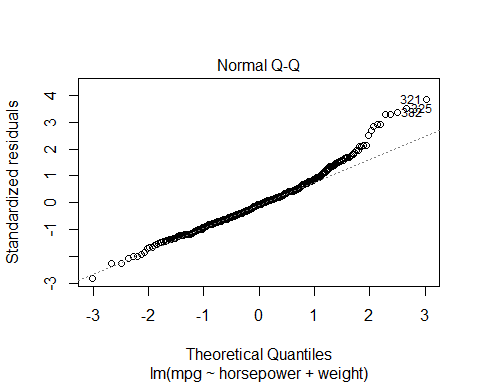
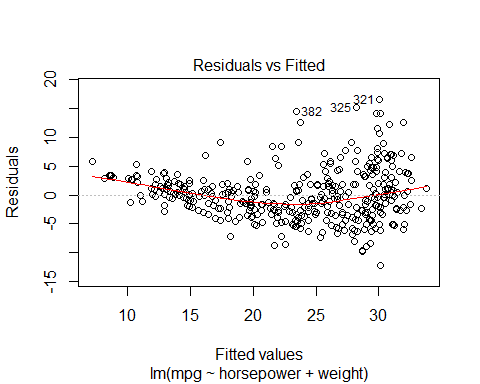
hist(error19,prob=T,breaks = 20)  
  
lines(density(error19),col='red')  
  
  
  
#Normal Curve for automodel19  
  
  
mean\_error19<-mean(error19)  
variance\_error19<-var(error19)  
sd\_error19<-sqrt(variance\_error19)  
x\_error19<-seq(-25,25,length=30)  
y\_error19<-dnorm(x\_error19,mean\_error19,sd\_error19)  
lines(x\_error19,y\_error19,col='blue')



############# Model 20 ##############  
  
  
autoModel20= lm(mpg ~ horsepower + weight, data = autompg)  
summary(autoModel20)

##   
## Call:  
## lm(formula = mpg ~ horsepower + weight, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -12.1550 -2.9020 -0.2561 2.0801 16.5724   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 45.2039502 1.2055267 37.497 <2e-16 \*\*\*  
## horsepower 0.0096676 0.0086237 1.121 0.263   
## weight -0.0074766 0.0002995 -24.965 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.331 on 389 degrees of freedom  
## Multiple R-squared: 0.6936, Adjusted R-squared: 0.692   
## F-statistic: 440.3 on 2 and 389 DF, p-value: < 2.2e-16

plot(autoModel20)



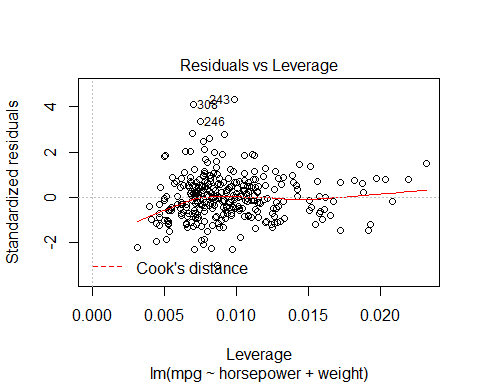
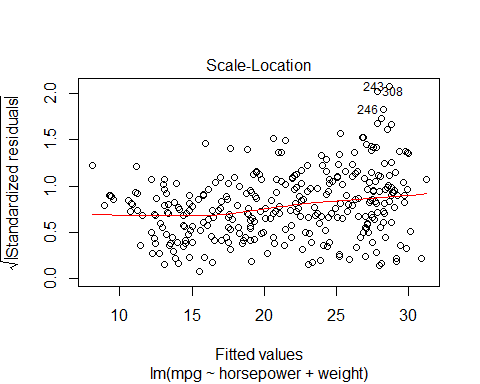
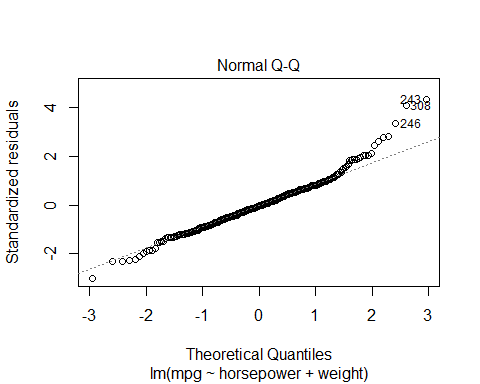
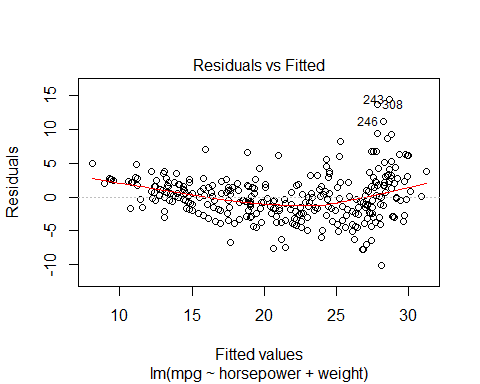
coefficients(autoModel20)

## (Intercept) horsepower weight   
## 45.203950234 0.009667566 -0.007476631

########### Regression Model #########  
  
autoModel20\_20 <- lm(mpg ~ horsepower + weight , data = training)  
summary(autoModel20\_20)

##   
## Call:  
## lm(formula = mpg ~ horsepower + weight, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -10.0796 -2.0537 -0.1372 1.8815 14.4464   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 41.0435823 1.0385086 39.52 <2e-16 \*\*\*  
## horsepower 0.0093445 0.0073592 1.27 0.205   
## weight -0.0064819 0.0002521 -25.71 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.36 on 315 degrees of freedom  
## Multiple R-squared: 0.7522, Adjusted R-squared: 0.7506   
## F-statistic: 478 on 2 and 315 DF, p-value: < 2.2e-16

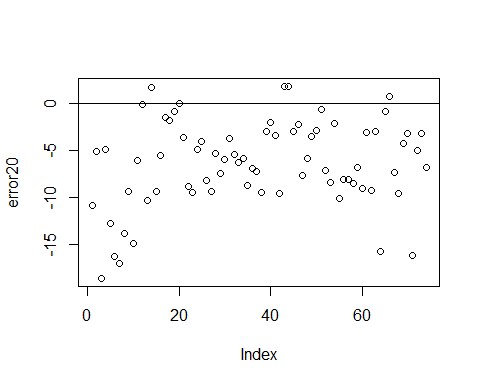
plot(autoModel20\_20)



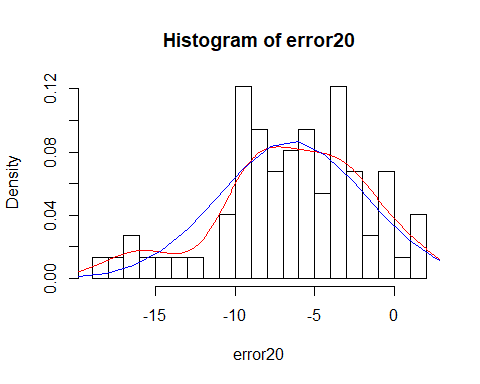
### use model on test dataset ###  
  
B0<-autoModel20\_20$coefficients[1]  
  
B1<-autoModel20\_20$coefficients[2]  
  
B2<-autoModel20\_20$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted20 <- B0 + B1 \* test$horsepower+ B2 \*test$weight  
  
  
  
  
# #calculating error ##  
  
error20<- y\_predicted20- test$mpg  
  
summary(error20)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -18.654 -9.241 -6.047 -6.438 -3.139 1.794

plot(error20)  
  
abline(0,0)



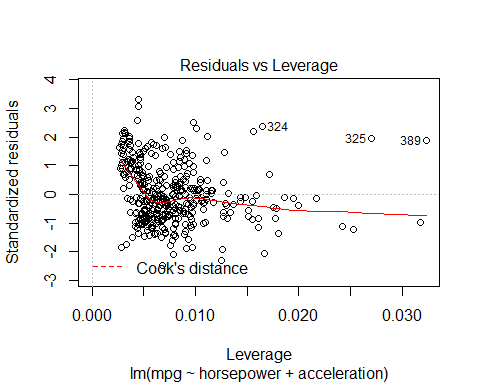
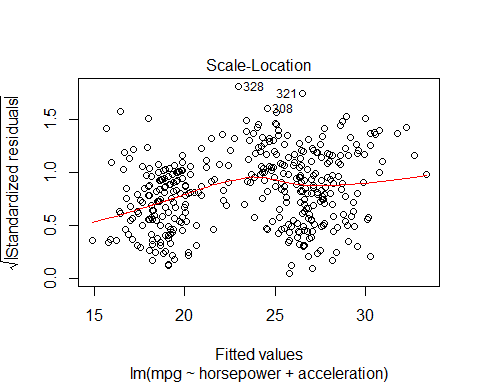
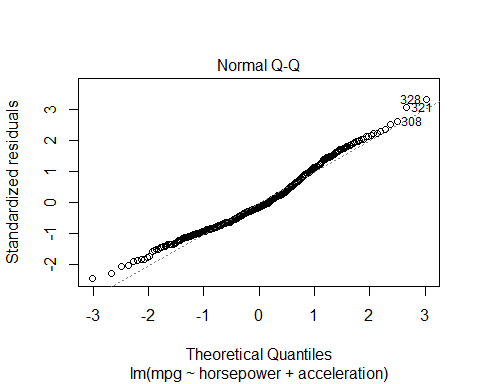
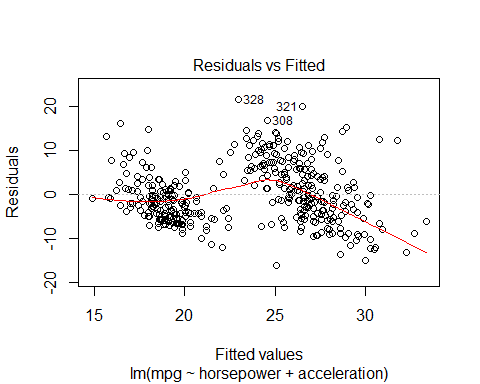
hist(error20,prob=T,breaks = 20)  
  
lines(density(error20),col='red')  
  
  
  
#Normal Curve for automodel20  
  
  
mean\_error20<-mean(error20)  
variance\_error20<-var(error20)  
sd\_error20<-sqrt(variance\_error20)  
x\_error20<-seq(-25,25,length=30)  
y\_error20<-dnorm(x\_error20,mean\_error20,sd\_error20)  
lines(x\_error20,y\_error20,col='blue')



############### Model 21 #####################  
  
  
autoModel21= lm(mpg ~ horsepower + acceleration , data = autompg)  
summary(autoModel21)

##   
## Call:  
## lm(formula = mpg ~ horsepower + acceleration, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -16.081 -4.748 -1.047 4.154 21.615   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.31594 1.89751 2.275 0.0235 \*   
## horsepower 0.09517 0.01171 8.127 5.93e-15 \*\*\*  
## acceleration 0.91151 0.12520 7.280 1.86e-12 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 6.555 on 389 degrees of freedom  
## Multiple R-squared: 0.2983, Adjusted R-squared: 0.2947   
## F-statistic: 82.7 on 2 and 389 DF, p-value: < 2.2e-16

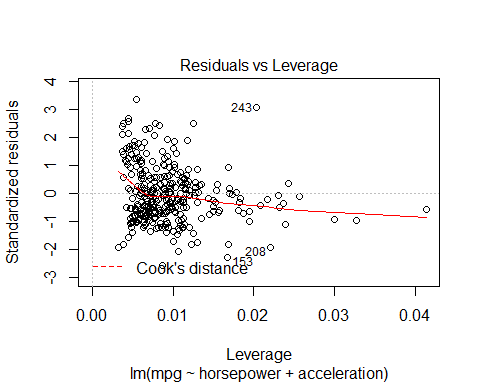
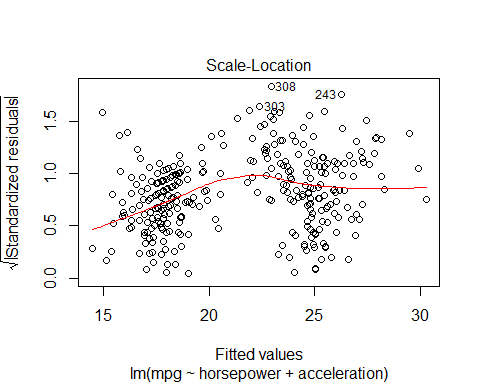
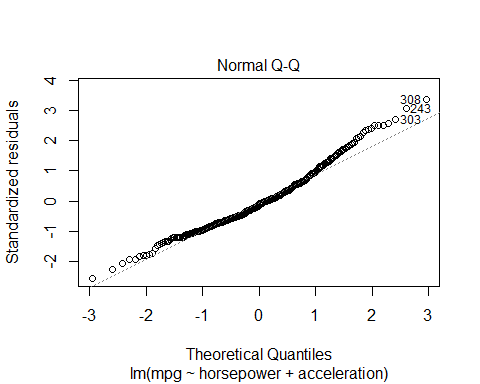
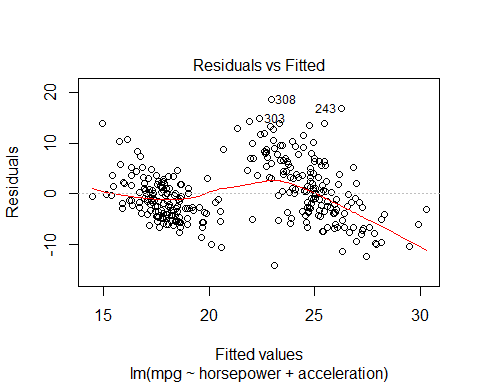
plot(autoModel21)



########## Regression Model ##########  
  
autoModel21\_21<- lm(mpg ~ horsepower + acceleration , data = training)  
summary(autoModel21\_21)

##   
## Call:  
## lm(formula = mpg ~ horsepower + acceleration, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -14.0812 -3.8495 -0.6738 3.1572 18.5685   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 5.33833 1.76983 3.016 0.00277 \*\*   
## horsepower 0.08865 0.01082 8.197 6.29e-15 \*\*\*  
## acceleration 0.76261 0.11840 6.441 4.43e-10 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.56 on 315 degrees of freedom  
## Multiple R-squared: 0.3215, Adjusted R-squared: 0.3172   
## F-statistic: 74.64 on 2 and 315 DF, p-value: < 2.2e-16

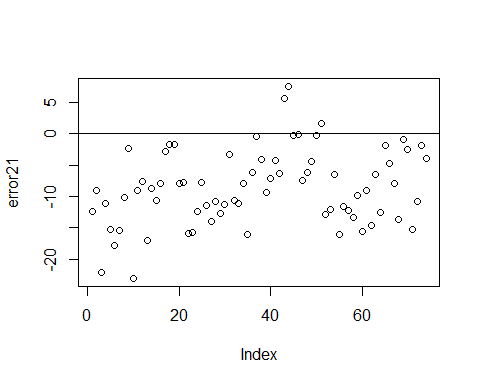
plot(autoModel21\_21)



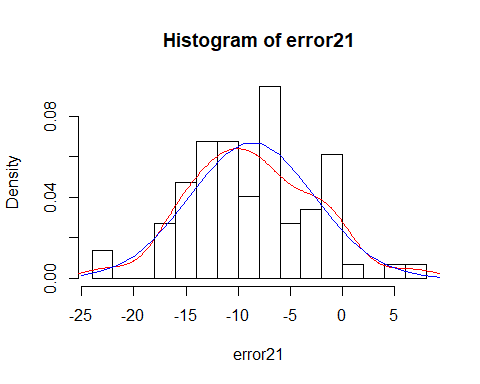
### use model on test dataset ###  
  
B0<-autoModel21\_21$coefficients[1]  
  
B1<-autoModel21\_21$coefficients[2]  
  
B2<-autoModel21\_21$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted21 <- B0 + B1 \* test$horsepower+ B2 \*test$acceleration  
  
  
  
  
# #calculating error ##  
  
error21<- y\_predicted21- test$mpg  
  
summary(error21)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -23.064 -12.495 -8.995 -8.643 -4.319 7.578

plot(error21)  
  
abline(0,0)



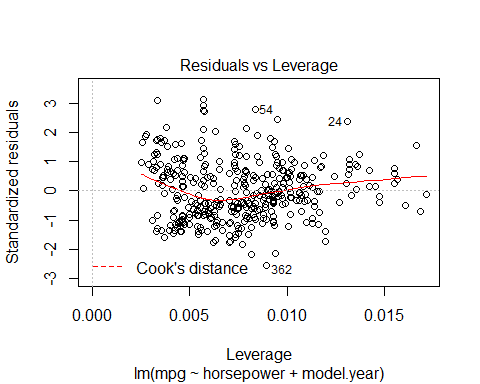
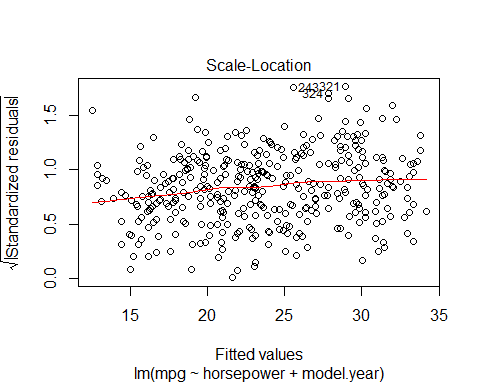
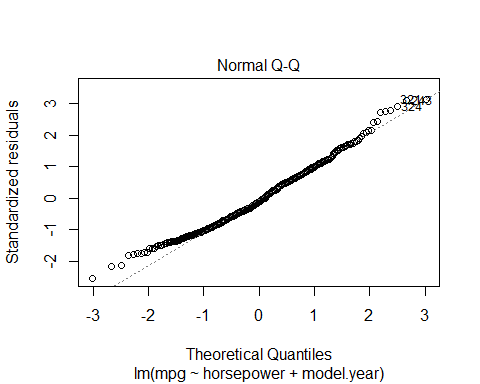
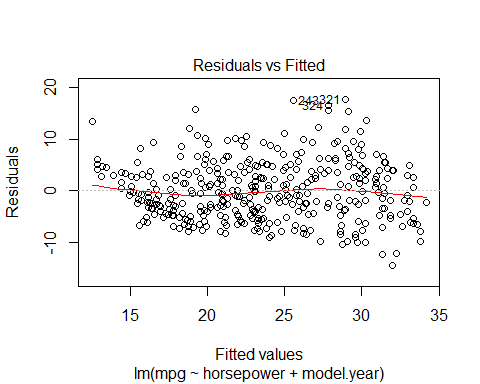
hist(error21,prob=T,breaks = 20)  
  
lines(density(error21),col='red')  
  
  
  
#Normal Curve for automodel21  
  
  
mean\_error21<-mean(error21)  
variance\_error21<-var(error21)  
sd\_error21<-sqrt(variance\_error21)  
x\_error21<-seq(-25,25,length=30)  
y\_error21<-dnorm(x\_error21,mean\_error21,sd\_error21)  
lines(x\_error21,y\_error21,col='blue')



############### Model 22 #################   
  
  
  
autoModel22= lm(mpg ~ horsepower + model.year , data = autompg)  
summary(autoModel22)

##   
## Call:  
## lm(formula = mpg ~ horsepower + model.year, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -14.3458 -4.2946 -0.5907 3.7365 17.6705   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -66.829941 5.919937 -11.29 <2e-16 \*\*\*  
## horsepower 0.099825 0.009799 10.19 <2e-16 \*\*\*  
## model.year 1.119629 0.078467 14.27 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.661 on 389 degrees of freedom  
## Multiple R-squared: 0.4766, Adjusted R-squared: 0.474   
## F-statistic: 177.1 on 2 and 389 DF, p-value: < 2.2e-16

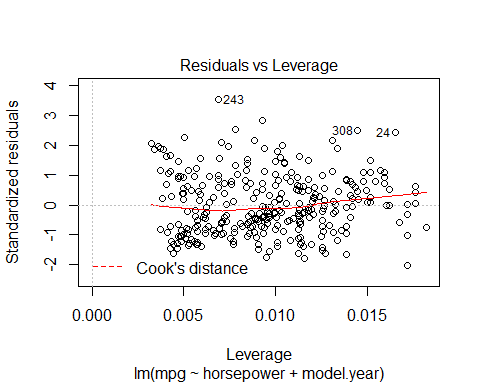
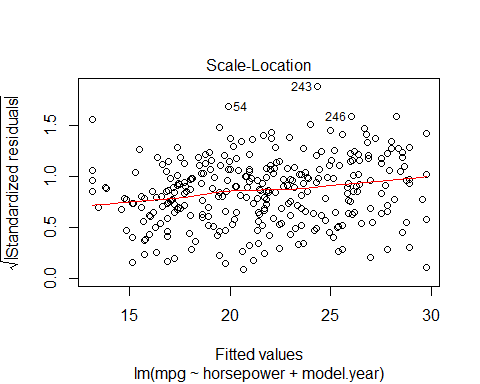
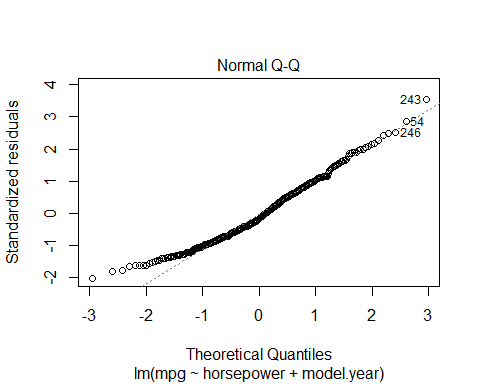
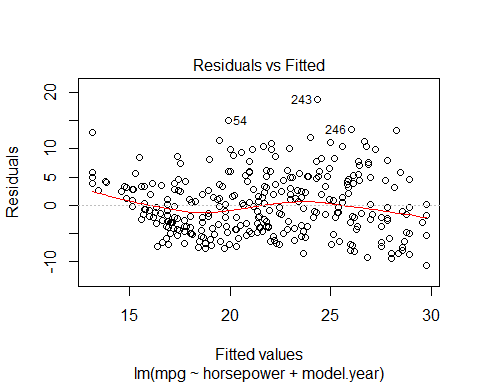
plot(autoModel22)



######### regression model ######  
  
  
autoModel22\_22 <- lm(mpg ~ horsepower + model.year , data = training)  
summary(autoModel22\_22)

##   
## Call:  
## lm(formula = mpg ~ horsepower + model.year, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -10.6450 -4.0245 -0.9147 3.7218 18.7712   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -48.034511 7.568762 -6.346 7.67e-10 \*\*\*  
## horsepower 0.105684 0.009977 10.592 < 2e-16 \*\*\*  
## model.year 0.858634 0.101108 8.492 8.13e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.336 on 315 degrees of freedom  
## Multiple R-squared: 0.3752, Adjusted R-squared: 0.3713   
## F-statistic: 94.59 on 2 and 315 DF, p-value: < 2.2e-16

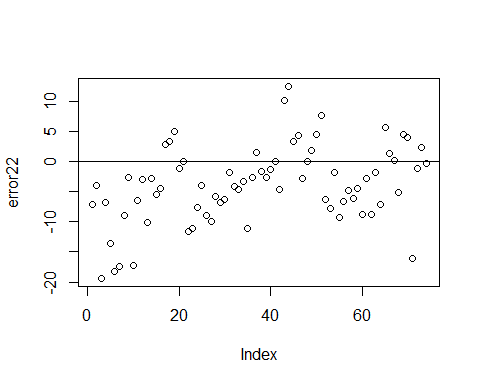
plot(autoModel22\_22)



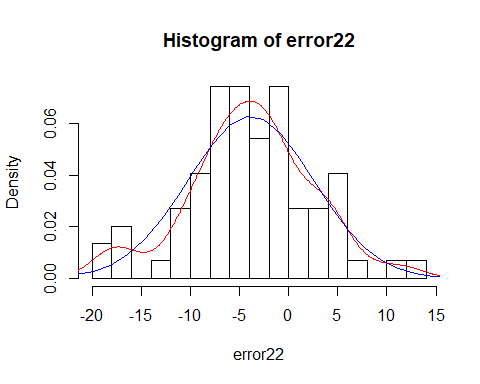
### use model on test dataset ###  
  
B0<-autoModel22\_22$coefficients[1]  
  
B1<-autoModel22\_22$coefficients[2]  
  
B2<-autoModel22\_22$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted22 <- B0 + B1 \* test$horsepower+ B2 \*test$model.year  
  
  
  
  
# #calculating error ##  
  
error22<- y\_predicted22- test$mpg  
  
summary(error22)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -19.39138 -7.06049 -4.03080 -3.86286 -0.04628 12.47524

plot(error22)  
  
abline(0,0)



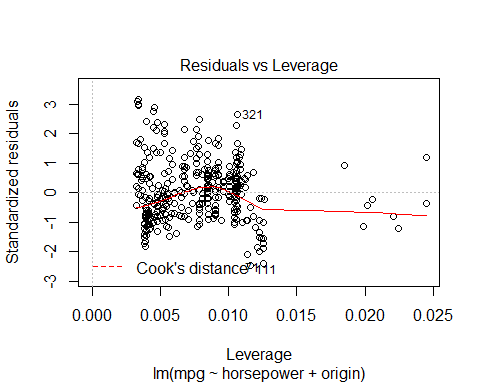
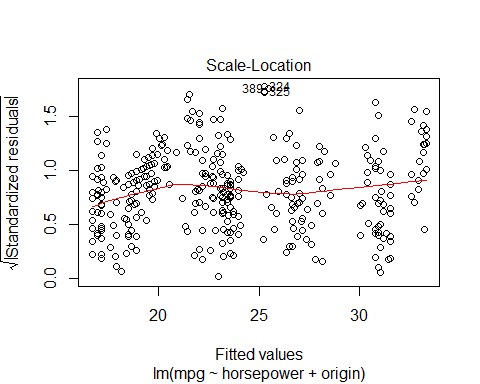
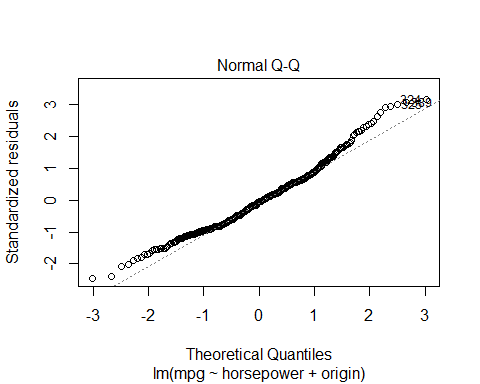
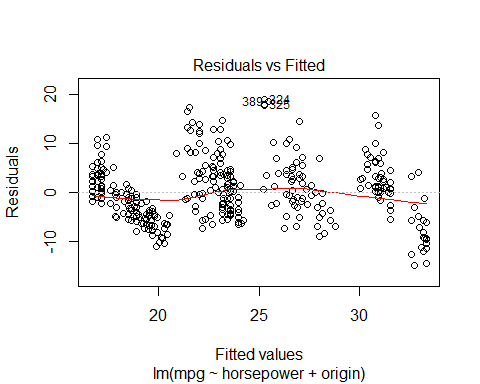
hist(error22,prob=T,breaks = 22)  
  
lines(density(error22),col='red')  
  
  
  
#Normal Curve for automodel22  
  
  
mean\_error22<-mean(error22)  
variance\_error22<-var(error22)  
sd\_error22<-sqrt(variance\_error22)  
x\_error22<-seq(-25,25,length=30)  
y\_error22<-dnorm(x\_error22,mean\_error22,sd\_error22)  
lines(x\_error22,y\_error22,col='blue')



############### Model 23 ##################  
  
  
autoModel23= lm(mpg ~ horsepower + origin , data = autompg)  
summary(autoModel23)

##   
## Call:  
## lm(formula = mpg ~ horsepower + origin, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -14.734 -4.629 -0.299 3.430 19.009   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 11.95959 0.76925 15.547 < 2e-16 \*\*\*  
## horsepower 0.08173 0.01083 7.545 3.23e-13 \*\*\*  
## origin 4.58173 0.39666 11.551 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 6.029 on 389 degrees of freedom  
## Multiple R-squared: 0.4063, Adjusted R-squared: 0.4033   
## F-statistic: 133.1 on 2 and 389 DF, p-value: < 2.2e-16

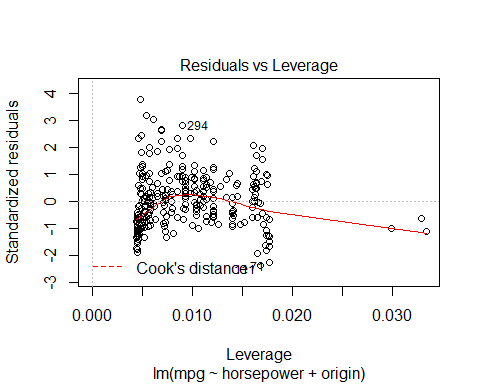
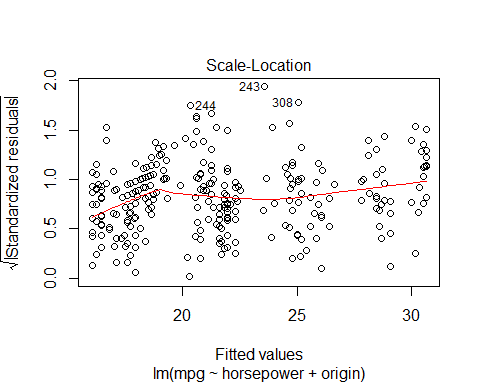
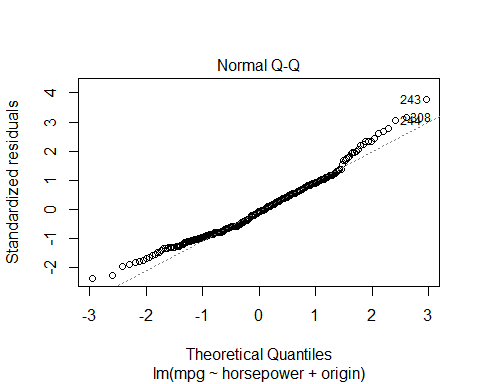
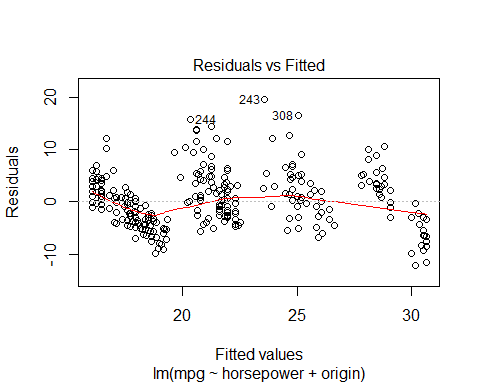
plot(autoModel23)



##### Regression Model ##########  
  
autoModel23\_23 <- lm(mpg ~ horsepower + origin , data = training)  
summary(autoModel23\_23)

##   
## Call:  
## lm(formula = mpg ~ horsepower + origin, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -12.1321 -3.9600 -0.3323 3.1620 19.4991   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 11.87176 0.70973 16.727 < 2e-16 \*\*\*  
## horsepower 0.07017 0.01045 6.715 8.76e-11 \*\*\*  
## origin 4.07521 0.41963 9.711 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.189 on 315 degrees of freedom  
## Multiple R-squared: 0.4091, Adjusted R-squared: 0.4053   
## F-statistic: 109 on 2 and 315 DF, p-value: < 2.2e-16

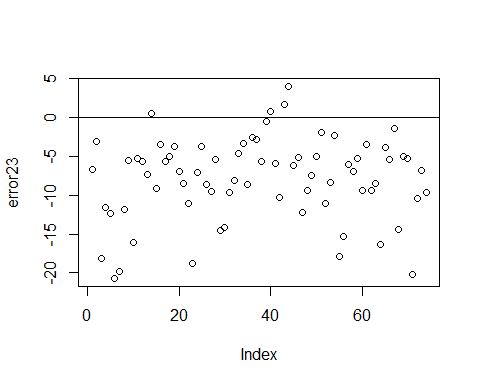
plot(autoModel23\_23)



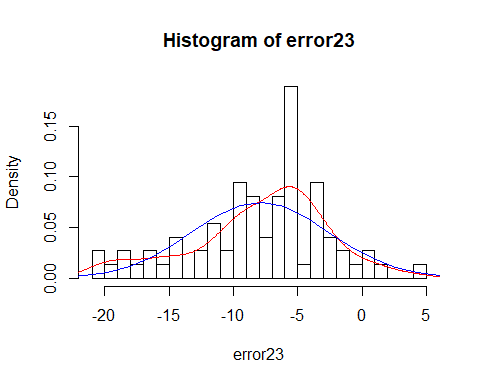
### use model on test dataset ###  
  
B0<-autoModel23\_23$coefficients[1]  
  
B1<-autoModel23\_23$coefficients[2]  
  
B2<-autoModel23\_23$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted23 <- B0 + B1 \* test$horsepower+ B2 \*test$origin  
  
  
# #calculating error ##  
  
error23<- y\_predicted23- test$mpg  
  
summary(error23)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -20.699 -10.405 -6.956 -7.889 -5.018 4.031

plot(error23)  
  
abline(0,0)



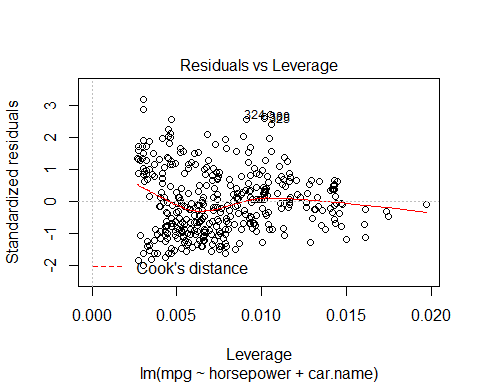
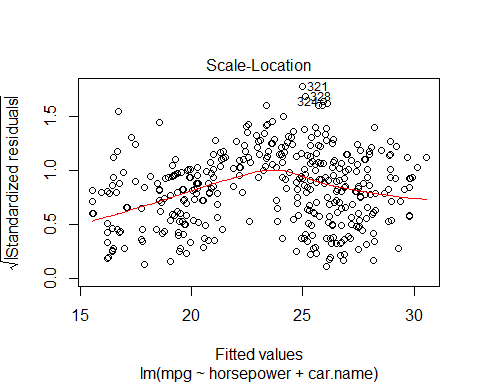
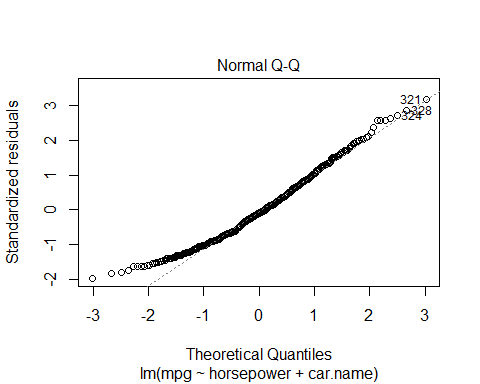
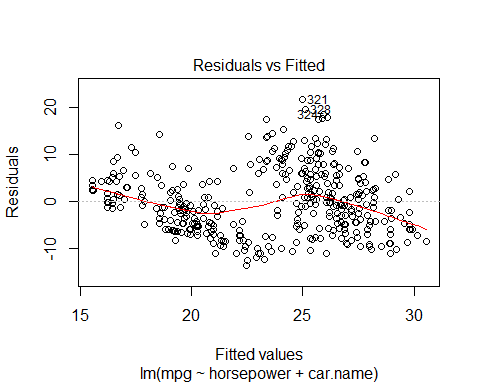
hist(error23,prob=T,breaks = 20)  
  
lines(density(error23),col='red')  
  
  
  
#Normal Curve for automodel23  
  
  
mean\_error23<-mean(error23)  
variance\_error23<-var(error23)  
sd\_error23<-sqrt(variance\_error23)  
x\_error23<-seq(-25,25,length=30)  
y\_error23<-dnorm(x\_error23,mean\_error23,sd\_error23)  
lines(x\_error23,y\_error23,col='blue')



################## Model 24 #################  
  
  
autoModel24= lm(mpg ~ horsepower + car.name , data = autompg)  
summary(autoModel24)

##   
## Call:  
## lm(formula = mpg ~ horsepower + car.name, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -13.4645 -5.2850 -0.6943 4.4648 21.6022   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 15.173578 0.847421 17.906 < 2e-16 \*\*\*  
## horsepower 0.110162 0.011884 9.270 < 2e-16 \*\*\*  
## car.name 0.016916 0.003914 4.322 1.97e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 6.825 on 389 degrees of freedom  
## Multiple R-squared: 0.2393, Adjusted R-squared: 0.2354   
## F-statistic: 61.17 on 2 and 389 DF, p-value: < 2.2e-16

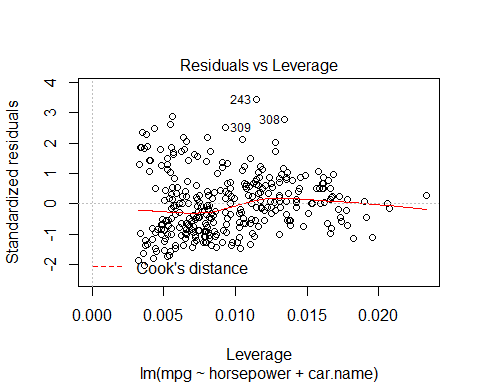
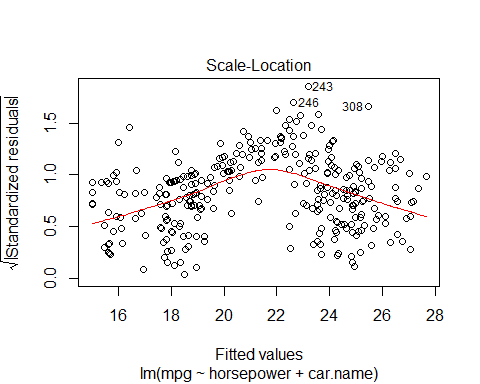
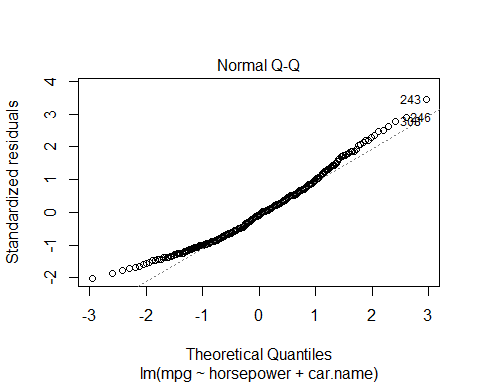
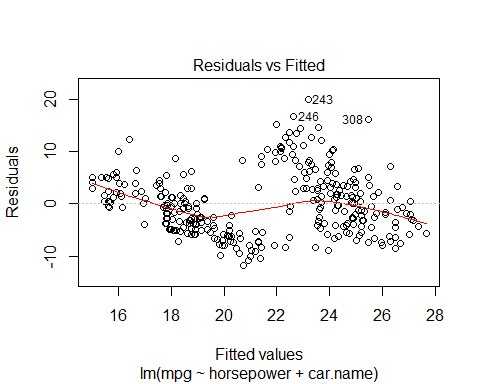
plot(autoModel24)



##### Regression Model ##########  
  
autoModel24\_24 <- lm(mpg ~ horsepower + car.name , data = training)  
summary(autoModel24\_24)

##   
## Call:  
## lm(formula = mpg ~ horsepower + car.name, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -11.7434 -4.4841 -0.5211 3.4456 19.9237   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 14.717434 0.770676 19.097 < 2e-16 \*\*\*  
## horsepower 0.102177 0.011059 9.239 < 2e-16 \*\*\*  
## car.name 0.011200 0.003677 3.046 0.00252 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.83 on 315 degrees of freedom  
## Multiple R-squared: 0.2541, Adjusted R-squared: 0.2494   
## F-statistic: 53.67 on 2 and 315 DF, p-value: < 2.2e-16

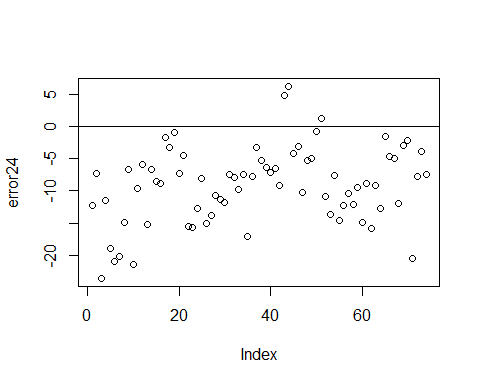
plot(autoModel24\_24)



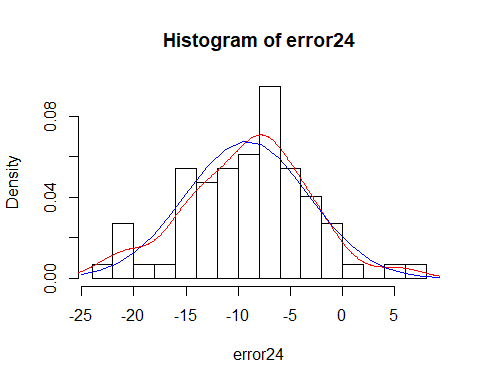
### use model on test dataset ###  
  
B0<-autoModel24\_24$coefficients[1]  
  
B1<-autoModel24\_24$coefficients[2]  
  
B2<-autoModel24\_24$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted24 <- B0 + B1 \* test$horsepower+ B2 \*test$car.name  
  
  
# #calculating error ##  
  
error24<- y\_predicted24- test$mpg  
  
summary(error24)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -23.565 -12.570 -8.669 -9.145 -5.246 6.211

plot(error24)  
  
abline(0,0)



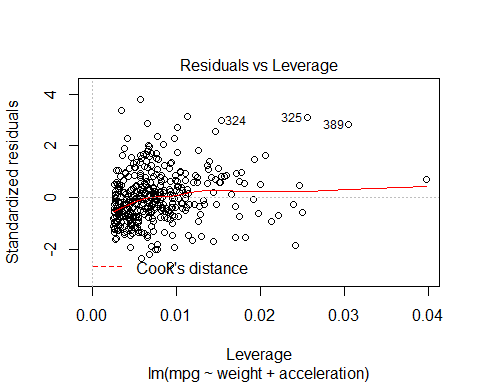
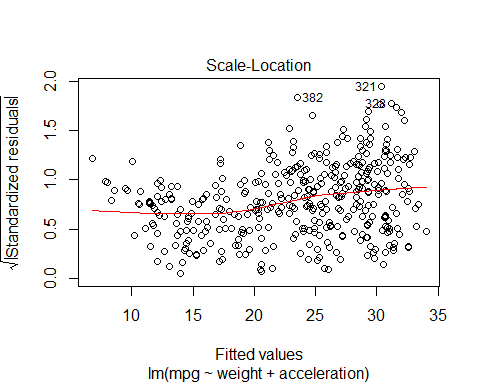
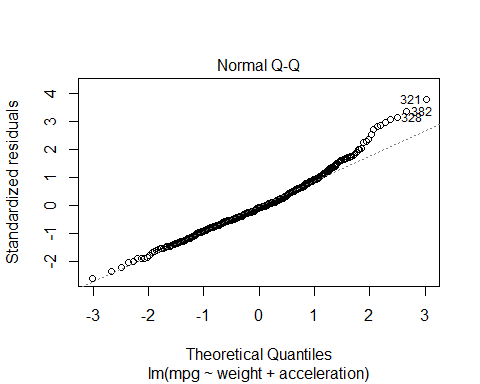
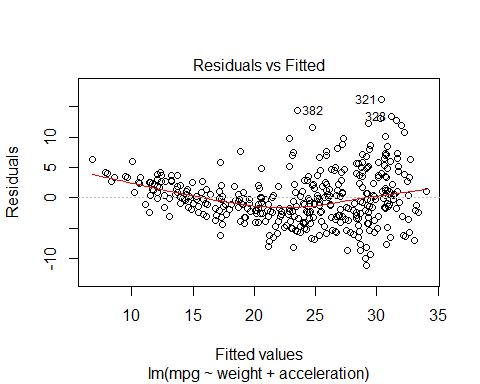
hist(error24,prob=T,breaks = 20)  
  
lines(density(error24),col='red')  
  
  
  
#Normal Curve for automodel24  
  
  
mean\_error24<-mean(error24)  
variance\_error24<-var(error24)  
sd\_error24<-sqrt(variance\_error24)  
x\_error24<-seq(-25,25,length=30)  
y\_error24<-dnorm(x\_error24,mean\_error24,sd\_error24)  
lines(x\_error24,y\_error24,col='blue')



################ Model 25 #################  
  
  
autoModel25= lm(mpg ~ weight + acceleration , data = autompg)  
summary(autoModel25)

##   
## Call:  
## lm(formula = mpg ~ weight + acceleration, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -11.1371 -2.7860 -0.3355 2.4192 16.2096   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 41.0953288 1.8680355 21.999 < 2e-16 \*\*\*  
## weight -0.0072931 0.0002809 -25.966 < 2e-16 \*\*\*  
## acceleration 0.2616504 0.0864755 3.026 0.00265 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.288 on 389 degrees of freedom  
## Multiple R-squared: 0.6997, Adjusted R-squared: 0.6982   
## F-statistic: 453.2 on 2 and 389 DF, p-value: < 2.2e-16

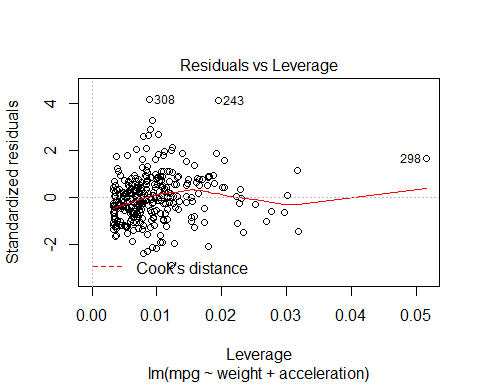
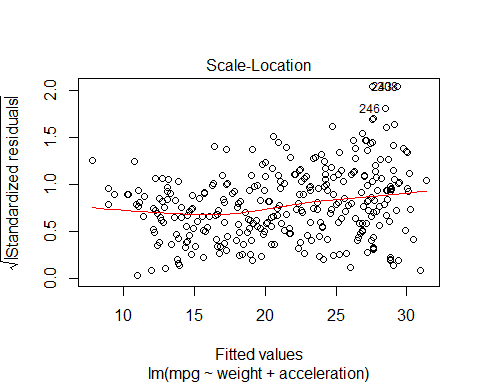
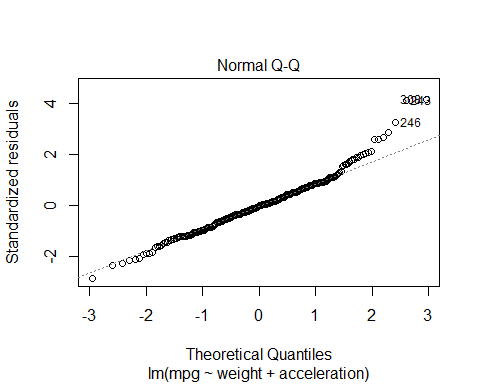
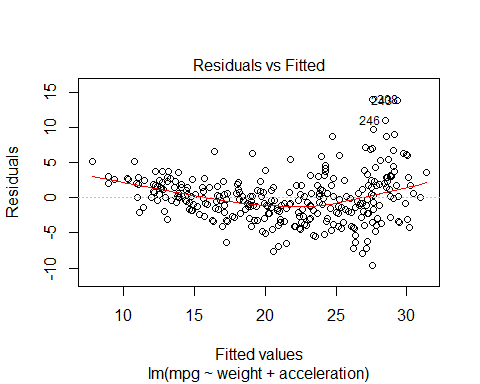
plot(autoModel25)



##### Regression Model ##########  
  
autoModel25\_25 <- lm(mpg ~ weight + acceleration , data = training)  
summary(autoModel25\_25)

##   
## Call:  
## lm(formula = mpg ~ weight + acceleration, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -9.574 -2.131 -0.053 1.836 13.931   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 39.9795733 1.6591217 24.097 <2e-16 \*\*\*  
## weight -0.0065032 0.0002404 -27.057 <2e-16 \*\*\*  
## acceleration 0.1042437 0.0768013 1.357 0.176   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.359 on 315 degrees of freedom  
## Multiple R-squared: 0.7524, Adjusted R-squared: 0.7508   
## F-statistic: 478.5 on 2 and 315 DF, p-value: < 2.2e-16

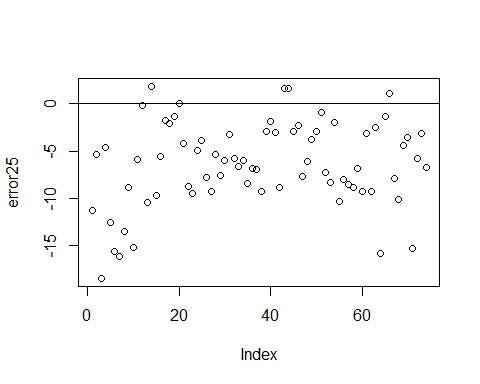
plot(autoModel25\_25)



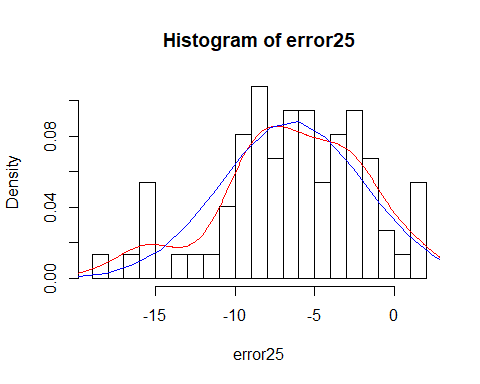
### use model on test dataset ###  
  
B0<-autoModel25\_25$coefficients[1]  
  
B1<-autoModel25\_25$coefficients[2]  
  
B2<-autoModel25\_25$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted25 <- B0 + B1 \* test$weight + B2 \*test$acceleration  
  
  
# #calculating error ##  
  
error25<- y\_predicted25- test$mpg  
  
summary(error25)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -18.476 -8.879 -6.054 -6.410 -3.024 1.845

plot(error25)  
  
abline(0,0)



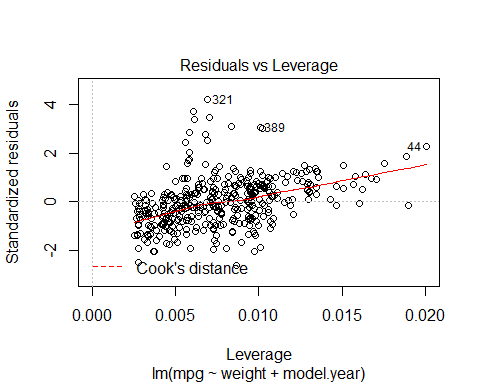
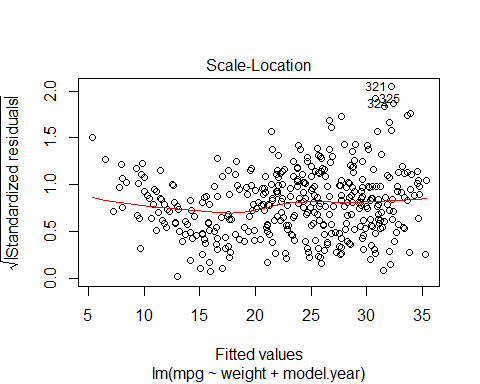
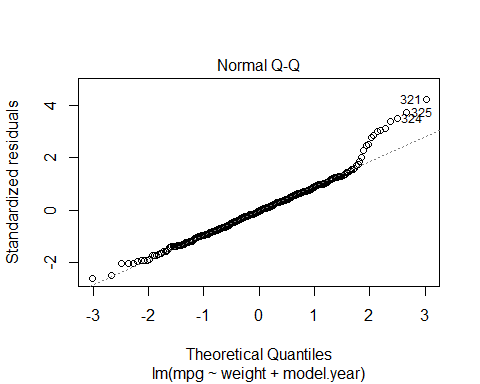
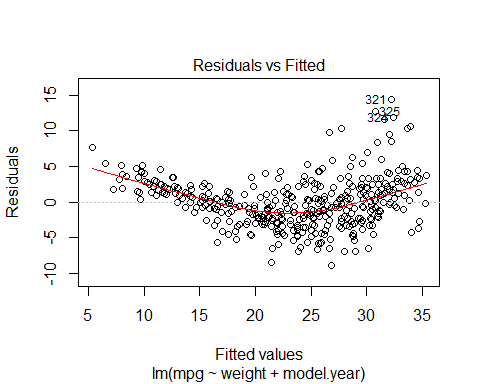
hist(error25,prob=T,breaks = 20)  
  
lines(density(error25),col='red')  
  
  
  
#Normal Curve for automodel25  
  
  
mean\_error25<-mean(error25)  
variance\_error25<-var(error25)  
sd\_error25<-sqrt(variance\_error25)  
x\_error25<-seq(-25,25,length=30)  
y\_error25<-dnorm(x\_error25,mean\_error25,sd\_error25)  
lines(x\_error25,y\_error25,col='blue')



############# Model 26 ################  
  
  
autoModel26= lm(mpg ~ weight + model.year , data = autompg)  
summary(autoModel26)

##   
## Call:  
## lm(formula = mpg ~ weight + model.year, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -8.8505 -2.3014 -0.1167 2.0367 14.3555   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.435e+01 4.007e+00 -3.581 0.000386 \*\*\*  
## weight -6.632e-03 2.146e-04 -30.911 < 2e-16 \*\*\*  
## model.year 7.573e-01 4.947e-02 15.308 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.427 on 389 degrees of freedom  
## Multiple R-squared: 0.8082, Adjusted R-squared: 0.8072   
## F-statistic: 819.5 on 2 and 389 DF, p-value: < 2.2e-16

plot(autoModel26)



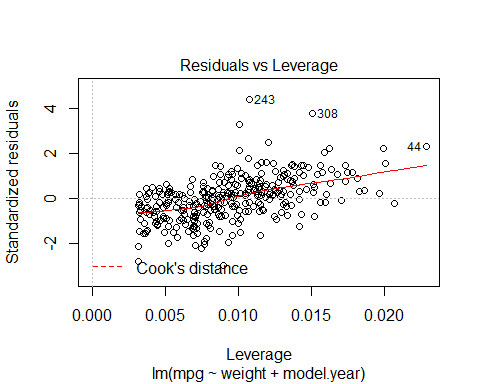
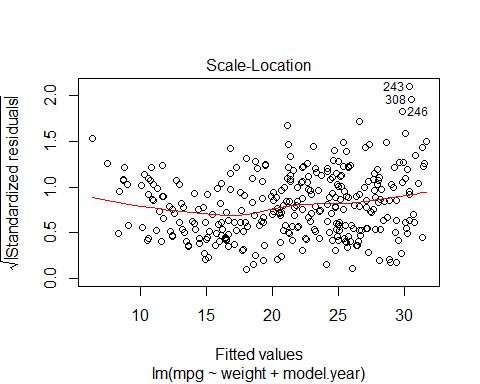
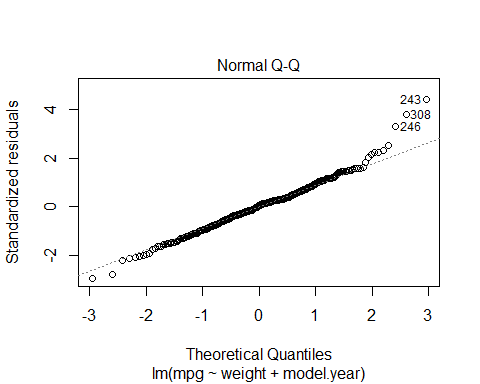
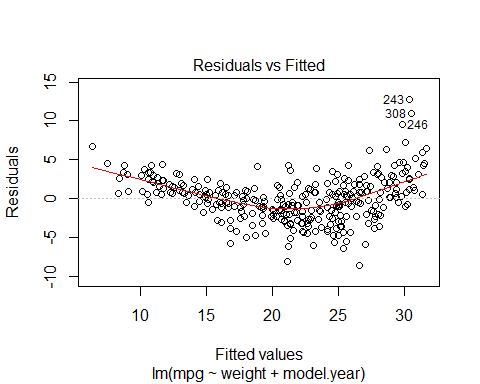
coefficients(autoModel26)

## (Intercept) weight model.year   
## -14.347253018 -0.006632075 0.757318281

##### Regression Model ##########  
  
autoModel26\_26 <- lm(mpg ~ weight + model.year , data = training)  
summary(autoModel26\_26)

##   
## Call:  
## lm(formula = mpg ~ weight + model.year, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -8.5611 -1.8108 0.1019 1.6390 12.7390   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -2.6280483 4.2986403 -0.611 0.541   
## weight -0.0063294 0.0001882 -33.628 <2e-16 \*\*\*  
## model.year 0.5840110 0.0556736 10.490 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.9 on 315 degrees of freedom  
## Multiple R-squared: 0.8154, Adjusted R-squared: 0.8142   
## F-statistic: 695.7 on 2 and 315 DF, p-value: < 2.2e-16

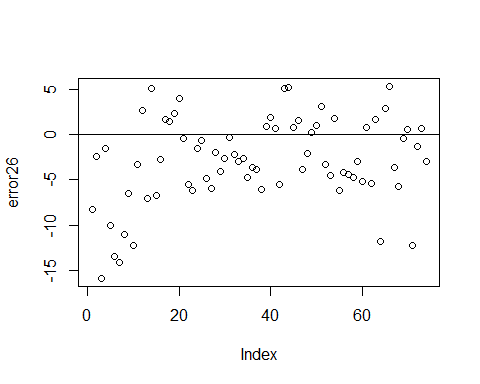
plot(autoModel26\_26)



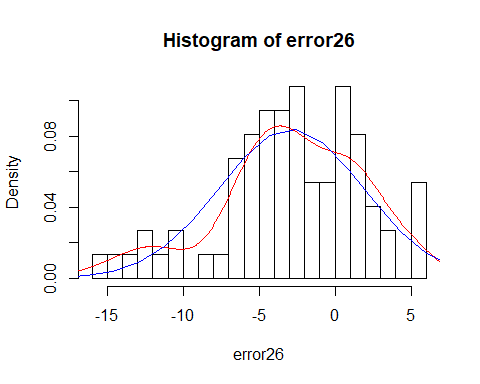
### use model on test dataset ###  
  
B0<-autoModel26\_26$coefficients[1]  
  
B1<-autoModel26\_26$coefficients[2]  
  
B2<-autoModel26\_26$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted26 <- B0 + B1 \* test$weight + B2 \*test$model.year  
  
  
# #calculating error ##  
  
error26<- y\_predicted26- test$mpg  
  
summary(error26)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -15.8622 -5.4667 -2.8750 -2.8994 0.7451 5.3170

plot(error26)  
  
abline(0,0)



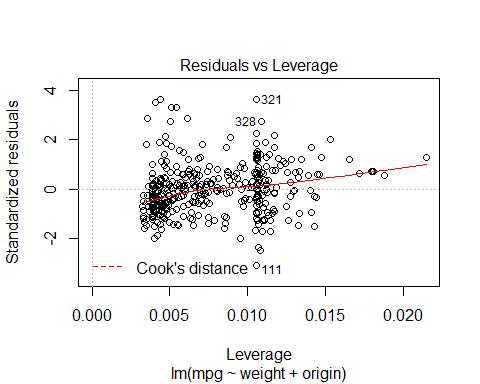
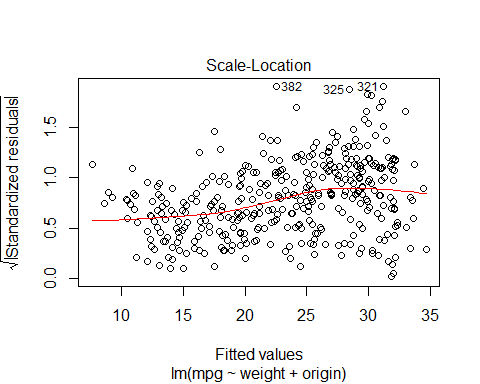
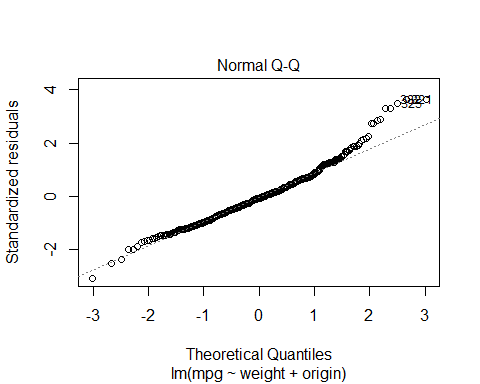
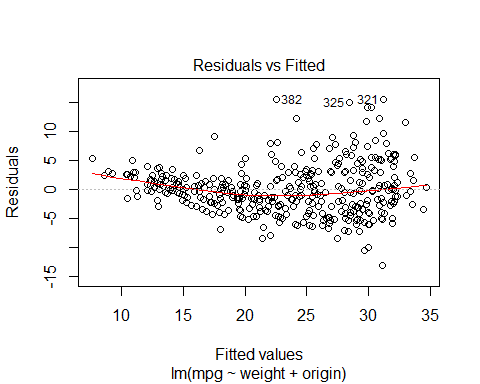
hist(error26,prob=T,breaks = 20)  
  
lines(density(error26),col='red')  
  
  
  
#Normal Curve for automodel26  
  
  
mean\_error26<-mean(error26)  
variance\_error26<-var(error26)  
sd\_error26<-sqrt(variance\_error26)  
x\_error26<-seq(-25,25,length=30)  
y\_error26<-dnorm(x\_error26,mean\_error26,sd\_error26)  
lines(x\_error26,y\_error26,col='blue')



########### Model 27 ###############  
  
autoModel27= lm(mpg ~ weight + origin, data = autompg)  
summary(autoModel27)

##   
## Call:  
## lm(formula = mpg ~ weight + origin, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -13.0698 -2.7888 -0.3122 2.4489 15.4816   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 42.4908175 1.3266161 32.03 < 2e-16 \*\*\*  
## weight -0.0070071 0.0003136 -22.34 < 2e-16 \*\*\*  
## origin 1.1540278 0.3306915 3.49 0.000539 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.272 on 389 degrees of freedom  
## Multiple R-squared: 0.702, Adjusted R-squared: 0.7004   
## F-statistic: 458.1 on 2 and 389 DF, p-value: < 2.2e-16

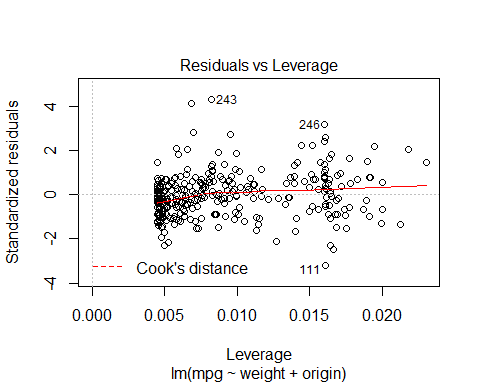
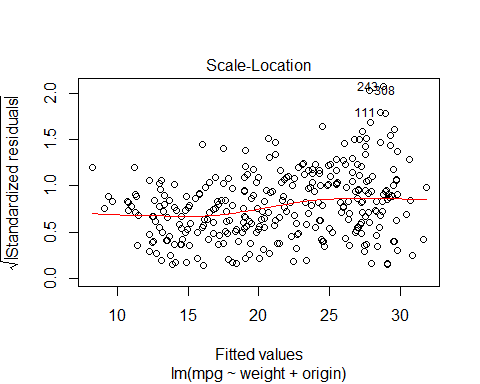
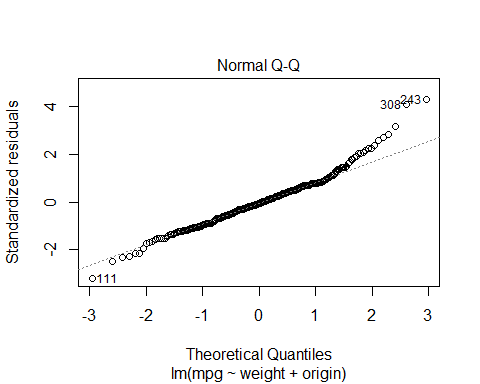
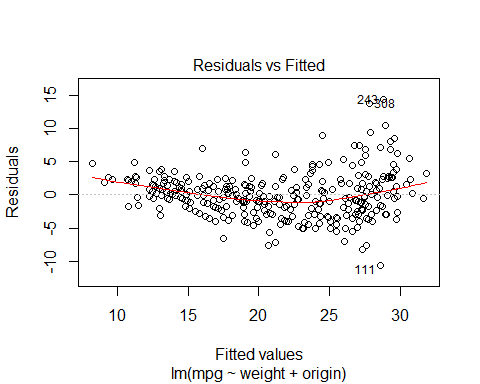
plot(autoModel27)



##### Regression Model ##########  
  
autoModel27\_27 <- lm(mpg ~ weight + origin , data = training)  
summary(autoModel27\_27)

##   
## Call:  
## lm(formula = mpg ~ weight + origin, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -10.586 -2.108 -0.126 1.780 14.312   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 39.9509642 1.1888230 33.605 <2e-16 \*\*\*  
## weight -0.0063024 0.0002683 -23.488 <2e-16 \*\*\*  
## origin 0.6739308 0.3145300 2.143 0.0329 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.345 on 315 degrees of freedom  
## Multiple R-squared: 0.7545, Adjusted R-squared: 0.7529   
## F-statistic: 484 on 2 and 315 DF, p-value: < 2.2e-16

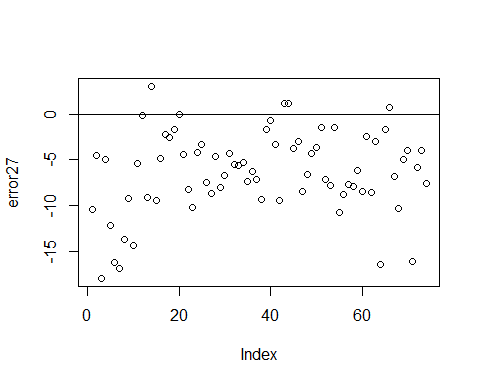
plot(autoModel27\_27)



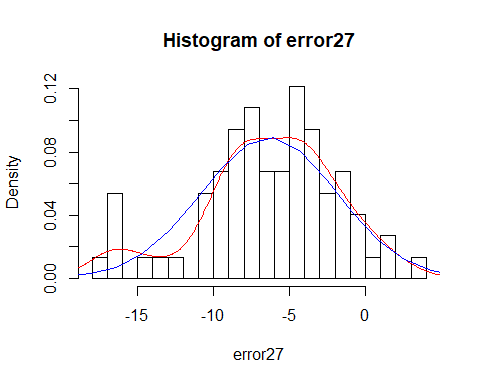
### use model on test dataset ###  
  
B0<-autoModel27\_27$coefficients[1]  
  
B1<-autoModel27\_27$coefficients[2]  
  
B2<-autoModel27\_27$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted27 <- B0 + B1 \* test$weight + B2 \*test$origin  
  
  
# #calculating error ##  
  
error27<- y\_predicted27- test$mpg  
  
summary(error27)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -17.925 -8.636 -6.008 -6.349 -3.409 3.021

plot(error27)  
  
abline(0,0)



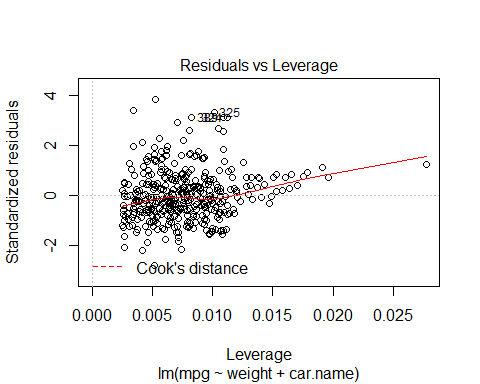
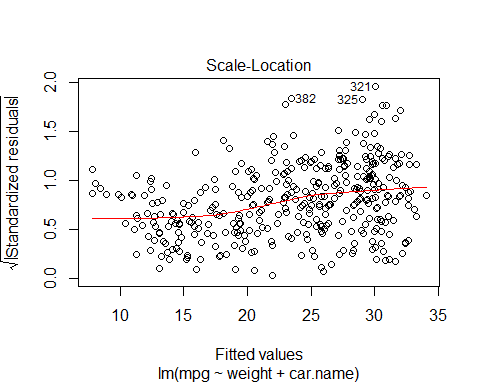
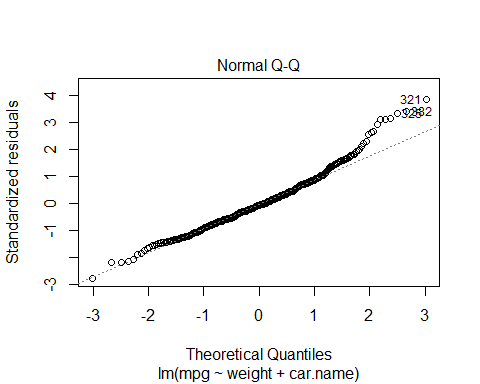
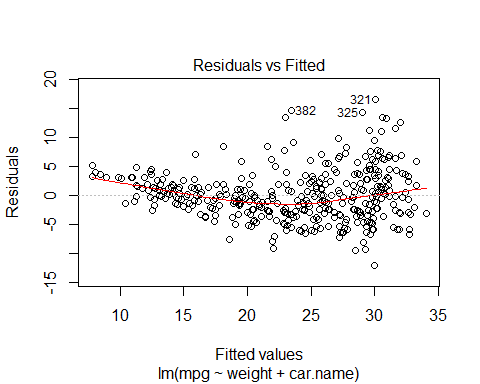
hist(error27,prob=T,breaks = 20)  
  
lines(density(error27),col='red')  
  
  
  
#Normal Curve for automodel27  
  
  
mean\_error27<-mean(error27)  
variance\_error27<-var(error27)  
sd\_error27<-sqrt(variance\_error27)  
x\_error27<-seq(-25,25,length=30)  
y\_error27<-dnorm(x\_error27,mean\_error27,sd\_error27)  
lines(x\_error27,y\_error27,col='blue')



##### Model 28 ################  
  
  
autoModel28= lm(mpg ~ weight + car.name, data = autompg)  
summary(autoModel28)

##   
## Call:  
## lm(formula = mpg ~ weight + car.name, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -11.9900 -2.8371 -0.3312 2.3735 16.4941   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 44.9890196 0.9806902 45.875 <2e-16 \*\*\*  
## weight -0.0075049 0.0002653 -28.289 <2e-16 \*\*\*  
## car.name 0.0053794 0.0025162 2.138 0.0331 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.313 on 389 degrees of freedom  
## Multiple R-squared: 0.6962, Adjusted R-squared: 0.6946   
## F-statistic: 445.7 on 2 and 389 DF, p-value: < 2.2e-16

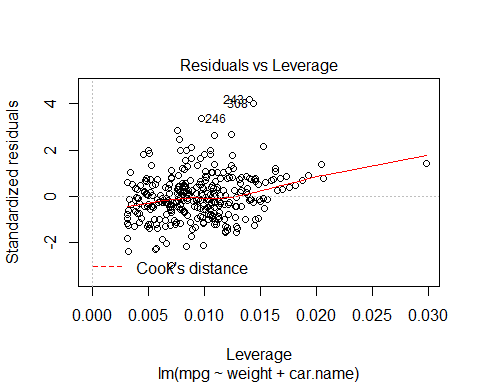
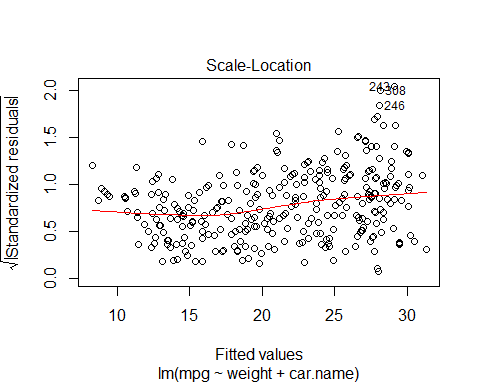
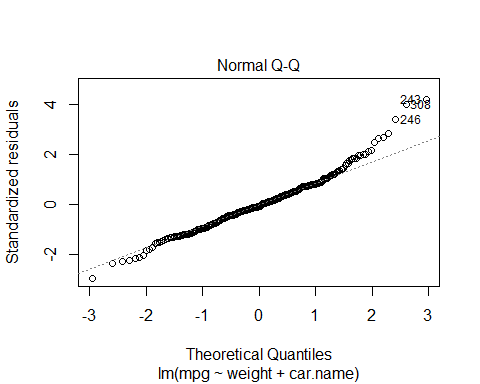
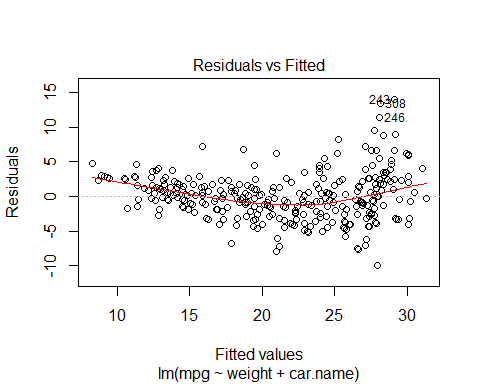
plot(autoModel28)



##### Regression Model ##########  
  
autoModel28\_28 <- lm(mpg ~ weight + car.name , data = training)  
summary(autoModel28\_28)

##   
## Call:  
## lm(formula = mpg ~ weight + car.name, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -9.9245 -2.0649 -0.2879 1.8109 13.9585   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 41.4589372 0.8338569 49.719 <2e-16 \*\*\*  
## weight -0.0065880 0.0002209 -29.824 <2e-16 \*\*\*  
## car.name 0.0026203 0.0021437 1.222 0.223   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.361 on 315 degrees of freedom  
## Multiple R-squared: 0.7521, Adjusted R-squared: 0.7505   
## F-statistic: 477.8 on 2 and 315 DF, p-value: < 2.2e-16

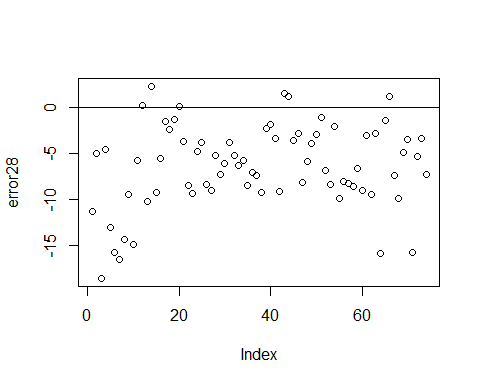
plot(autoModel28\_28)



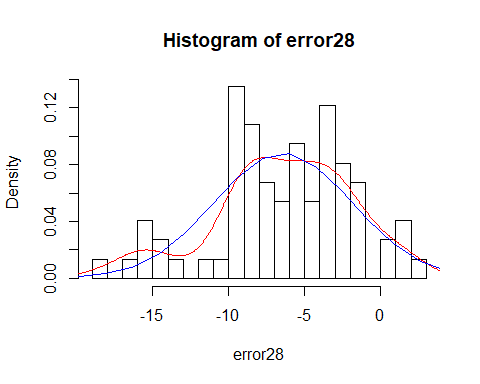
### use model on test dataset ###  
  
B0<-autoModel28\_28$coefficients[1]  
  
B1<-autoModel28\_28$coefficients[2]  
  
B2<-autoModel28\_28$coefficients[3]  
  
  
  
##### calculate predicted values #######  
  
y\_predicted28 <- B0 + B1 \* test$weight + B2 \*test$car.name  
  
  
# #calculating error ##  
  
error28<- y\_predicted28- test$mpg  
  
summary(error28)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -18.578 -9.039 -5.963 -6.421 -3.384 2.295

plot(error28)  
  
abline(0,0)



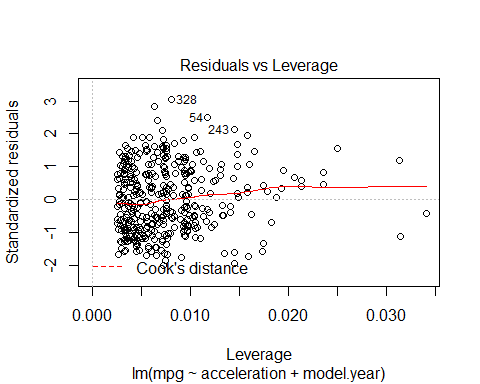
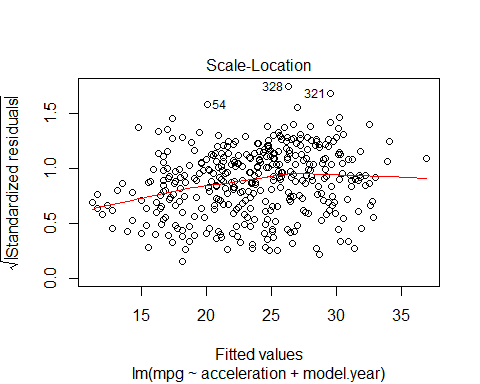
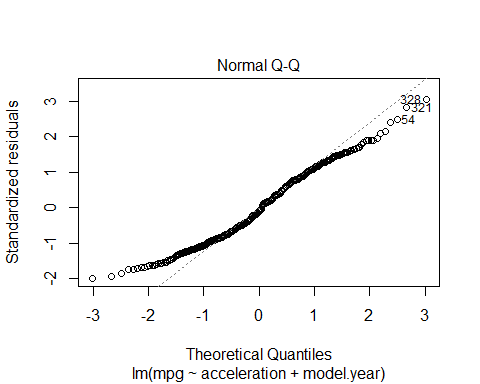
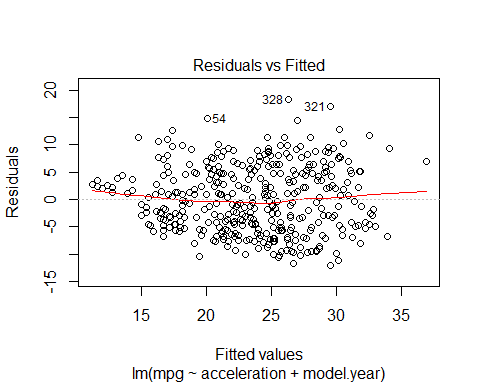
hist(error28,prob=T,breaks = 20)  
  
lines(density(error28),col='red')  
  
  
  
#Normal Curve for automodel28  
  
  
mean\_error28<-mean(error28)  
variance\_error28<-var(error28)  
sd\_error28<-sqrt(variance\_error28)  
x\_error28<-seq(-25,25,length=30)  
y\_error28<-dnorm(x\_error28,mean\_error28,sd\_error28)  
lines(x\_error28,y\_error28,col='blue')



######### Model 29 ##########  
  
autoModel29= lm(mpg ~ acceleration + model.year , data = autompg)  
summary(autoModel29)

##   
## Call:  
## lm(formula = mpg ~ acceleration + model.year, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -11.995 -5.022 -0.588 4.795 18.268   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -69.24131 6.28891 -11.010 < 2e-16 \*\*\*  
## acceleration 0.78716 0.11534 6.825 3.39e-11 \*\*\*  
## model.year 1.05889 0.08638 12.258 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 6.021 on 389 degrees of freedom  
## Multiple R-squared: 0.4079, Adjusted R-squared: 0.4049   
## F-statistic: 134 on 2 and 389 DF, p-value: < 2.2e-16

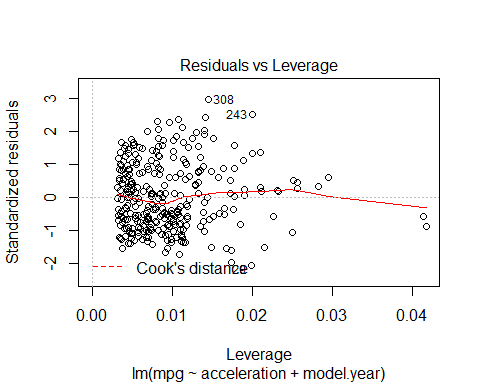
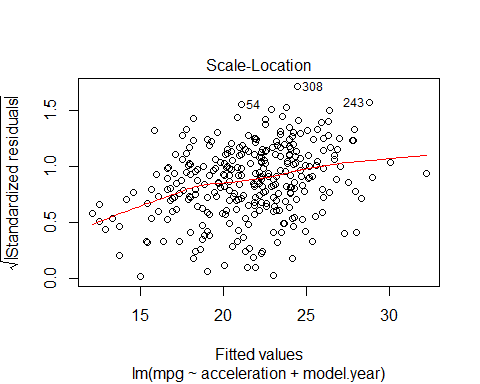
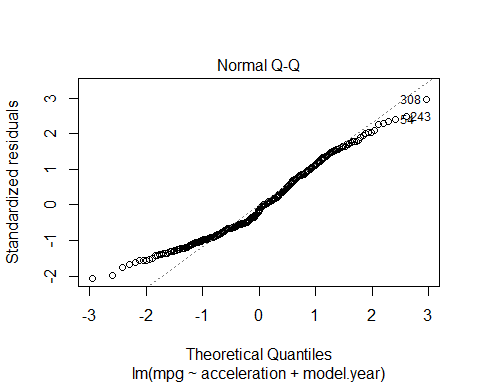
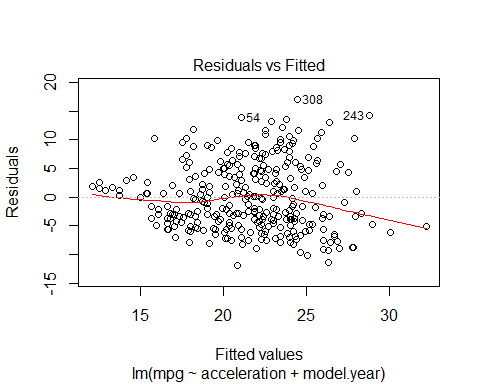
plot(autoModel29)



##### Regression Model ##########  
  
autoModel29\_29 <- lm(mpg ~ acceleration + model.year , data = training)  
summary(autoModel29\_29)

##   
## Call:  
## lm(formula = mpg ~ acceleration + model.year, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -11.8492 -4.5997 -0.8997 4.4277 17.0114   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -42.3051 8.2361 -5.137 4.91e-07 \*\*\*  
## acceleration 0.8354 0.1232 6.781 5.90e-11 \*\*\*  
## model.year 0.6814 0.1140 5.975 6.23e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.804 on 315 degrees of freedom  
## Multiple R-squared: 0.2606, Adjusted R-squared: 0.2559   
## F-statistic: 55.51 on 2 and 315 DF, p-value: < 2.2e-16

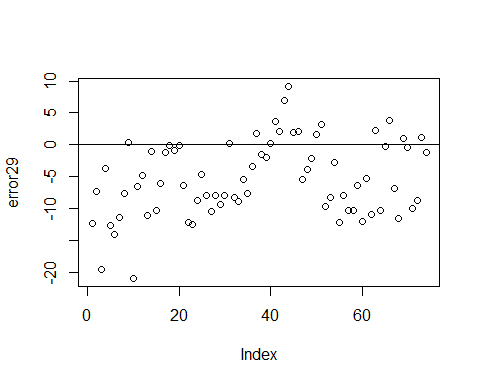
plot(autoModel29\_29)



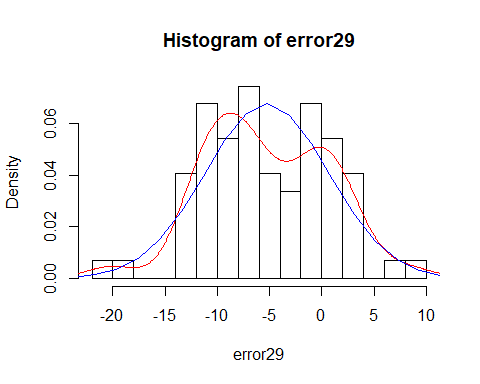
### use model on test dataset ###  
  
B0<-autoModel29\_29$coefficients[1]  
  
B1<-autoModel29\_29$coefficients[2]  
  
B2<-autoModel29\_29$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted29 <- B0 + B1 \*test$acceleration + B2 \* test$model.year  
  
  
# #calculating error ##  
  
error29<-y\_predicted29- test$mpg  
  
summary(error29)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -20.8633 -9.8201 -6.1503 -5.2654 -0.3359 9.1573

plot(error29)  
  
abline(0,0)



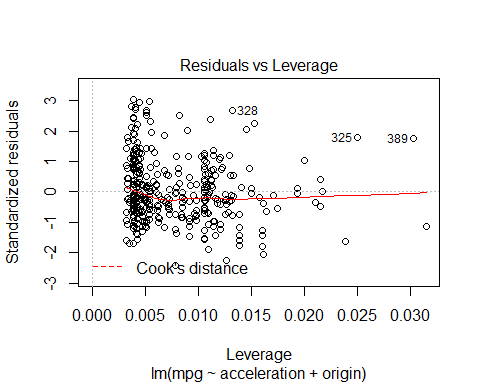
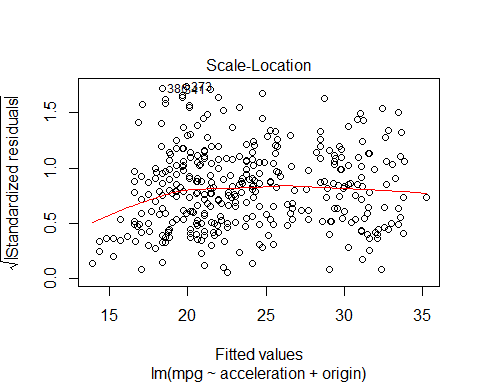
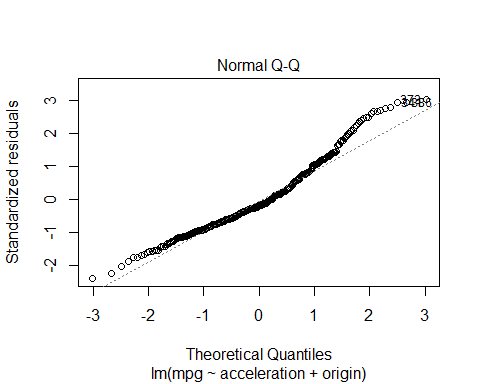
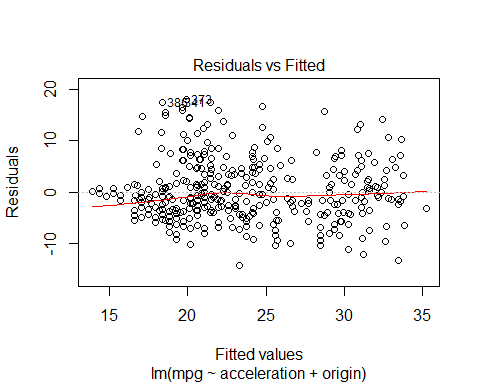
hist(error29,prob=T,breaks = 20)  
  
lines(density(error29),col='red')  
  
  
  
#Normal Curve for automodel29  
  
  
mean\_error29<-mean(error29)  
variance\_error29<-var(error29)  
sd\_error29<-sqrt(variance\_error29)  
x\_error29<-seq(-30,30,length=30)  
y\_error29<-dnorm(x\_error29,mean\_error29,sd\_error29)  
lines(x\_error29,y\_error29,col='blue')



############## Model 30 ##########  
  
  
autoModel30= lm(mpg ~ acceleration + origin , data = autompg)  
summary(autoModel30)

##   
## Call:  
## lm(formula = mpg ~ acceleration + origin, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -14.323 -4.064 -1.151 3.382 18.090   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.8861 1.7465 1.080 0.281   
## acceleration 0.8981 0.1122 8.003 1.41e-14 \*\*\*  
## origin 4.8222 0.3843 12.547 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.982 on 389 degrees of freedom  
## Multiple R-squared: 0.4157, Adjusted R-squared: 0.4127   
## F-statistic: 138.4 on 2 and 389 DF, p-value: < 2.2e-16

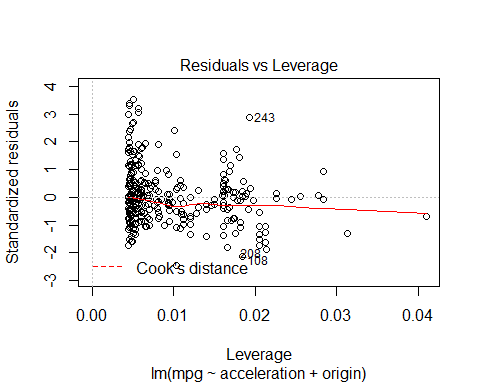
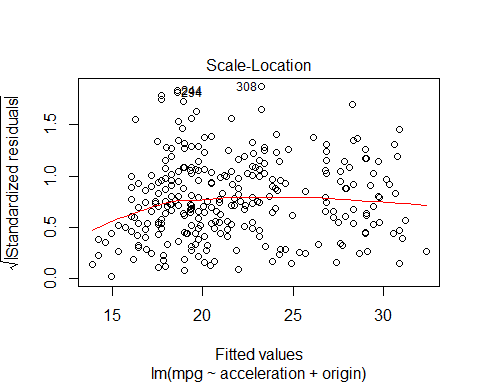
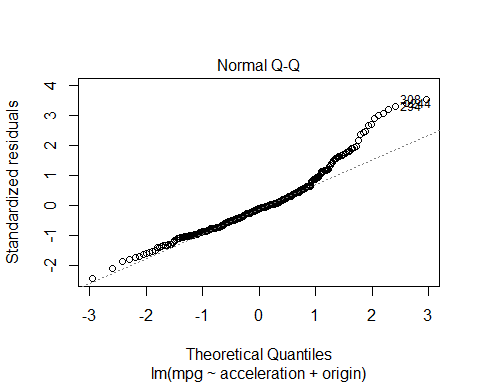
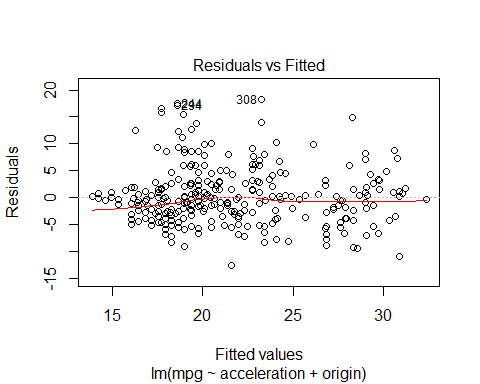
plot(autoModel30)



##### Regression Model ##########  
  
autoModel30\_30 <- lm(mpg ~ acceleration + origin , data = training)  
summary(autoModel30\_30)

##   
## Call:  
## lm(formula = mpg ~ acceleration + origin, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -12.5937 -3.5764 -0.5921 2.1684 18.2261   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.5730 1.6627 2.149 0.0324 \*   
## acceleration 0.7328 0.1096 6.687 1.04e-10 \*\*\*  
## origin 4.4647 0.4020 11.107 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.192 on 315 degrees of freedom  
## Multiple R-squared: 0.4085, Adjusted R-squared: 0.4047   
## F-statistic: 108.8 on 2 and 315 DF, p-value: < 2.2e-16

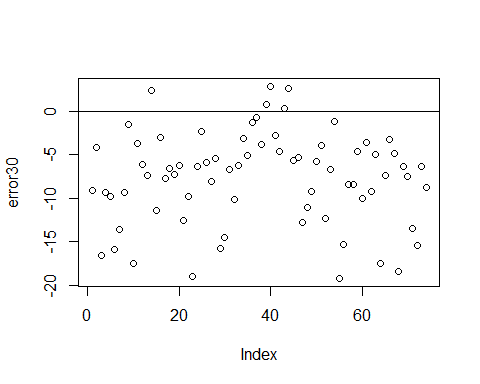
plot(autoModel30\_30)



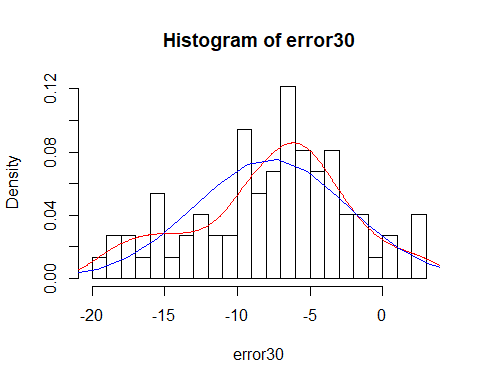
### use model on test dataset ###  
  
B0<-autoModel30\_30$coefficients[1]  
  
B1<-autoModel30\_30$coefficients[2]  
  
B2<-autoModel30\_30$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted30 <- B0 + B1 \*test$acceleration + B2 \* test$origin  
  
  
# #calculating error ##  
  
error30<-y\_predicted30- test$mpg  
  
summary(error30)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -19.191 -10.080 -6.695 -7.661 -4.231 2.879

plot(error30)  
  
abline(0,0)



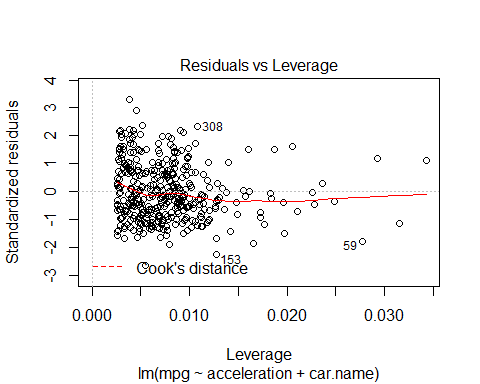
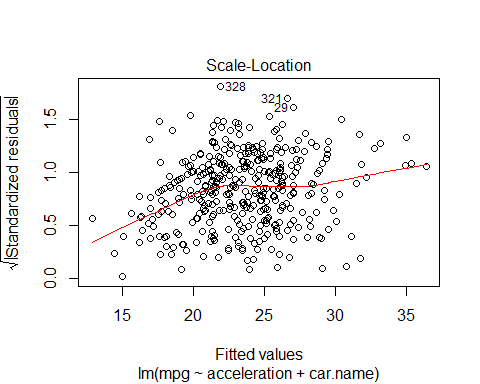
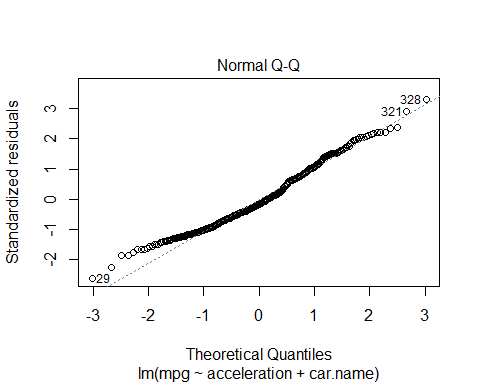
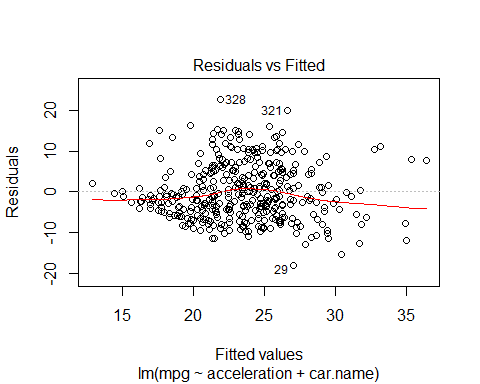
hist(error30,prob=T,breaks = 20)  
  
lines(density(error30),col='red')  
  
  
  
#Normal Curve for automodel30  
  
  
mean\_error30<-mean(error30)  
variance\_error30<-var(error30)  
sd\_error30<-sqrt(variance\_error30)  
x\_error30<-seq(-30,30,length=30)  
y\_error30<-dnorm(x\_error30,mean\_error30,sd\_error30)  
lines(x\_error30,y\_error30,col='blue')



########### Model 31 ############  
  
  
  
autoModel31= lm(mpg ~ acceleration + car.name , data = autompg)  
summary(autoModel31)

##   
## Call:  
## lm(formula = mpg ~ acceleration + car.name, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -18.021 -5.017 -1.084 4.748 22.694   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.320982 2.019744 1.644 0.101   
## acceleration 1.116048 0.127524 8.752 < 2e-16 \*\*\*  
## car.name 0.018616 0.003928 4.739 3.02e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 6.893 on 389 degrees of freedom  
## Multiple R-squared: 0.224, Adjusted R-squared: 0.22   
## F-statistic: 56.15 on 2 and 389 DF, p-value: < 2.2e-16

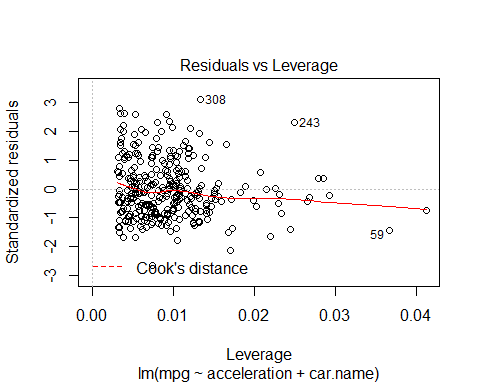
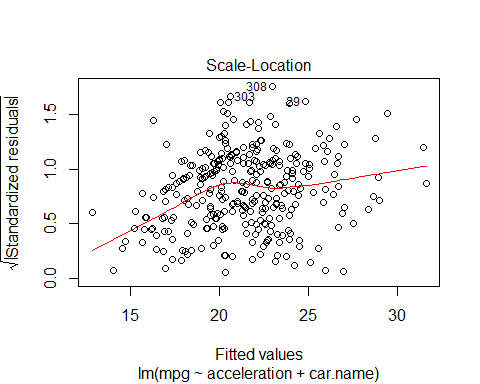
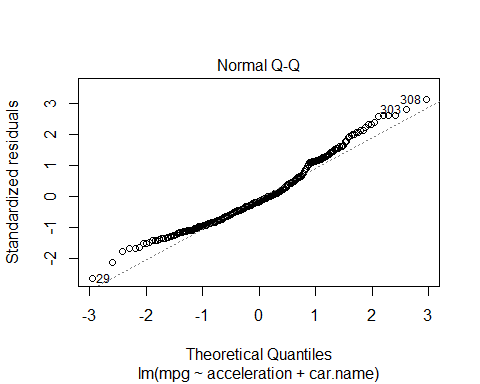
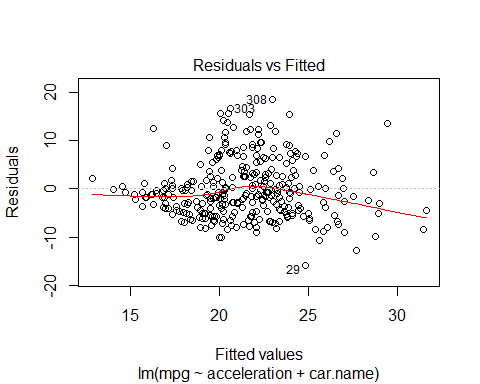
plot(autoModel31)



##### Regression Model ##########  
  
autoModel31\_31 <- lm(mpg ~ acceleration + car.name , data = training)  
summary(autoModel31\_31)

##   
## Call:  
## lm(formula = mpg ~ acceleration + car.name, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -15.820 -4.449 -1.009 3.405 18.517   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.470146 1.937495 2.307 0.02169 \*   
## acceleration 0.979687 0.123624 7.925 3.98e-14 \*\*\*  
## car.name 0.013571 0.003755 3.614 0.00035 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 6.001 on 315 degrees of freedom  
## Multiple R-squared: 0.2096, Adjusted R-squared: 0.2046   
## F-statistic: 41.77 on 2 and 315 DF, p-value: < 2.2e-16

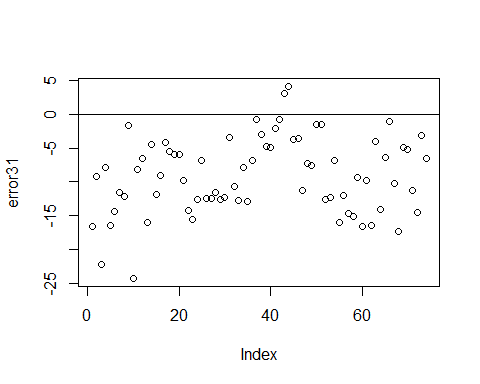
plot(autoModel31\_31)



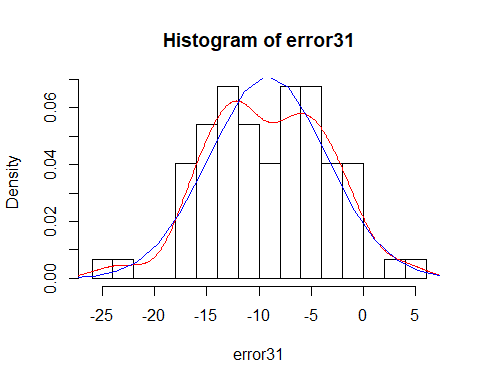
### use model on test dataset ###  
  
B0<-autoModel31\_31$coefficients[1]  
  
B1<-autoModel31\_31$coefficients[2]  
  
B2<-autoModel31\_31$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted31 <- B0 + B1 \*test$acceleration + B2 \* test$car.name  
  
  
# #calculating error ##  
  
error31<-y\_predicted31- test$mpg  
  
summary(error31)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -24.290 -12.582 -9.282 -9.130 -4.977 4.124

plot(error31)  
  
abline(0,0)



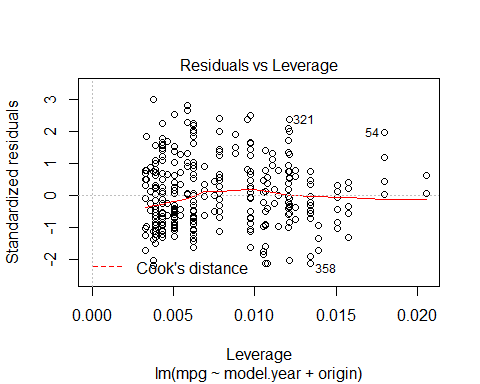
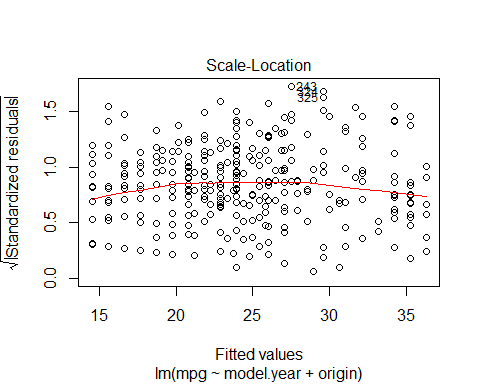
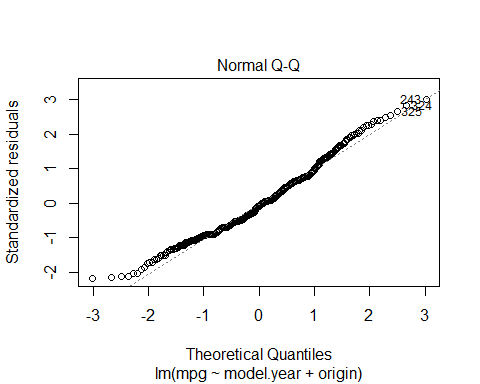
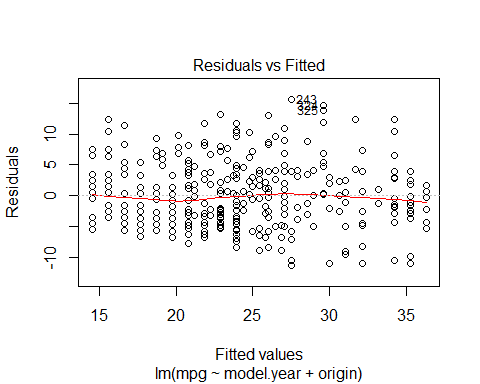
hist(error31,prob=T,breaks = 20)  
  
lines(density(error31),col='red')  
  
  
  
#Normal Curve for automodel31  
  
  
mean\_error31<-mean(error31)  
variance\_error31<-var(error31)  
sd\_error31<-sqrt(variance\_error31)  
x\_error31<-seq(-30,30,length=30)  
y\_error31<-dnorm(x\_error31,mean\_error31,sd\_error31)  
lines(x\_error31,y\_error31,col='blue')



#################### Model 32 ####################  
  
  
  
autoModel32= lm(mpg ~ model.year + origin, data = autompg)  
summary(autoModel32)

##   
## Call:  
## lm(formula = mpg ~ model.year + origin, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -11.3126 -3.7257 -0.4732 3.3893 15.5874   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -63.37982 5.46818 -11.59 <2e-16 \*\*\*  
## model.year 1.04715 0.07282 14.38 <2e-16 \*\*\*  
## origin 4.60725 0.33301 13.84 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.216 on 389 degrees of freedom  
## Multiple R-squared: 0.5557, Adjusted R-squared: 0.5534   
## F-statistic: 243.2 on 2 and 389 DF, p-value: < 2.2e-16

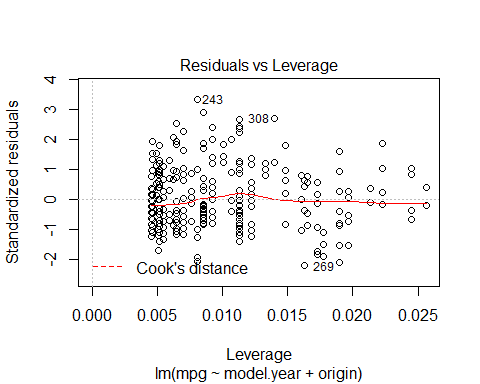
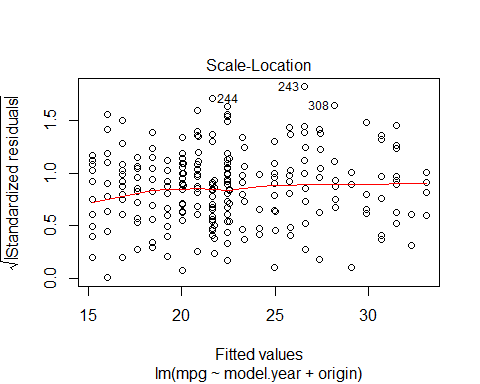
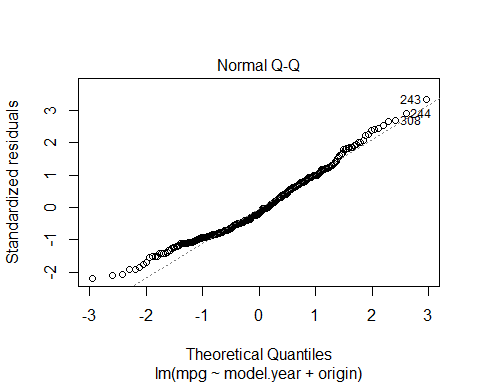
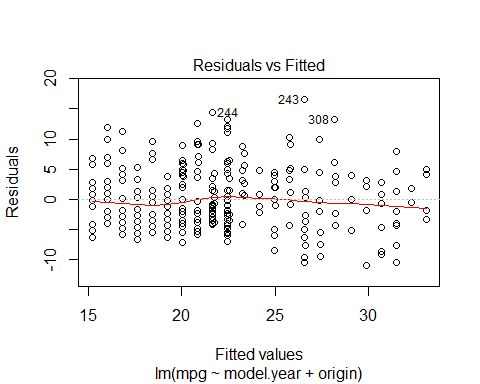
plot(autoModel32)



##### Regression Model ##########  
  
autoModel32\_32 <- lm(mpg ~ model.year + origin , data = training)  
summary(autoModel32\_32)

##   
## Call:  
## lm(formula = mpg ~ model.year + origin, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -10.8568 -3.8064 -0.9844 3.3540 16.5462   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -46.10736 7.09684 -6.497 3.20e-10 \*\*\*  
## model.year 0.80555 0.09504 8.476 9.12e-16 \*\*\*  
## origin 4.91414 0.37664 13.047 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.006 on 315 degrees of freedom  
## Multiple R-squared: 0.4499, Adjusted R-squared: 0.4465   
## F-statistic: 128.8 on 2 and 315 DF, p-value: < 2.2e-16

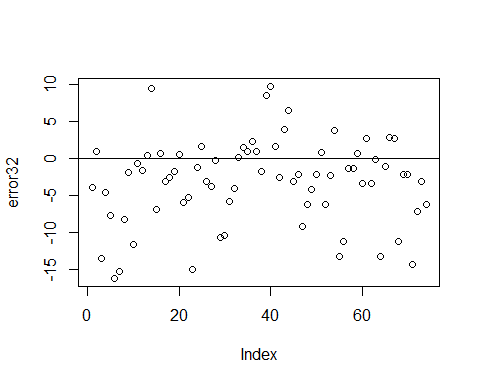
plot(autoModel32\_32)



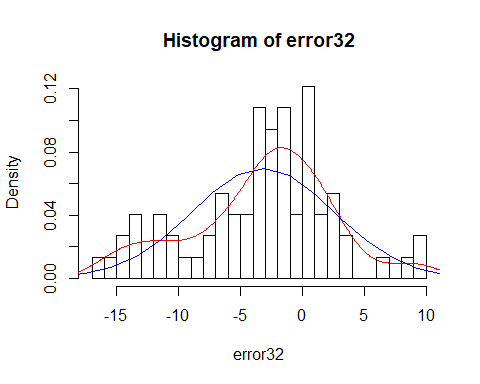
### use model on test dataset ###  
  
B0<-autoModel32\_32$coefficients[1]  
  
B1<-autoModel32\_32$coefficients[2]  
  
B2<-autoModel32\_32$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted32 <- B0 + B1 \* test$model.year + B2 \* test$origin   
  
  
# #calculating error ##  
  
error32<-y\_predicted32- test$mpg  
  
summary(error32)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -16.1351 -6.1381 -2.2240 -3.1818 0.6484 9.6846

plot(error32)  
  
abline(0,0)



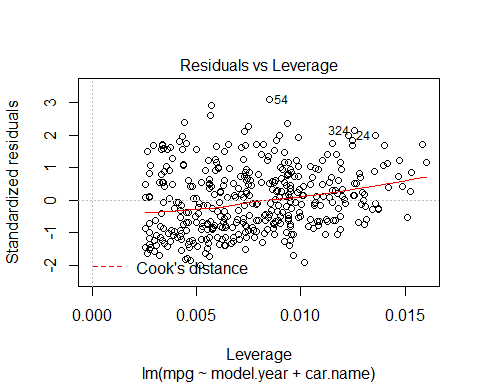
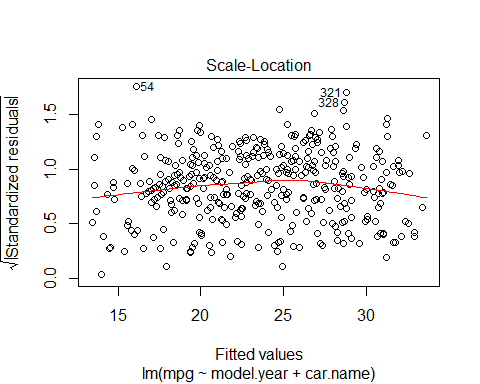
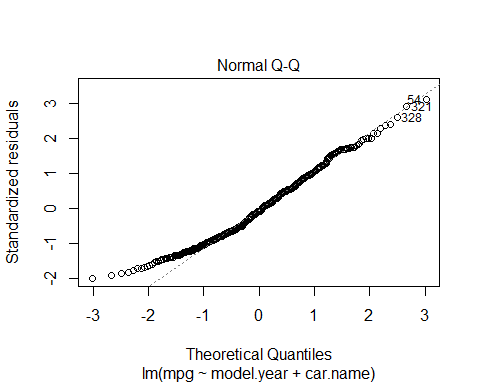
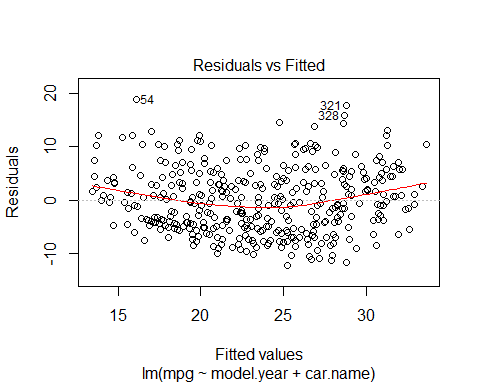
hist(error32,prob=T,breaks = 20)  
  
lines(density(error32),col='red')  
  
  
  
#Normal Curve for automodel32  
  
  
mean\_error32<-mean(error32)  
variance\_error32<-var(error32)  
sd\_error32<-sqrt(variance\_error32)  
x\_error32<-seq(-30,30,length=30)  
y\_error32<-dnorm(x\_error32,mean\_error32,sd\_error32)  
lines(x\_error32,y\_error32,col='blue')



############# Model 33 ################  
  
  
autoModel33= lm(mpg ~ model.year + car.name, data = autompg)  
summary(autoModel33)

##   
## Call:  
## lm(formula = mpg ~ model.year + car.name, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -12.2425 -4.8532 -0.5334 4.2004 18.9144   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -70.203372 6.397873 -10.973 < 2e-16 \*\*\*  
## model.year 1.194139 0.084346 14.158 < 2e-16 \*\*\*  
## car.name 0.019547 0.003469 5.634 3.38e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 6.126 on 389 degrees of freedom  
## Multiple R-squared: 0.387, Adjusted R-squared: 0.3839   
## F-statistic: 122.8 on 2 and 389 DF, p-value: < 2.2e-16

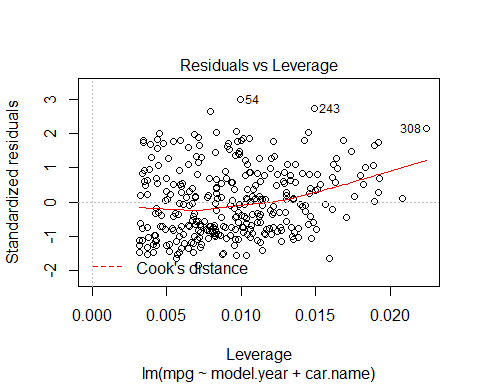
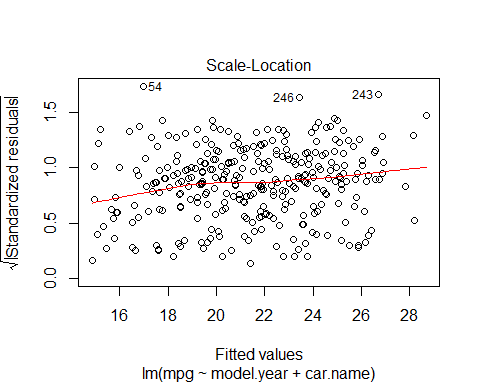
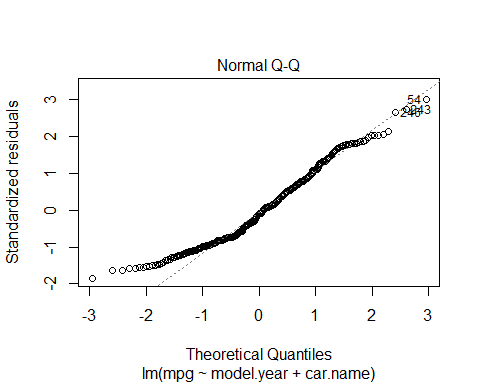
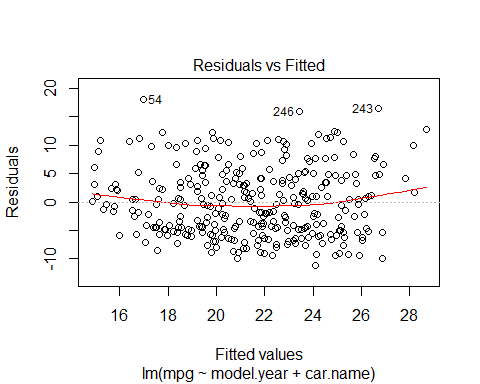
plot(autoModel33)



##### Regression Model ##########  
  
autoModel33\_33 <- lm(mpg ~ model.year + car.name , data = training)  
summary(autoModel33\_33)

##   
## Call:  
## lm(formula = mpg ~ model.year + car.name, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -11.1047 -4.8885 -0.5828 4.1602 18.0167   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -46.859847 8.554947 -5.478 8.83e-08 \*\*\*  
## model.year 0.881131 0.114182 7.717 1.59e-13 \*\*\*  
## car.name 0.016661 0.003746 4.447 1.21e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 6.027 on 315 degrees of freedom  
## Multiple R-squared: 0.2027, Adjusted R-squared: 0.1977   
## F-statistic: 40.05 on 2 and 315 DF, p-value: 3.177e-16

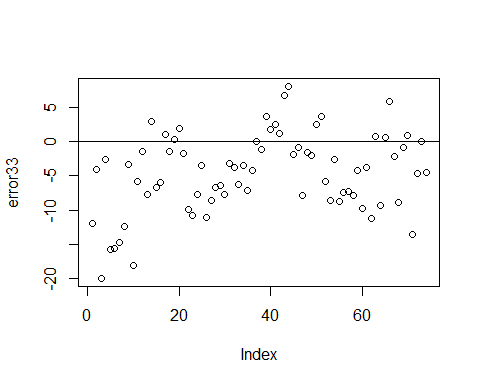
plot(autoModel33\_33)



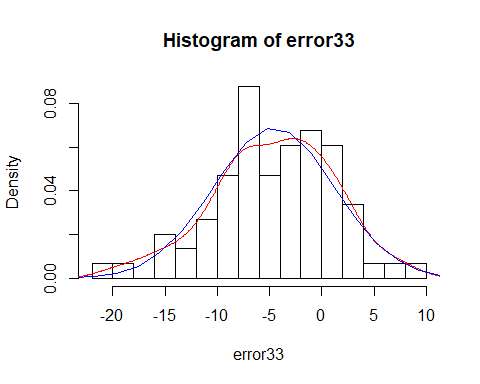
### use model on test dataset ###  
  
B0<-autoModel33\_33$coefficients[1]  
  
B1<-autoModel33\_33$coefficients[2]  
  
B2<-autoModel33\_33$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted33 <- B0 + B1 \* test$model.year + B2 \* test$car.name  
  
  
# #calculating error ##  
  
error33<-y\_predicted33- test$mpg  
  
summary(error33)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -20.0204 -7.8496 -4.1384 -4.6072 -0.8241 8.1280

plot(error33)  
  
abline(0,0)



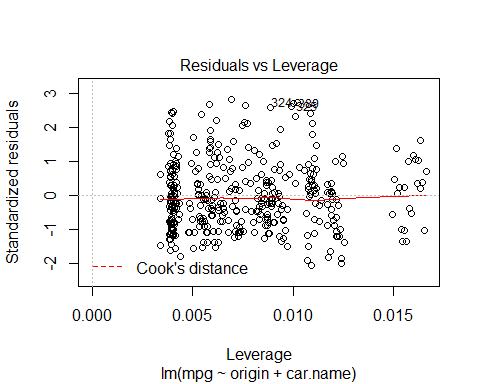
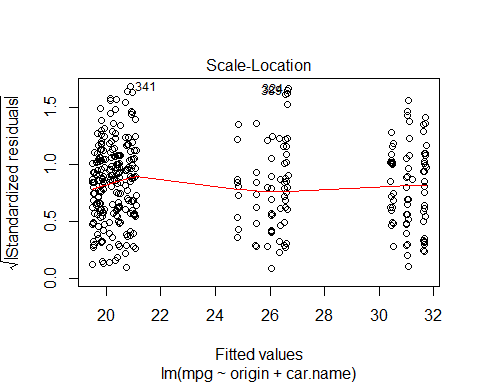
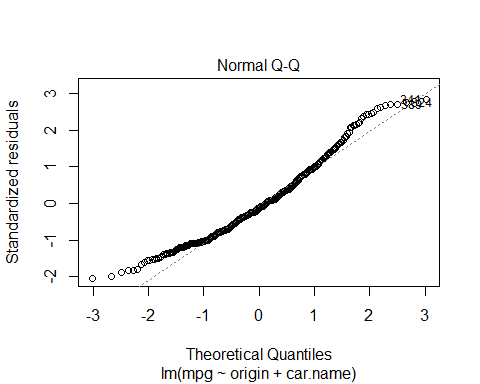
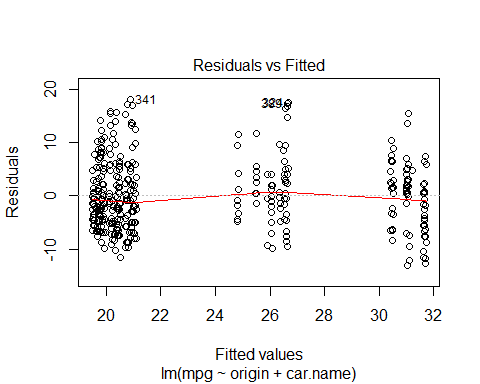
hist(error33,prob=T,breaks = 20)  
  
lines(density(error33),col='red')  
  
  
  
#Normal Curve for automodel33  
  
  
mean\_error33<-mean(error33)  
variance\_error33<-var(error33)  
sd\_error33<-sqrt(variance\_error33)  
x\_error33<-seq(-30,30,length=30)  
y\_error33<-dnorm(x\_error33,mean\_error33,sd\_error33)  
lines(x\_error33,y\_error33,col='blue')



########### Model 34 ###########  
  
autoModel34= lm(mpg ~ origin + car.name, data = autompg)  
summary(autoModel34)

##   
## Call:  
## lm(formula = mpg ~ origin + car.name, data = autompg)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -13.0417 -4.8501 -0.7349 3.9511 18.1201   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 14.252755 0.790209 18.04 <2e-16 \*\*\*  
## origin 5.219667 0.432488 12.07 <2e-16 \*\*\*  
## car.name 0.006457 0.003890 1.66 0.0978 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 6.433 on 389 degrees of freedom  
## Multiple R-squared: 0.3242, Adjusted R-squared: 0.3208   
## F-statistic: 93.33 on 2 and 389 DF, p-value: < 2.2e-16

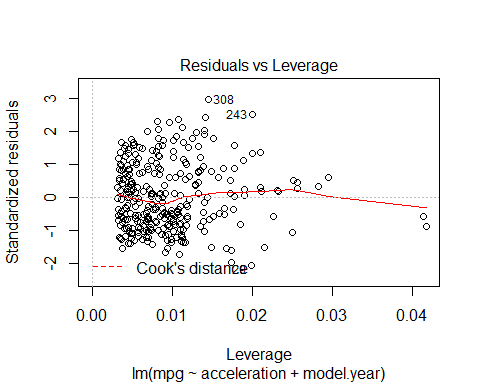
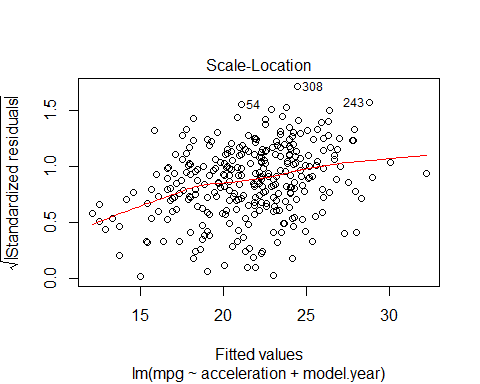
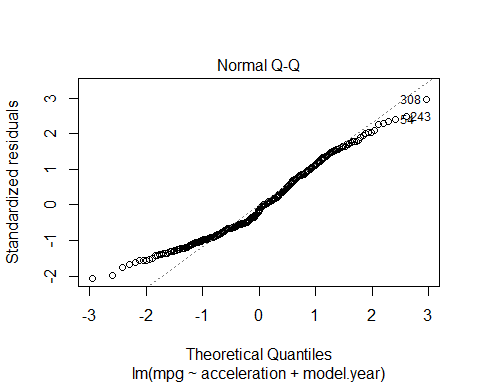
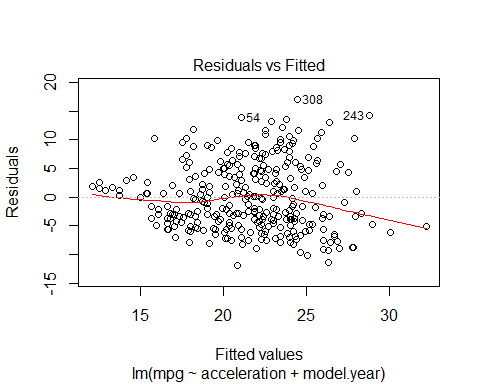
plot(autoModel34)



##### Regression Model ##########  
  
autoModel34\_34 <- lm(mpg ~ origin + car.name , data = training)  
summary(autoModel29\_29)

##   
## Call:  
## lm(formula = mpg ~ acceleration + model.year, data = training)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -11.8492 -4.5997 -0.8997 4.4277 17.0114   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -42.3051 8.2361 -5.137 4.91e-07 \*\*\*  
## acceleration 0.8354 0.1232 6.781 5.90e-11 \*\*\*  
## model.year 0.6814 0.1140 5.975 6.23e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.804 on 315 degrees of freedom  
## Multiple R-squared: 0.2606, Adjusted R-squared: 0.2559   
## F-statistic: 55.51 on 2 and 315 DF, p-value: < 2.2e-16

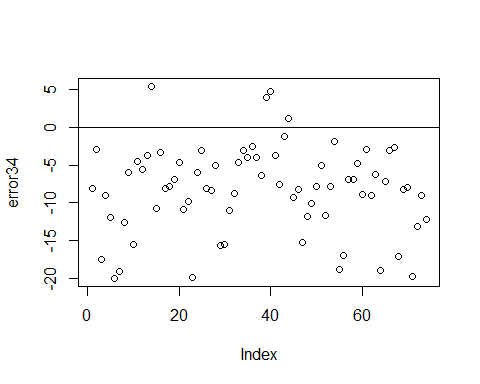
plot(autoModel29\_29)



### use model on test dataset ###  
  
B0<-autoModel34\_34$coefficients[1]  
  
B1<-autoModel34\_34$coefficients[2]  
  
B2<-autoModel34\_34$coefficients[3]  
  
  
  
  
##### calculate predicted values #######  
  
y\_predicted34 <- B0 + B1 \*test$origin + B2 \* test$car.name  
  
  
# #calculating error ##  
  
error34<-y\_predicted34- test$mpg  
  
summary(error34)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -19.976 -11.514 -7.943 -8.248 -4.537 5.427

plot(error34)  
  
abline(0,0)



hist(error34,prob=T,breaks = 20)  
  
lines(density(error34),col='red')  
  
  
  
#Normal Curve for automodel34  
  
  
mean\_error34<-mean(error34)  
variance\_error34<-var(error34)  
sd\_error34<-sqrt(variance\_error34)  
x\_error34<-seq(-30,30,length=30)  
y\_error34<-dnorm(x\_error34,mean\_error34,sd\_error34)  
lines(x\_error34,y\_error34,col='blue')

