

1.

Code:

```
1 getwd()
2 setwd("C:/Users/engbi/OneDrive/Documents/GitHub/IT365/week 6/Linear Regression using R")
3 car <- read.csv("CarPrice.csv")
4
5 wheelBP <- car
6 owheel <- boxplot(wheelBP$wheelbase, plot = FALSE)$out
7 wheelBP <- wheelBP[-which(wheelBP$wheelbase %in% owheel),]
8 png("Nwheel.png")
9 boxplot(wheelBP$wheelbase)
10 owheel <- boxplot(wheelBP$wheelbase, plot = FALSE)$out
11 wheelBP <- wheelBP[-which(wheelBP$wheelbase %in% owheel),]
12 png("Nwheel.png")
13 boxplot(wheelBP$wheelbase)
14 owheel <- boxplot(wheelBP$wheelbase, plot = FALSE)$out
15 wheelBP <- wheelBP[-which(wheelBP$wheelbase %in% owheel),]
16 png("Nwheel.png")
17 boxplot(wheelBP$wheelbase)
18 dev.off()
19
20 OLength <- boxplot(wheelBP$carlength, plot = FALSE)$out
21 wheelBP <- wheelBP[-which(wheelBP$carlength %in% OLength),]
22 png("NLength.png")
23 boxplot(wheelBP$carlength)
24 dev.off()
25
26 owidth <- boxplot(wheelBP$carwidth, plot = FALSE)$out
27 wheelBP <- wheelBP[-which(wheelBP$carwidth %in% owidth),]
28 png("Nwidth.png")
29 boxplot(wheelBP$carwidth)
30 dev.off()
```

```

32 png("OHeight.png")
33 boxplot(wheelBP$carheight)
34 dev.off()
35
36 png("Oweight.png")
37 boxplot(wheelBP$curbweight)
38 dev.off()
39
40 OSize <- boxplot(wheelBP$enginesize, plot = FALSE)$out
41 wheelBP <- wheelBP[-which(wheelBP$enginesize %in% OSize),]
42 png("NSize.png")
43 boxplot(wheelBP$enginesize)
44 dev.off()
45
46 png("ORatio.png")
47 boxplot(wheelBP$boreratio)
48 dev.off()
49
50 Ostroke <- boxplot(wheelBP$stroke, plot = FALSE)$out
51 wheelBP <- wheelBP[-which(wheelBP$stroke %in% Ostroke),]
52 png("NStroke.png")
53 boxplot(wheelBP$stroke)
54 Ostroke <- boxplot(wheelBP$stroke, plot = FALSE)$out
55 wheelBP <- wheelBP[-which(wheelBP$stroke %in% Ostroke),]
56 png("NStroke.png")
57 boxplot(wheelBP$stroke)
58 dev.off()
59
60 OCompress <- boxplot(wheelBP$compressionratio, plot = FALSE)$out
61 wheelBP <- wheelBP[-which(wheelBP$compressionratio %in% OCompress),]
62 png("NCompress.png")
63 boxplot(wheelBP$compressionratio)
64 OCompress <- boxplot(wheelBP$compressionratio, plot = FALSE)$out
65 wheelBP <- wheelBP[-which(wheelBP$compressionratio %in% OCompress),]
66 png("NCompress.png")
67 boxplot(wheelBP$compressionratio)
68 dev.off()
69
70 OHorsepower <- boxplot(wheelBP$horsepower, plot = FALSE)$out
71 wheelBP <- wheelBP[-which(wheelBP$horsepower %in% OHorsepower),]
72 png("NHorsepower.png")
73 boxplot(wheelBP$horsepower)
74 dev.off()
75
76 OPeak <- boxplot(wheelBP$peakrpm, plot = FALSE)$out
77 wheelBP <- wheelBP[-which(wheelBP$peakrpm %in% OPeak),]
78 png("NPeak.png")
79 boxplot(wheelBP$peakrpm)
80 dev.off()
81
82 png("NCity.png")
83 boxplot(wheelBP$citympg)
84 dev.off()
85
86 png("NHighway.png")
87 boxplot(wheelBP$highwaympg)
88 dev.off()

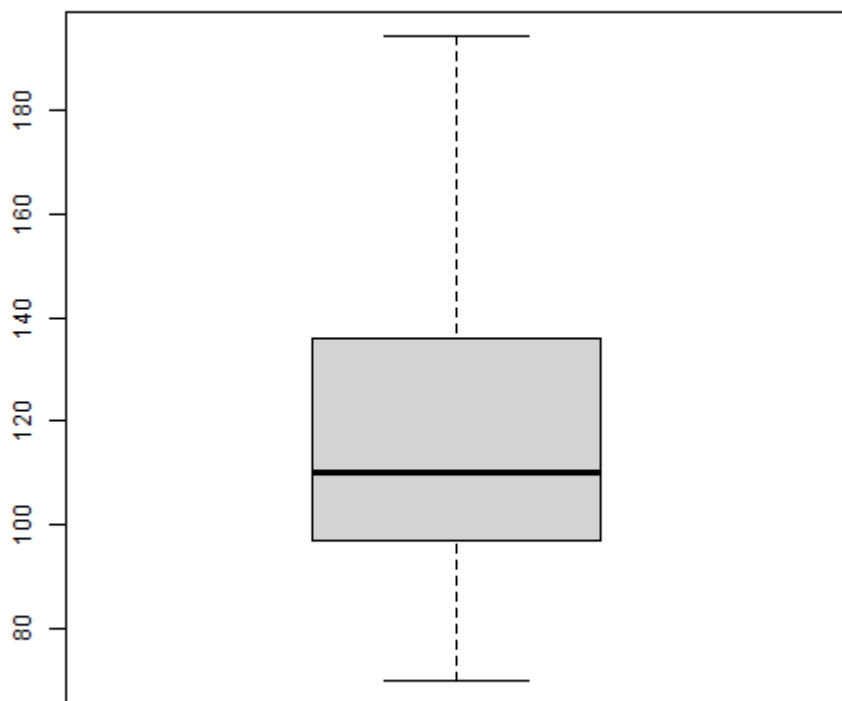
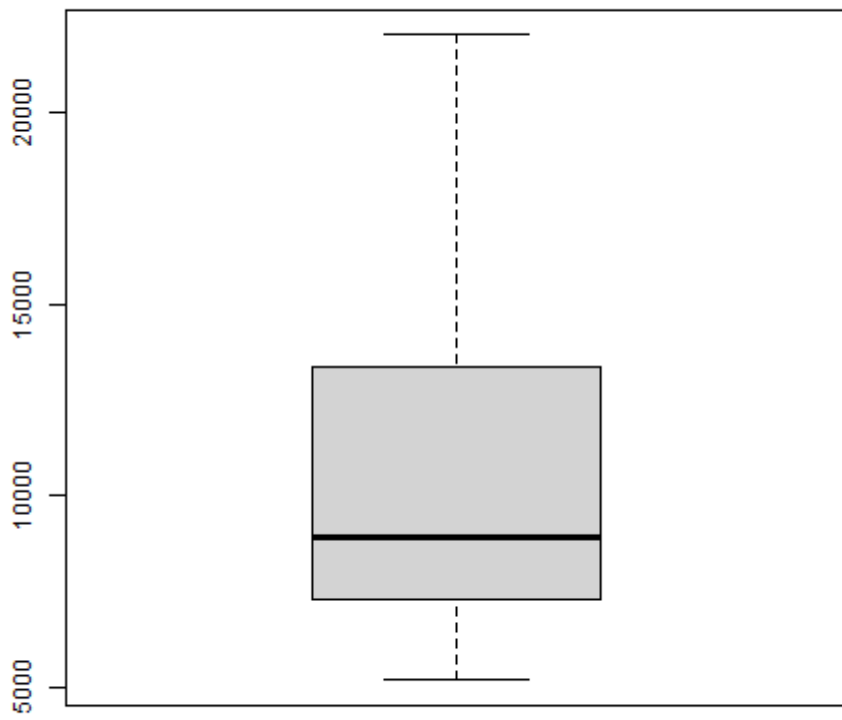
```

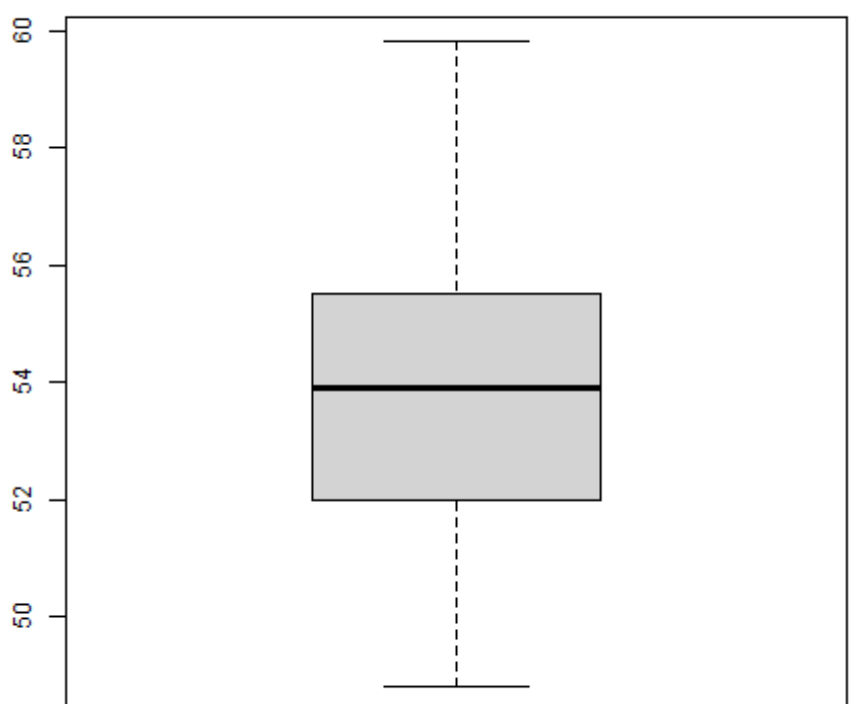
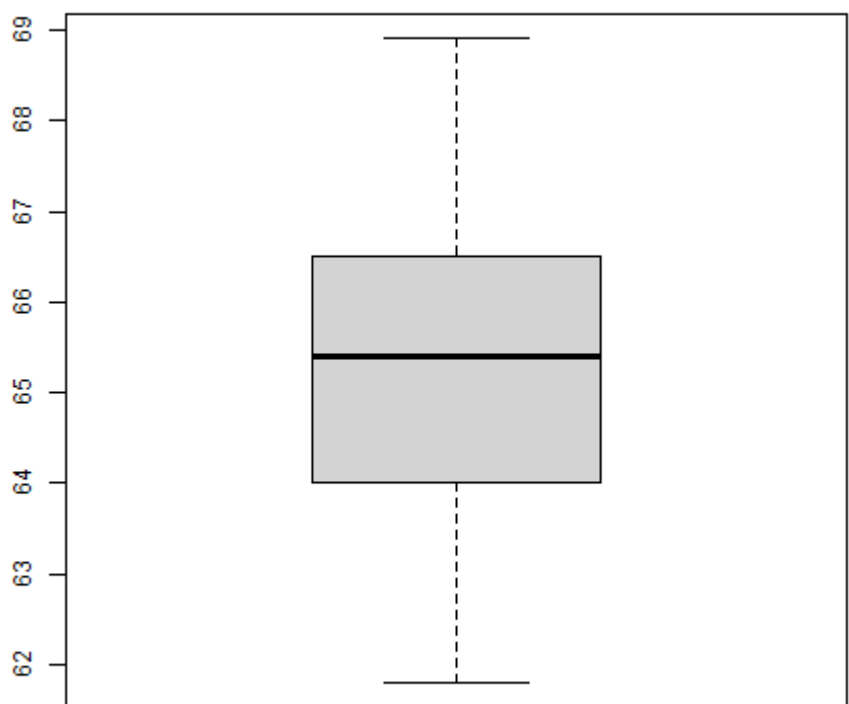
```

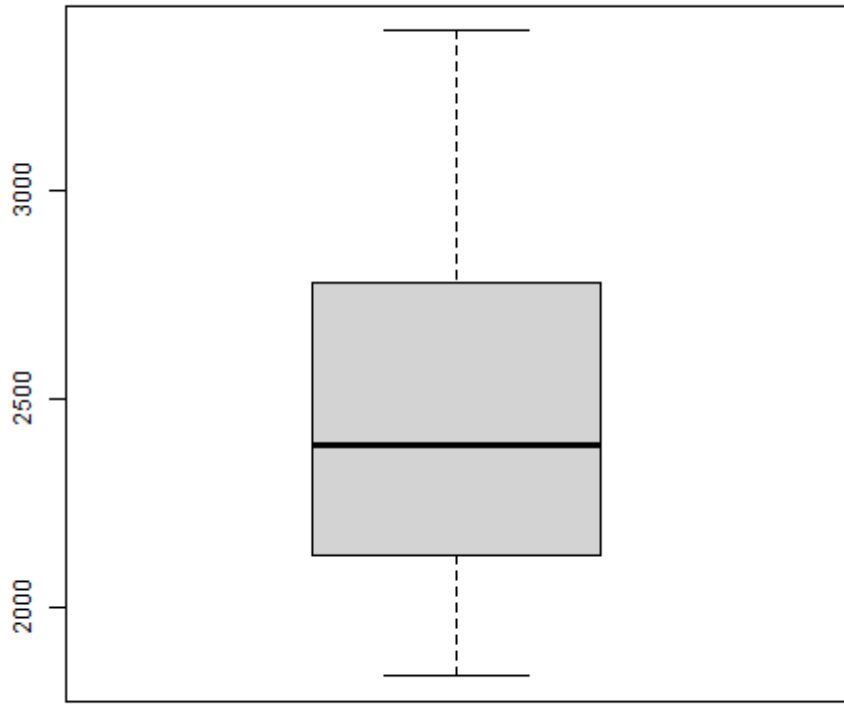
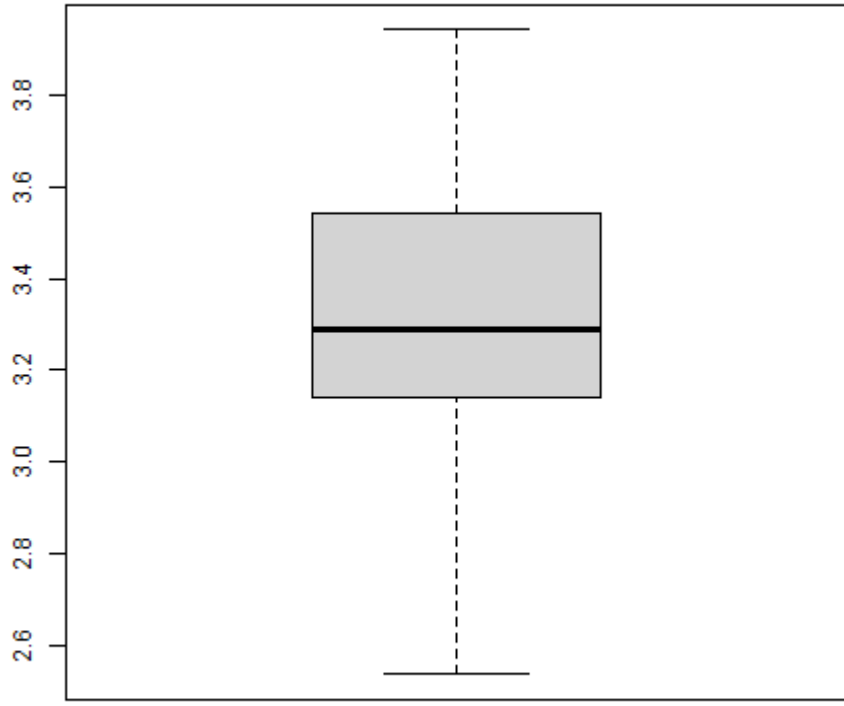
90 OPrice <- boxplot(wheelBP$price, plot = FALSE)$out
91 wheelBP <- wheelBP[-which(wheelBP$price %in% OPrice),]
92 png("NPrice.png")
93 boxplot(wheelBP$price)
94 OPrice <- boxplot(wheelBP$price, plot = FALSE)$out
95 wheelBP <- wheelBP[-which(wheelBP$price %in% OPrice),]
96 png("NPrice.png")
97 boxplot(wheelBP$price)
98 dev.off()

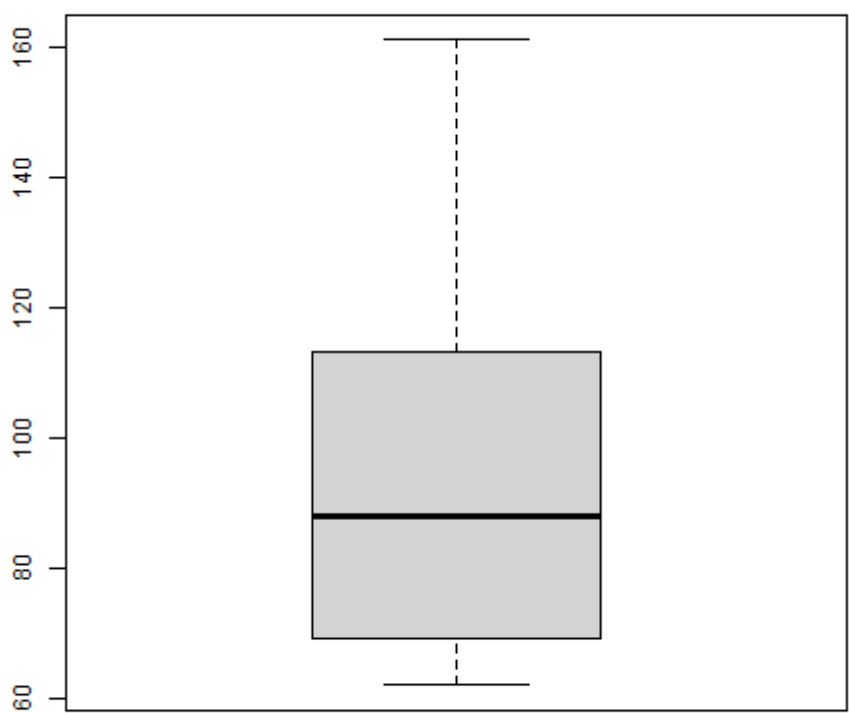
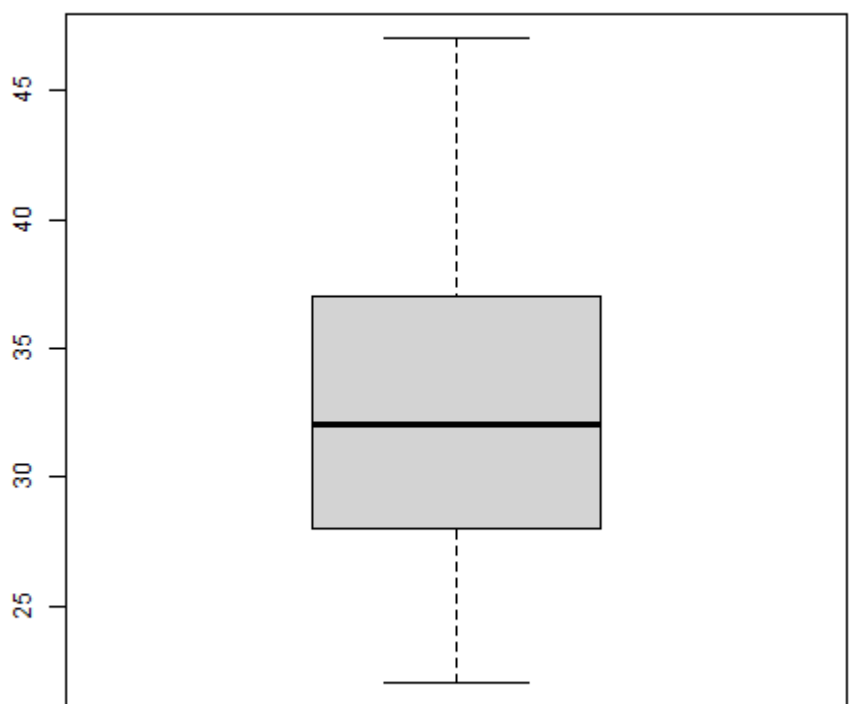
```

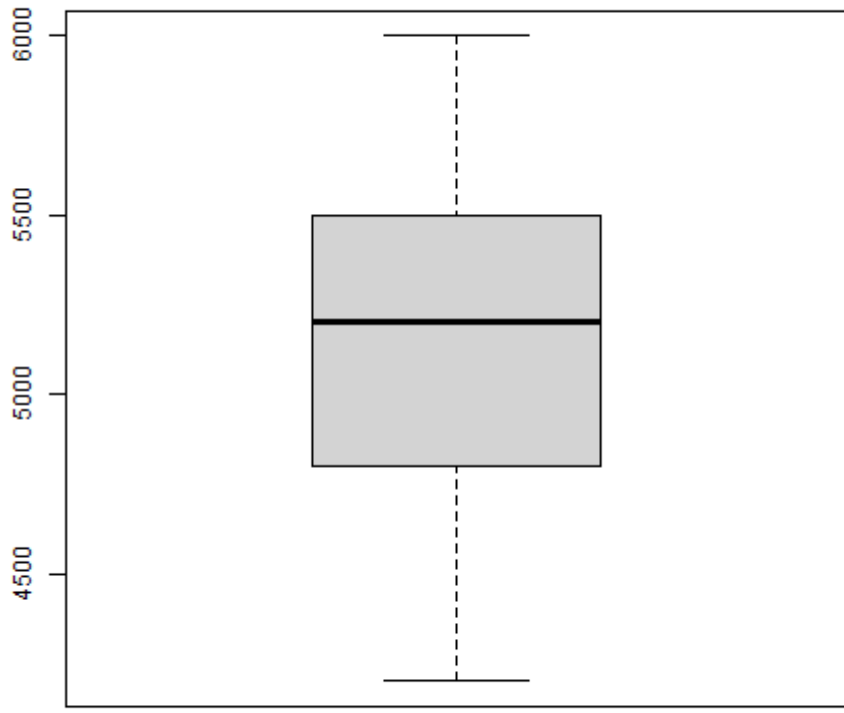
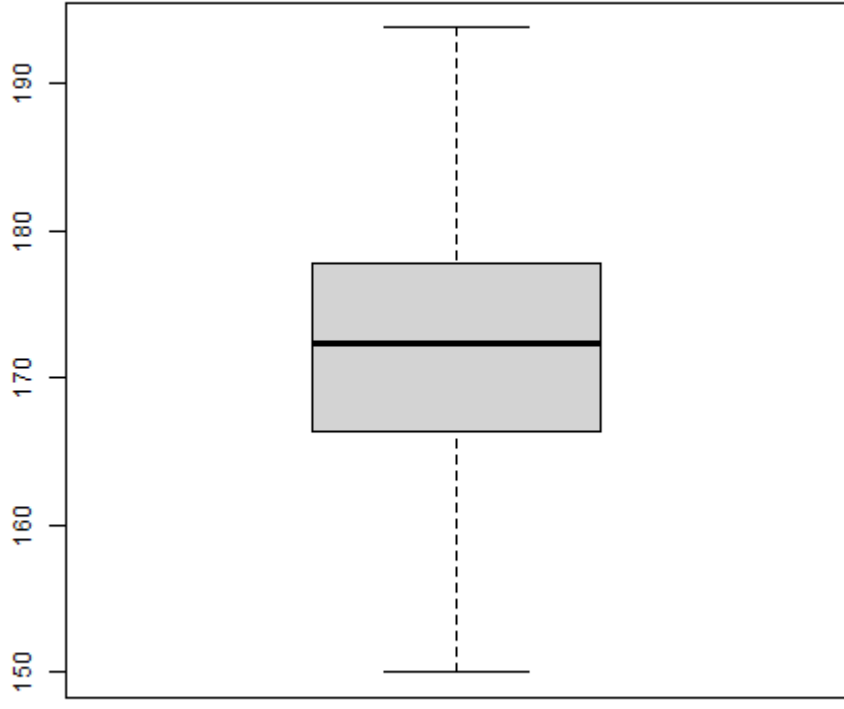
Boxplots:











2.

Independent Variable/s:

- wheelbase
- carlength
- carwidth
- carheight
- curbweight
- enginesize
- boreratio
- stroke
- compressionratio
- horsepower
- peakrpm
- citympg
- highwaympg

Dependent Variable:

- Price

3.

Code:

```
'''
```

```
regModel <- lm(wheelBP$price~wheelBP$wheelbase + wheelBP$carlength +  
               wheelBP$carwidth + wheelBP$carheight + wheelBP$curbweight +  
               wheelBP$enginesize + wheelBP$boreratio + wheelBP$stroke +  
               wheelBP$compressionratio + wheelBP$horsepower +  
               wheelBP$peakrpm + wheelBP$citympg +  
               wheelBP$highwaympg)
```

```
summary(regModel)
```

```
'''
```


Display:

```
Residuals:
    Min       1Q   Median       3Q      Max
-3936.2 -1114.9  -43.7    866.0  6162.7

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   -7671.4809  15605.3753  -0.492  0.624024
wheelBP$wheelbase    253.5403    95.1159   2.666  0.008888 **
wheelBP$carlength   -53.7180    47.6994  -1.126  0.262633
wheelBP$carwidth     94.4068   242.6847   0.389  0.698049
wheelBP$carheight   -58.4899   103.3447  -0.566  0.572611
wheelBP$curbweight    1.2918     1.3946   0.926  0.356392
wheelBP$enginesize    7.6696    16.5084   0.465  0.643182
wheelBP$boreratio    729.1847  1243.5224   0.586  0.558863
wheelBP$stroke      -4529.6048  1078.7842  -4.199  5.6e-05 ***
wheelBP$compressionratio  392.3718   635.7259   0.617  0.538425
wheelBP$horsepower    63.3833    17.8491   3.551  0.000574 ***
wheelBP$peakrpm        0.3837     0.5537   0.693  0.489862
wheelBP$citympg      -437.7216   180.8922  -2.420  0.017231 *
wheelBP$highwaympg    252.7915   173.7896   1.455  0.148739
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1861 on 106 degrees of freedom
Multiple R-squared:  0.8235,    Adjusted R-squared:  0.8018
F-statistic: 38.03 on 13 and 106 DF,  p-value: < 2.2e-16
```

4.

Code:

Remove carlength

'''

```
regModel <- lm(wheelBP$price~wheelBP$wheelbase +
               wheelBP$carwidth + wheelBP$carheight + wheelBP$curbweight +
               wheelBP$enginesize + wheelBP$boreratio + wheelBP$stroke +
               wheelBP$compressionratio + wheelBP$horsepower +
               wheelBP$peakrpm + wheelBP$citympg +
               wheelBP$highwaympg)
```

```
summary(regModel)
```

'''

```

Residuals:
    Min       1Q   Median       3Q      Max
-3849.5 -1043.1  -102.5   821.4  6344.9

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -6668.1721 15599.4436  -0.427 0.669902
wheelBP$wheelbase    215.3469    88.9756   2.420 0.017193 *
wheelBP$carwidth     67.9820   241.8504   0.281 0.779183
wheelBP$carheight   -93.3240    98.7308  -0.945 0.346668
wheelBP$curbweight    0.8937     1.3507   0.662 0.509623
wheelBP$enginesize    8.0026    16.5265   0.484 0.629212
wheelBP$boreratio    279.2724  1179.0767   0.237 0.813221
wheelBP$stroke      -4660.7091  1073.8281  -4.340 3.24e-05 ***
wheelBP$compressionratio 347.1080   635.2490   0.546 0.585920
wheelBP$horsepower    62.9050    17.8664   3.521 0.000633 ***
wheelBP$peakrpm        0.4353     0.5525   0.788 0.432566
wheelBP$citympg      -382.9230   174.4432  -2.195 0.030316 *
wheelBP$highwaympg     211.1883   170.0306   1.242 0.216930
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1863 on 107 degrees of freedom
Multiple R-squared:  0.8213,    Adjusted R-squared:  0.8013
F-statistic: 40.99 on 12 and 107 DF,  p-value: < 2.2e-16

```

Remove carwidth

```

...

regModel <- lm(wheelBP$price~wheelBP$wheelbase + wheelBP$carheight + wheelBP$curbweight +
               wheelBP$enginesize + wheelBP$boreratio + wheelBP$stroke +
               wheelBP$compressionratio + wheelBP$horsepower +
               wheelBP$peakrpm + wheelBP$citympg +
               wheelBP$highwaympg)

summary(regModel)

...

```

```

Residuals:
    Min       1Q   Median       3Q      Max
-3868.3 -1017.2  -118.7   803.7  6249.6

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -3512.5735  10784.8844  -0.326  0.745286
wheelBP$wheelbase    227.3850    77.6576   2.928  0.004161 **
wheelBP$carheight   -100.5361    94.9313  -1.059  0.291944
wheelBP$curbweight    0.9387     1.3354   0.703  0.483617
wheelBP$enginesize    7.7758    16.4362   0.473  0.637104
wheelBP$boreratio    363.6906   1135.3157   0.320  0.749327
wheelBP$stroke     -4628.2638   1063.0446  -4.354  3.05e-05 ***
wheelBP$compressionratio 335.0503    631.0908   0.531  0.596574
wheelBP$horsepower    63.7170    17.5560   3.629  0.000436 ***
wheelBP$peakrpm        0.4459     0.5489   0.812  0.418331
wheelBP$citympg     -386.3535   173.2722  -2.230  0.027833 *
wheelBP$highwaympg    214.2284   168.9613   1.268  0.207555
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1855 on 108 degrees of freedom
Multiple R-squared:  0.8212,    Adjusted R-squared:  0.803
F-statistic: 45.1 on 11 and 108 DF,  p-value: < 2.2e-16

```

Remove carheight

```
'''
```

```

regModel <- lm(wheelBP$price~wheelBP$wheelbase + wheelBP$curbweight +
              wheelBP$enginesize + wheelBP$boreratio + wheelBP$stroke +
              wheelBP$compressionratio + wheelBP$horsepower +
              wheelBP$peakrpm + wheelBP$citympg +
              wheelBP$highwaympg)

```

```
summary(regModel)
```

```
'''
```

```

Residuals:
    Min       1Q   Median       3Q      Max
-4024.8 -1027.3  -128.9   782.6  6271.4

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   -7119.9962  10238.5330  -0.695  0.488277
wheelBP$wheelbase    196.2064    71.9008   2.729  0.007411 **
wheelBP$curbweight     0.6262     1.3032   0.481  0.631802
wheelBP$enginesize     8.4017    16.4347   0.511  0.610234
wheelBP$boreratio    472.9017   1131.2529   0.418  0.676745
wheelBP$stroke   -4517.3293   1058.4609  -4.268  4.23e-05 ***
wheelBP$compressionratio  425.7726    625.5983   0.681  0.497578
wheelBP$horsepower    67.9313    17.1086   3.971  0.000129 ***
wheelBP$peakrpm       0.4554     0.5491   0.829  0.408673
wheelBP$citympg     -395.7218    173.1427  -2.286  0.024217 *
wheelBP$highwaympg    219.8573    168.9718   1.301  0.195952
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1856 on 109 degrees of freedom
Multiple R-squared:  0.8194,    Adjusted R-squared:  0.8028
F-statistic: 49.44 on 10 and 109 DF,  p-value: < 2.2e-16

```

Remove curbweight

```
...
```

```

regModel <- lm(wheelBP$price~wheelBP$wheelbase +
               wheelBP$enginesize + wheelBP$boreratio + wheelBP$stroke +
               wheelBP$compressionratio + wheelBP$horsepower +
               wheelBP$peakrpm + wheelBP$citympg +
               wheelBP$highwaympg)

```

```
summary(regModel)
```

```
...
```

```

Residuals:
    Min       1Q   Median       3Q      Max
-4058.5 -1048.6  -134.4   792.2  6176.8

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   -6612.3359  10148.2225  -0.652  0.51603
wheelBP$wheelbase    207.8956    67.4242   3.083  0.00259 **
wheelBP$enginesize    11.3794    15.1685   0.750  0.45474
wheelBP$boreratio    540.6292   1118.5084   0.483  0.62981
wheelBP$stroke   -4631.9586   1027.6202  -4.507  1.65e-05 ***
wheelBP$compressionratio  458.3881    619.7278   0.740  0.46108
wheelBP$horsepower    68.8397    16.9443   4.063  9.12e-05 ***
wheelBP$peakrpm       0.4275     0.5441   0.786  0.43374
wheelBP$citympg     -400.4382    172.2589  -2.325  0.02193 *
wheelBP$highwaympg    205.6806    165.7939   1.241  0.21740
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1850 on 110 degrees of freedom
Multiple R-squared:  0.819,    Adjusted R-squared:  0.8042
F-statistic: 55.29 on 9 and 110 DF,  p-value: < 2.2e-16

```

Remove the enginesize

```
...  
  
regModel <- lm(wheelBP$price~wheelBP$wheelbase + wheelBP$boreratio + wheelBP$stroke +  
               wheelBP$compressionratio + wheelBP$horsepower +  
               wheelBP$peakrpm + wheelBP$citympg +  
               wheelBP$highwaympg)  
  
summary(regModel)  
...
```

```
Residuals:  
    Min       1Q   Median       3Q      Max   
-3805.4 -1028.2  -166.5    880.6   6346.0  
  
Coefficients:  
                Estimate Std. Error t value Pr(>|t|)      
(Intercept)   -6684.4156 10127.7629  -0.660  0.51062      
wheelBP$wheelbase    216.0282    66.4158   3.253  0.00151 **    
wheelBP$boreratio    652.5192   1106.3344   0.590  0.55652      
wheelBP$stroke   -4449.8354    996.5640  -4.465 1.93e-05 ***   
wheelBP$compressionratio  405.4470    614.4830   0.660  0.51074      
wheelBP$horsepower    78.2314     11.3961   6.865 4.01e-10 ***   
wheelBP$peakrpm      0.2352      0.4790   0.491  0.62440      
wheelBP$citympg    -375.8625    168.7819  -2.227  0.02797 *     
wheelBP$highwaympg    192.6831    164.5611   1.171  0.24415      
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 1846 on 111 degrees of freedom  
Multiple R-squared:  0.818,    Adjusted R-squared:  0.8049  
F-statistic: 62.38 on 8 and 111 DF,  p-value: < 2.2e-16
```

Remove boreratio

```
...  
  
regModel <- lm(wheelBP$price~wheelBP$wheelbase + wheelBP$stroke +  
               wheelBP$compressionratio + wheelBP$horsepower +  
               wheelBP$peakrpm + wheelBP$citympg +  
               wheelBP$highwaympg)  
  
summary(regModel)  
...
```

```

Residuals:
    Min       1Q   Median       3Q      Max
-3849.5 -1072.9  -169.4   884.0  6178.6

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   -6186.3006  10063.0634  -0.615  0.539965
wheelBP$wheelbase    233.1071    59.5966   3.911  0.000158 ***
wheelBP$stroke   -4515.6743   987.4051  -4.573  1.25e-05 ***
wheelBP$compressionratio  473.3206   601.8512   0.786  0.433270
wheelBP$horsepower    78.7380    11.3306   6.949  2.57e-10 ***
wheelBP$peakrpm      0.1777     0.4676   0.380  0.704666
wheelBP$citympg    -400.3316   163.1267  -2.454  0.015663 *
wheelBP$highwaympg   206.8536   162.3231   1.274  0.205184
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1841 on 112 degrees of freedom
Multiple R-squared:  0.8175,    Adjusted R-squared:  0.8061
F-statistic: 71.66 on 7 and 112 DF,  p-value: < 2.2e-16

```

Remove compressionratio

```

'''

regModel <- lm(wheelBP$price~wheelBP$wheelbase + wheelBP$stroke + wheelBP$horsepower +
              wheelBP$peakrpm + wheelBP$citympg +
              wheelBP$highwaympg)

summary(regModel)

'''

```

```

Residuals:
    Min       1Q   Median       3Q      Max
-3880.4 -1059.4  -199.3   857.5  6297.6

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -2340.7056   8780.2769  -0.267  0.790273
wheelBP$wheelbase    226.8236    58.9588   3.847  0.000198 ***
wheelBP$stroke   -4663.1036   967.8084  -4.818  4.54e-06 ***
wheelBP$horsepower    82.0357    10.5083   7.807  3.24e-12 ***
wheelBP$peakrpm      0.3167     0.4322   0.733  0.465202
wheelBP$citympg   -379.1988   160.6264  -2.361  0.019951 *
wheelBP$highwaympg   205.7071    162.0424   1.269  0.206884
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1838 on 113 degrees of freedom
Multiple R-squared:  0.8165,    Adjusted R-squared:  0.8067
F-statistic: 83.78 on 6 and 113 DF,  p-value: < 2.2e-16

```

Remove peakrpm

...

```
regModel <- lm(wheelBP$price~wheelBP$wheelbase + wheelBP$stroke + wheelBP$horsepower + wheelBP$citympg +
```

```
              wheelBP$highwaympg)
```

```
summary(regModel)
```

...

```
Residuals:
    Min       1Q   Median       3Q      Max
-3943.2 -1080.3  -157.1   908.4  6025.6

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    1076.21   7424.70    0.145  0.885006
wheelBP$wheelbase    212.81    55.66    3.824  0.000215 ***
wheelBP$stroke   -4703.86    964.24   -4.878  3.50e-06 ***
wheelBP$horsepower    82.41    10.47    7.867  2.28e-12 ***
wheelBP$citympg   -361.67    158.51   -2.282  0.024366 *
wheelBP$highwaympg   182.00    158.46    1.149  0.253137
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1834 on 114 degrees of freedom
Multiple R-squared:  0.8156,    Adjusted R-squared:  0.8075
F-statistic: 100.8 on 5 and 114 DF,  p-value: < 2.2e-16
```

Remove highwaympg

...

```
regModel <- lm(wheelBP$price~wheelBP$wheelbase + wheelBP$stroke + wheelBP$horsepower + wheelBP$citympg)
```

```
summary(regModel)
```

...

```
Residuals:
    Min       1Q   Median       3Q      Max
-4123.5  -967.0  -243.5   925.1  6245.4

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    4580.89   6778.10    0.676  0.500501
wheelBP$wheelbase    194.92    53.51    3.643  0.000406 ***
wheelBP$stroke   -4809.30    961.20   -5.003  2.04e-06 ***
wheelBP$horsepower    81.90    10.48    7.815  2.87e-12 ***
wheelBP$citympg   -191.19    55.72   -3.431  0.000835 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1837 on 115 degrees of freedom
Multiple R-squared:  0.8135,    Adjusted R-squared:  0.807
F-statistic: 125.4 on 4 and 115 DF,  p-value: < 2.2e-16
```