Actividad Integradora 2

José Carlos Sánchez Gómez

2024-09-06

```
# Carga de datos
data = read.csv("C:\\Users\\jcsg6\\Downloads\\precios_autos.csv")
group_data = subset(data, select = c("symboling", "CarName", "fueltype",
"wheelbase" ,"horsepower", "price"))
group_data = data.frame(group_data)

# Gas = 1, Diesel = 0
group_data$fueltype = ifelse(group_data$fueltype == "gas", 1, 0)
group_data
```

```
##     symboling                          CarName fueltype wheelbase
horsepower
## 1           3               alfa-romero giulia        1      88.6
111
## 2           3              alfa-romero stelvio        1      88.6
111
## 3           1         alfa-romero Quadrifoglio        1      94.5
154
## 4           2                      audi 100 ls        1      99.8
102
## 5           2                       audi 100ls        1      99.4
115
## 6           2                         audi fox        1      99.8
110
## 7           1                       audi 100ls        1     105.8
110
## 8           1                        audi 5000        1     105.8
110
## 9           1                        audi 4000        1     105.8
140
## 10          0              audi 5000s (diesel)        1      99.5
160
## 11          2                         bmw 320i        1     101.2
101
## 12          0                         bmw 320i        1     101.2
101
## 13          0                           bmw x1        1     101.2
121
## 14          0                           bmw x3        1     101.2
121
## 15          1                           bmw z4        1     103.5
121
```

```
## 16          0                     bmw x4        1     103.5
182
## 17          0                     bmw x5        1     103.5
182
## 18          0                     bmw x3        1     110.0
182
## 19          2          chevrolet impala        1      88.4
48
## 20          1      chevrolet monte carlo        1      94.5
70
## 21          0        chevrolet vega 2300        1      94.5
70
## 22          1             dodge rampage        1      93.7
68
## 23          1        dodge challenger se        1      93.7
68
## 24          1                dodge d200        1      93.7
102
## 25          1          dodge monaco (sw)        1      93.7
68
## 26          1          dodge colt hardtop        1      93.7
68
## 27          1             dodge colt (sw)        1      93.7
68
## 28          1        dodge coronet custom        1      93.7
102
## 29         -1           dodge dart custom        1     103.3
88
## 30          3   dodge coronet custom (sw)        1      95.9
145
## 31          2                honda civic        1      86.6
58
## 32          2           honda civic cvcc        1      86.6
76
## 33          1                honda civic        1      93.7
60
## 34          1          honda accord cvcc        1      93.7
76
## 35          1           honda civic cvcc        1      93.7
76
## 36          0            honda accord lx        1      96.5
76
## 37          0         honda civic 1500 gl        1      96.5
76
## 38          0               honda accord        1      96.5
86
## 39          0           honda civic 1300        1      96.5
86
## 40          0              honda prelude        1      96.5
86
```

```
## 41          0              honda accord      1     96.5
86
## 42          0               honda civic      1     96.5
101
## 43          1        honda civic (auto)      1     96.5
100
## 44          0                isuzu MU-X      1     94.3
78
## 45          1               isuzu D-Max      1     94.5
70
## 46          0       isuzu D-Max V-Cross      1     94.5
70
## 47          2               isuzu D-Max      1     96.0
90
## 48          0                 jaguar xj      1    113.0
176
## 49          0                 jaguar xf      1    113.0
176
## 50          0                 jaguar xk      1    102.0
262
## 51          1                 maxda rx3      1     93.1
68
## 52          1          maxda glc deluxe      1     93.1
68
## 53          1           mazda rx2 coupe      1     93.1
68
## 54          1                mazda rx-4      1     93.1
68
## 55          1          mazda glc deluxe      1     93.1
68
## 56          3                 mazda 626      1     95.3
101
## 57          3                 mazda glc      1     95.3
101
## 58          3            mazda rx-7 gs      1     95.3
101
## 59          3               mazda glc 4      1     95.3
135
## 60          1                 mazda 626      1     98.8
84
## 61          0        mazda glc custom l      1     98.8
84
## 62          1          mazda glc custom      1     98.8
84
## 63          0                mazda rx-4      1     98.8
84
## 64          0          mazda glc deluxe      0     98.8
64
## 65          0                 mazda 626      1     98.8
84
```

```
## 66          0                       mazda glc        1     104.9
120
## 67          0                   mazda rx-7 gs        0     104.9
72
## 68         -1         buick electra 225 custom        0     110.0
123
## 69         -1         buick century luxus (sw)        0     110.0
123
## 70          0                   buick century        0     106.7
123
## 71         -1                   buick skyhawk        0     115.6
123
## 72         -1          buick opel isuzu deluxe        1     115.6
155
## 73          3                   buick skylark        1      96.6
155
## 74          0           buick century special        1     120.9
184
## 75          1 buick regal sport coupe (turbo)        1     112.0
184
## 76          1                  mercury cougar        1     102.7
175
## 77          2               mitsubishi mirage        1      93.7
68
## 78          2               mitsubishi lancer        1      93.7
68
## 79          2            mitsubishi outlander        1      93.7
68
## 80          1                   mitsubishi g4        1      93.0
102
## 81          3            mitsubishi mirage g4        1      96.3
116
## 82          3                   mitsubishi g4        1      96.3
88
## 83          3            mitsubishi outlander        1      95.9
145
## 84          3                   mitsubishi g4        1      95.9
145
## 85          3            mitsubishi mirage g4        1      95.9
145
## 86          1              mitsubishi montero        1      96.3
88
## 87          1               mitsubishi pajero        1      96.3
88
## 88          1            mitsubishi outlander        1      96.3
116
## 89         -1            mitsubishi mirage g4        1      96.3
116
## 90          1                    Nissan versa        1      94.5
69
```

```
## 91            1                  nissan gt-r        0      94.5
55
## 92            1                  nissan rogue       1      94.5
69
## 93            1                  nissan latio       1      94.5
69
## 94            1                  nissan titan       1      94.5
69
## 95            1                   nissan leaf       1      94.5
69
## 96            1                   nissan juke       1      94.5
69
## 97            1                  nissan latio       1      94.5
69
## 98            1                   nissan note       1      94.5
69
## 99            2                nissan clipper       1      95.1
69
## 100           0                  nissan rogue       1      97.2
97
## 101           0                  nissan nv200       1      97.2
97
## 102           0                   nissan dayz       1     100.4
152
## 103           0                   nissan fuga       1     100.4
152
## 104           0                   nissan otti       1     100.4
152
## 105           3                  nissan teana       1      91.3
160
## 106           3                  nissan kicks       1      91.3
200
## 107           1                nissan clipper       1      99.2
160
## 108           0                   peugeot 504       1     107.9
97
## 109           0                   peugeot 304       0     107.9
95
## 110           0              peugeot 504 (sw)       1     114.2
97
## 111           0                   peugeot 504       0     114.2
95
## 112           0                   peugeot 504       1     107.9
95
## 113           0                 peugeot 604sl       0     107.9
95
## 114           0                   peugeot 504       1     114.2
95
## 115           0     peugeot 505s turbo diesel       0     114.2
95
```

```
## 116          0                      peugeot 504        1    107.9
97
## 117          0                      peugeot 504        0    107.9
95
## 118          0                    peugeot 604sl        1    108.0
142
## 119          1                plymouth fury iii        1     93.7
68
## 120          1                 plymouth cricket        1     93.7
102
## 121          1                plymouth fury iii        1     93.7
68
## 122          1  plymouth satellite custom (sw)        1     93.7
68
## 123          1         plymouth fury gran sedan        1     93.7
68
## 124         -1                 plymouth valiant        1    103.3
88
## 125          3                  plymouth duster        1     95.9
145
## 126          3                    porsche macan        1     94.5
143
## 127          3                porcshce panamera        1     89.5
207
## 128          3                  porsche cayenne        1     89.5
207
## 129          3                   porsche boxter        1     89.5
207
## 130          1                  porsche cayenne        1     98.4
288
## 131          0                     renault 12tl        1     96.1
90
## 132          2                    renault 5 gtl        1     96.1
90
## 133          3                         saab 99e        1     99.1
110
## 134          2                        saab 99le        1     99.1
110
## 135          3                        saab 99le        1     99.1
110
## 136          2                       saab 99gle        1     99.1
110
## 137          3                       saab 99gle        1     99.1
160
## 138          2                         saab 99e        1     99.1
160
## 139          2                           subaru        1     93.7
69
## 140          2                        subaru dl        1     93.7
73
```

```
## 141          2                 subaru dl        1      93.3
73
## 142          0                    subaru        1      97.2
82
## 143          0                subaru brz        1      97.2
82
## 144          0               subaru baja        1      97.2
94
## 145          0                 subaru r1        1      97.0
82
## 146          0                 subaru r2        1      97.0
111
## 147          0             subaru trezia        1      97.0
82
## 148          0            subaru tribeca        1      97.0
94
## 149          0                 subaru dl        1      96.9
82
## 150          0                 subaru dl        1      96.9
111
## 151          1    toyota corona mark ii        1      95.7
62
## 152          1             toyota corona        1      95.7
62
## 153          1       toyota corolla 1200        1      95.7
62
## 154          0     toyota corona hardtop        1      95.7
62
## 155          0 toyota corolla 1600 (sw)        1      95.7
62
## 156          0             toyota carina        1      95.7
62
## 157          0            toyota mark ii        1      95.7
70
## 158          0       toyota corolla 1200        1      95.7
70
## 159          0             toyota corona        0      95.7
56
## 160          0            toyota corolla        0      95.7
56
## 161          0             toyota corona        1      95.7
70
## 162          0            toyota corolla        1      95.7
70
## 163          0            toyota mark ii        1      95.7
70
## 164          1   toyota corolla liftback        1      94.5
70
## 165          1             toyota corona        1      94.5
70
```

```
## 166          1        toyota celica gt liftback        1      94.5
112
## 167          1            toyota corolla tercel        1      94.5
112
## 168          2           toyota corona liftback        1      98.4
116
## 169          2                   toyota corolla        1      98.4
116
## 170          2                   toyota starlet        1      98.4
116
## 171          2                    toyota tercel        1      98.4
116
## 172          2                   toyota corolla        1      98.4
116
## 173          2                  toyota cressida        1      98.4
116
## 174         -1                   toyota corolla        1     102.4
92
## 175         -1                 toyota celica gt        0     102.4
73
## 176         -1                    toyota corona        1     102.4
92
## 177         -1                   toyota corolla        1     102.4
92
## 178         -1                   toyota mark ii        1     102.4
92
## 179          3          toyota corolla liftback        1     102.9
161
## 180          3                    toyota corona        1     102.9
161
## 181         -1                   toyota starlet        1     104.5
156
## 182         -1                   toyouta tercel        1     104.5
156
## 183          2                 vokswagen rabbit        0      97.3
52
## 184          2   volkswagen 1131 deluxe sedan        1      97.3
85
## 185          2             volkswagen model 111        0      97.3
52
## 186          2                volkswagen type 3        1      97.3
85
## 187          2              volkswagen 411 (sw)        1      97.3
85
## 188          2          volkswagen super beetle        0      97.3
68
## 189          2                volkswagen dasher        1      97.3
100
## 190          3                        vw dasher        1      94.5
90
```

```
## 191          3                         vw rabbit        1       94.5
90
## 192          0                volkswagen rabbit        1      100.4
110
## 193          0         volkswagen rabbit custom        0      100.4
68
## 194          0                volkswagen dasher        1      100.4
88
## 195         -2                  volvo 145e (sw)        1      104.3
114
## 196         -1                      volvo 144ea        1      104.3
114
## 197         -2                      volvo 244dl        1      104.3
114
## 198         -1                        volvo 245        1      104.3
114
## 199         -2                      volvo 264gl        1      104.3
162
## 200         -1                     volvo diesel        1      104.3
162
## 201         -1                  volvo 145e (sw)        1      109.1
114
## 202         -1                      volvo 144ea        1      109.1
160
## 203         -1                      volvo 244dl        1      109.1
134
## 204         -1                        volvo 246        0      109.1
106
## 205         -1                      volvo 264gl        1      109.1
114
##         price
## 1   13495.00
## 2   16500.00
## 3   16500.00
## 4   13950.00
## 5   17450.00
## 6   15250.00
## 7   17710.00
## 8   18920.00
## 9   23875.00
## 10  17859.17
## 11  16430.00
## 12  16925.00
## 13  20970.00
## 14  21105.00
## 15  24565.00
## 16  30760.00
## 17  41315.00
## 18  36880.00
## 19   5151.00
```

```
## 20   6295.00
## 21   6575.00
## 22   5572.00
## 23   6377.00
## 24   7957.00
## 25   6229.00
## 26   6692.00
## 27   7609.00
## 28   8558.00
## 29   8921.00
## 30  12964.00
## 31   6479.00
## 32   6855.00
## 33   5399.00
## 34   6529.00
## 35   7129.00
## 36   7295.00
## 37   7295.00
## 38   7895.00
## 39   9095.00
## 40   8845.00
## 41  10295.00
## 42  12945.00
## 43  10345.00
## 44   6785.00
## 45   8916.50
## 46   8916.50
## 47  11048.00
## 48  32250.00
## 49  35550.00
## 50  36000.00
## 51   5195.00
## 52   6095.00
## 53   6795.00
## 54   6695.00
## 55   7395.00
## 56  10945.00
## 57  11845.00
## 58  13645.00
## 59  15645.00
## 60   8845.00
## 61   8495.00
## 62  10595.00
## 63  10245.00
## 64  10795.00
## 65  11245.00
## 66  18280.00
## 67  18344.00
## 68  25552.00
## 69  28248.00
```

```
## 70  28176.00
## 71  31600.00
## 72  34184.00
## 73  35056.00
## 74  40960.00
## 75  45400.00
## 76  16503.00
## 77   5389.00
## 78   6189.00
## 79   6669.00
## 80   7689.00
## 81   9959.00
## 82   8499.00
## 83  12629.00
## 84  14869.00
## 85  14489.00
## 86   6989.00
## 87   8189.00
## 88   9279.00
## 89   9279.00
## 90   5499.00
## 91   7099.00
## 92   6649.00
## 93   6849.00
## 94   7349.00
## 95   7299.00
## 96   7799.00
## 97   7499.00
## 98   7999.00
## 99   8249.00
## 100  8949.00
## 101  9549.00
## 102 13499.00
## 103 14399.00
## 104 13499.00
## 105 17199.00
## 106 19699.00
## 107 18399.00
## 108 11900.00
## 109 13200.00
## 110 12440.00
## 111 13860.00
## 112 15580.00
## 113 16900.00
## 114 16695.00
## 115 17075.00
## 116 16630.00
## 117 17950.00
## 118 18150.00
## 119  5572.00
```

```
## 120  7957.00
## 121  6229.00
## 122  6692.00
## 123  7609.00
## 124  8921.00
## 125 12764.00
## 126 22018.00
## 127 32528.00
## 128 34028.00
## 129 37028.00
## 130 31400.50
## 131  9295.00
## 132  9895.00
## 133 11850.00
## 134 12170.00
## 135 15040.00
## 136 15510.00
## 137 18150.00
## 138 18620.00
## 139  5118.00
## 140  7053.00
## 141  7603.00
## 142  7126.00
## 143  7775.00
## 144  9960.00
## 145  9233.00
## 146 11259.00
## 147  7463.00
## 148 10198.00
## 149  8013.00
## 150 11694.00
## 151  5348.00
## 152  6338.00
## 153  6488.00
## 154  6918.00
## 155  7898.00
## 156  8778.00
## 157  6938.00
## 158  7198.00
## 159  7898.00
## 160  7788.00
## 161  7738.00
## 162  8358.00
## 163  9258.00
## 164  8058.00
## 165  8238.00
## 166  9298.00
## 167  9538.00
## 168  8449.00
## 169  9639.00
```

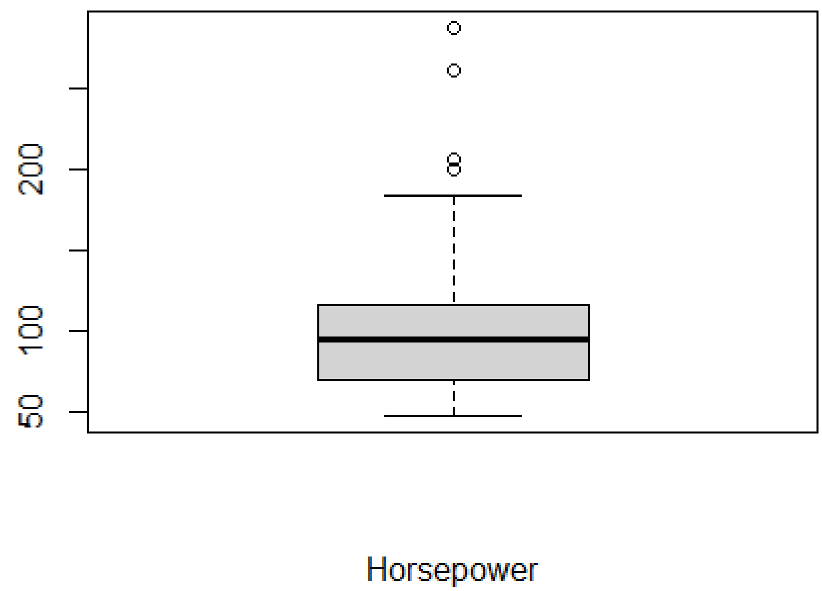
```
## 170  9989.00
## 171 11199.00
## 172 11549.00
## 173 17669.00
## 174  8948.00
## 175 10698.00
## 176  9988.00
## 177 10898.00
## 178 11248.00
## 179 16558.00
## 180 15998.00
## 181 15690.00
## 182 15750.00
## 183  7775.00
## 184  7975.00
## 185  7995.00
## 186  8195.00
## 187  8495.00
## 188  9495.00
## 189  9995.00
## 190 11595.00
## 191  9980.00
## 192 13295.00
## 193 13845.00
## 194 12290.00
## 195 12940.00
## 196 13415.00
## 197 15985.00
## 198 16515.00
## 199 18420.00
## 200 18950.00
## 201 16845.00
## 202 19045.00
## 203 21485.00
## 204 22470.00
## 205 22625.00
```

Exploracion de los datos

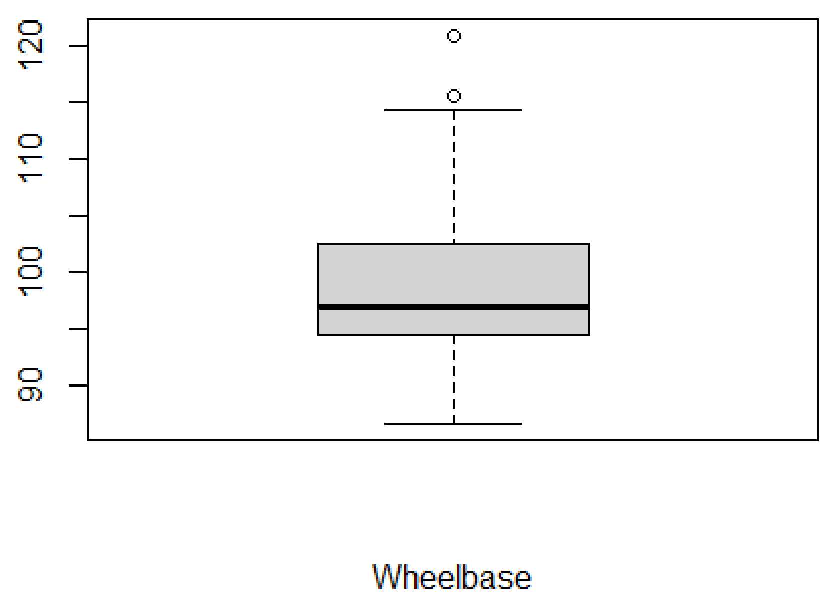
Medias estadisticas del modelo

```
cat("Datos sobre fueltype", "\n")

## Datos sobre fueltype

table(group_data$fueltype)

## 
##   0   1 
##  20 185

cat("\n", "Datos sobre horsepower: ", "\n")
```

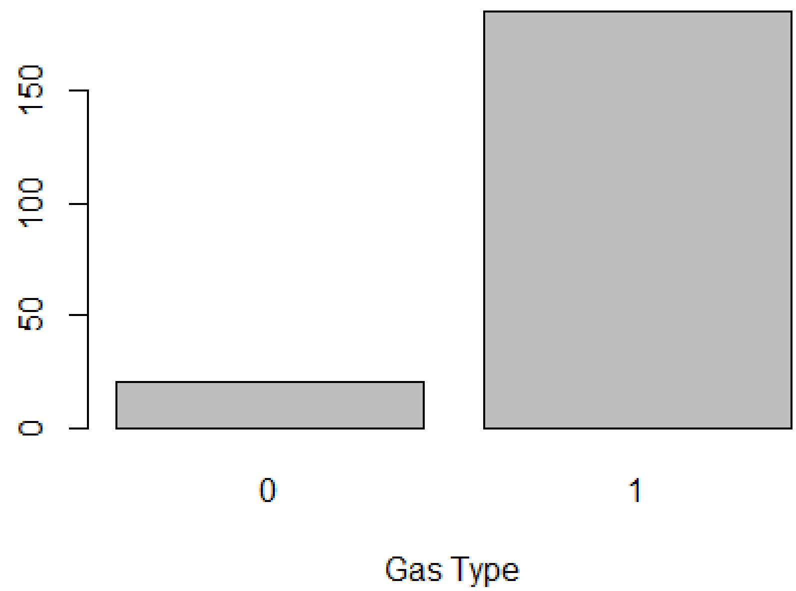
```
##
##  Datos sobre horsepower:

summary(group_data$horsepower)

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    48.0    70.0    95.0   104.1   116.0   288.0

cat("\n", "Datos sobre wheelbase", "\n")

##
##  Datos sobre wheelbase

summary(group_data$wheelbase)

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   86.60   94.50   97.00   98.76  102.40  120.90

cat("Correlacion entre los datos de gas", "\n", "\n")

## Correlacion entre los datos de gas
##

numeric_data = subset(group_data, select = c("symboling", "wheelbase",
"horsepower", "price"))

cor(numeric_data, use = "complete.obs")

##              symboling  wheelbase horsepower       price
## symboling   1.00000000 -0.5319537 0.07087272 -0.07997822
## wheelbase  -0.53195368  1.0000000 0.35329448  0.57781560
## horsepower  0.07087272  0.3532945 1.00000000  0.80813882
## price      -0.07997822  0.5778156 0.80813882  1.00000000

boxplot(group_data$horsepower, xlab = "Horsepower")
```




```
boxplot(group_data$wheelbase, xlab = "Wheelbase")
```




```
barplot(table(group_data$fueltype), xlab = "Gas Type")
```



Modelacion y Verificacion del Modelo

```
M1 = lm(price ~ horsepower * fueltype, data = group_data)
M2 = lm(price ~ wheelbase * horsepower, data = group_data)
```

Analisis del Modelo 1

```
#valor frontera del modelo
abs( qt(0.04/2,(length(M1))))

## [1] 2.302722
```

Hipotesis

vf = 2.30 $h_0 < vf $ El modelo no es significativo $h_1 >= vf $ El modelo es significativo

```
summary(M1)

##
## Call:
## lm(formula = price ~ horsepower * fueltype, data = group_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11904.3  -1776.2   -381.8   1458.9  19435.5
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)         -7731.37    3298.65  -2.344  0.02006 *
## horsepower            279.09      37.42   7.459 2.56e-12 ***
## fueltype             3016.83    3414.35   0.884  0.37798
## horsepower:fueltype  -112.36      38.21  -2.940  0.00366 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4234 on 201 degrees of freedom
## Multiple R-squared:  0.7233, Adjusted R-squared:  0.7191
## F-statistic: 175.1 on 3 and 201 DF,  p-value: < 2.2e-16
```

```r
anova(M1)
```

```
## Analysis of Variance Table
##
## Response: price
##                      Df     Sum Sq    Mean Sq  F value    Pr(>F)
## horsepower            1 8502974873 8502974873 474.3725 < 2.2e-16 ***
## fueltype              1  758833988  758833988  42.3346 5.993e-10 ***
## horsepower:fueltype   1  154969609  154969609   8.6456  0.003663 **
## Residuals           201 3602860891   17924681
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Se rechaza h_0 porque al menos un valor f sobrepasa nuestro alfa.

Significancia de Bi

$h_0 < afa\ h_1 > alfa$

```r
summary(M1)$coefficients
```

```
##                      Estimate Std. Error    t value     Pr(>|t|)
## (Intercept)         -7731.3655 3298.64586 -2.3437998 2.006495e-02
## horsepower            279.0943   37.41715  7.4589949 2.564780e-12
## fueltype             3016.8270 3414.34521  0.8835741 3.779818e-01
## horsepower:fueltype  -112.3606   38.21346 -2.9403400 3.662756e-03
```

Se acepta h_0 porque todos los valores p de los coeficientes son menores a alfa.

Porcentaje de variacion del modelo

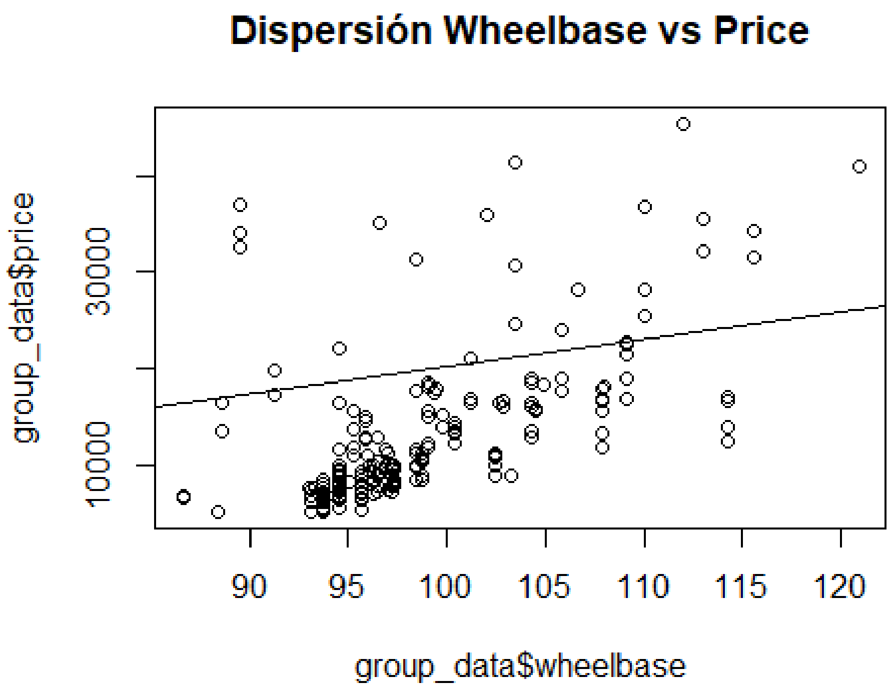
```r
summary(M1)$r.squared
```

```
## [1] 0.7232749
```

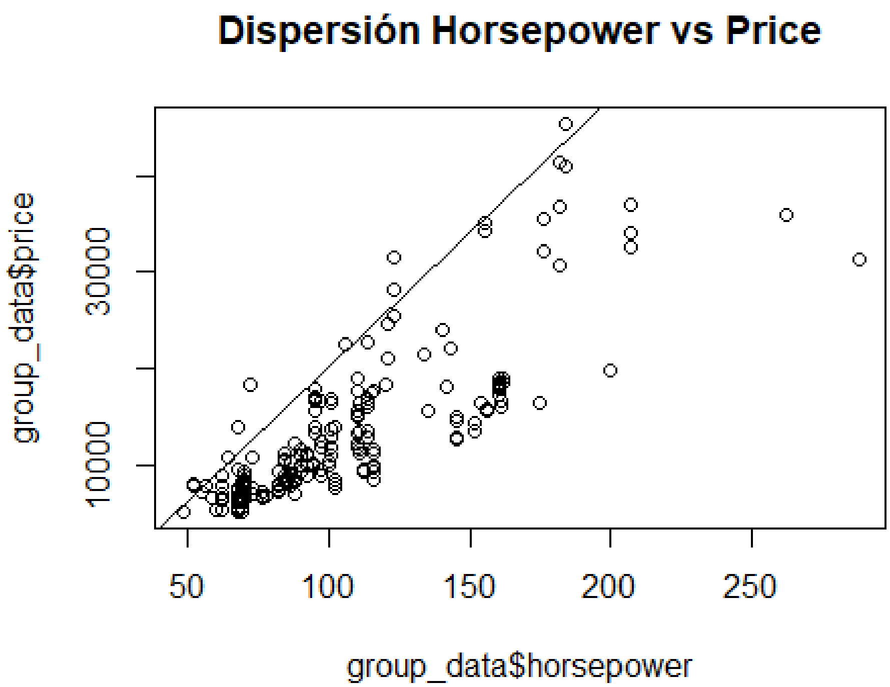
Diagrama de dispersion

```r
plot(group_data$wheelbase, group_data$price, main = "Dispersión Wheelbase
vs Price")
abline(M1)
```

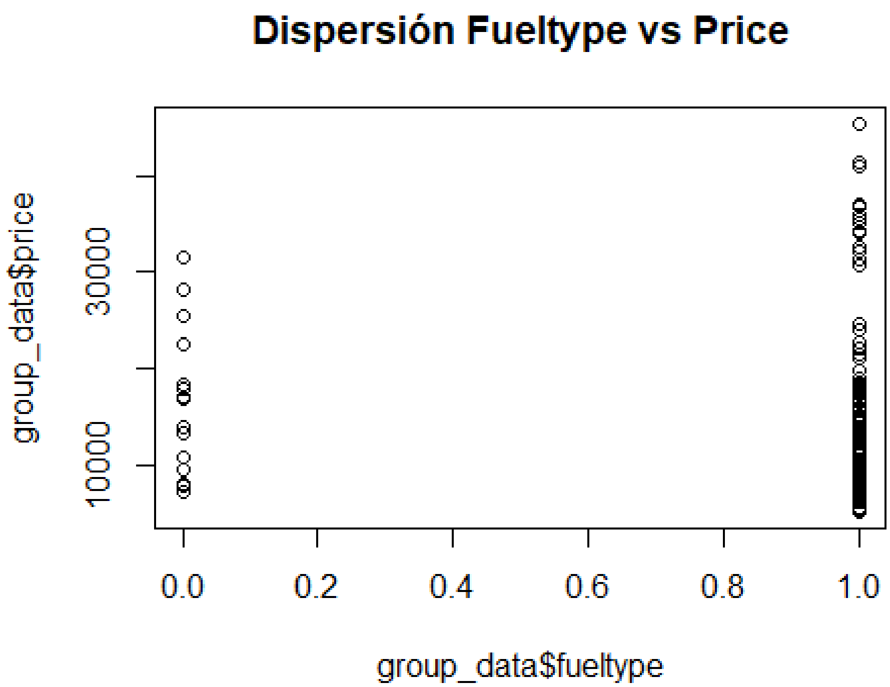
```
## Warning in abline(M1): only using the first two of 4 regression
coefficients
```




```
plot(group_data$horsepower, group_data$price, main = "Dispersión
Horsepower vs Price")
abline(M1)

## Warning in abline(M1): only using the first two of 4 regression
coefficients
```



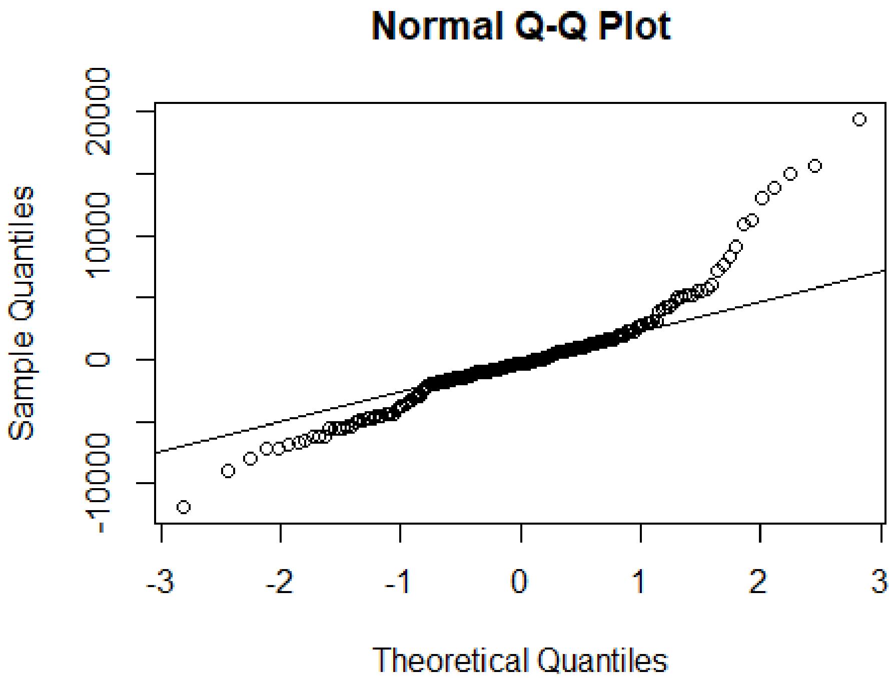

```
plot(group_data$fueltype, group_data$price, main = "Dispersión Fueltype
vs Price")
abline(M1)

## Warning in abline(M1): only using the first two of 4 regression
coefficients
```



Se analiza la posible relación entre wheelbase y price, horsepower y price, y fueltype y price. Podemos observar que el precio aumenta considerablemente si el fueltype es diesel. De la misma manera con horsepower, mientras este aumenta, de igual manera lo hace el precio, y exactamente el mismo comportamiento sucede con wheelbase.

Validez del modelo propuesto

```
library(nortest)
ad.test(M1$residuals)

##
##  Anderson-Darling normality test
##
## data:  M1$residuals
## A = 4.7325, p-value = 9.266e-12

qqnorm(M1$residuals)
qqline(M1$residuals)
```



Se rechaza h_0 porque el valor p es menor que alfa (0.04).

Verificacion de media 0

```
t.test(M1$residuals)

##
##  One Sample t-test
##
## data:  M1$residuals
## t = -1.5791e-15, df = 204, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -578.714  578.714
## sample estimates:
##     mean of x
## -4.634905e-13
```

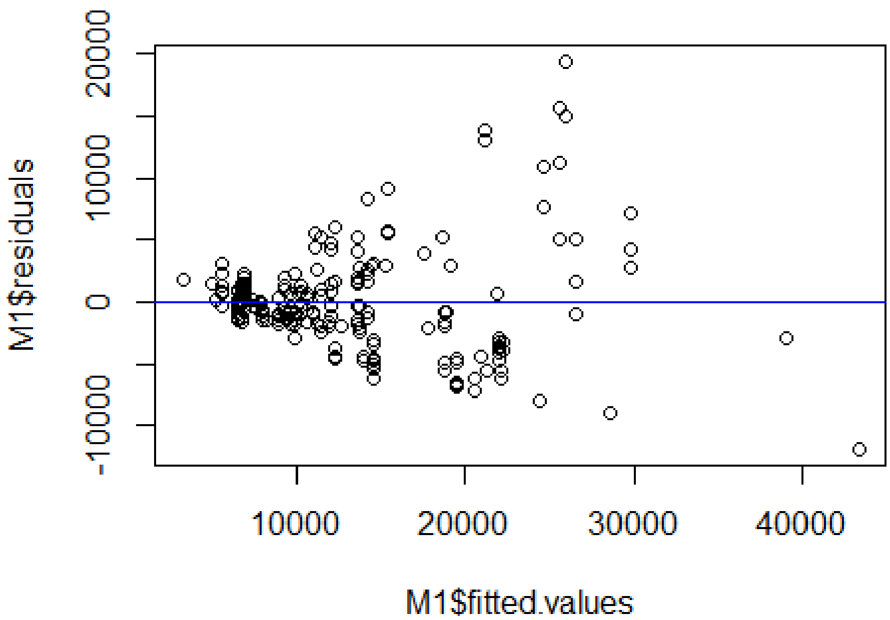
Se rechaza h_0 porque el valor p de la media cero no es igual a 0, por lo que se acepta h_1

Homocedasticidad

```
library(lmtest)

## Cargando paquete requerido: zoo

##
## Adjuntando el paquete: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```r
plot(M1$fitted.values, M1$residuals)
abline(h=0, col="blue")
```




```r
bptest(M1)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  M1
## BP = 62.878, df = 3, p-value = 1.426e-13
```

```r
dwtest(M1)
```

```
##
##  Durbin-Watson test
##
## data:  M1
## DW = 1.0589, p-value = 3.272e-12
## alternative hypothesis: true autocorrelation is greater than 0
```

Se acepta h_0 debido a que el valor p que nos da la prueba de BP es mucho menor a 0.05, por lo que no hay suficiente evidencia de que existe heterocedasticidad.

Debido a que el valor de la prueba de Durbin-Watson es cercano a 2, podemos entender que los residuos son independientes, por lo que aceptamos h_0 ya que los errores no estan correlacionados.

```
library(lmtest)
bptest(M1)

##
##  studentized Breusch-Pagan test
##
## data:  M1
## BP = 62.878, df = 3, p-value = 1.426e-13
```

Analisis del Modelo 2

```
#valor frontera del modelo
abs( qt(0.04/2,(length(M2))))

## [1] 2.302722
```

Hipotesis

vf = 2.30 \$h_0 < vf \$ El modelo no es significativo \$h_1 >= vf \$ El modelo es significativo

```
summary(M2)

##
## Call:
## lm(formula = price ~ wheelbase * horsepower, data = group_data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -8847  -2050   -177   1350  15889
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)           -17059.574  14377.287  -1.187   0.2368
## wheelbase                155.900    148.256   1.052   0.2943
## horsepower               -89.721    111.777  -0.803   0.4231
## wheelbase:horsepower       2.342      1.140   2.055   0.0412 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3977 on 201 degrees of freedom
## Multiple R-squared:  0.7558, Adjusted R-squared:  0.7522
## F-statistic: 207.4 on 3 and 201 DF,  p-value: < 2.2e-16

anova(M2)

## Analysis of Variance Table
##
## Response: price
```

```
##                       Df     Sum Sq    Mean Sq  F value Pr(>F)    
## wheelbase              1 4346878264 4346878264 274.8574 <2e-16 ***
## horsepower             1 5427172318 5427172318 343.1654 <2e-16 ***
## wheelbase:horsepower   1   66767095   66767095   4.2217 0.0412 *  
## Residuals            201 3178821685   15815033                    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Se rechaza h_0 porque al menos un valor f sobrepasa nuestro alfa.

Significancia de Bi

$h_0 < afa \; h_1 > alfa$

```
summary(M2)$coefficients
```

```
##                          Estimate   Std. Error    t value   Pr(>|t|)
## (Intercept)          -17059.573669 14377.287334 -1.1865641 0.23680017
## wheelbase               155.899761   148.256466  1.0515545 0.29426654
## horsepower              -89.720766   111.776866 -0.8026774 0.42310928
## wheelbase:horsepower      2.342437     1.140044  2.0546894 0.04120263
```

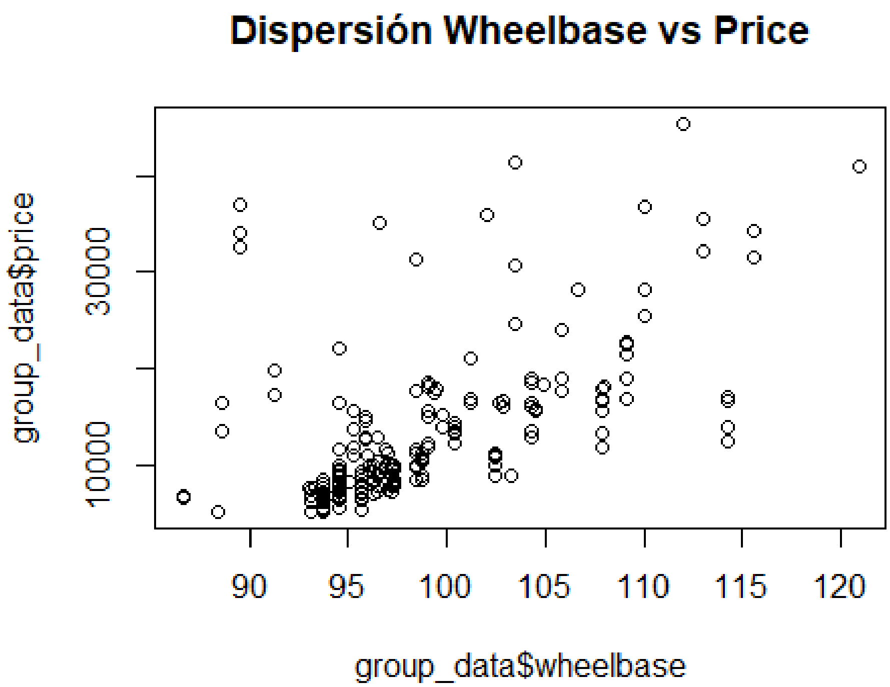
Se acepta h_0 porque todos los valores p de los coeficientes son menores a alfa.

Porcentaje de variacion del modelo

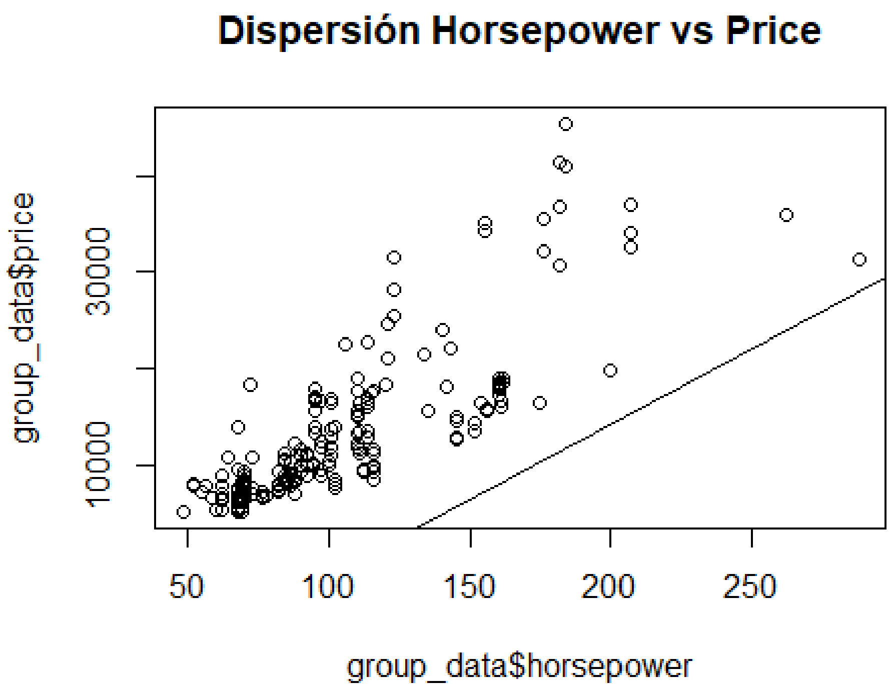
```
summary(M2)$r.squared

## [1] 0.7558441
```

Diagrama de dispersion

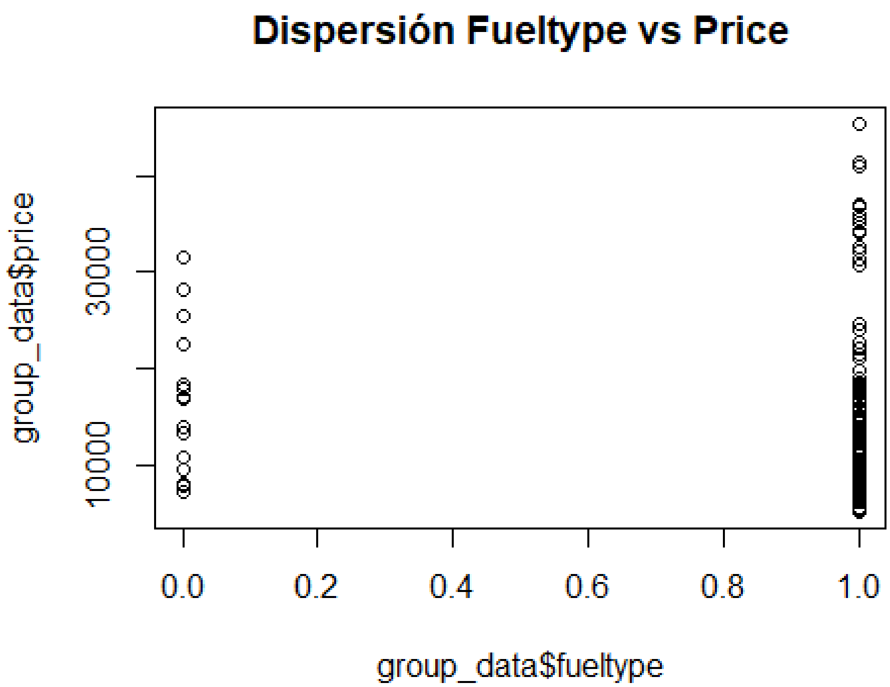
```
plot(group_data$wheelbase, group_data$price, main = "Dispersión Wheelbase
vs Price")
abline(M2)

## Warning in abline(M2): only using the first two of 4 regression
coefficients
```




```
plot(group_data$horsepower, group_data$price, main = "Dispersión
Horsepower vs Price")
abline(M2)

## Warning in abline(M2): only using the first two of 4 regression
coefficients
```



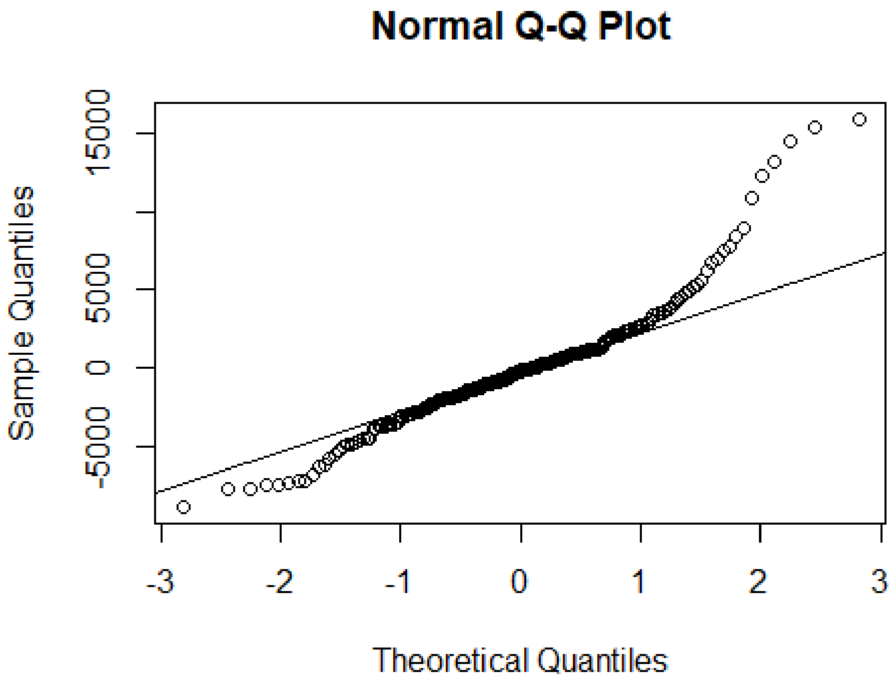

```
plot(group_data$fueltype, group_data$price, main = "Dispersión Fueltype
vs Price")
abline(M2)

## Warning in abline(M2): only using the first two of 4 regression
coefficients
```



Se analiza la posible relación entre wheelbase y price, horsepower y price, y fueltype y price. Podemos observar que el precio aumenta considerablemente si el fueltype es diesel. De la misma manera con horsepower, mientras este aumenta, de igual manera lo hace el precio, y exactamente el mismo comportamiento sucede con wheelbase.

Validez del modelo propuesto

```
library(nortest)
ad.test(M2$residuals)

##
##  Anderson-Darling normality test
##
## data:  M2$residuals
## A = 3.6742, p-value = 3.374e-09

qqnorm(M2$residuals)
qqline(M2$residuals)
```



Se rechaza h_0 porque el valor p es menor que alfa (0.04).

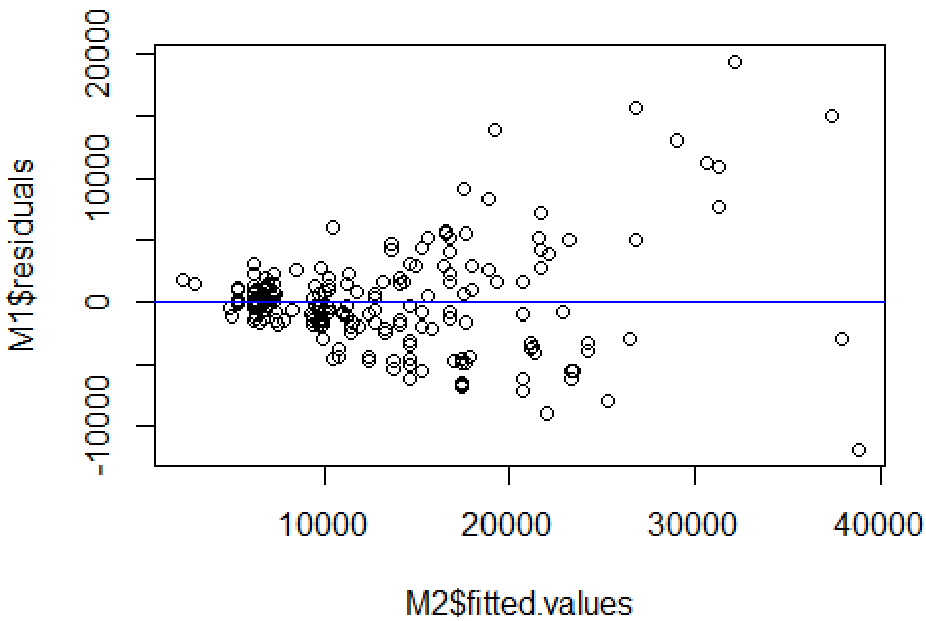
Verificacion de media 0

```
t.test(M2$residuals)

##
##  One Sample t-test
##
## data:  M2$residuals
## t = -5.7484e-17, df = 204, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -543.5923  543.5923
## sample estimates:
##     mean of x
## -1.584857e-14
```

Se rechaza h_0 porque el valor p de la media cero no es igual a 0, por lo que se acepta h_1

Homocedasticidad

```
library(lmtest)
plot(M2$fitted.values, M1$residuals)
abline(h=0, col="blue")
```




```r
bptest(M2)

##
##  studentized Breusch-Pagan test
##
## data:  M2
## BP = 60.863, df = 3, p-value = 3.845e-13

dwtest(M2)

##
##  Durbin-Watson test
##
## data:  M2
## DW = 1.0509, p-value = 1.575e-12
## alternative hypothesis: true autocorrelation is greater than 0
```

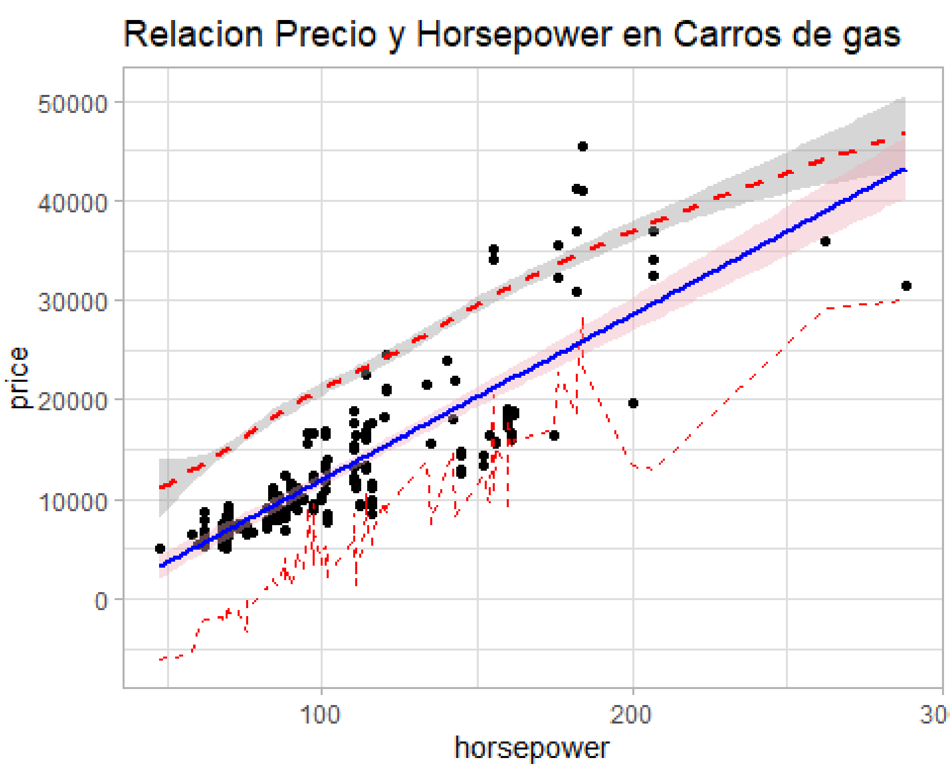
Se acepta h_0 debido a que el valor p que nos da la prueba de BP es mucho menor a 0.05, por lo que no hay suficiente evidencia de que existe heterocedasticidad.

Debido a que el valor de la prueba de Durbin-Watson es cercano a 2, podemos entender que los residuos son independientes, por lo que aceptamos h_0 ya que los errores no estan correlacionados.

Tras ver los resultados de las pruebas de análisis, podemos concluir que el mejor modelo es el segundo, pues nos ofrece un mejor porcentaje de variacion de los datos, además de esto, los coeficientes de este modelo presentan una mayor significancia para el modelo.

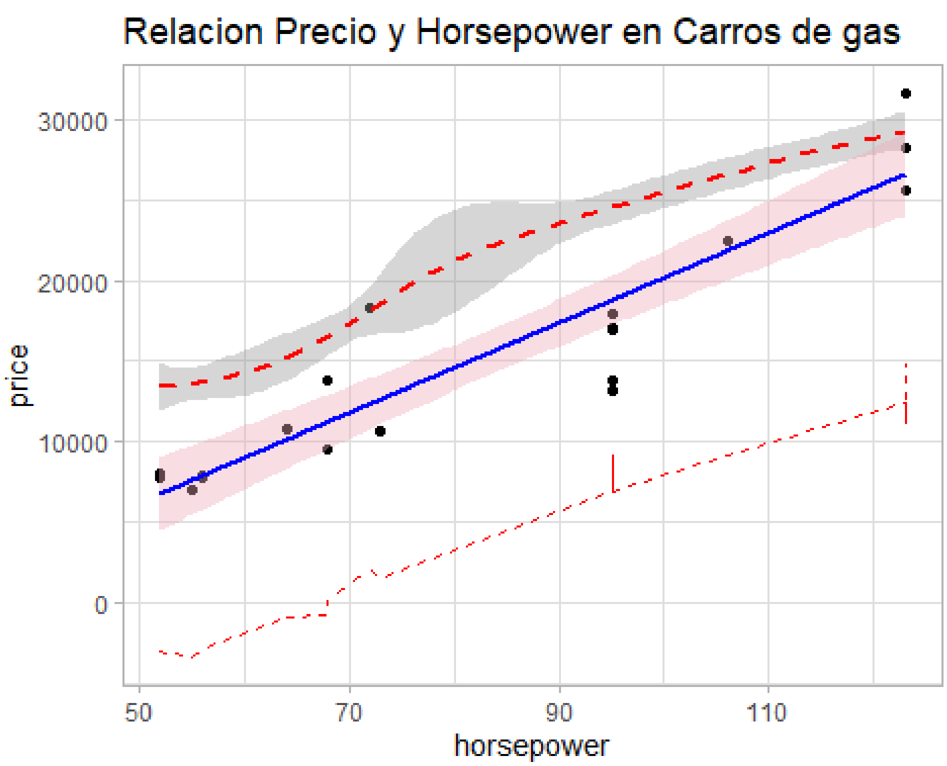
Intervalos de prediccion y significancia

```
A = M2
Ip=predict(object=A,interval="prediction",level=0.96)

## Warning in predict.lm(object = A, interval = "prediction", level =
0.96): predictions on current data refer to _future_ responses

M2=cbind(group_data,Ip)
M2g = subset(M2, fueltype == 1)
M2d = subset(M2, fueltype == 0)

library(ggplot2)
ggplot(M2g,aes(x= horsepower ,y= price))+
  ggtitle("Relacion Precio y Horsepower en Carros de gas")+
  geom_point()+
  geom_line(aes(y=lwr), color="red", linetype="dashed")+
  geom_smooth(aes(y=upr), color="red", linetype="dashed")+
  geom_smooth(method=lm, formula=y~x, se=TRUE, level=0.96, col="blue",
fill="pink2")+
  theme_light()

## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```




```
library(ggplot2)
ggplot(M2d,aes(x= horsepower ,y= price))+
```

```
  ggtitle("Relacion Precio y Horsepower en Carros de gas")+
  geom_point()+
  geom_line(aes(y=lwr), color="red", linetype="dashed")+
  geom_smooth(aes(y=upr), color="red", linetype="dashed")+
  geom_smooth(method=lm, formula=y~x, se=TRUE, level=0.96, col="blue",
fill="pink2")+
  theme_light()

## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```



Debido a problemas con el paquete de graficas no se pudo hacer unas graficas de prediccion y confianza que explique correctamente el comportamiento de las variables predecidas. Sin embargo, debido al análisis estadístico anterior podemos decir que este Modelo, es el que más nos puede servir para predecir el precio de un automovil, con relacion a las variables de caballos de fuerza y wheelbase. Si bien, no se puede apreciar correctamente los limites, podemos ver que la recta que se genera, atraveza la mayoria de los puntos, lo que confirma nuestra probabilidad de varianza que tiene un valor de 0.75.

Analisis final

```
quantitative_data = data[, c('wheelbase', 'carlength', 'carwidth',
'carheight', 'enginesize',
                          'stroke', 'compressionratio', 'horsepower',
'peakrpm',
                          'citympg', 'highwaympg', 'curbweight',
```

```
'price')]
cor(quantitative_data)

##                   wheelbase  carlength   carwidth   carheight
enginesize
## wheelbase         1.0000000  0.8745875  0.7951436  0.58943476
0.56932868
## carlength         0.8745875  1.0000000  0.8411183  0.49102946
0.68335987
## carwidth          0.7951436  0.8411183  1.0000000  0.27921032
0.73543340
## carheight         0.5894348  0.4910295  0.2792103  1.00000000
0.06714874
## enginesize        0.5693287  0.6833599  0.7354334  0.06714874
1.00000000
## stroke            0.1609590  0.1295326  0.1829417 -0.05530667
0.20312859
## compressionratio  0.2497858  0.1584137  0.1811286  0.26121423
0.02897136
## horsepower        0.3532945  0.5526230  0.6407321 -0.10880206
0.80976865
## peakrpm          -0.3604687 -0.2872422 -0.2200123 -0.32041072 -
0.24465983
## citympg          -0.4704136 -0.6709087 -0.6427043 -0.04863963 -
0.65365792
## highwaympg       -0.5440819 -0.7046616 -0.6772179 -0.10735763 -
0.67746991
## curbweight        0.7763863  0.8777285  0.8670325  0.29557173
0.85059407
## price             0.5778156  0.6829200  0.7593253  0.11933623
0.87414480
##                       stroke compressionratio  horsepower     peakrpm
## wheelbase         0.16095905       0.24978585  0.35329448 -0.36046875
## carlength         0.12953261       0.15841371  0.55262297 -0.28724220
## carwidth          0.18294169       0.18112863  0.64073208 -0.22001230
## carheight        -0.05530667       0.26121423 -0.10880206 -0.32041072
## enginesize        0.20312859       0.02897136  0.80976865 -0.24465983
## stroke            1.00000000       0.18611011  0.08093954 -0.06796375
## compressionratio  0.18611011       1.00000000 -0.20432623 -0.43574051
## horsepower        0.08093954      -0.20432623  1.00000000  0.13107251
## peakrpm          -0.06796375      -0.43574051  0.13107251  1.00000000
## citympg          -0.04214475       0.32470142 -0.80145618 -0.11354438
## highwaympg       -0.04393093       0.26520139 -0.77054389 -0.05427481
## curbweight        0.16879004       0.15136174  0.75073925 -0.26624318
## price             0.07944308       0.06798351  0.80813882 -0.08526715
##                      citympg  highwaympg curbweight       price
## wheelbase        -0.47041361 -0.54408192  0.7763863  0.57781560
## carlength        -0.67090866 -0.70466160  0.8777285  0.68292002
## carwidth         -0.64270434 -0.67721792  0.8670325  0.75932530
## carheight        -0.04863963 -0.10735763  0.2955717  0.11933623
```

```
## enginesize       -0.65365792 -0.67746991  0.8505941  0.87414480
## stroke           -0.04214475 -0.04393093  0.1687900  0.07944308
## compressionratio  0.32470142  0.26520139  0.1513617  0.06798351
## horsepower       -0.80145618 -0.77054389  0.7507393  0.80813882
## peakrpm          -0.11354438 -0.05427481 -0.2662432 -0.08526715
## citympg           1.00000000  0.97133704 -0.7574138 -0.68575134
## highwaympg        0.97133704  1.00000000 -0.7974648 -0.69759909
## curbweight       -0.75741378 -0.79746479  1.0000000  0.83530488
## price            -0.68575134 -0.69759909  0.8353049  1.00000000
```

Considero que un buen grupo de datos para poder predecir el precio serian horsepower junto con enginesize puesto que entre estas dos variables existe una gran correlación, además de que cada una de ellas tienen una gran correlación con el precio de un auto.