Reporte Final: El precio de los autos

TC3004B.104 Inteligencia Artificial Avanzada para la Ciencia de Datos I

Luis Ángel Guzmán Iribe - A017471757

12 de Septiembre de 2023

Índice

1 Resumen

En este trabajo se realiza un análisis sobre un conjunto de datos de características de automóviles con 205 instancias que contienen variables cuantitativas y cualitativas que describen las características del automóvil. Se realizaron análisis sobre estas variables como distribución de frecuencia con histogramas, colinealidad, distribución de cuartiles y valores atípicos en el caso de las variables cuantitativas; en el caso de las categóricas, se realizaron gráficas de histograma, pastel y la relación que cada categoría tiene con el precio.

Las variables seleccionadas para el modelo fueron: *carwidth*, *curbweight*, *enginesize*, *horsepower*, *citympg*, *highwaympg* y *carbody*. Se encontró una alta significancia del modelo, con un valor de R cuadrada ajustada de 0.88, así como normalidad y homocedasticidad en los residuos.

2 Introducción

Bajo el escenario de esta actividad, una empresa automovilística china pretende entrar al mercado americano, para esto, contratan los servicios de una consultora para realizar un análisis de mercado con la finalidad de determinar qué variables de un carro son más significativas al momento de estimar el precio de un automóvil, así como generar un modelo predictivo que tome en cuentas estas características para predecir el precio de un automóvil.

Este estudio se apoya en un conjunto de datos compuesto por 205 registros que contiene información de una variedad de carros en el mercado, que nos permiten generar un modelo de regresión que utilice las variables que se describirán a continuación que nos permite conocer las variables más influenciables sobre el precio de un auto, así como la capacidad que tiene nuestro modelo de realizar predicciones sobre el valor de un auto dadas sus características.

El proceso de análisis se realiza utilizando el lenguaje y herramienta de análisis estadístico R, que se utiliza con la finalidad de facilitar el análisis de la información, limpieza del conjunto de datos, y la generación del modelo de regresión deseado para la solución del problema, así como el análisis de la validez del mismo.

La correcta implementación de las técnicas de análisis y predicción que se describen a continuación pueden resultar vitales para el éxito de la hipotética empresa china que consulta el estudio, dado que influenciados en los resultados obtenidos, se tomarían decisiones estratégicas que prueben ser instrumentales en su incursión en el mercado automovilístico estadounidense.

A medida que avancemos, detallaremos en profundidad la metodología utilizada para llevar a cabo este análisis y posteriormente presentaremos los resultados y conclusiones que se desprenden de esta evaluación.

3 Análisis de variables

Antes de realizar el modelo que se utilizará para ejecutar predicciones sobre el precio de un determinado vehículo, es necesario analizar qué variables muestran correlaciones significativas que hagan que merezca la pena incluirlas en el modelo preliminar. El conjunto de datos con el que se trabaja contiene 21 variables que describen diversas características de un carro, estas variables son las siguientes:

Variable	Descripción	Tipo
Symboling	Su calificación de riesgo asegurado asignada. Un valor de +3 indica que el auto tiene un alto riesgo, -3 que probablemente es bastante seguro.	Categórico
CarName	Nombre de la compañía automotriz	Categórico
fueltype	Tipo de combustible del auto, es decir, gasolina o diésel	Categórico
carbody	Cuerpo del auto	Categórico
drivewheel	Tipo de tracción	Categórico
enginelocation	Ubicación del motor del auto	Categórico
wheelbase	Distancia entre ejes del auto	Numérico
carlength	Longitud del auto	Numérico

Variable	Descripción	Tipo
carwidth	Ancho del auto	Numérico
carheight	Altura del auto	Numérico
curbweight	El peso de un auto sin ocupantes o equipaje.	Numérico
enginetype	Tipo de motor.	Categórico
cylindernumber	Cilindro ubicado en el auto	Categórico
enginesize	Tamaño del auto	Numérico
stroke	Carrera o volumen dentro del motor	Numérico
compressionratio	Relación de compresión del auto	Numérico
horsepower	Potencia en caballos de fuerza	Numérico
peakrpm	RPM máximo del auto	Numérico
citympg	Consumo de combustible en ciudad	Numérico
highwaympg	Consumo de combustible en carretera	Numérico
price	Precio del auto (Variable dependiente)	Numérico

Como podemos observar, tenemos una variedad de variables tanto numéricas como categóricas (también conocidas como cuantitativas y cualitativas, respectivamente), el objetivo de esta fase del escrito es encontrar aquellas variables que muestran altas correlaciones con nuestra variable objetivo (price).

Para decidir qué variables es pertinente incluir se consideran 3 criterios principales, en los que se basará todo nuestro análisis de variables. Estos criterios son los siguientes:

1. Todas estas variables presentan altos índices de correlación (>0.75) con relación al precio del vehículo.
2. En el caso de las variables cualitativas, no existe una categoría que domine considerablemente la frecuencia en comparación a otras variables, además, es claro que la media del promedio por categoría varía considerablemente.
3. Cuentan con pocos valores atípicos apreciables en sus boxplots, por lo que retirar estos valores no afectará a la relevancia estadística de estas variables.

Una vez establecidos estos criterios, podemos proceder al análisis para elegir nuestras variables candidatas para la implementación del modelo de regresión. A continuación se encuentran 2 secciones correspondientes para el análisis de variables numéricas y categóricas.

3.1 Variables numéricas

Ver anexo

3.1.1 Histogramas

El uso de histogramas permite hacernos una idea general de la distribución de los datos, permitiéndonos conocer si estos cuentan con una distribución normal, sesgada, y explorar valores atípicos de manera visual e intuitiva. Realizamos este análisis únicamente sobre las variables numéricas ya que no es posible calcular la distribución estadística de variables categóricas, pero sí de variables reales (continuas y discretas). Conocer la distribución de los datos también nos permite elegir candidatos para transformaciones (cómo box cox), en caso de que sea necesario normalizar los datos para encontrar mejores resultados en la distribución de residuos del modelo.

Ver anexo

3.1.2 Diagramas de caja y bigote

Los diagramas de caja y bigote (o box plot) cumplen en ciertos aspectos la misma función que los histogramas, presentándonos la distribución con cuartiles de los datos, lo que nos permite detectar sesgos de manera sencilla, al igual que los valores atípicos.

Ver anexo

3.1.3 Diagramas de dispersión

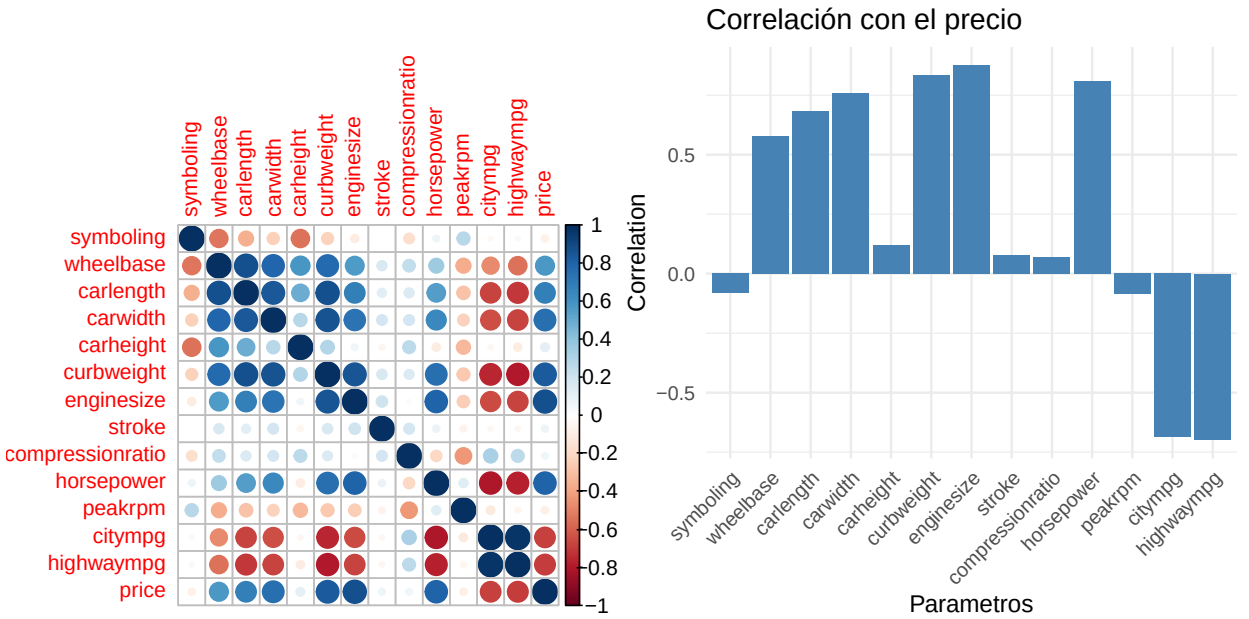
Los diagramas de dispersión se generan tomando como eje x a la variable independiente y a la el eje y toma el precio. Esto nos permite apreciar de manera visual la manera en la que se relaciona cada variable numérica con el precio de un vehículo.

Ver anexo

3.1.4 Matriz de correlación

La matriz de correlación genera un indicador numérico que indica el nivel de colinealidad que existe entre cada variable, esto nos permite conocer 2 cosas, una cantidad cuantitativa que describe el comportamiento observado en los diagramas de dispersión, permitiendo realizar evaluaciones más rigurosas sobre las variables que se consideran relevantes para el modelo. En segundo lugar, esta matriz nos permite conocer variables independientes que resulten redundantes, ya que si estas resultan ser colineales, podemos descartar una de las 2 y quedarnos con la que mejor se ajuste al precio, reduciendo de este modo la complejidad del modelo.

En la gráficas que se muestran a continuación, podemos apreciar los indices de correlación entre las variables, y más importante, las correlaciones con el precio representadas como histograma. Estos valores erán de suma importancia al momento de decidir sobre que variables incluir en el modelo.



3.2 Variables categóricas

Ver anexo

3.2.1 Gráficas de pastel

Las gráficas de pastel nos permiten conocer las distribuciones de las categorías de las variables cuantitativas, esto nos permite descartar variables dominadas por una sola categoría que pudieran resultar no significativas al afectar tan solo una pequeña porción de los datos.

Ver anexo

3.2.2 Media de precio por categoría

Graficado como histograma, este cálculo facilita la identificación de categorías significativas, si por ejemplo encontramos que una variable muestra diferencias considerables de precio por categoría, podría ser relevante

para incluirlo en nuestro análisis.

Ver anexo

3.3 Selección de variables significativas

Con base en el análisis y criterios mencionados anteriormente, se seleccionan 7 variables para el modelo preliminar.

Variables cuantitativas seleccionadas:

- Car Width
- Curb Weigth
- Enginesize
- Horsepower
- Citympg
- Highwaympg

Todas estas variables cuentan con altos índices de correlación con el precio del vehículo, y cuentan con pocos o ningún valor atípico que pudiera ocasionar sesgo en los datos; sin embargo, es posible apreciar en los histogramas que en la práctica totalidad de estas variables existen considerables sesgos o desviaciones de normalidad, por lo que se tratará de reducir estos efectos con transformaciones.

Variables cualitativas seleccionadas:

- Carbody

Esta variable se seleccionó porque existe una distribución relativamente balanceada de los tipos de cuerpo de vehículo existentes en el dataset, sin existir ninguna clase decisivamente dominante sobre las demás, también se muestran diferencias considerables de media de precio entre las diferentes categorías, lo que sugiere que esta variable es significativa para determinar el precio de un auto. Es importante señalar que será necesario transformar esta variable en variables dummies, para facilitar la integración con el resto de variables cuantiativas.

Estas variables seleccionadas constituyen la base del modelo preliminar y se espera que proporcionen una representación precisa y confiable de los factores que influyen en el precio de los vehículos en nuestro conjunto de datos. Es importante destacar que el proceso de selección de variables es iterativo y, en etapas posteriores, se continuará evaluando la relevancia de estas variables y se considerarán ajustes si es necesario.

4 Transformación de variables

En iteraciones pasadas del trabajo, se encontró que el uso de las variables en su estado actual producen modelos estadísticamente significativos pero con altos niveles de heterocedasticidad en los residuos. Con la finalidad de reducir esto, se emplean transformaciones de box-cox, buscando incrementar la homocedasticidad de los residuos, además de reducir el impacto que tienen los valores atípicos sobre el modelo.

En este caso podemos aplicar la ya mencionada transformación box-cox debido a que todos los valores de nuestras variables numéricas son superiores a 0, cayendo perfectamente en el rango efectivo de la transformación.

Ver anexo

5 Modelo de regresión lineal

Una vez decididas las variables finales para la implementación del modelo, generamos una regresión lineal utilizando la función *lm* de R. Posteriormente, se hace uso de la fucnión *step*, una función que de manera iterativa retira variables consideradas como no significativas para el modelo. Sorprendentemente, y a diferencia

de pasadas iteraciones de este trabajo, la función determinó que todas las variables presentes en el modelo son estadisticamente significativas para el modelo.

5.1 Verificación del modelo

Posteriormente, se genera el resumen del modelo, con la finalidad de encontrar más en detalle la significancia del modelo. El modelo en general tiene un p-value de 0, un excelente valor que confirma su significacia estadistica, a su vez, contamos con un valor de R cuadrada ajustada de 0.88, un excelente indicador de la capacidad del modelo de explicar la varianza presente en los resultados.

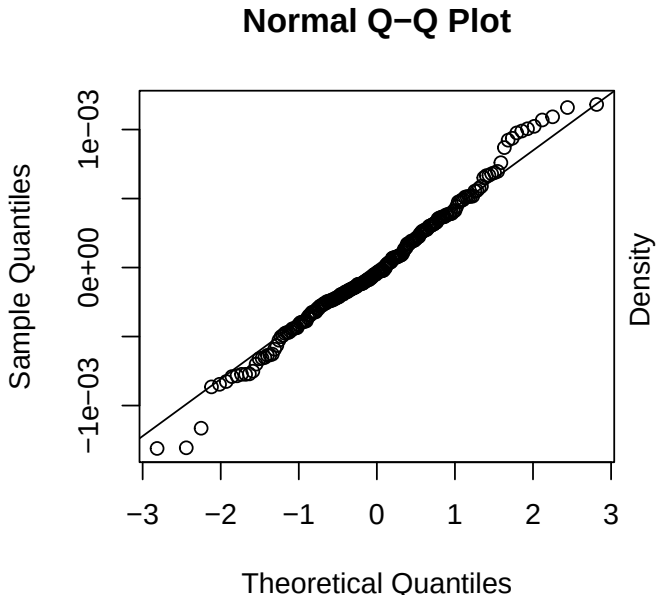
Al analizar la significacia de los coeficientes de manera individual encontramos que todos tienen valores p cercanos a 0, con el más elevado siedo 0.07 de la variable *enginesize*, que si bien no entra dentro del marjen de confainza buscado de 0.05, considero que la diferencia no es suficiente como para retirarlo del modelo. En general, observamos una excelente significancia estadistica de manera general como individual en cada coeficiente.

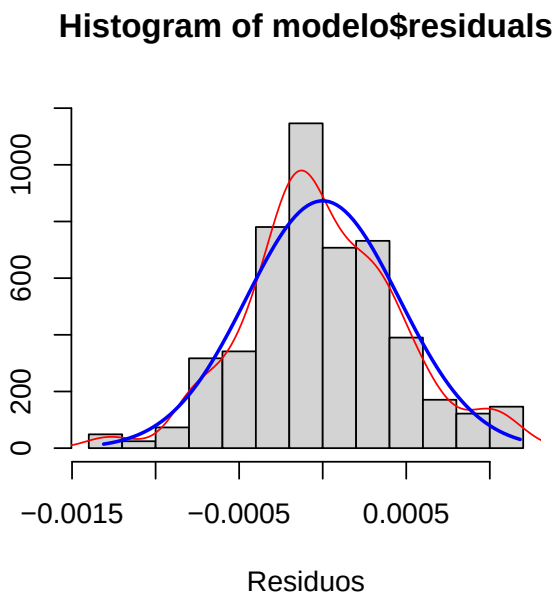
```
##
## Call:
## lm(formula = price ~ carwidth + curbweight + enginesize + horsepower +
##     citympg + highwaympg + carbody, data = precio_autos_final)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0013109 -0.0002588 -0.0000396  0.0002997  0.0011818
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -1.267e+01  4.590e+00  -2.761  0.00631 **
## carwidth          2.717e+01  9.282e+00   2.927  0.00383 **
## curbweight        4.196e-01  5.358e-02   7.831 3.09e-13 ***
## enginesize       -5.157e-02  2.741e-02  -1.881  0.06145 .
## horsepower        1.883e-02  4.155e-03   4.533 1.02e-05 ***
## citympg          -1.608e-03  6.235e-04  -2.579  0.01066 *
## highwaympg        9.951e-04  3.319e-04   2.998  0.00307 **
## carbodyhardtop   -4.875e-04  2.562e-04  -1.903  0.05853 .
## carbodyhatchback -9.656e-04  2.091e-04  -4.617 7.07e-06 ***
## carbodysedan     -6.615e-04  2.057e-04  -3.216  0.00153 **
## carbodywagon     -1.010e-03  2.231e-04  -4.528 1.04e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0004686 on 194 degrees of freedom
## Multiple R-squared:  0.8882, Adjusted R-squared:  0.8824
## F-statistic: 154.1 on 10 and 194 DF,  p-value: < 2.2e-16
```

5.2 Validación del modelo

5.2.1 Normalidad de los residuos

Analizando las gráficas de qqplot e histograma de recursos, creo que es seguro afirmar que los residuos cuentan con una distribución satisfactoriamente normal, aunque cabe recalcar algo gruesa en las colas, lo que se conoce como una distribución platicurtica.





5.2.2 Media 0

- Paso 1. Definir la hipotesis

  $H_0 : \mu = 0$

  $H_1 : \mu \neq 0$

- Paso 2. Regla de decisión

  Nivel de confianza = 0.95

  $\alpha = 0.05$

- Paso 3. Análisis del resultado

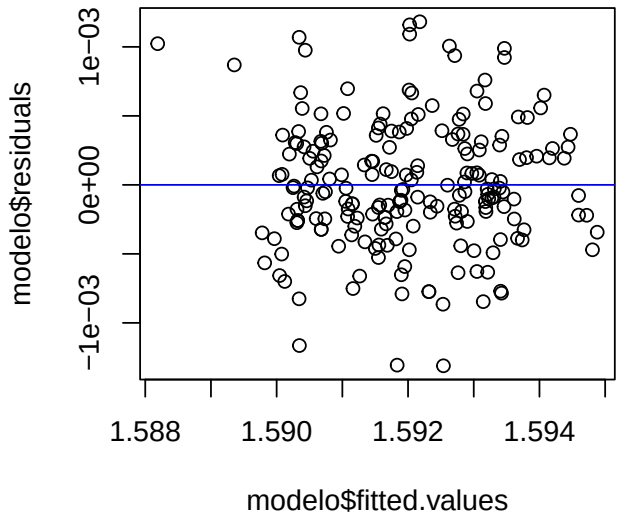
```
##
##  One Sample t-test
##
## data:  modelo$residuals
## t = 5.1367e-16, df = 204, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -6.293162e-05  6.293162e-05
## sample estimates:
##    mean of x
## 1.639525e-20
```

$$\alpha < p$$

- Paso 4. Conclusion

Como mi valor p es mayor que alfa ($1 > 0.05$) no puedo rechazar la hipotesis nula, lo que me permite afirmar con un alto grado de seguridad, que la hipotesis nula es correcta, y la media de los residuos es 0. La demás información del t-test también siguere fuertemente que la media de los residuos sea 0, por ejemplo, el t-value, la cantidad de desviaciones estandar de la que se encuentra la media de la muestra de la media teorizada es de 8.34e-16, es decir, prácticamente 0, lo que indica una cercanía casi absoluta de la media de la muestra a la media teorizada.

5.2.3 Homocedasticidad e independencia



Considero que se puede apreciar claramente en la gráfica la homocedasticidad de los residuos, presentando una distribución casi uniforme a lo largo del rango de la gráfica, salvo por algunos valores atipicos. La gráfica también presenta simetría, otro favorable indicador de variana uniforme entre los residuos.

6 Conclusión

Se generó un modelo de regresión lineal adecuado para las necesidades de la actividad, el cual cumple con los factores de verificación y validez discutidos a lo largo del curso. Este modelo además se sustenta en una serie de análisis y procedimientos correspondientes tratados en el curso como transformaciones, análisis de frecuencia, transformación de variables categoricas en dummies, etc.

Respondiendo a la pregunta detotante de la actividad: ¿Que valores son más influyentes en el precio de un vehiculo? Podemos regresar al modelo y análizar que coeficientes cuentan con valores absolutos más elevados, puesto que son estos los que tendrían un mayor peso sobre la predicción del modelo. Estas variables son *Carwidth* y *Carweight*, con coeficientes de 3.41 y 2.02 respectivamente; tal parece que los valores más importantes para determinar el precio de un carro, son su peso y tamaño, posiblemente devido a la cantidad de material requeridos para la fabricación. De nuevo, es importante resaltar que todos las variables seleccionadas para el modelo final son consideradas como estadisticamente significativas, por lo que son importante para la realización de predicciones precisas.

7 Anexos

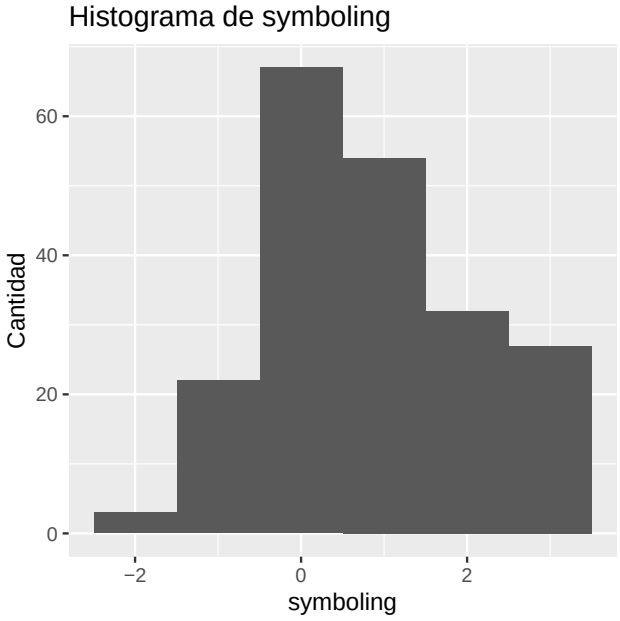
7.1 Análisis de variables

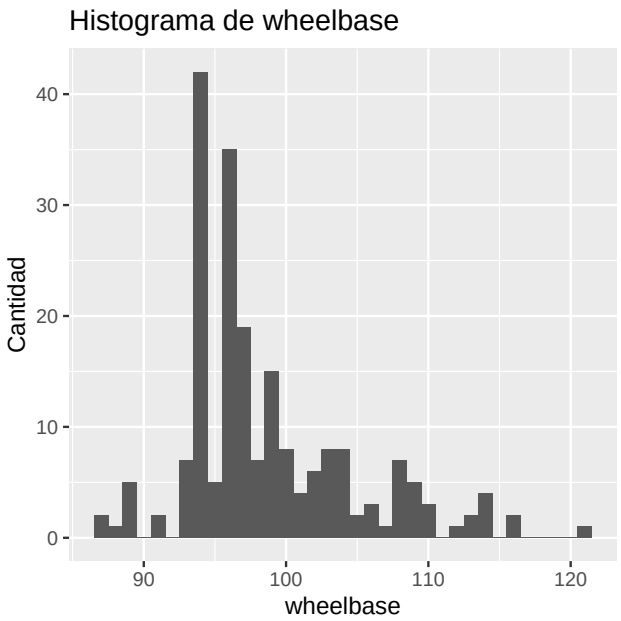
7.1.1 Variables numéricas

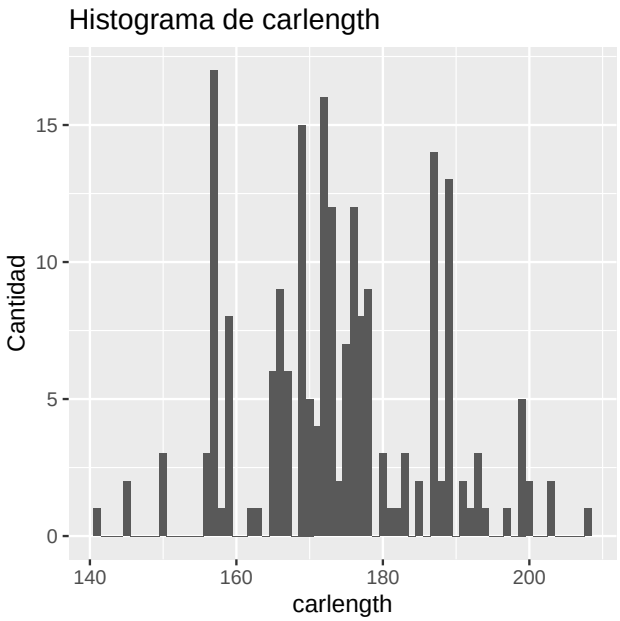
```
##    symboling         wheelbase        carlength        carwidth
##  Min.   :-2.0000   Min.   : 86.60   Min.   :141.1   Min.   :60.30
##  1st Qu.: 0.0000   1st Qu.: 94.50   1st Qu.:166.3   1st Qu.:64.10
##  Median : 1.0000   Median : 97.00   Median :173.2   Median :65.50
##  Mean   : 0.8341   Mean   : 98.76   Mean   :174.0   Mean   :65.91
##  3rd Qu.: 2.0000   3rd Qu.:102.40   3rd Qu.:183.1   3rd Qu.:66.90
##  Max.   : 3.0000   Max.   :120.90   Max.   :208.1   Max.   :72.30
##    carheight       curbweight     enginesize        stroke
##  Min.   :47.80   Min.   :1488   Min.   : 61.0   Min.   :2.070
##  1st Qu.:52.00   1st Qu.:2145   1st Qu.: 97.0   1st Qu.:3.110
```

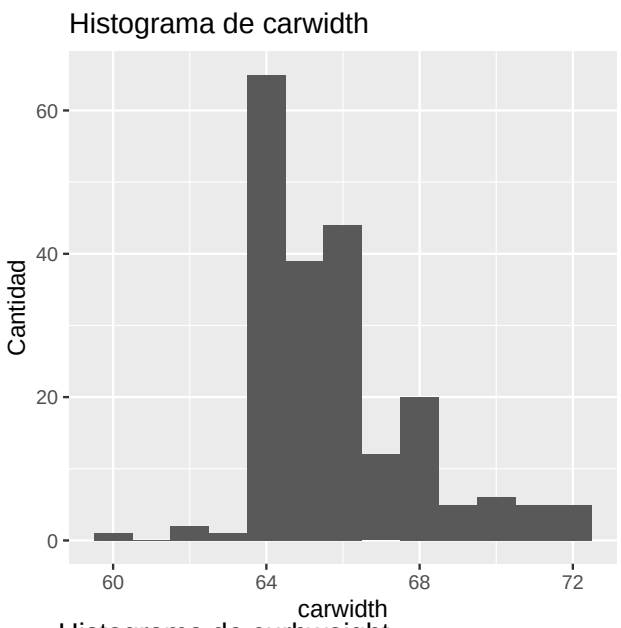
```
##  Median :54.10   Median :2414   Median :120.0   Median :3.290
##  Mean   :53.72   Mean   :2556   Mean   :126.9   Mean   :3.255
##  3rd Qu.:55.50   3rd Qu.:2935   3rd Qu.:141.0   3rd Qu.:3.410
##  Max.   :59.80   Max.   :4066   Max.   :326.0   Max.   :4.170
##  compressionratio   horsepower       peakrpm        citympg
##  Min.   : 7.00    Min.   : 48.0   Min.   :4150   Min.   :13.00
##  1st Qu.: 8.60    1st Qu.: 70.0   1st Qu.:4800   1st Qu.:19.00
##  Median : 9.00    Median : 95.0   Median :5200   Median :24.00
##  Mean   :10.14    Mean   :104.1   Mean   :5125   Mean   :25.22
##  3rd Qu.: 9.40    3rd Qu.:116.0   3rd Qu.:5500   3rd Qu.:30.00
##  Max.   :23.00    Max.   :288.0   Max.   :6600   Max.   :49.00
##    highwaympg        price
##  Min.   :16.00   Min.   : 5118
##  1st Qu.:25.00   1st Qu.: 7788
##  Median :30.00   Median :10295
##  Mean   :30.75   Mean   :13277
##  3rd Qu.:34.00   3rd Qu.:16503
##  Max.   :54.00   Max.   :45400
```

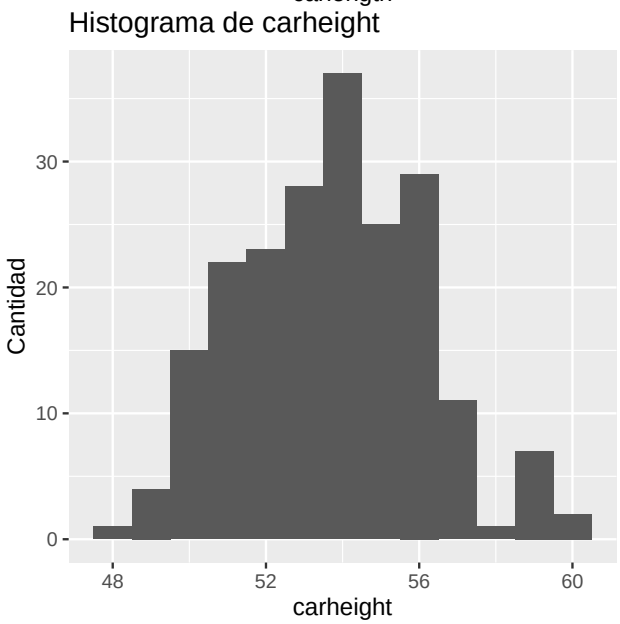
7.1.1.1 Histogramas

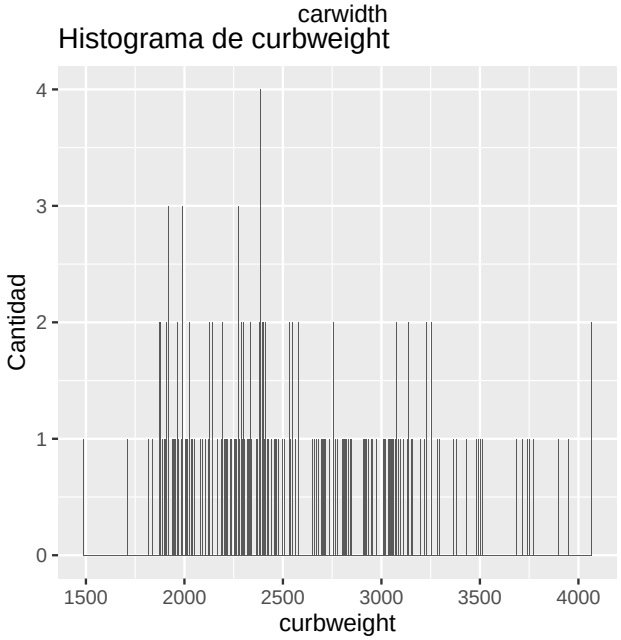


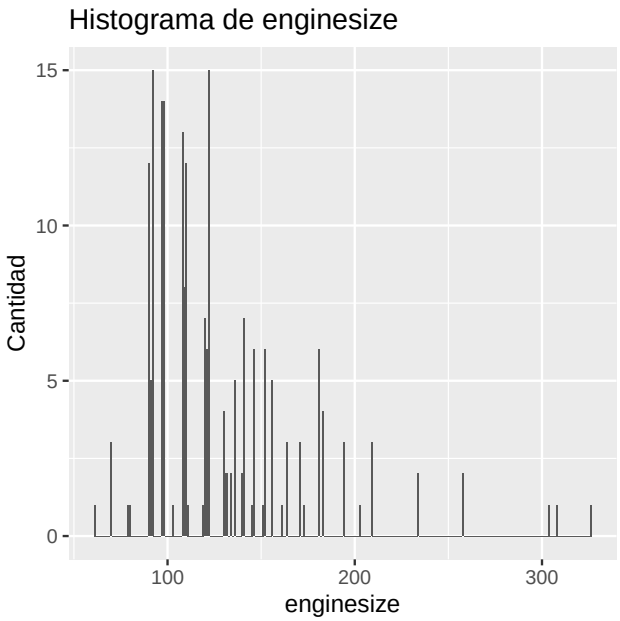


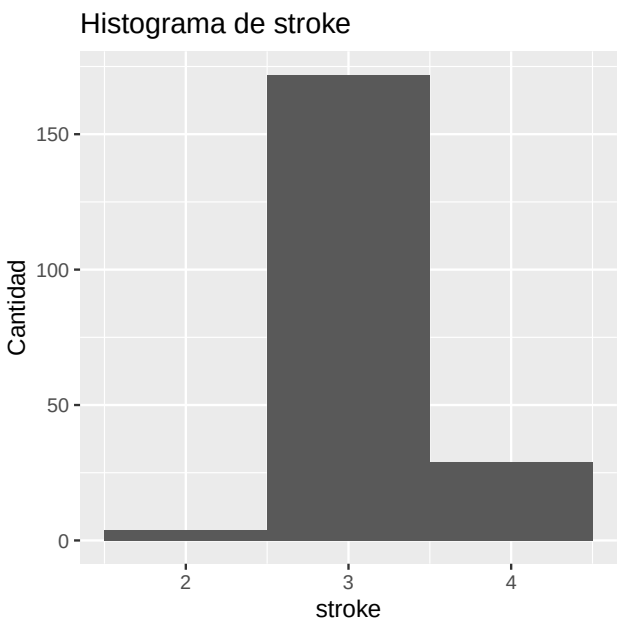

Histograma de carlength
Cantidad
15
10
5
0
140
160
180
200
carlength

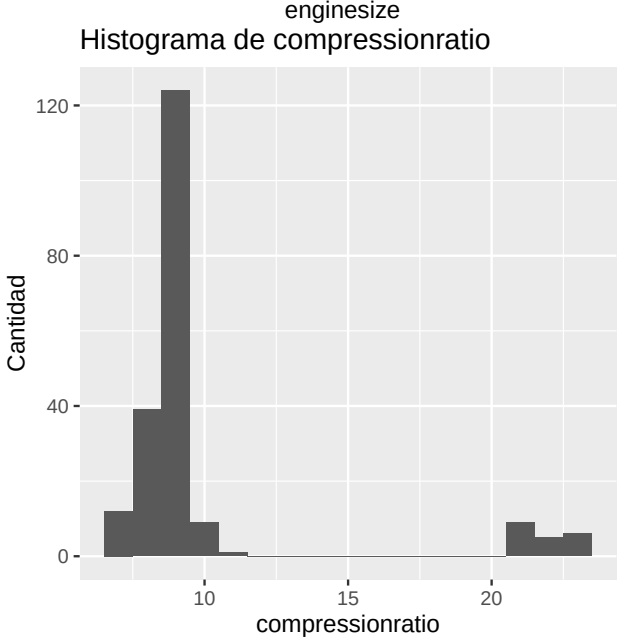

Histograma de carwidth
Cantidad
60
40
20
0
60
64
68
72
carwidth

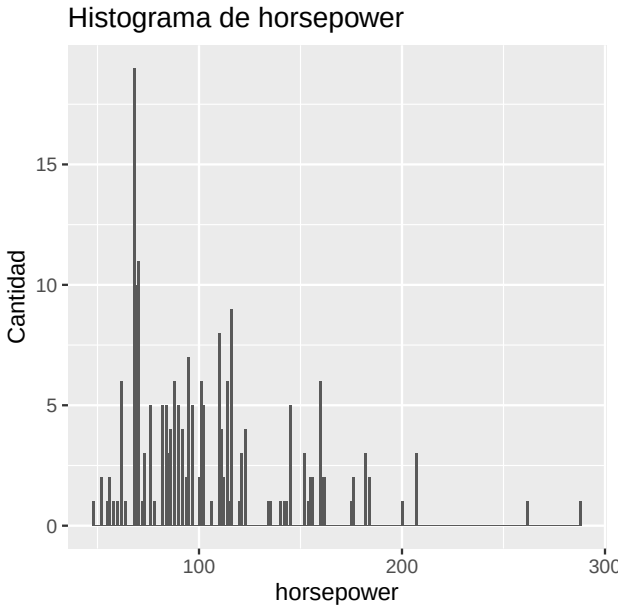

Histograma de carheight
Cantidad
30
20
10
0
48
52
56
60
carheight

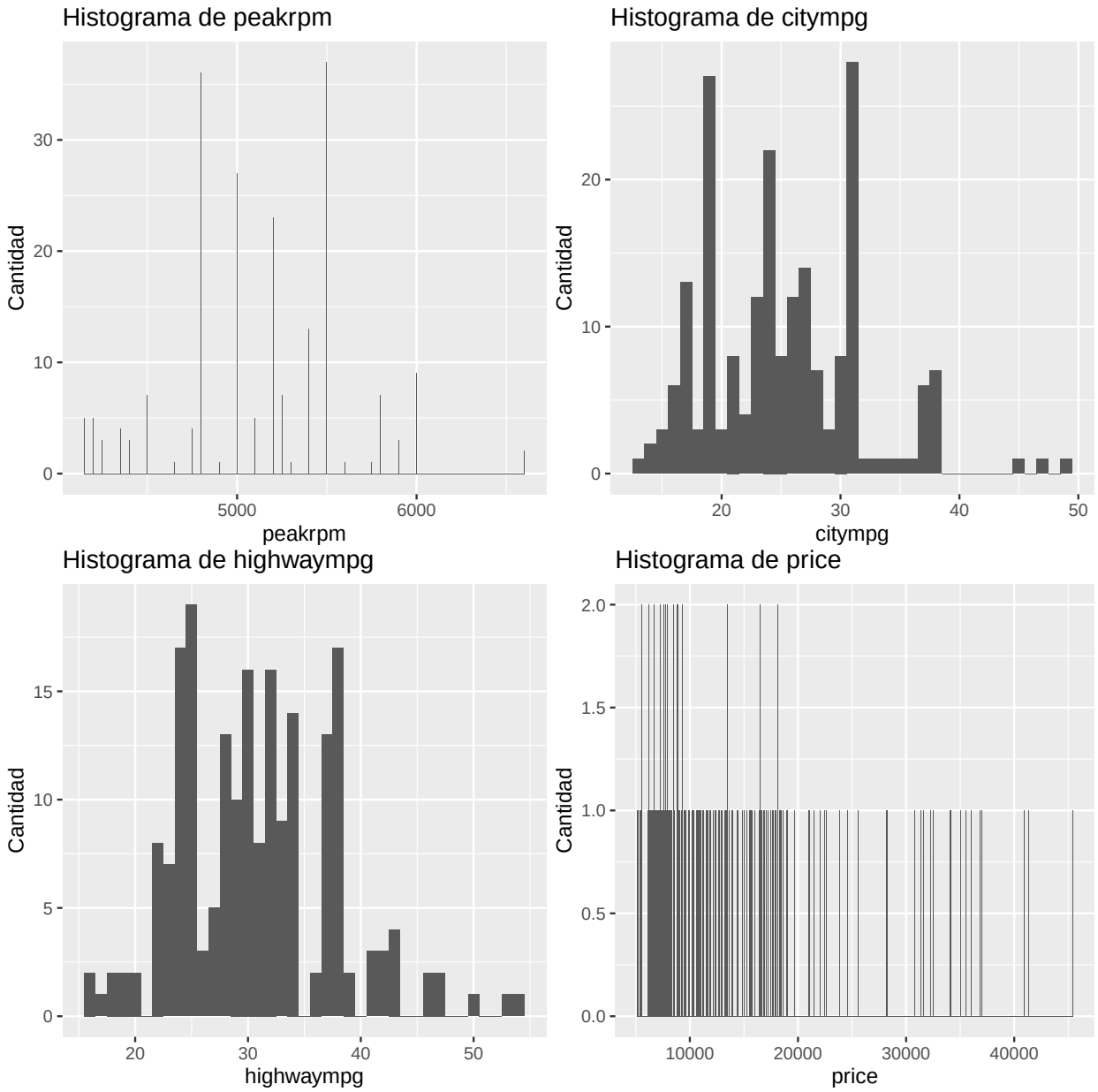

Histograma de curbweight
Cantidad
4
3
2
1
0
1500
2000
2500
3000
3500
4000
curbweight


Histograma de enginesize
Cantidad
15
10
5
0
100
200
300
enginesize

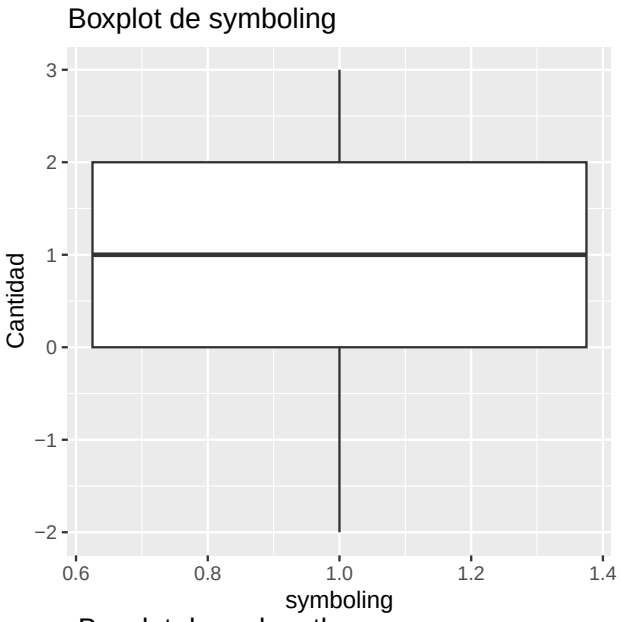

Histograma de stroke
Cantidad
150
100
50
0
2
3
4
stroke

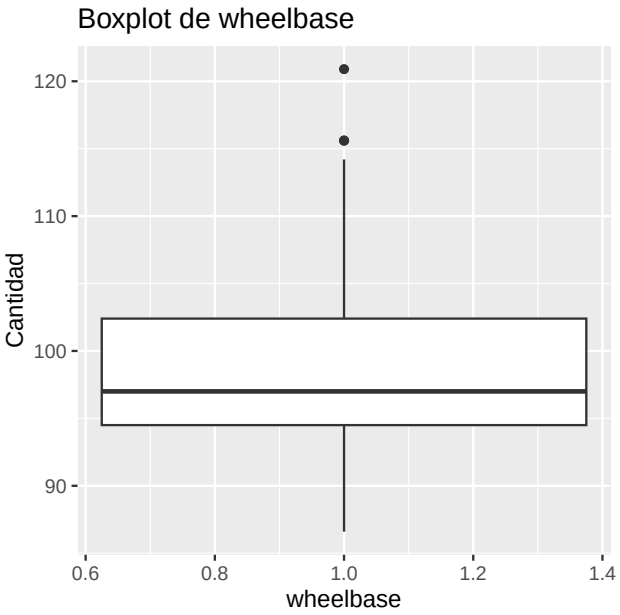

Histograma de compressionratio
Cantidad
120
80
40
0
10
15
20
compressionratio

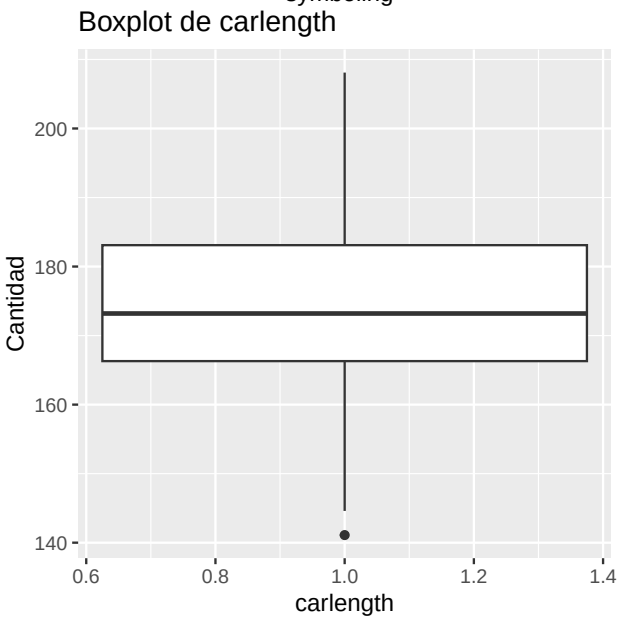

Histograma de horsepower
Cantidad
15
10
5
0
100
200
300
horsepower

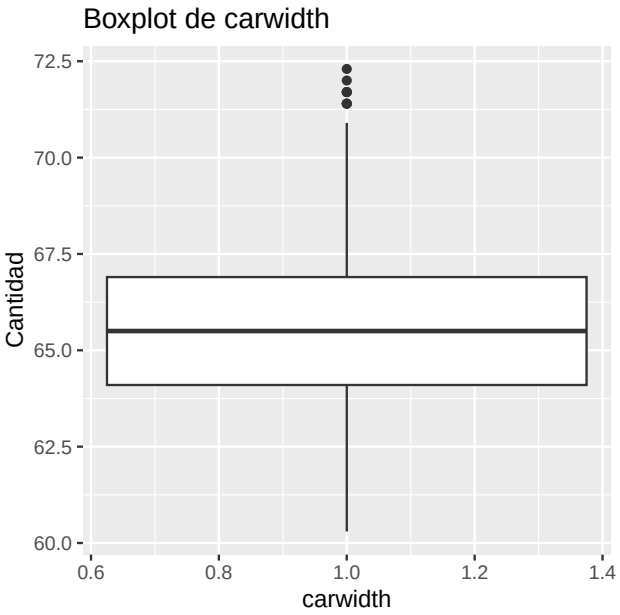


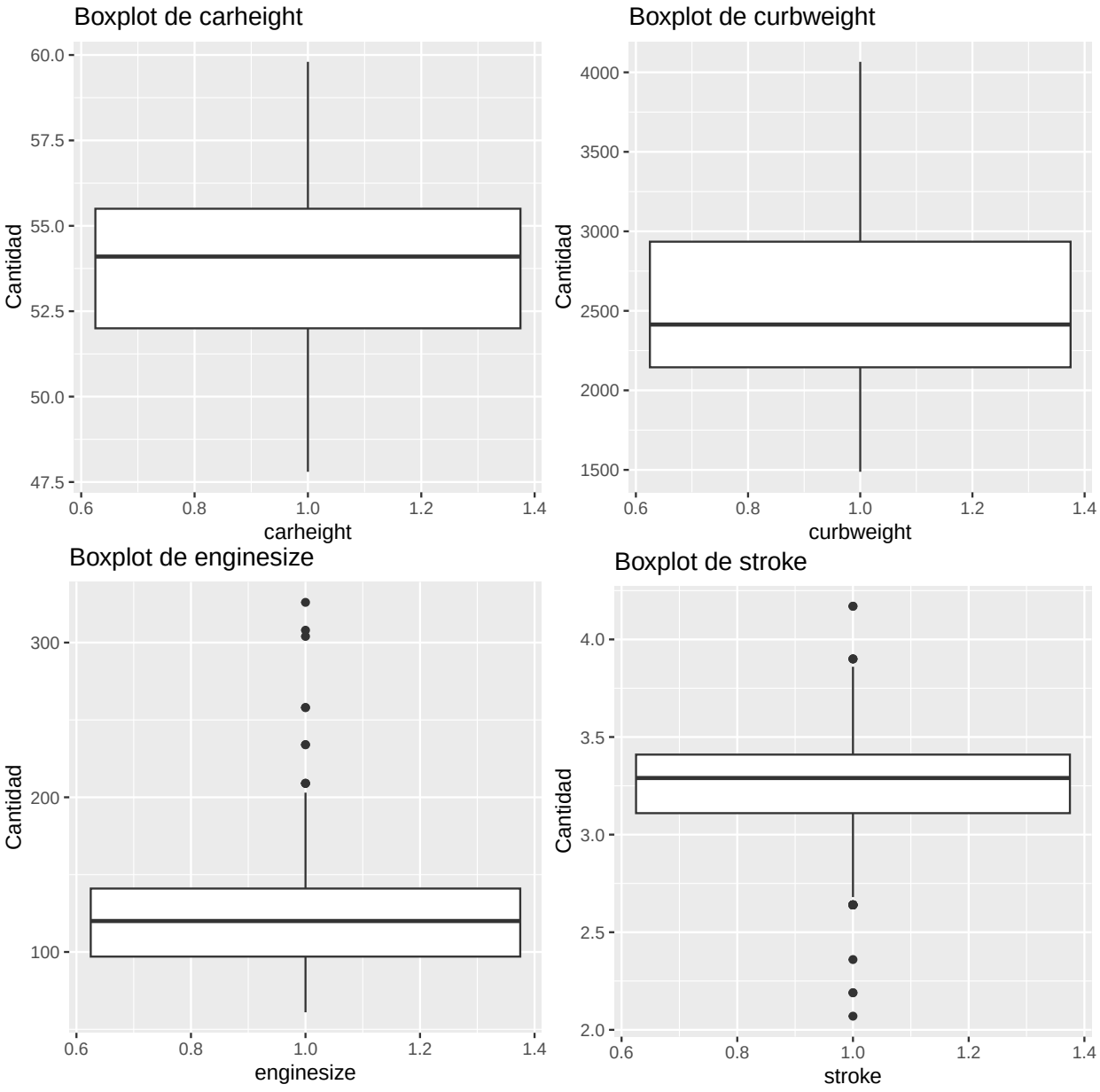
7.1.1.2 Diagramas de caja y bigote

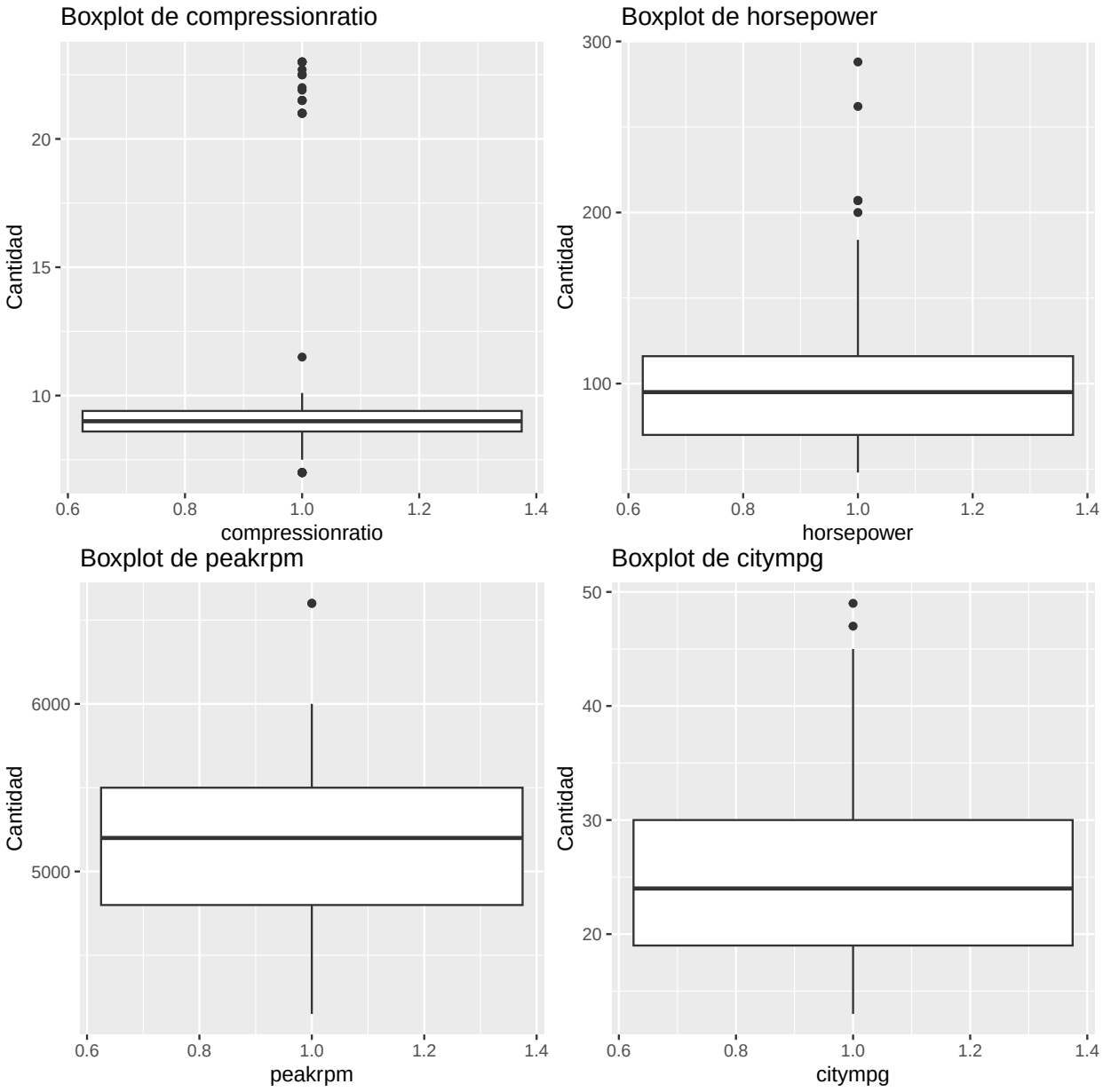

Boxplot de symboling
Cantidad
3
2
1
0
−1
−2
0.6
0.8
1.0
1.2
1.4
symboling

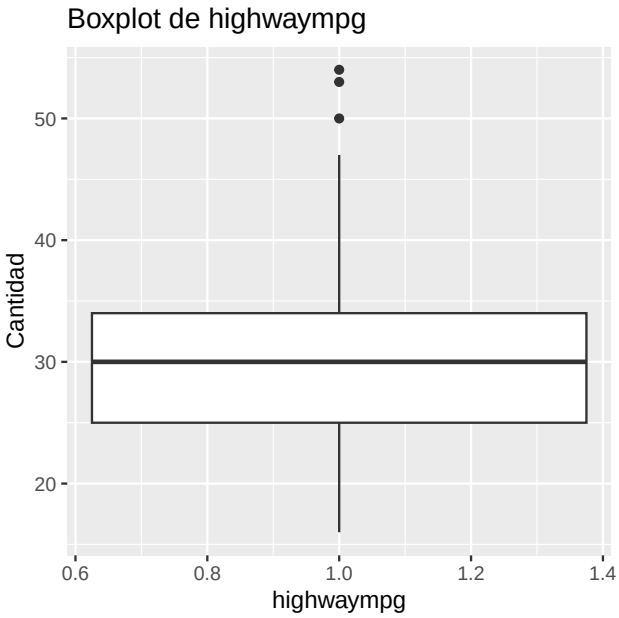

Boxplot de wheelbase
Cantidad
120
110
100
90
0.6
0.8
1.0
1.2
1.4
wheelbase

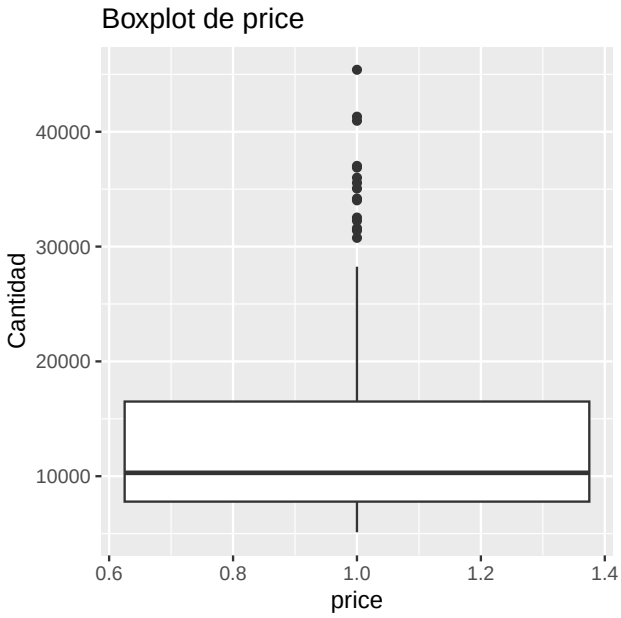

Boxplot de carlength
Cantidad
200
180
160
140
0.6
0.8
1.0
1.2
1.4
carlength


Boxplot de carwidth
Cantidad
72.5
70.0
67.5
65.0
62.5
60.0
0.6
0.8
1.0
1.2
1.4
carwidth

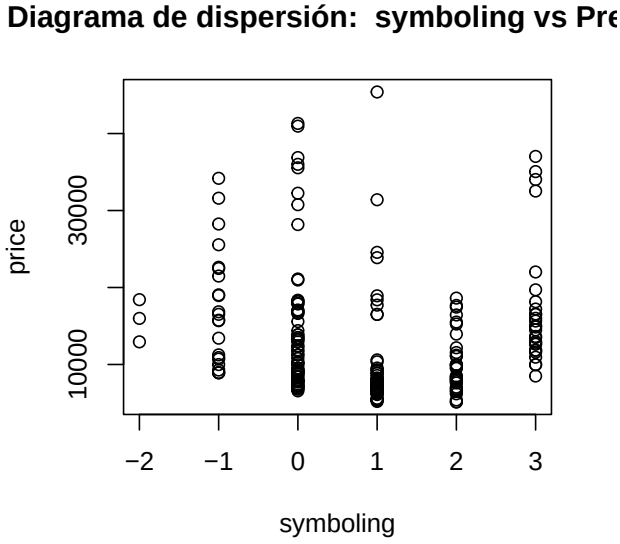

Boxplot de carheight
Cantidad
carheight
60.0
57.5
55.0
52.5
50.0
47.5
0.6
0.8
1.0
1.2
1.4
Boxplot de curbweight
Cantidad
curbweight
4000
3500
3000
2500
2000
1500
Boxplot de enginesize
Cantidad
enginesize
300
200
100
Boxplot de stroke
Cantidad
stroke
4.0
3.5
3.0
2.5
2.0


Boxplot de compressionratio
Cantidad
compressionratio
Boxplot de horsepower
Cantidad
horsepower
Boxplot de peakrpm
Cantidad
peakrpm
Boxplot de citympg
Cantidad
citympg





7.1.1.3 Diagramas de dispersión



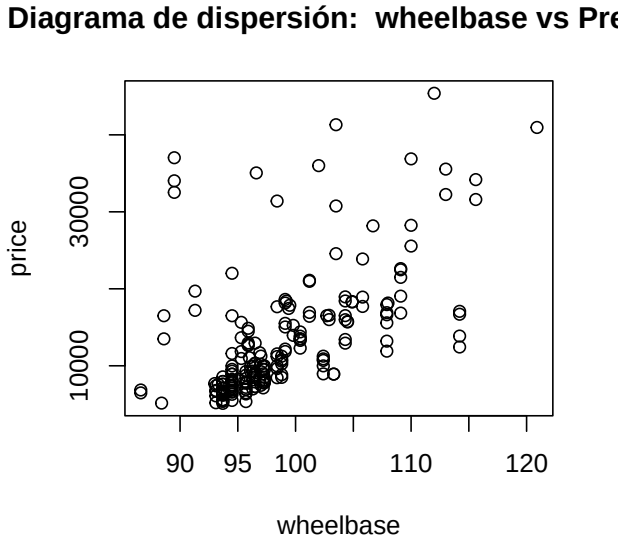


Diagrama de dispersión: carlength vs Pre

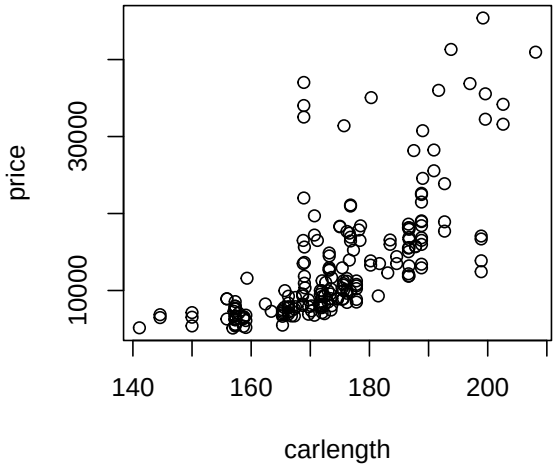


Diagrama de dispersión: carwidth vs Pre

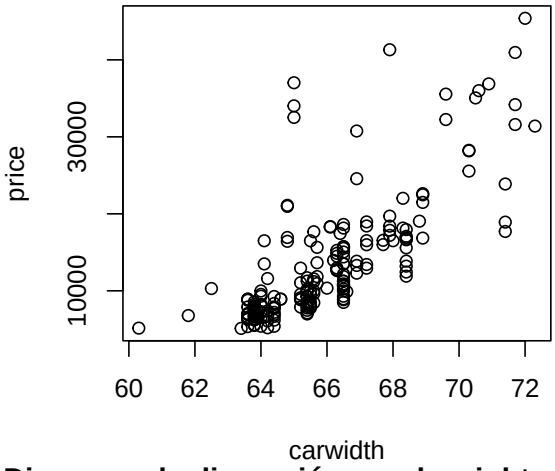


Diagrama de dispersión: carheight vs Pre

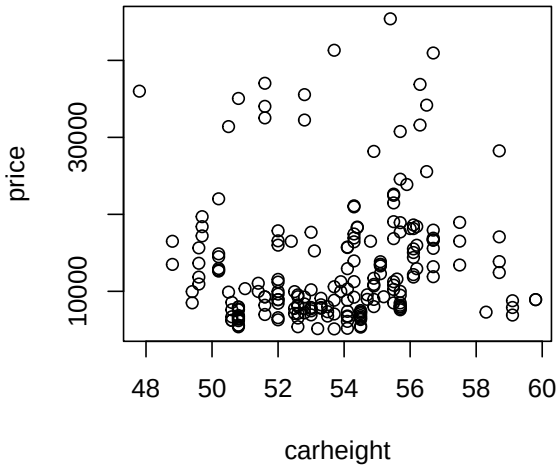


Diagrama de dispersión: curbweight vs Pr

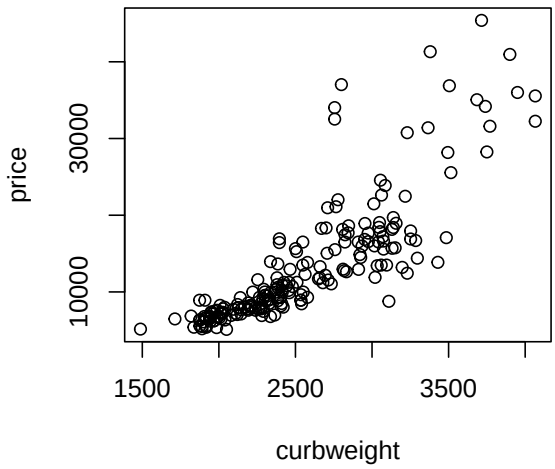


Diagrama de dispersión: enginesize vs Pre

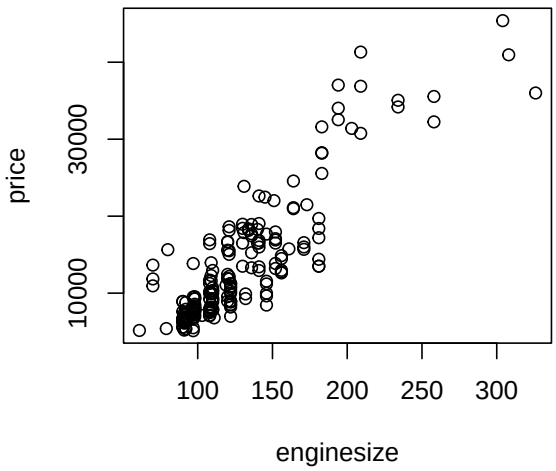
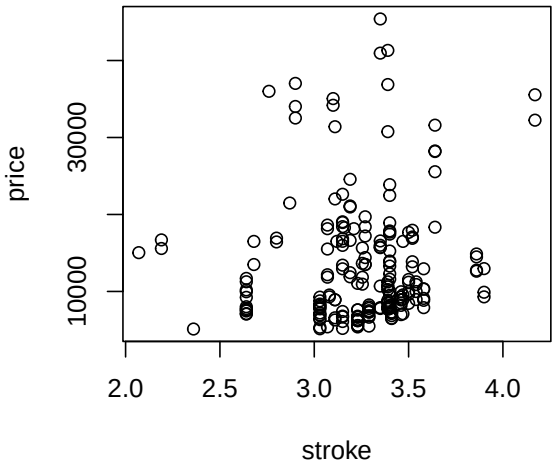
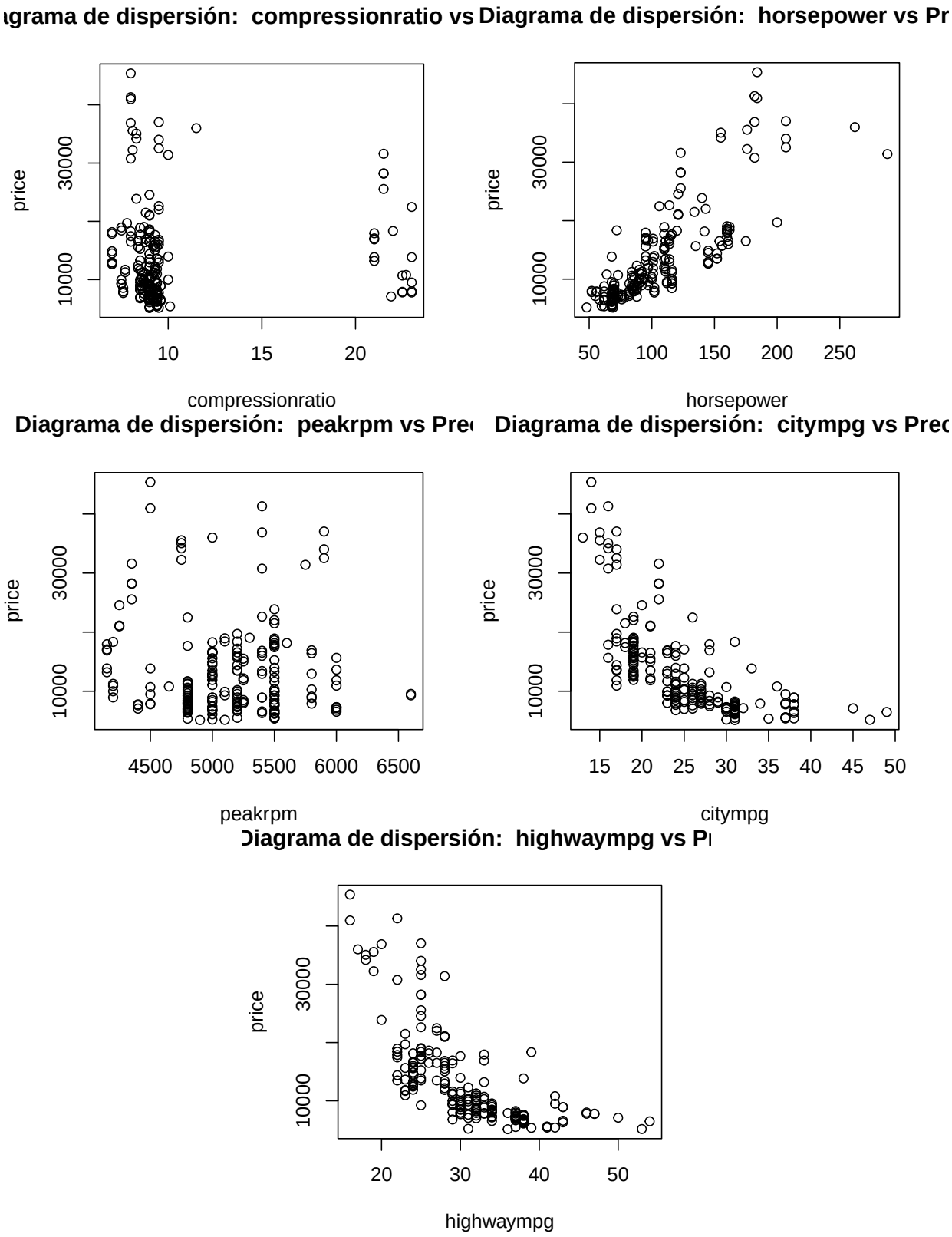
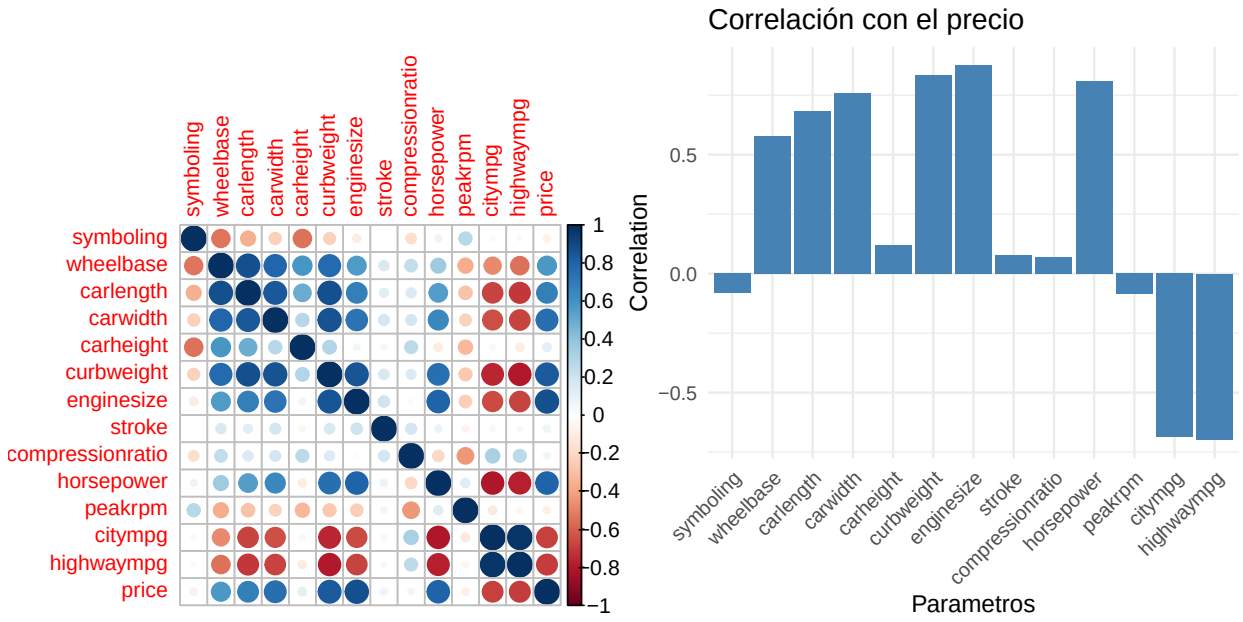


Diagrama de dispersión: stroke vs Preci





7.1.1.4 Matriz de correlación



7.1.2 Variables categóricas

```
##
## diesel    gas
##     20    185
```

```
##
## convertible     hardtop   hatchback       sedan       wagon
##           6           8          70          96          25
```

```
##
## 4wd fwd rwd
##   9 120  76
```

```
##
## front  rear
##   202     3
```

```
##
##  dohc dohcv     l   ohc  ohcf  ohcv rotor
##    12     1    12   148    15    13     4
```

```
##
##  eight   five   four    six  three twelve    two
##      5     11    159     24      1      1      4
```

7.1.2.1 Gráficas de pastel

Diagrama de pastel: fueltype

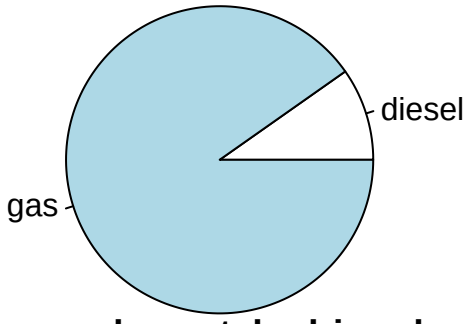


Diagrama de pastel: carbody

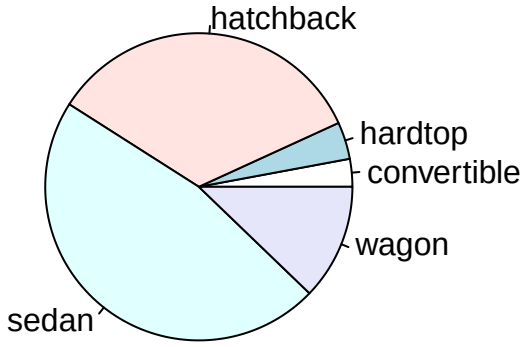


Diagrama de pastel: drivewheel

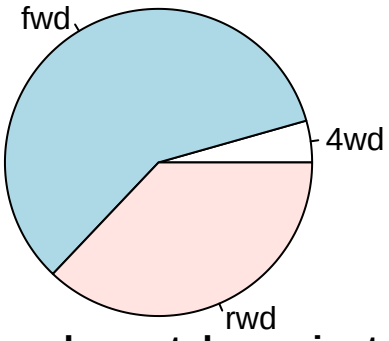


Diagrama de pastel: enginelocation

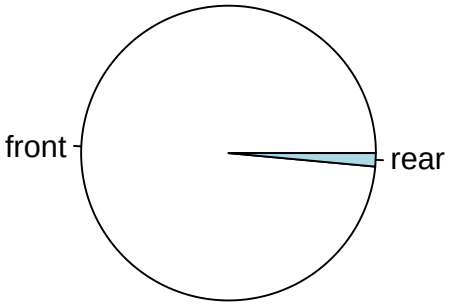


Diagrama de pastel: enginetype

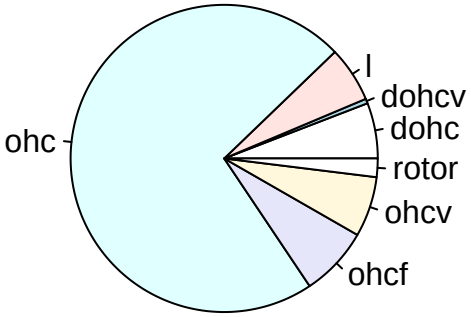
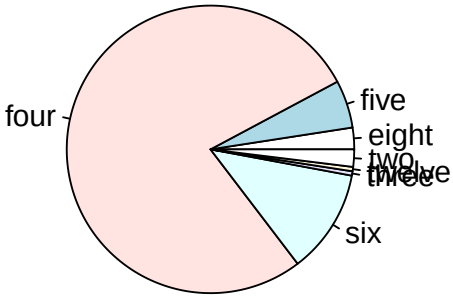
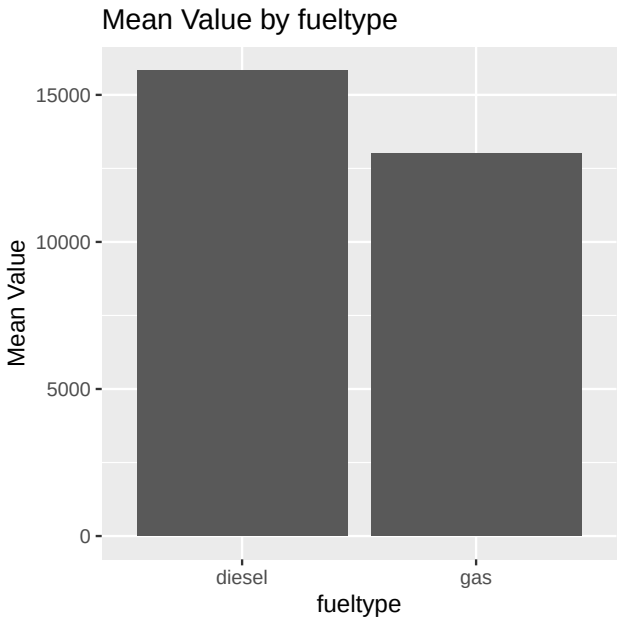
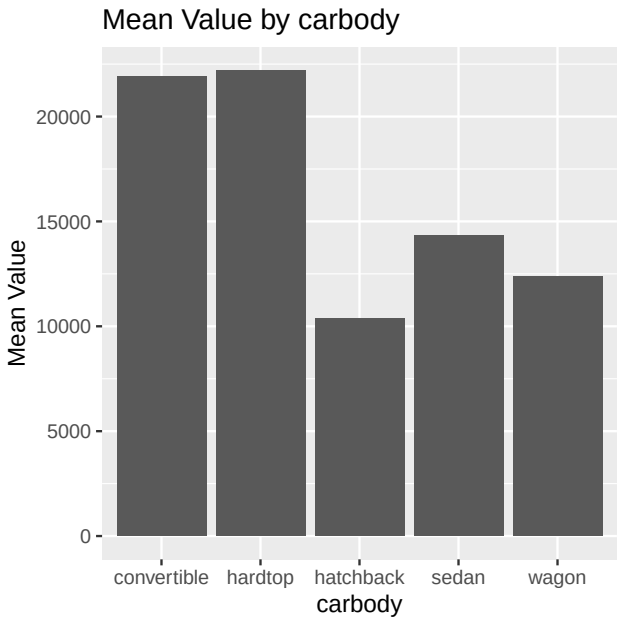


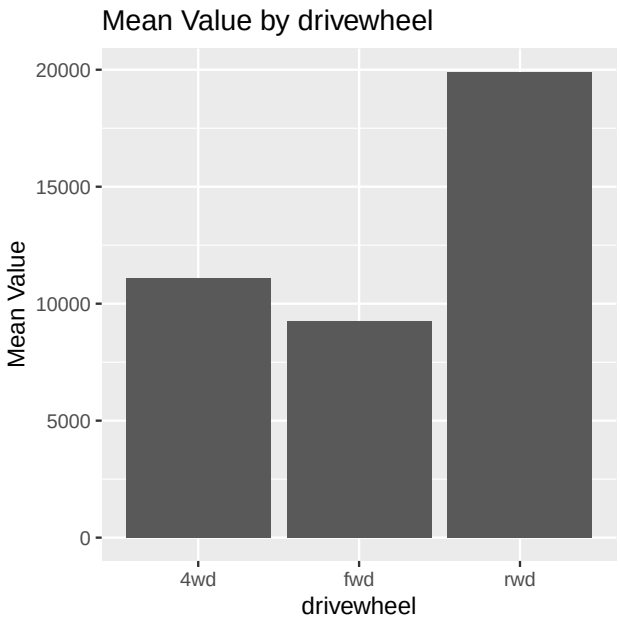
Diagrama de pastel: cylindernumber

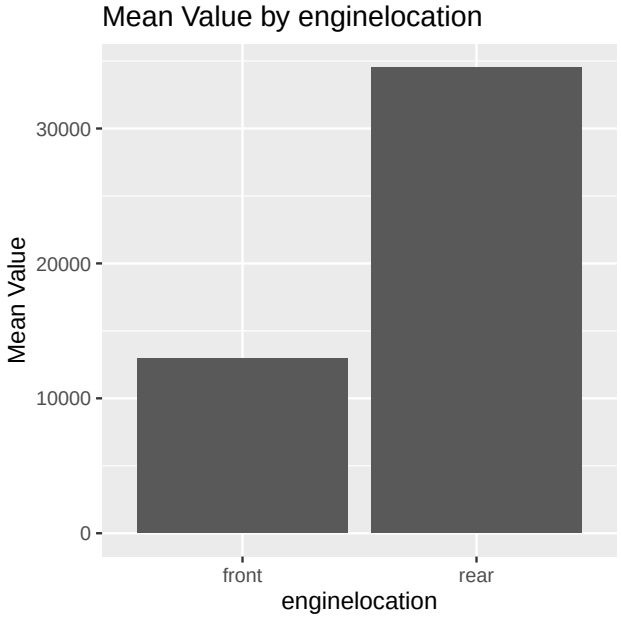


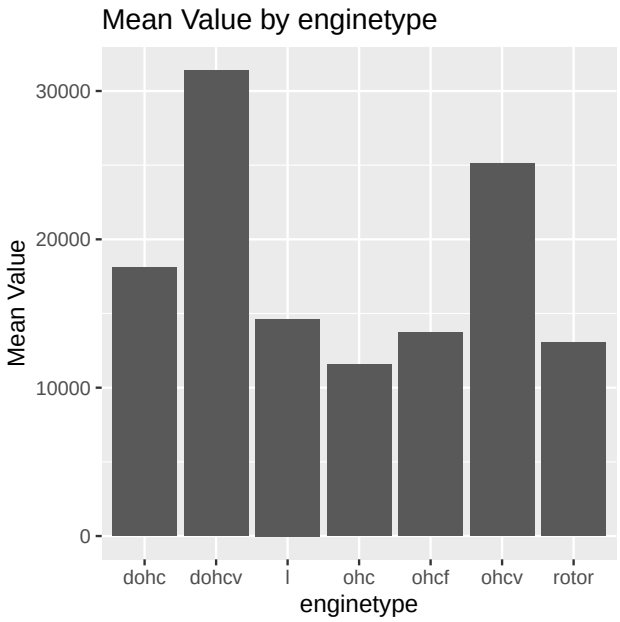
7.1.2.2 Media de precio por categoría

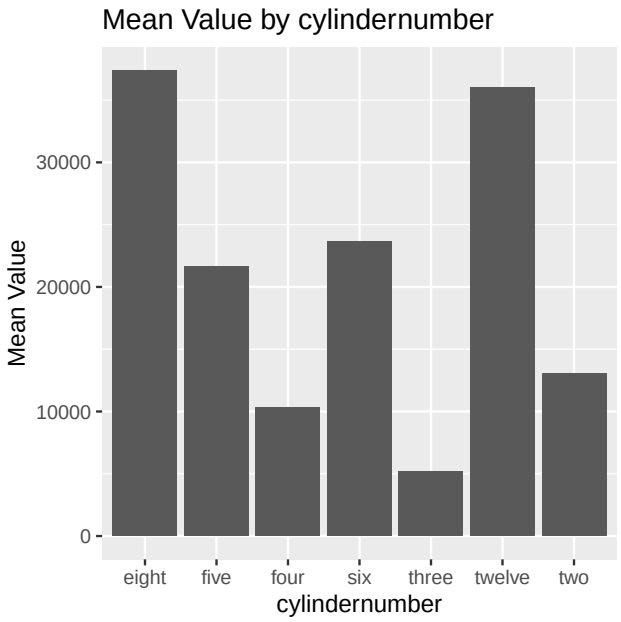

Mean Value by fueltype
15000
10000
5000
0
Mean Value
diesel
gas
fueltype


Mean Value by carbody
20000
15000
10000
5000
0
Mean Value
convertible
hardtop
hatchback
sedan
wagon
carbody


Mean Value by drivewheel
20000
15000
10000
5000
0
Mean Value
4wd
fwd
rwd
drivewheel

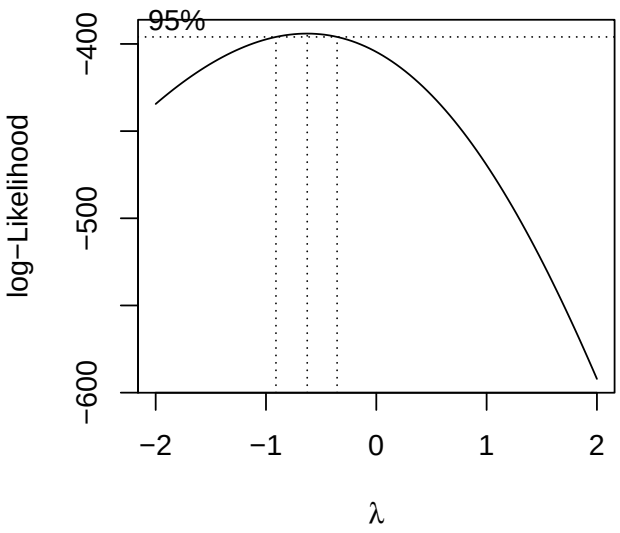

Mean Value by enginelocation
30000
20000
10000
0
Mean Value
front
rear
enginelocation

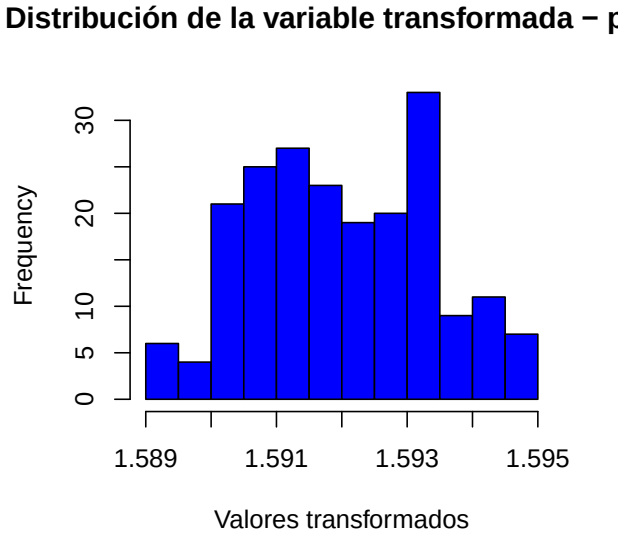


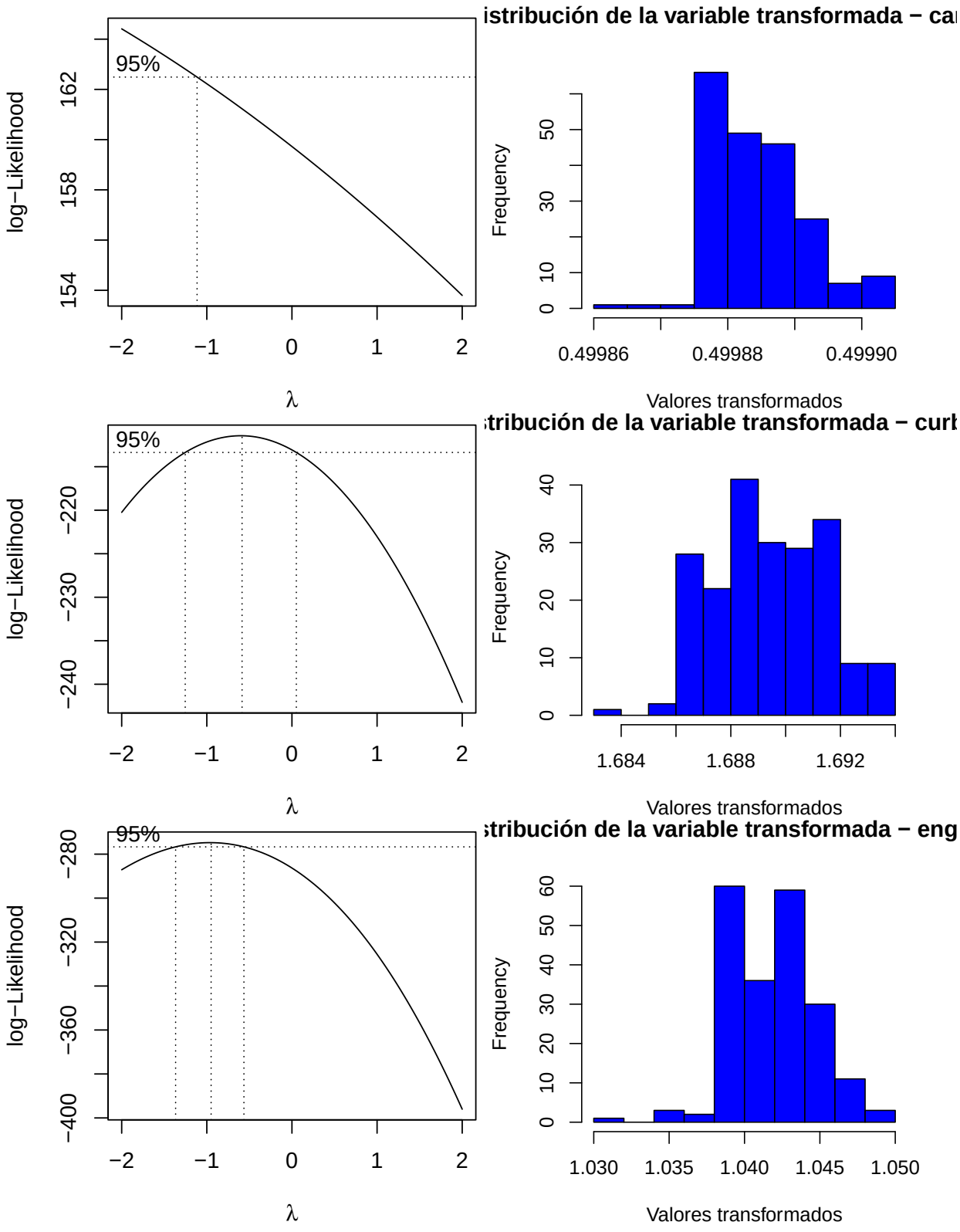


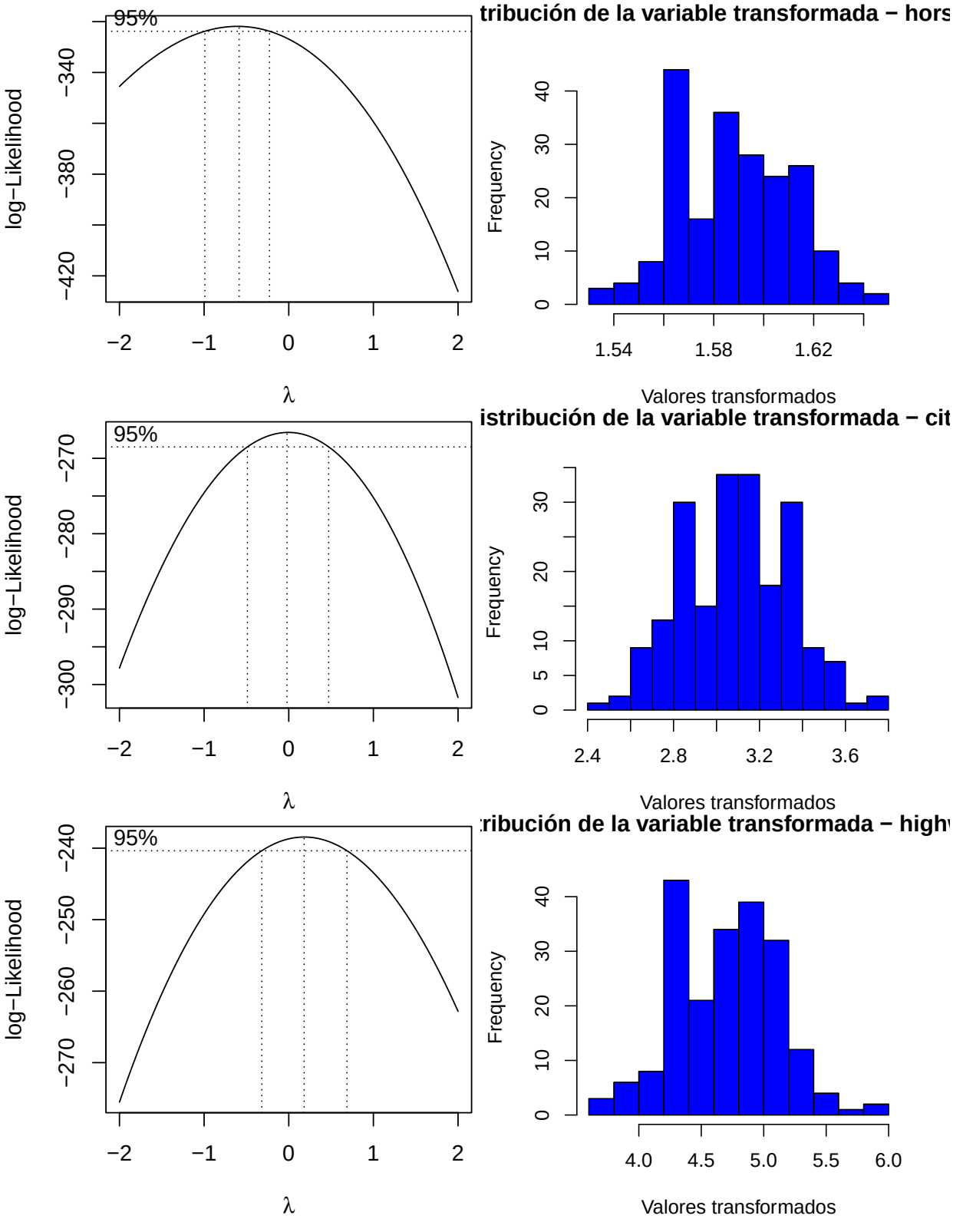
7.2 Transformación de variables

```
## Lamda optima -  price :  -0.6262626
## Lamda optima -  carwidth :  -2
## Lamda optima -  curbweight :  -0.5858586
## Lamda optima -  enginesize :  -0.9494949
## Lamda optima -  horsepower :  -0.5858586
## Lamda optima -  citympg :  -0.02020202
## Lamda optima -  highwaympg :  0.1818182
```






95%
log−Likelihood
154
158
162
−2
−1
0
1
2
λ
istribución de la variable transformada − ca
Frequency
0
10
30
50
0.49986
0.49988
0.49990
Valores transformados
95%
log−Likelihood
−220
−230
−240
−2
−1
0
1
2
λ
tribución de la variable transformada − curb
Frequency
0
10
20
30
40
1.684
1.688
1.692
Valores transformados
95%
log−Likelihood
−280
−320
−360
−400
−2
−1
0
1
2
λ
tribución de la variable transformada − eng
Frequency
0
10
20
30
40
50
60
1.030
1.035
1.040
1.045
1.050
Valores transformados


95%
log−Likelihood
−340
−380
−420
−2
−1
0
1
2
λ
tribución de la variable transformada − hors
Frequency
0
10
20
30
40
1.54
1.58
1.62
Valores transformados
95%
log−Likelihood
−270
−280
−290
−300
−2
−1
0
1
2
λ
istribución de la variable transformada − cit
Frequency
0
5
10
20
30
2.4
2.8
3.2
3.6
Valores transformados
95%
log−Likelihood
−240
−250
−260
−270
−2
−1
0
1
2
λ
ribución de la variable transformada − high
Frequency
0
10
20
30
40
4.0
4.5
5.0
5.5
6.0
Valores transformados