DSC 520 Final Project

Bilal Kudaimi

2020-11-07

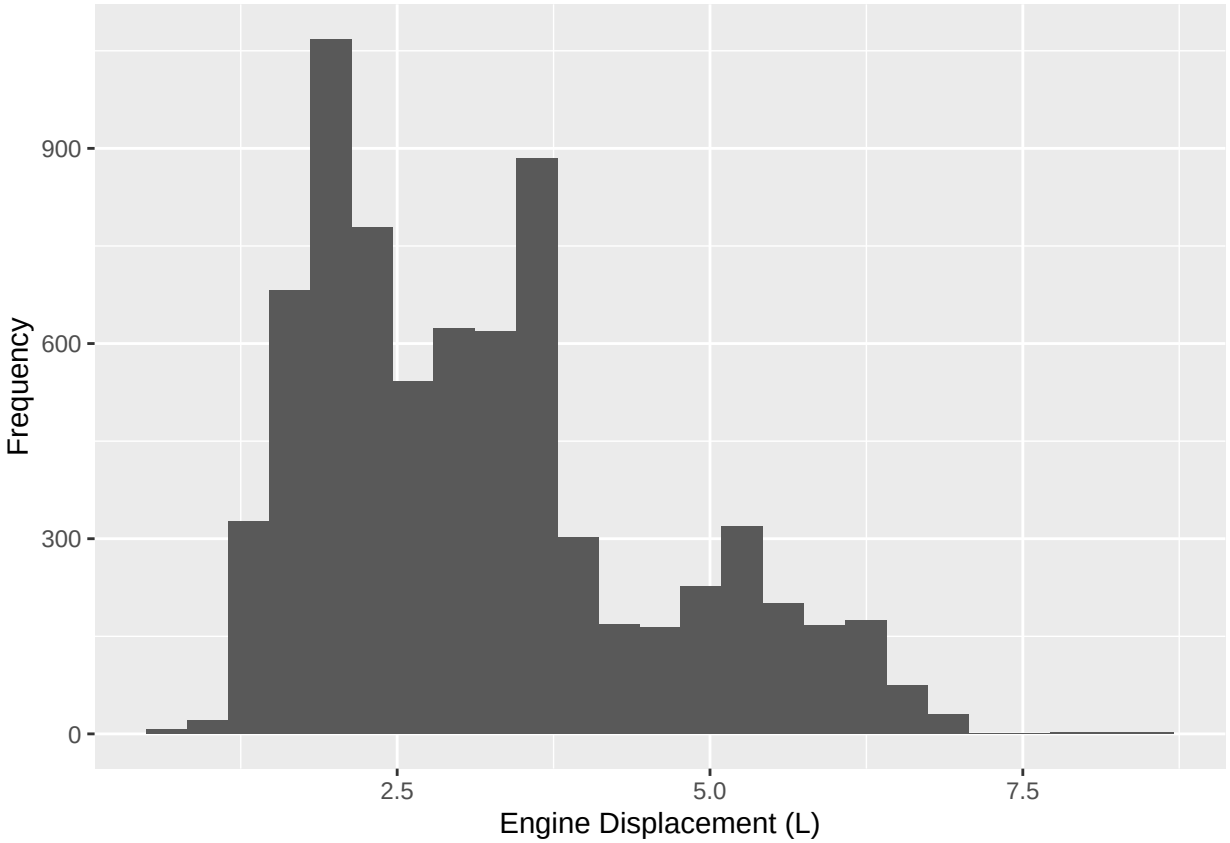
Summary of the data

```
## Warning: package 'tidyr' was built under R version 4.0.3

## Warning: package 'dplyr' was built under R version 4.0.3

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

## Loading required package: carData

##
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':
##
##     recode

## Warning: package 'tm' was built under R version 4.0.3

## Loading required package: NLP

## Warning: package 'NLP' was built under R version 4.0.3

##
## Attaching package: 'NLP'

## The following object is masked from 'package:ggplot2':
##
##     annotate
```

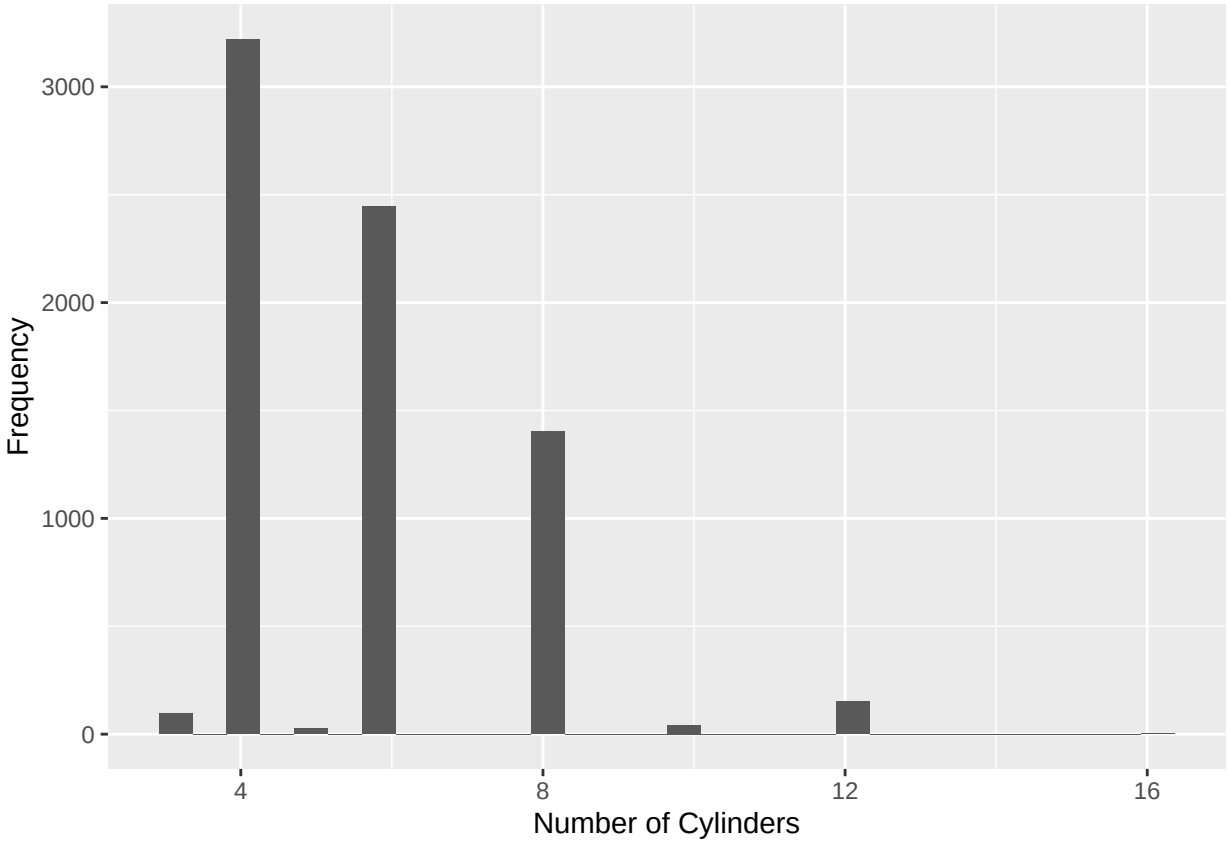
```
##             Make         Model                Vehicle.Class  Engine.Size.L.
##  FORD         : 628   Length:7385        SUV - SMALL   :1217   Min.   :0.6758
##  CHEVROLET    : 588   Class :character   MID-SIZE      :1133   1st Qu.:2.0503
##  BMW          : 527   Mode  :character   COMPACT       :1022   Median :2.9263
##  MERCEDES-BENZ: 419                      SUV - STANDARD: 735   Mean   :3.1618
##  PORSCHE      : 376                      FULL-SIZE     : 639   3rd Qu.:3.7733
##  TOYOTA       : 330                      SUBCOMPACT    : 606   Max.   :8.5624
##  (Other)      :4517                      (Other)       :2033
##    Cylinders      Transmission Fuel.Type Fuel.Consumption.City..L.100.km.
##  Min.   : 3.000   A :1851      D: 175    Min.   : 4.20
##  1st Qu.: 4.000   AM: 646      E: 370    1st Qu.:10.10
##  Median : 6.000   AS:3127      N:   1    Median :12.10
##  Mean   : 5.615   AV: 576      X:3637    Mean   :12.56
##  3rd Qu.: 6.000   M :1185      Z:3202    3rd Qu.:14.60
##  Max.   :16.000                          Max.   :30.60
##
##  Fuel.Consumption.Hwy..L.100.km. Fuel.Consumption.Comb..L.100.km.
##  Min.   : 4.000                  Min.   : 4.10
##  1st Qu.: 7.500                  1st Qu.: 8.90
##  Median : 8.700                  Median :10.60
##  Mean   : 9.042                  Mean   :10.98
##  3rd Qu.:10.200                  3rd Qu.:12.60
##  Max.   :20.600                  Max.   :26.10
##
##  Fuel.Consumption.Comb..mpg. CO2.Emissions.g.km.
##  Min.   :11.00               Min.   : 95.99
##  1st Qu.:22.00               1st Qu.:207.80
##  Median :27.00               Median :245.72
##  Mean   :27.48               Mean   :250.59
##  3rd Qu.:32.00               3rd Qu.:288.31
##  Max.   :69.00               Max.   :521.84
##
```

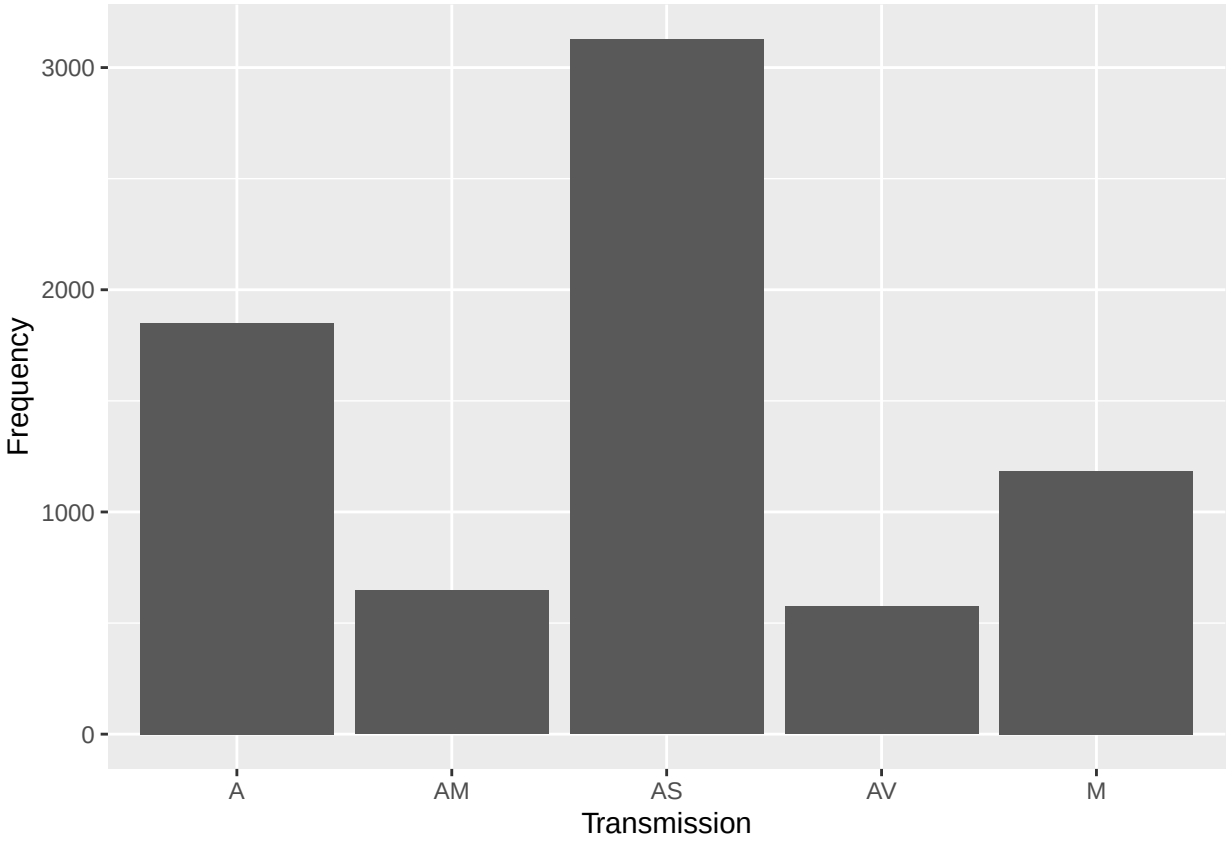
```
## 'data.frame':    7385 obs. of  12 variables:
##  $ Make                           : Factor w/ 42 levels "ACURA","ALFA ROMEO",..: 1 1 1 1 1 1 1 1 1
##  $ Model                          : chr  "ILX" "ILX" "ILX HYBRID" "MDX 4WD" ...
##  $ Vehicle.Class                  : Factor w/ 16 levels "COMPACT","FULL-SIZE",..: 1 1 1 12 12 3 3 3
##  $ Engine.Size.L.                 : num  1.91 2.55 1.29 3.65 3.52 ...
##  $ Cylinders                      : int  4 4 4 6 6 6 6 6 6 4 ...
##  $ Transmission                   : Factor w/ 5 levels "A","AM","AS",..: 3 5 4 3 3 3 3 3 5 3 ...
##  $ Fuel.Type                      : Factor w/ 5 levels "D","E","N","X",..: 5 5 5 5 5 5 5 5 5 5 ...
##  $ Fuel.Consumption.City..L.100.km.: num  9.9 11.2 6 12.7 12.1 11.9 11.8 12.8 13.4 10.6 ...
##  $ Fuel.Consumption.Hwy..L.100.km. : num  6.7 7.7 5.8 9.1 8.7 7.7 8.1 9 9.5 7.5 ...
##  $ Fuel.Consumption.Comb..L.100.km.: num  8.5 9.6 5.9 11.1 10.6 10 10.1 11.1 11.6 9.2 ...
##  $ Fuel.Consumption.Comb..mpg.     : int  33 29 48 25 27 28 28 25 24 31 ...
##  $ CO2.Emissions.g.km.             : num  196 221 136 255 244 ...
```

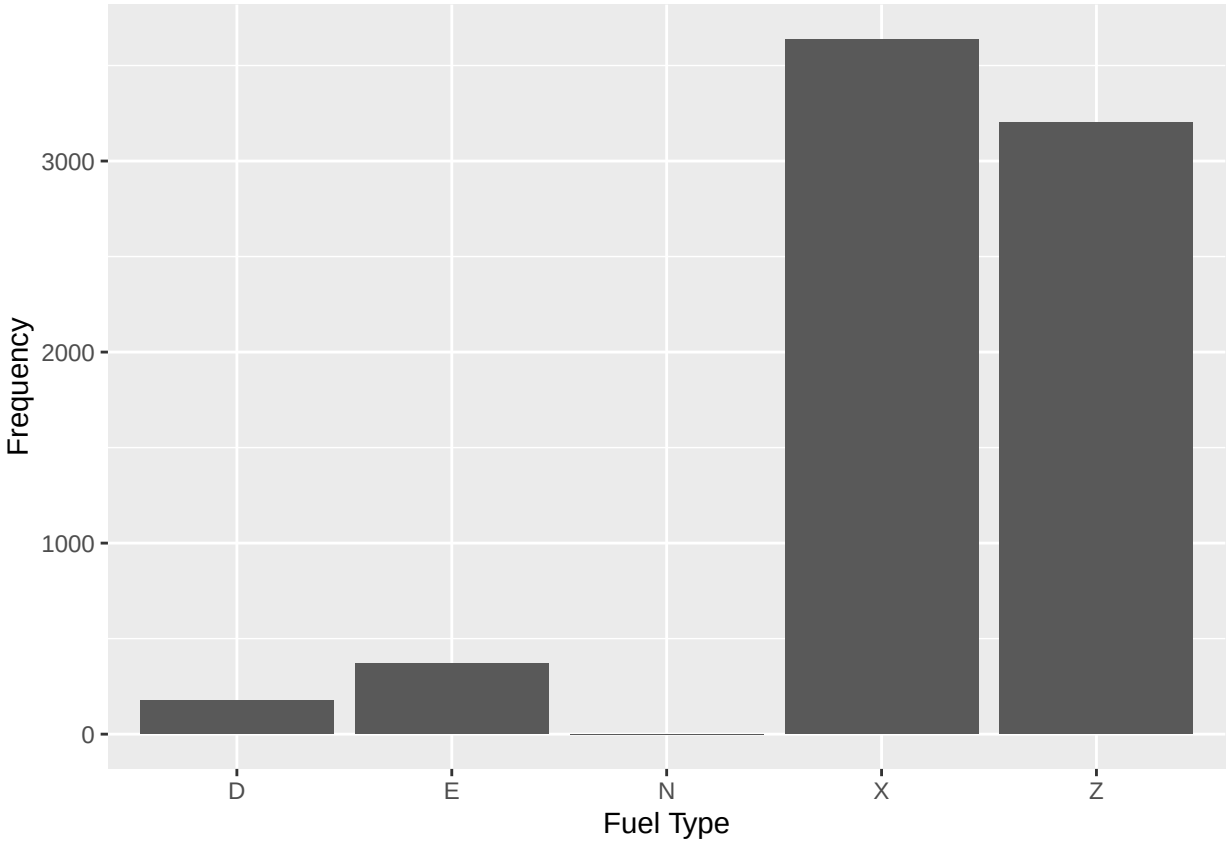
Histograms and bar charts of engine specifications to find the most common engine type



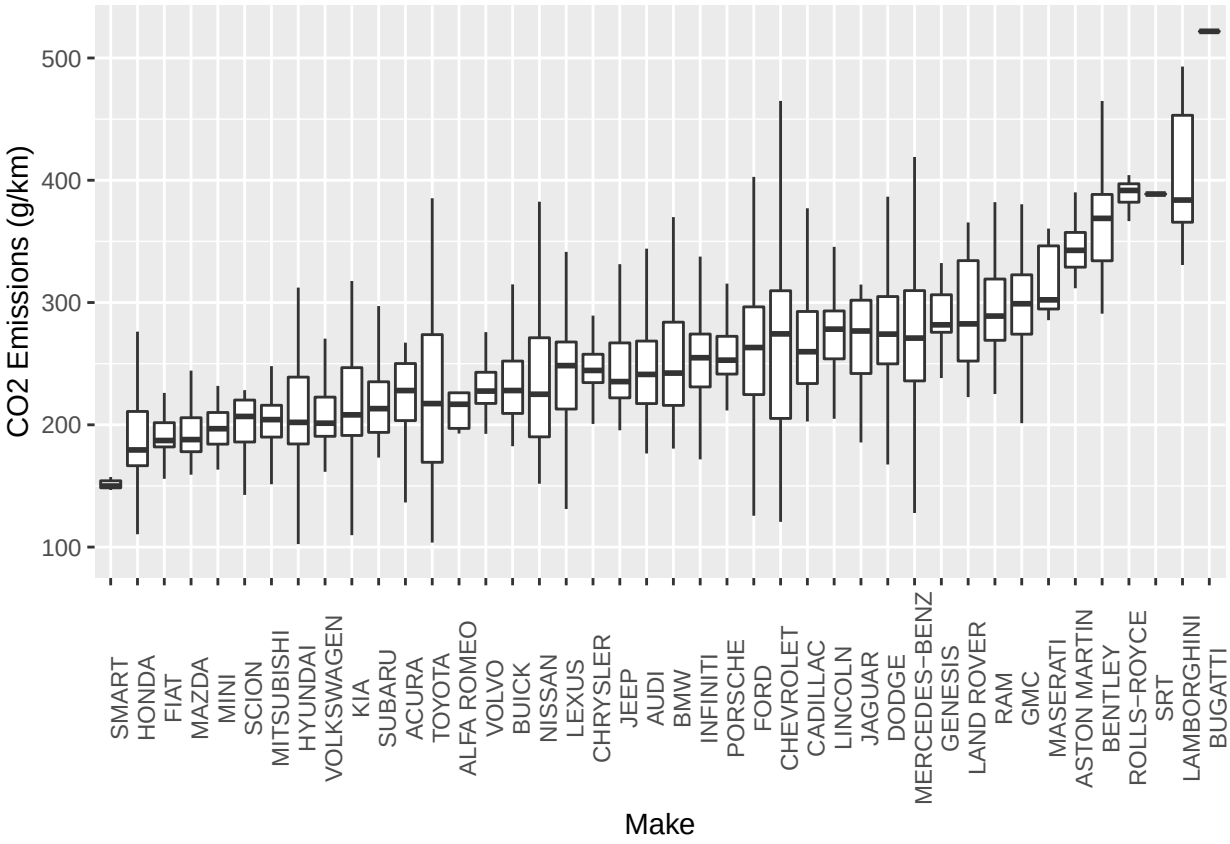

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

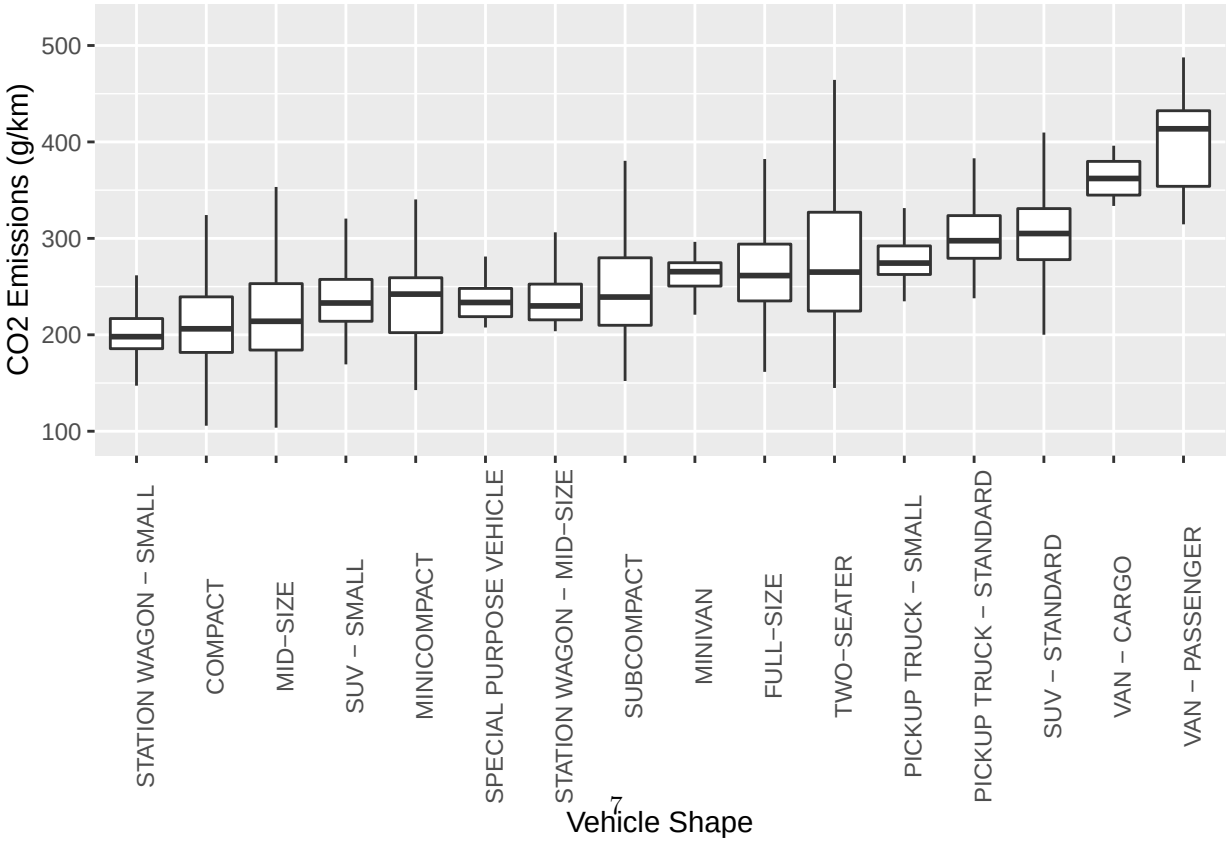

3000
2000
1000
0
Frequency
4
8
12
16
Number of Cylinders


3000
2000
1000
0
Frequency
A
AM
AS
AV
M
Transmission

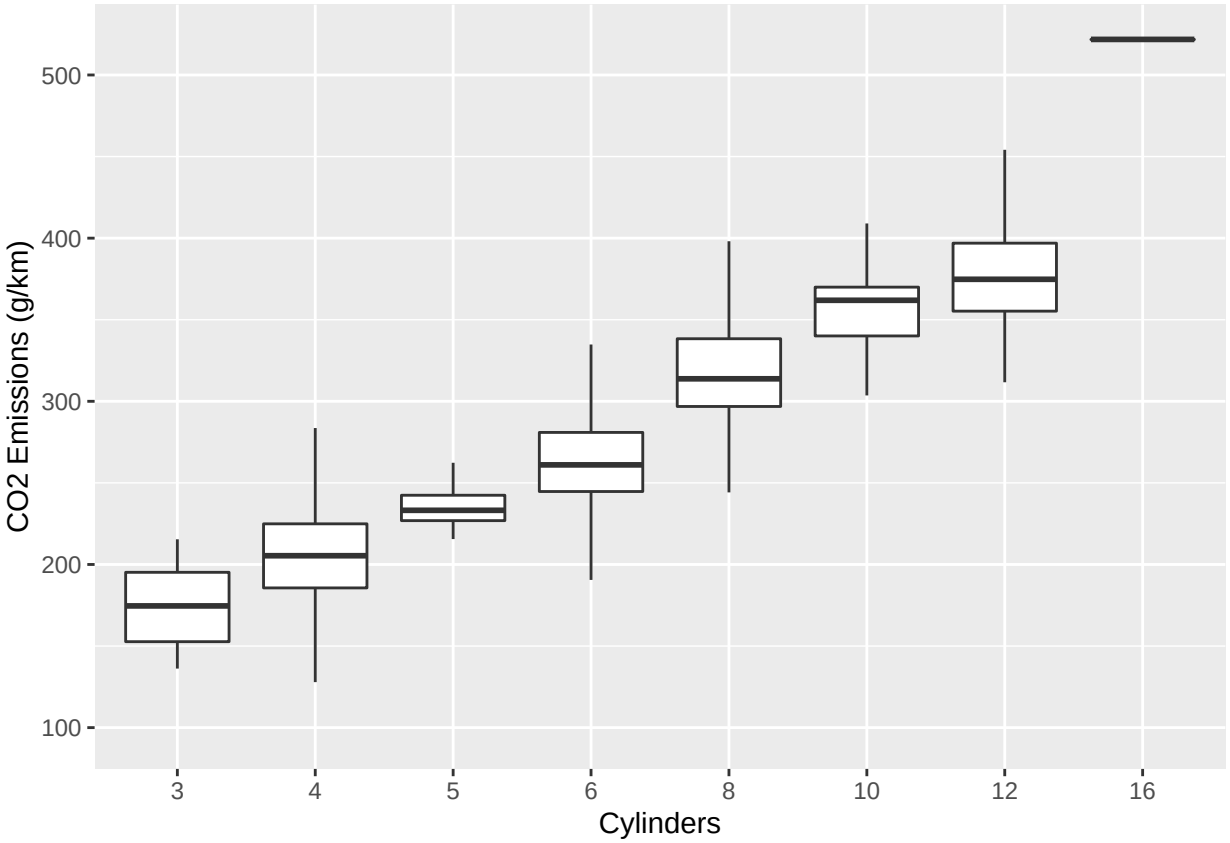

Frequency
3000
2000
1000
0
D
E
N
X
Z
Fuel Type

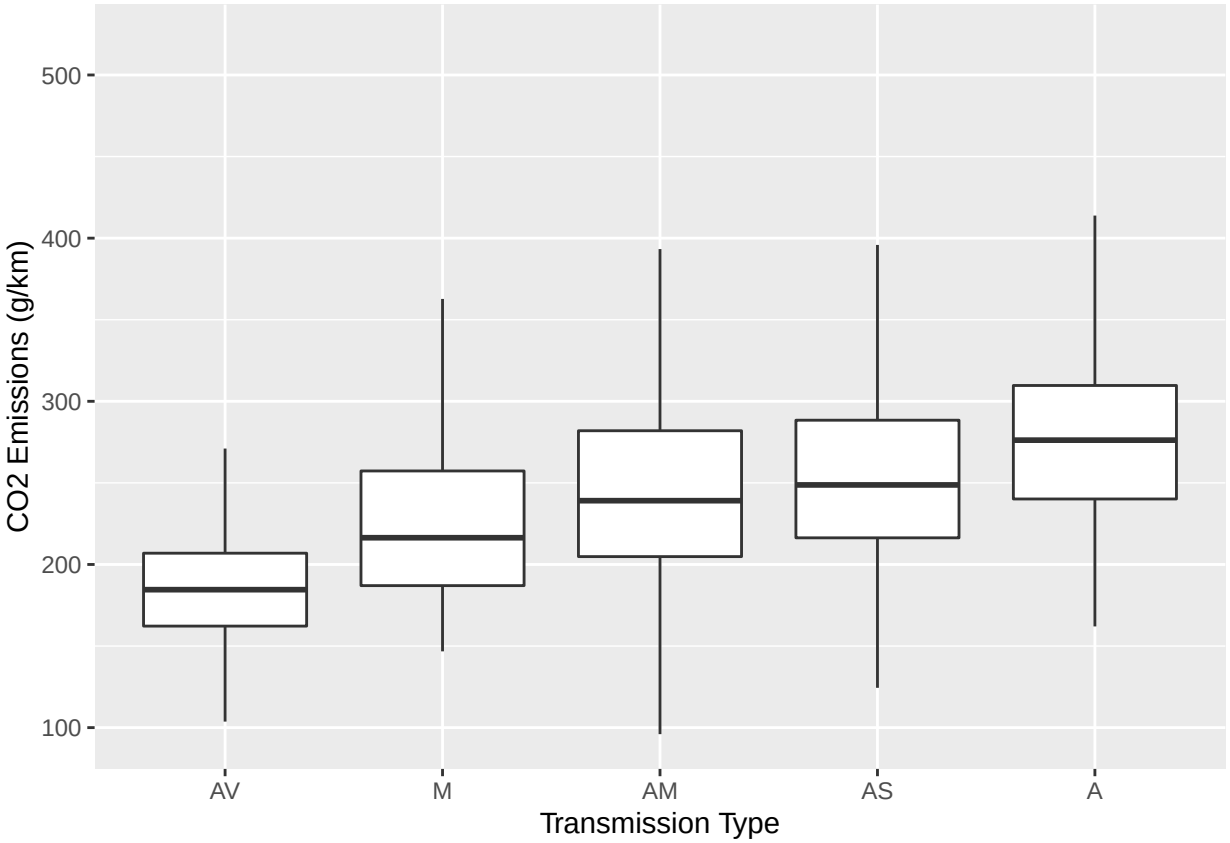
Scatter plots and box plots to find what the correlation is between each variable and CO2 emissions

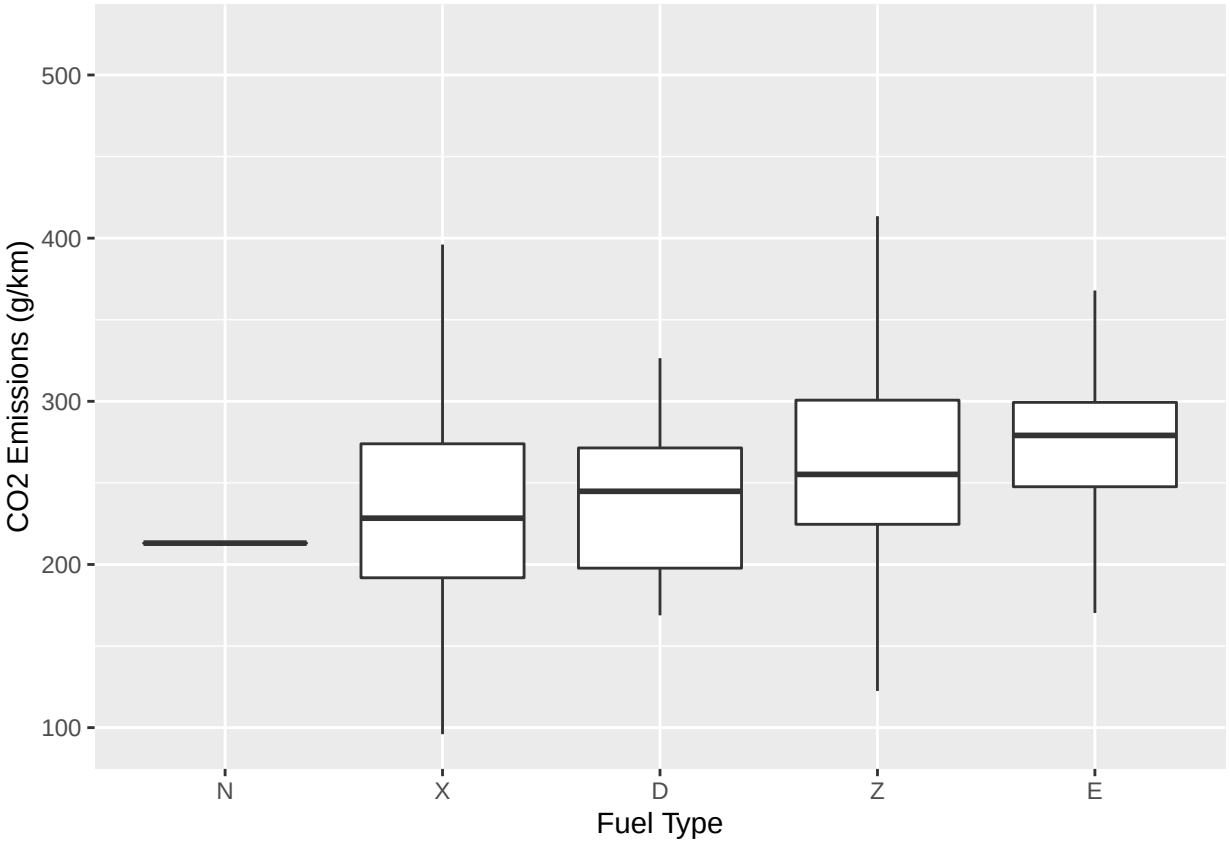


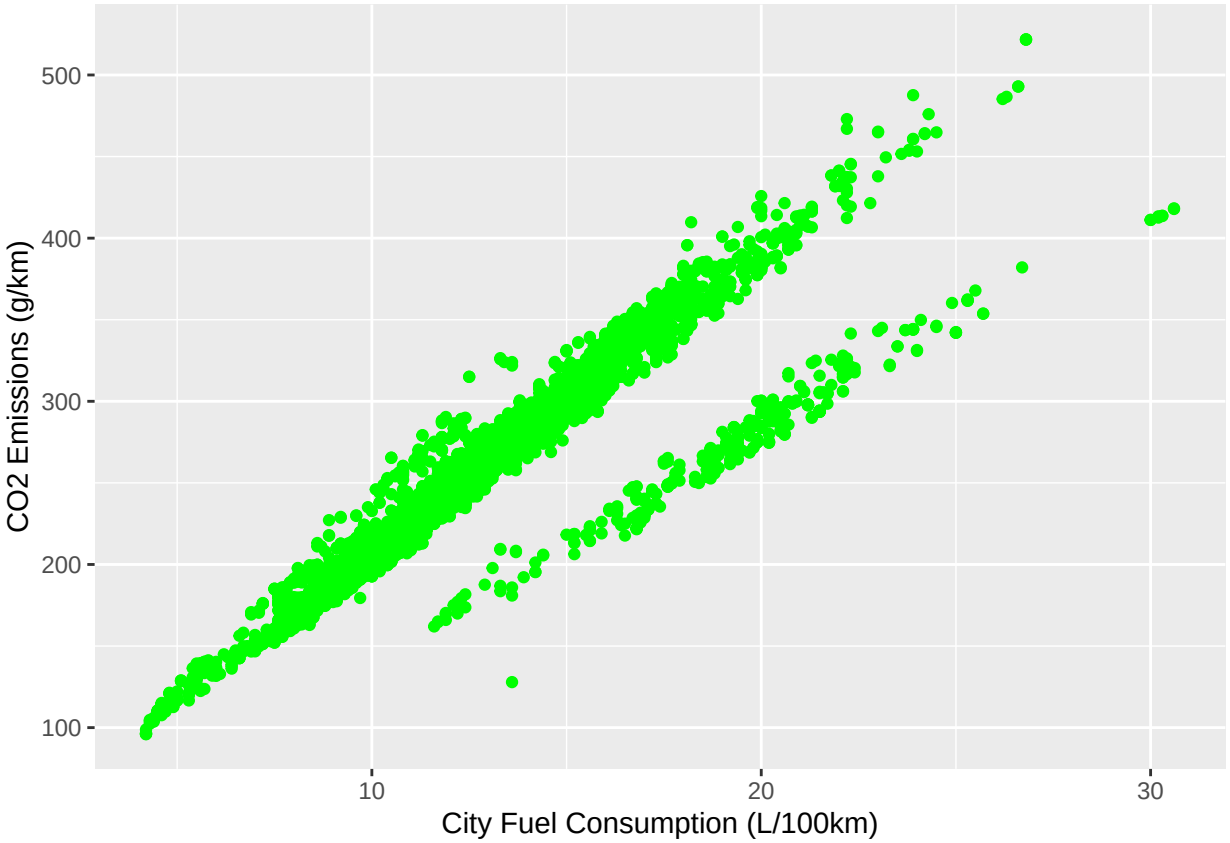


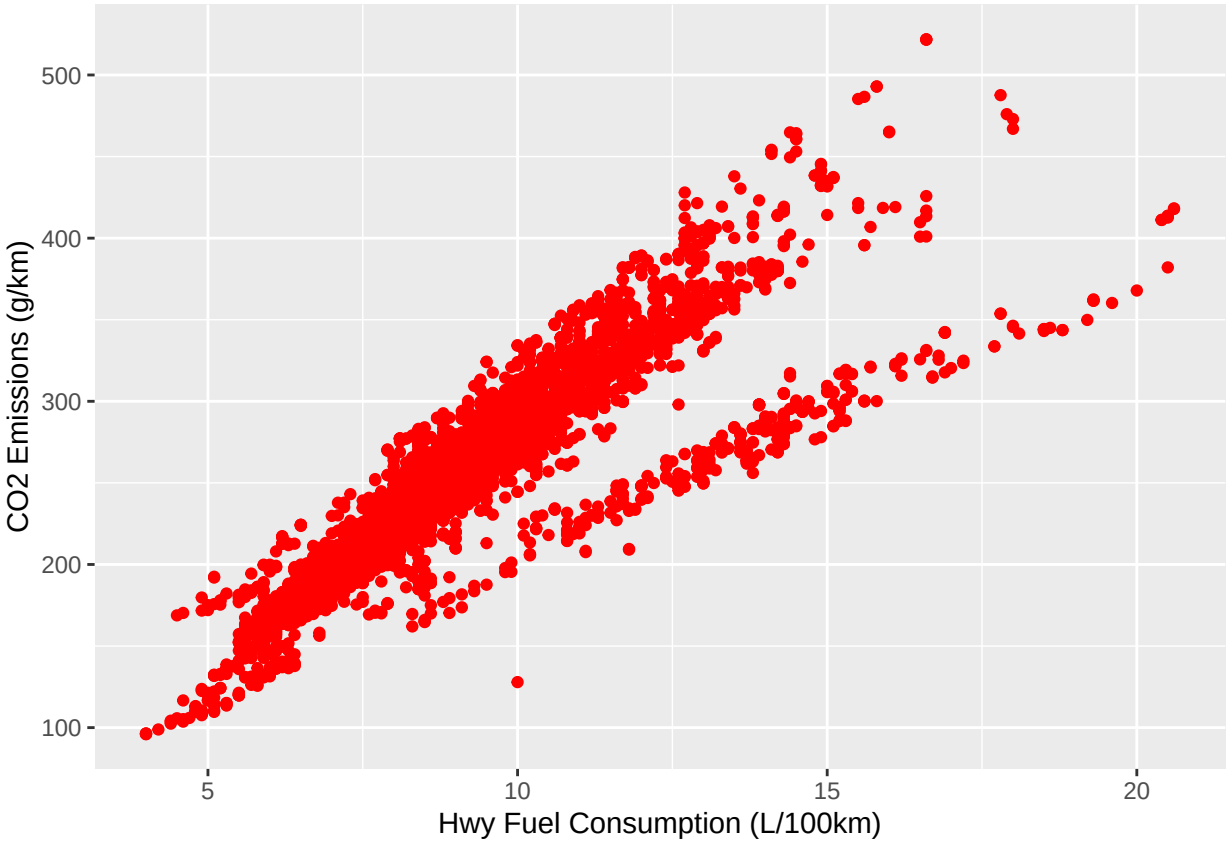

CO2 Emissions (g/km)
500
400
300
200
100
2
4
6
8
Engine Displacement (L)

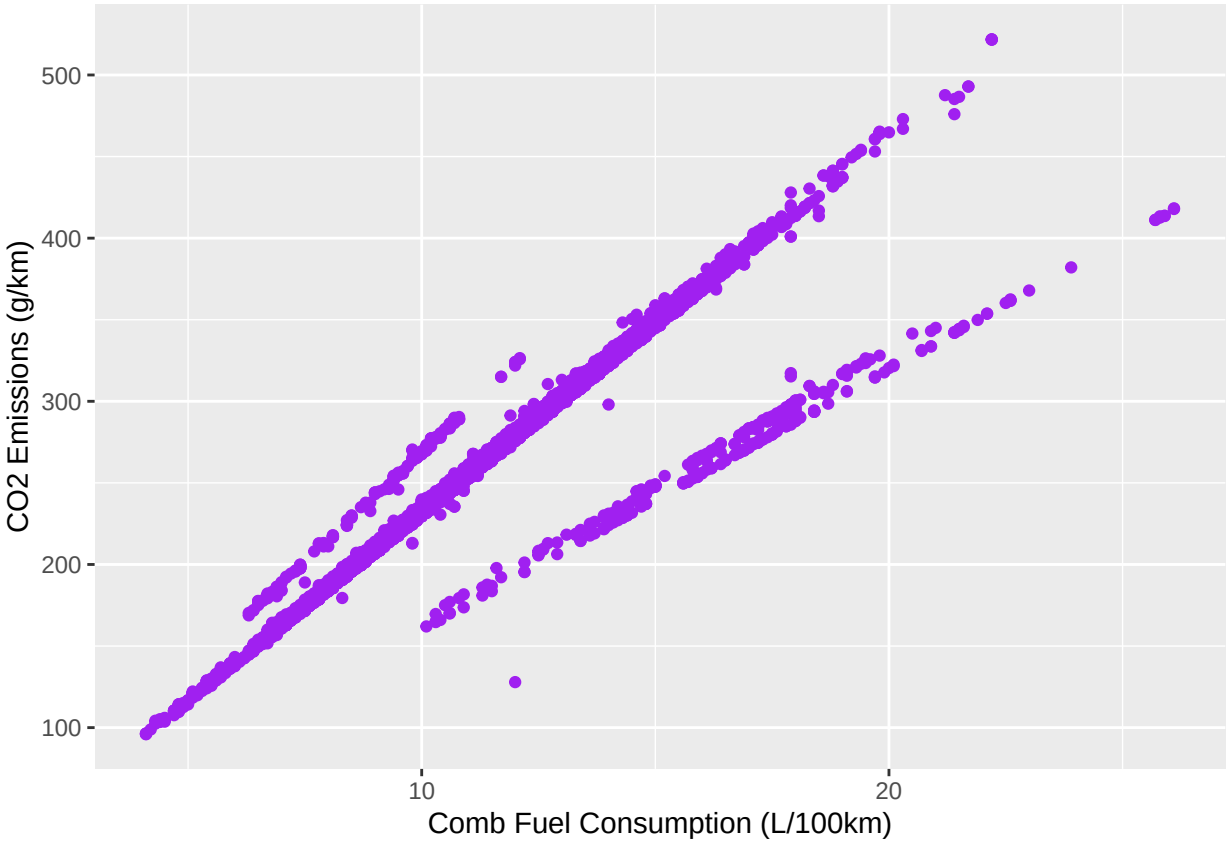

500
400
300
200
100
CO2 Emissions (g/km)
3
4
5
6
8
10
12
16
Cylinders


500
400
300
200
100
CO2 Emissions (g/km)
AV
M
AM
AS
A
Transmission Type


500
400
300
200
100
CO2 Emissions (g/km)
N
X
D
Z
E
Fuel Type

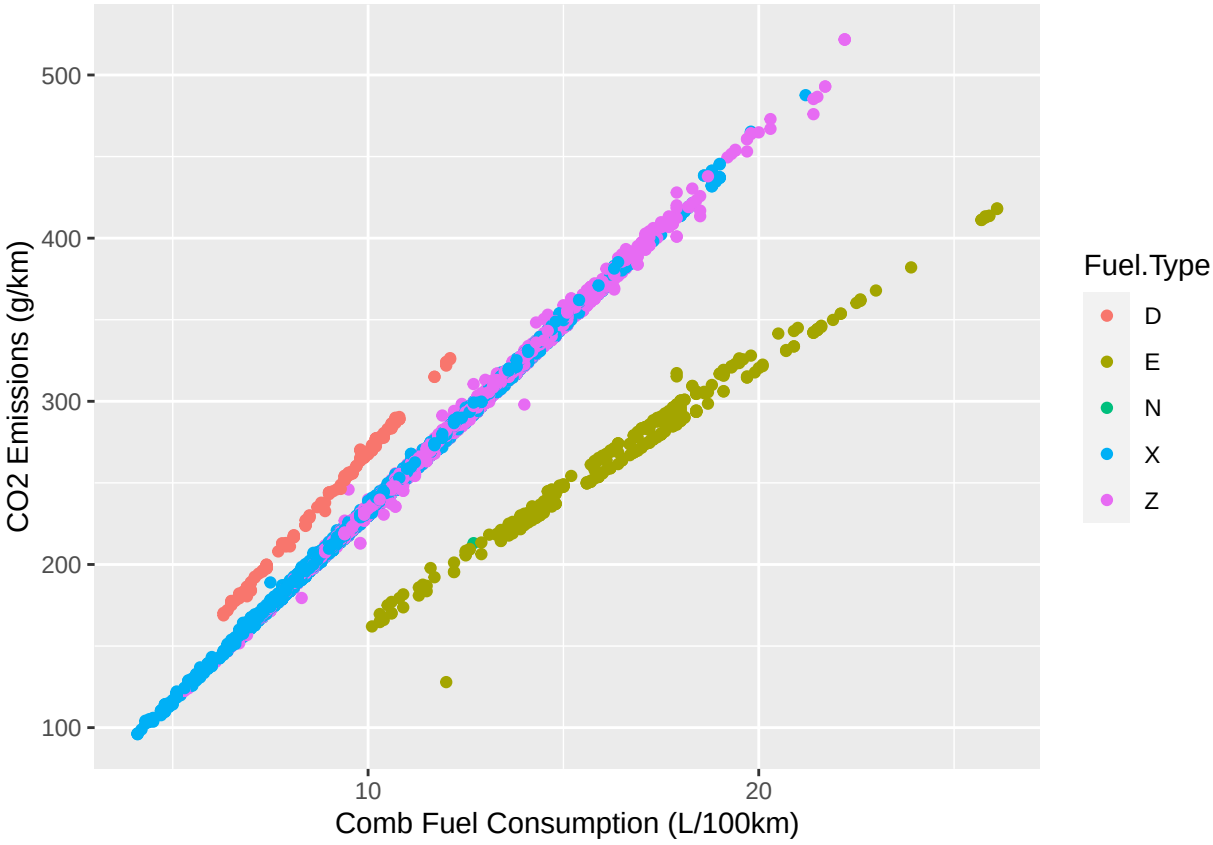

500
400
300
200
100
CO2 Emissions (g/km)
10
20
30
City Fuel Consumption (L/100km)


500
400
300
200
100
CO2 Emissions (g/km)
5
10
15
20
Hwy Fuel Consumption (L/100km)


500
400
300
200
100
CO2 Emissions (g/km)
10
20
Comb Fuel Consumption (L/100km)

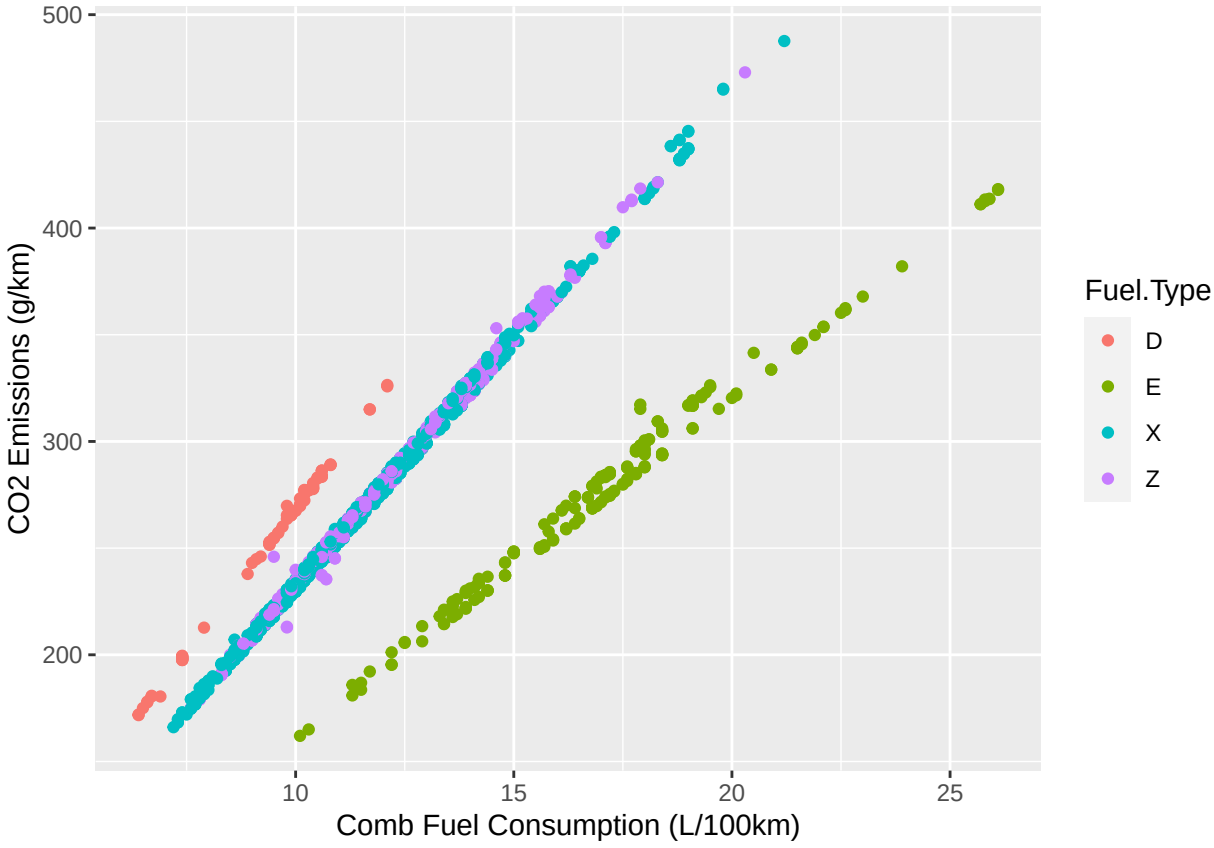
The emissions vs fuel consumption plot displayed multiple lines. Let's separate that data to see if we can see any interesting information.

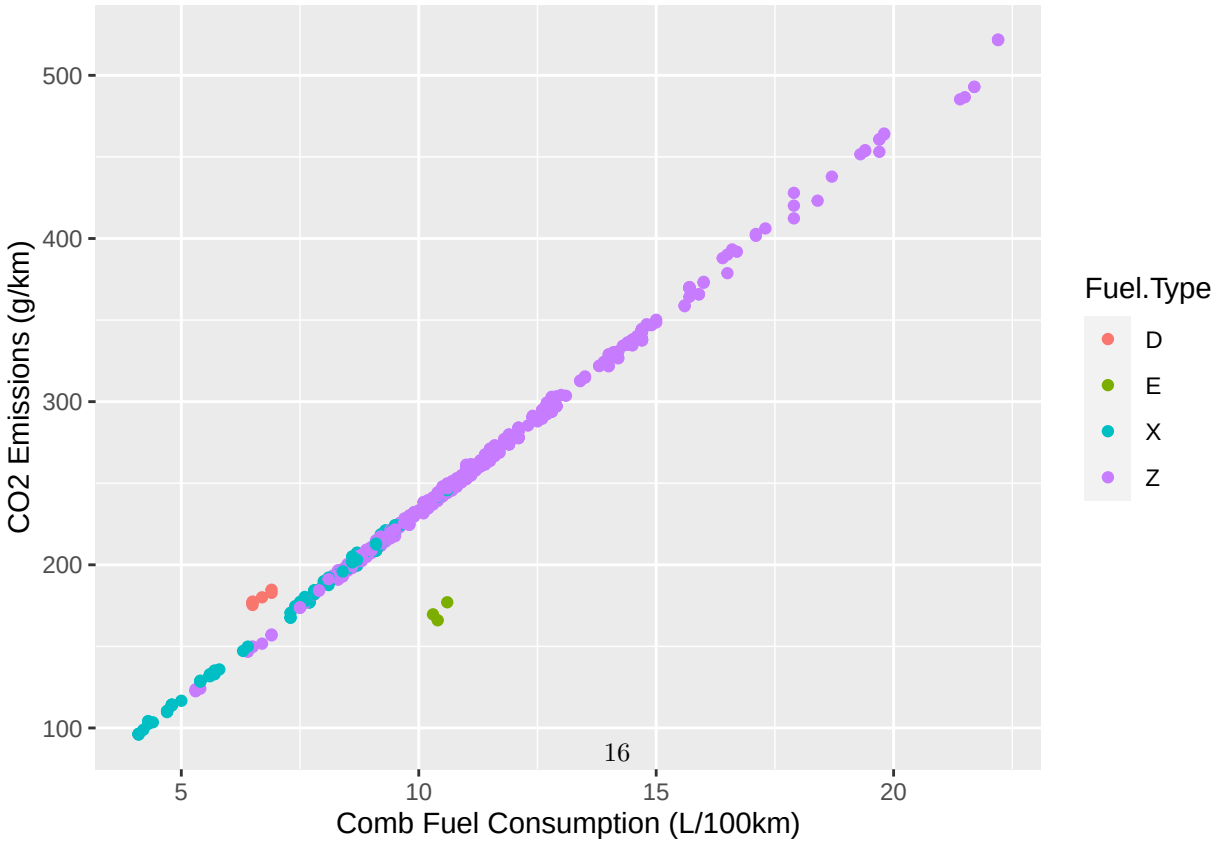
Starting with separating by fuel type:

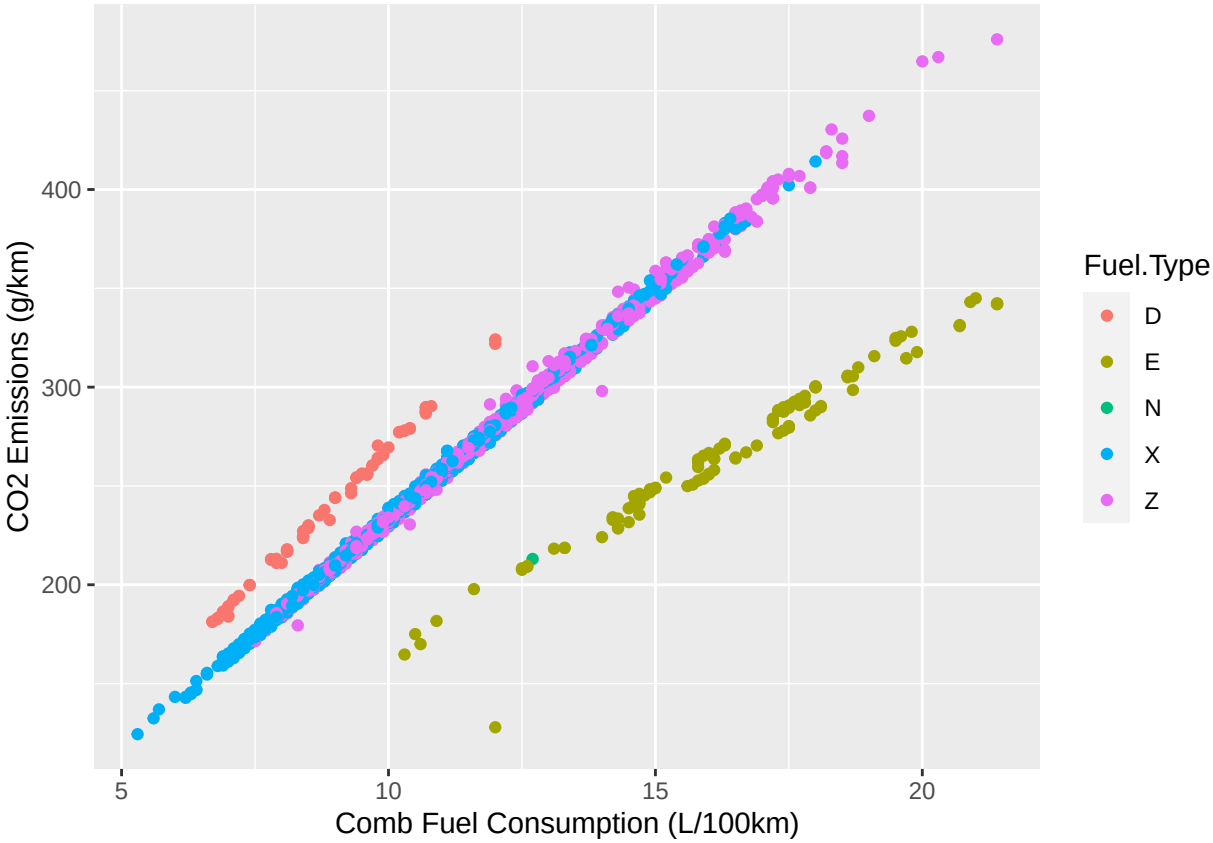


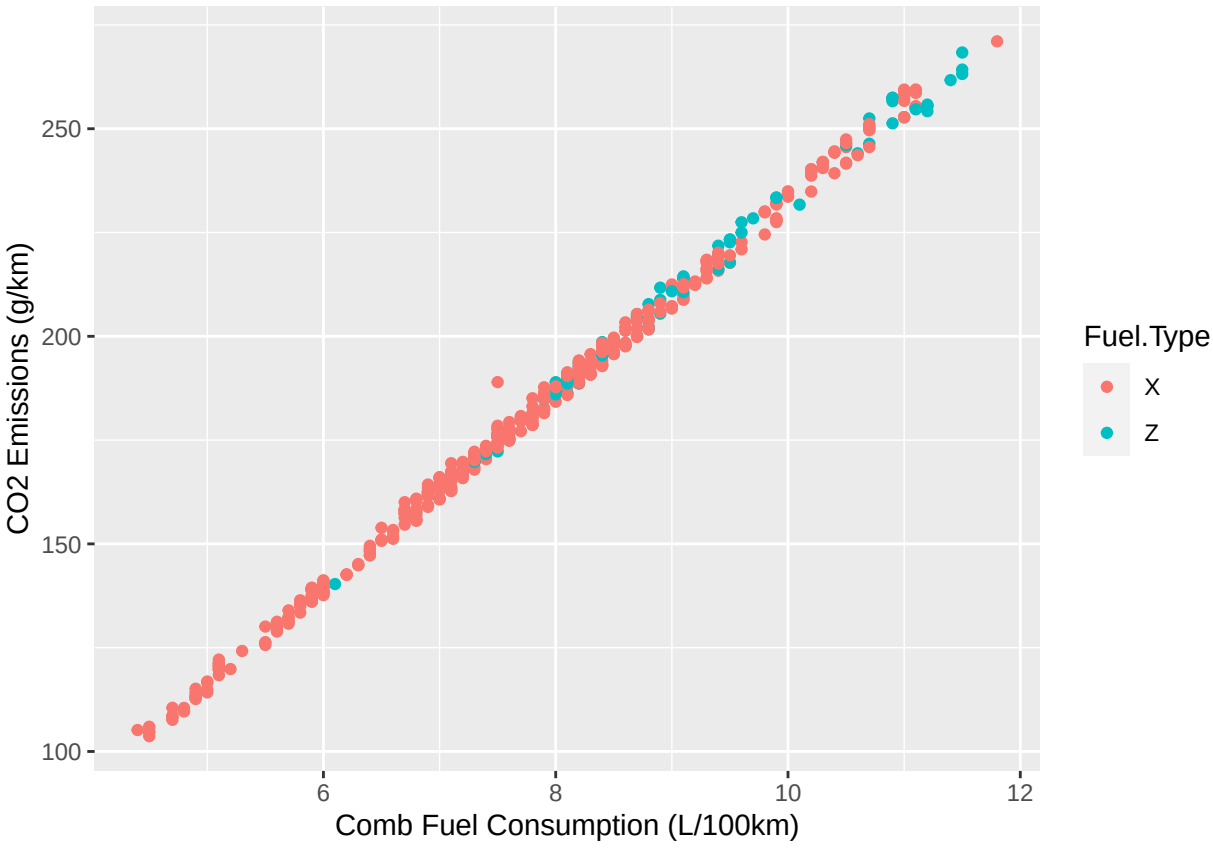
We can clearly see that each of the lines seen in the original plot is due to a different fuel type.

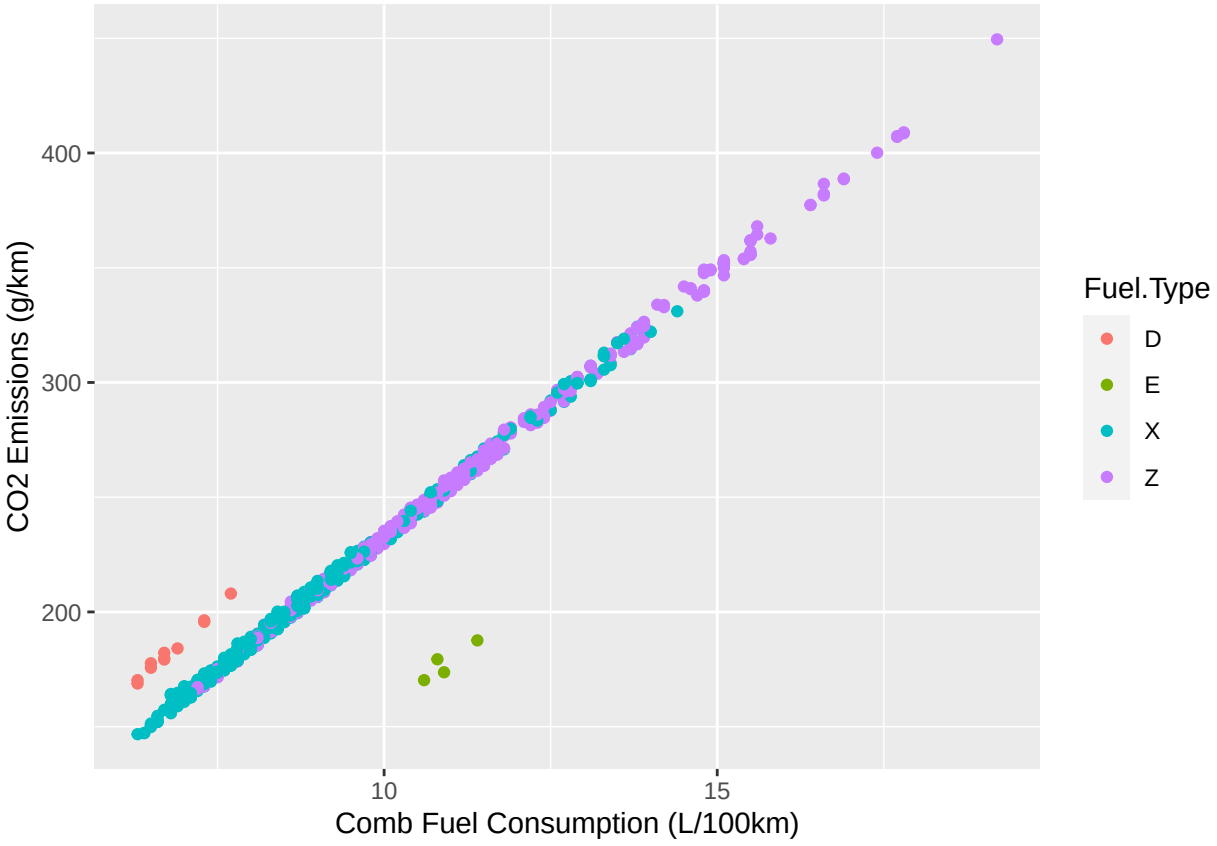
Now I will separate by transmissions and keep the fuel type color coding.






CO2 Emissions (g/km)
400
300
200
5
10
15
20
Comb Fuel Consumption (L/100km)
Fuel.Type
D
E
N
X
Z


CO2 Emissions (g/km)
250
200
150
100
6
8
10
12
Comb Fuel Consumption (L/100km)
Fuel.Type
X
Z



This separation reveals that some transmissions appear to be more common with some fuels.

Automatic and hybrid transmissions can be found on all car types, manual and automated manual transmissions are more common in premium gas cars, and CVTs are more common in regular gas cars. This may be because sports cars, which use premium gas, tend to use manually controlled transmissions due to their powerful engines, whereas CVTs are meant to reduce pollution and fuel consumption so cars that have them use regular gas, which emits the least CO2 of the fuels.

Now I conduct two sample t tests on the following:

The combined fuel consumptions of gas and premium gas cars to see if there is a difference

The CO2 emissions of gas and diesel cars to determine if diesel pollutes more

The CO2 emissions of gas ethanol cars to determine if ethanol is more eco-friendly

Based on the t tests, there is a significant difference between gas and premium gas fuel consumption. There is also a significant difference between gas and diesel CO2 emissions, and gas and ethanol CO2 emissions. More data on diesel and ethanol cars will need to be obtained in the future, since their sample sizes are very small.

```
##
##  Welch Two Sample t-test
##
## data:  cars_gas$Fuel.Consumption.Comb..L.100.km. and cars_premium_gas$Fuel.Consumption.Comb..L.100.km
## t = -22.542, df = 6745.9, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.454563 -1.221821
## sample estimates:
## mean of x mean of y
##  10.08458  11.42277
```

```
##
##  Welch Two Sample t-test
##
## data:  cars_gas$CO2.Emissions.g.km. and cars_diesel$CO2.Emissions.g.km.
## t = -0.73599, df = 206.87, p-value = 0.4626
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -8.943088  4.081010
## sample estimates:
## mean of x mean of y
##  235.1194  237.5505
```

```
##
##  Welch Two Sample t-test
##
## data:  cars_gas$CO2.Emissions.g.km. and cars_ethanol$CO2.Emissions.g.km.
## t = -15.206, df = 487.91, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -45.09777 -34.77653
## sample estimates:
## mean of x mean of y
##  235.1194  275.0566
```

The scatter plots and box plots revealed that all the variables are linearly correlated with CO2 emissions. Therefore, I will set up a multiple linear regression model to predict CO2 emissions.

```
##
## Call:
## lm(formula = CO2.Emissions.g.km. ~ Make + Vehicle.Class + Transmission +
##     Fuel.Type + Cylinders + Engine.Size.L. + Fuel.Consumption.City..L.100.km. +
```

```
##     Fuel.Consumption.Hwy..L.100.km. + Fuel.Consumption.Comb..L.100.km.,
##     data = cars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -52.187  -2.221   0.161   1.896  43.034
##
## Coefficients:
##                                    Estimate Std. Error  t value Pr(>|t|)
## (Intercept)                       3.467e+01  8.358e-01   41.484  < 2e-16
## MakeALFA ROMEO                    2.342e+00  1.114e+00    2.103 0.035521
## MakeASTON MARTIN                  3.019e+00  1.025e+00    2.946 0.003229
## MakeAUDI                          1.167e+00  6.833e-01    1.707 0.087795
## MakeBENTLEY                       3.550e+00  1.012e+00    3.506 0.000458
## MakeBMW                           8.826e-01  6.486e-01    1.361 0.173578
## MakeBUGATTI                       1.098e+01  3.059e+00    3.590 0.000333
## MakeBUICK                         2.622e+00  8.041e-01    3.261 0.001114
## MakeCADILLAC                      2.079e+00  7.399e-01    2.809 0.004979
## MakeCHEVROLET                     2.618e+00  6.795e-01    3.853 0.000118
## MakeCHRYSLER                      7.394e+00  8.831e-01    8.373  < 2e-16
## MakeDODGE                         4.874e+00  7.446e-01    6.545 6.35e-11
## MakeFIAT                          2.135e+00  9.252e-01    2.307 0.021070
## MakeFORD                          4.580e+00  6.875e-01    6.662 2.90e-11
## MakeGENESIS                       5.495e+00  1.191e+00    4.615 4.00e-06
## MakeGMC                           2.419e+00  7.287e-01    3.320 0.000904
## MakeHONDA                         1.766e+00  7.403e-01    2.385 0.017114
## MakeHYUNDAI                       2.597e+00  7.270e-01    3.572 0.000357
## MakeINFINITI                      1.645e+00  7.763e-01    2.119 0.034118
## MakeJAGUAR                        1.252e+00  7.467e-01    1.677 0.093661
## MakeJEEP                          4.395e+00  7.340e-01    5.988 2.23e-09
## MakeKIA                           2.888e+00  7.251e-01    3.983 6.87e-05
## MakeLAMBORGHINI                   3.496e+00  1.096e+00    3.189 0.001434
## MakeLAND ROVER                   -5.373e-01  8.319e-01   -0.646 0.518388
## MakeLEXUS                         7.417e-01  7.185e-01    1.032 0.301918
## MakeLINCOLN                       3.582e+00  8.387e-01    4.271 1.97e-05
## MakeMASERATI                      4.165e+00  9.182e-01    4.535 5.84e-06
## MakeMAZDA                         1.203e+00  7.412e-01    1.622 0.104763
## MakeMERCEDES-BENZ                 1.733e+00  6.644e-01    2.609 0.009104
## MakeMINI                          3.819e-01  7.175e-01    0.532 0.594619
## MakeMITSUBISHI                    1.377e+00  8.331e-01    1.653 0.098425
## MakeNISSAN                        3.229e+00  7.202e-01    4.483 7.48e-06
## MakePORSCHE                       1.798e+00  6.960e-01    2.583 0.009812
## MakeRAM                           3.978e+00  8.820e-01    4.511 6.56e-06
## MakeROLLS-ROYCE                   2.594e+00  1.006e+00    2.578 0.009952
## MakeSCION                         6.946e-01  1.256e+00    0.553 0.580355
## MakeSMART                        -4.142e+00  2.021e+00   -2.050 0.040443
## MakeSRT                           1.126e+00  3.635e+00    0.310 0.756689
## MakeSUBARU                        2.638e+00  7.769e-01    3.395 0.000690
## MakeTOYOTA                        1.646e+00  7.034e-01    2.340 0.019319
## MakeVOLKSWAGEN                    1.131e+00  7.220e-01    1.566 0.117303
## MakeVOLVO                         9.389e-01  7.731e-01    1.214 0.224600
## Vehicle.ClassFULL-SIZE            7.159e-01  2.883e-01    2.483 0.013050
## Vehicle.ClassMID-SIZE            -2.536e-01  2.320e-01   -1.093 0.274480
## Vehicle.ClassMINICOMPACT         -6.085e-01  4.132e-01   -1.473 0.140870
```

```
## Vehicle.ClassMINIVAN                    1.016e+00  6.202e-01    1.638 0.101535
## Vehicle.ClassPICKUP TRUCK - SMALL       4.863e+00  5.214e-01    9.326  < 2e-16
## Vehicle.ClassPICKUP TRUCK - STANDARD    2.225e+00  4.199e-01    5.299 1.20e-07
## Vehicle.ClassSPECIAL PURPOSE VEHICLE    3.658e+00  6.389e-01    5.726 1.07e-08
## Vehicle.ClassSTATION WAGON - MID-SIZE   7.914e-01  7.250e-01    1.092 0.275070
## Vehicle.ClassSTATION WAGON - SMALL      3.325e-02  3.694e-01    0.090 0.928277
## Vehicle.ClassSUBCOMPACT                -8.453e-04  2.740e-01   -0.003 0.997539
## Vehicle.ClassSUV - SMALL                1.505e+00  2.593e-01    5.804 6.75e-09
## Vehicle.ClassSUV - STANDARD             1.795e+00  3.329e-01    5.391 7.21e-08
## Vehicle.ClassTWO-SEATER                 2.213e-01  3.365e-01    0.658 0.510761
## Vehicle.ClassVAN - CARGO               -1.447e+01  1.202e+00  -12.043  < 2e-16
## Vehicle.ClassVAN - PASSENGER           -8.390e+00  8.070e-01  -10.397  < 2e-16
## TransmissionAM                          7.479e-01  3.148e-01    2.376 0.017533
## TransmissionAS                          7.036e-01  2.160e-01    3.257 0.001131
## TransmissionAV                         -1.240e+00  3.509e-01   -3.533 0.000414
## TransmissionM                          -1.379e-02  2.436e-01   -0.057 0.954870
## Fuel.TypeE                             -1.416e+02  6.122e-01 -231.353  < 2e-16
## Fuel.TypeN                             -1.094e+02  5.050e+00  -21.657  < 2e-16
## Fuel.TypeX                             -2.988e+01  4.243e-01  -70.423  < 2e-16
## Fuel.TypeZ                             -2.806e+01  4.484e-01  -62.574  < 2e-16
## Cylinders                               7.196e-01  1.151e-01    6.252 4.28e-10
## Engine.Size.L.                          5.512e-02  1.547e-01    0.356 0.721630
## Fuel.Consumption.City..L.100.km.        6.513e+00  7.714e-01    8.442  < 2e-16
## Fuel.Consumption.Hwy..L.100.km.         4.881e+00  6.443e-01    7.576 4.01e-14
## Fuel.Consumption.Comb..L.100.km.        1.062e+01  1.397e+00    7.600 3.32e-14
##
## (Intercept)                            ***
## MakeALFA ROMEO                         *
## MakeASTON MARTIN                       **
## MakeAUDI                               .
## MakeBENTLEY                            ***
## MakeBMW
## MakeBUGATTI                            ***
## MakeBUICK                              **
## MakeCADILLAC                           **
## MakeCHEVROLET                          ***
## MakeCHRYSLER                           ***
## MakeDODGE                              ***
## MakeFIAT                               *
## MakeFORD                               ***
## MakeGENESIS                            ***
## MakeGMC                                ***
## MakeHONDA                              *
## MakeHYUNDAI                            ***
## MakeINFINITI                           *
## MakeJAGUAR                             .
## MakeJEEP                               ***
## MakeKIA                                ***
## MakeLAMBORGHINI                        **
## MakeLAND ROVER
## MakeLEXUS
## MakeLINCOLN                            ***
## MakeMASERATI                           ***
## MakeMAZDA
```

```
## MakeMERCEDES-BENZ                     **
## MakeMINI
## MakeMITSUBISHI                        .
## MakeNISSAN                            ***
## MakePORSCHE                           **
## MakeRAM                               ***
## MakeROLLS-ROYCE                       **
## MakeSCION
## MakeSMART                             *
## MakeSRT
## MakeSUBARU                            ***
## MakeTOYOTA                            *
## MakeVOLKSWAGEN
## MakeVOLVO
## Vehicle.ClassFULL-SIZE                *
## Vehicle.ClassMID-SIZE
## Vehicle.ClassMINICOMPACT
## Vehicle.ClassMINIVAN
## Vehicle.ClassPICKUP TRUCK - SMALL     ***
## Vehicle.ClassPICKUP TRUCK - STANDARD  ***
## Vehicle.ClassSPECIAL PURPOSE VEHICLE  ***
## Vehicle.ClassSTATION WAGON - MID-SIZE
## Vehicle.ClassSTATION WAGON - SMALL
## Vehicle.ClassSUBCOMPACT
## Vehicle.ClassSUV - SMALL              ***
## Vehicle.ClassSUV - STANDARD           ***
## Vehicle.ClassTWO-SEATER
## Vehicle.ClassVAN - CARGO              ***
## Vehicle.ClassVAN - PASSENGER          ***
## TransmissionAM                        *
## TransmissionAS                        **
## TransmissionAV                        ***
## TransmissionM
## Fuel.TypeE                            ***
## Fuel.TypeN                            ***
## Fuel.TypeX                            ***
## Fuel.TypeZ                            ***
## Cylinders                             ***
## Engine.Size.L.
## Fuel.Consumption.City..L.100.km.      ***
## Fuel.Consumption.Hwy..L.100.km.       ***
## Fuel.Consumption.Comb..L.100.km.      ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.019 on 7315 degrees of freedom
## Multiple R-squared:  0.9927, Adjusted R-squared:  0.9926
## F-statistic: 1.444e+04 on 69 and 7315 DF,  p-value: < 2.2e-16

## [1] 0.9927103
```

Calculating the RMSE of the model

```
## [1] 1.001919
```

All Cook's distances are less than 1, so nothing is having an undue influence on the model.

The following formula was used to find the bounds the covariance ratio should be in:

$1 \pm [3(k+1)/n]$

About 5% of ratios are outside the boundaries (shown below), but because all Cook's distances are below 1, there is little cause for alarm.

```
## numeric(0)
```

```
##   [1] 1.033958e+00 1.031494e+00 1.031734e+00 1.032396e+00 1.032949e+00
##   [6] 1.032674e+00 1.031762e+00 1.033122e+00 1.033840e+00 1.043775e+00
##  [11] 8.153864e-01 5.812110e-01 9.582068e-01 9.664204e-01 9.355363e-01
##  [16] 9.321586e-01 9.519017e-01 9.407011e-01 9.408905e-01 8.362781e-01
##  [21] 8.357893e-01 7.151504e-01 3.883741e-01 9.679663e-01 3.675793e-01
##  [26] 9.689546e-01 9.157657e-01 9.140320e-01 9.183702e-01 9.192156e-01
##  [31] 9.662442e-01 8.037696e-01 8.680502e-01 8.166528e-01 6.364571e-01
##  [36] 5.376701e-01 6.208568e-01 7.973309e-01 7.760842e-01 9.454154e-01
##  [41] 9.401526e-01 7.738993e-01 7.448032e-01 5.949734e-01 6.125347e-01
##  [46] 6.148263e-01 5.821756e-01 9.543482e-01 9.655901e-01 9.369796e-01
##  [51] 9.362383e-01 9.452452e-01 9.401493e-01 9.328592e-01 8.512033e-01
##  [56] 8.358471e-01 7.252770e-01 3.927456e-01 3.819141e-01 9.178493e-01
##  [61] 9.261272e-01 7.930650e-01 5.775486e-01 9.181862e-01 9.249883e-01
##  [66] 1.032664e+00 1.031511e+00 1.034558e+00 1.036199e+00 1.035452e+00
##  [71] 7.067547e-01 7.793618e-01 6.993413e-01 6.273237e-01 9.400769e-01
##  [76] 9.230092e-01 1.038476e+00 1.058314e+00 1.061254e+00 1.063653e+00
##  [81] 1.060069e+00 1.059947e+00 1.062835e+00 1.061905e+00 1.061340e+00
##  [86] 1.061221e+00 1.177576e+00 1.177987e+00 2.019241e+00 2.019241e+00
##  [91] 1.032661e+00 1.032207e+00 1.032634e+00 1.032299e+00 1.033310e+00
##  [96] 1.033169e+00 1.033226e+00 1.033449e+00 1.033422e+00 1.041308e+00
## [101] 9.005479e-01 9.483290e-01 3.589798e-01 3.581225e-01 9.614628e-01
## [106] 9.675579e-01 8.827807e-01 7.613831e-01 7.394867e-01 9.262350e-01
## [111] 8.499044e-01 6.320205e-01 6.411618e-01 6.827140e-01 9.489531e-01
## [116] 3.596854e-01 9.669683e-01 3.556359e-01 9.683426e-01 9.005797e-01
## [121] 8.832859e-01 8.913638e-01 9.398742e-01 8.282028e-01 1.032525e+00
## [126] 1.032063e+00 7.017997e-01 6.999088e-01 7.705121e-01 9.265933e-01
## [131] 9.657433e-01 1.038763e+00 1.058728e+00 1.062070e+00 1.063584e+00
## [136] 1.059526e+00 1.060509e+00 1.062093e+00 1.062535e+00 1.178107e+00
## [141] 1.177904e+00 1.046157e+00 1.032676e+00 1.036861e+00 1.034906e+00
## [146] 1.035235e+00 1.035541e+00 1.035725e+00 1.031233e+00 1.033517e+00
## [151] 1.034133e+00 1.032160e+00 1.034664e+00 1.031164e+00 1.031818e+00
## [156] 9.560866e-01 8.088924e-01 8.451365e-01 8.702845e-01 8.565983e-01
## [161] 9.182477e-01 5.349301e+13 6.784635e-01 8.895753e-01 9.446861e-01
## [166] 9.233527e-01 1.031502e+00 1.032529e+00 8.918791e-01 9.530481e-01
```

```
## [171] 6.844982e-01 1.031875e+00 9.426857e-01 9.538881e-01 5.347814e-01
## [176] 5.250326e-01 5.400693e-01 9.536736e-01 9.487190e-01 9.594187e-01
## [181] 9.281019e-01 9.450255e-01 9.667730e-01 8.705707e-01 8.483052e-01
## [186] 8.111550e-01 7.894167e-01 8.528841e-01 7.390968e-01 1.031953e+00
## [191] 1.036089e+00 1.035407e+00 1.032757e+00 1.033607e+00 1.032695e+00
## [196] 9.180677e-01 9.556943e-01 4.992145e-01 1.031364e+00 9.496193e-01
## [201] 7.779419e-01 1.031176e+00 1.032399e+00 1.033813e+00 1.062070e+00
## [206] 1.057654e+00 1.059374e+00 1.058944e+00 1.057481e+00 1.058880e+00
## [211] 1.177671e+00 1.178081e+00 1.033669e+00 1.045754e+00 1.045688e+00
## [216] 1.045641e+00 1.044310e+00 1.039413e+00 1.035514e+00 1.036460e+00
## [221] 1.034812e+00 1.033512e+00 1.033480e+00 1.034088e+00 1.033578e+00
## [226] 1.034631e+00 1.032763e+00 1.033958e+00 1.037271e+00 9.687928e-01
## [231] 8.328159e-01 8.754878e-01 8.566686e-01 8.745499e-01 9.232445e-01
## [236] 9.187885e-01 7.103916e-01 9.091783e-01 9.552082e-01 1.031426e+00
## [241] 9.565079e-01 8.311965e-01 9.537518e-01 9.614287e-01 9.414140e-01
## [246] 5.379039e-01 5.588465e-01 6.247252e-01 9.452242e-01 9.424194e-01
## [251] 9.622379e-01 9.397712e-01 9.352204e-01 1.052720e+00 1.052076e+00
## [256] 1.050629e+00 1.052437e+00 8.558420e-01 8.675878e-01 8.928609e-01
## [261] 8.189982e-01 7.575772e-01 1.032138e+00 1.036157e+00 1.035325e+00
## [266] 1.033874e+00 1.031424e+00 9.192694e-01 9.509740e-01 9.220170e-01
## [271] 8.121454e-01 6.042878e-01 1.031594e+00 1.032898e+00 1.033502e+00
## [276] 1.032123e+00 1.176902e+00 1.034581e+00 1.036685e+00 1.036455e+00
## [281] 1.036201e+00 1.046973e+00 1.046644e+00 1.045488e+00 1.044664e+00
## [286] 1.039699e+00 1.045895e+00 1.043751e+00 1.034756e+00 1.035850e+00
## [291] 1.035545e+00 1.033038e+00 1.035362e+00 1.032825e+00 1.033807e+00
## [296] 1.034015e+00 1.032090e+00 1.514448e+00 9.373044e-01 9.214710e-01
## [301] 8.771878e-01 9.403146e-01 1.031970e+00 9.538899e-01 9.121969e-01
## [306] 9.556068e-01 8.378735e-01 9.500063e-01 9.521154e-01 9.623203e-01
## [311] 9.405867e-01 9.538980e-01 9.368646e-01 9.601523e-01 1.052315e+00
## [316] 1.053165e+00 1.052534e+00 1.051262e+00 1.052470e+00 8.354517e-01
## [321] 8.834155e-01 7.642731e-01 1.035787e+00 1.035771e+00 1.035930e+00
## [326] 1.036003e+00 1.033940e+00 1.034427e+00 1.032331e+00 1.031869e+00
## [331] 1.032710e+00 1.032970e+00 1.034296e+00 1.035960e+00 1.036608e+00
## [336] 1.036626e+00 1.046135e+00 1.045592e+00 1.045753e+00 1.045141e+00
## [341] 1.041398e+00 1.046024e+00 1.045927e+00 1.042231e+00 1.034587e+00
## [346] 1.036659e+00 1.036144e+00 1.035016e+00 1.032679e+00 1.034880e+00
## [351] 1.514421e+00 9.443431e-01 8.786163e-01 9.490181e-01 1.032075e+00
## [356] 9.147589e-01 9.607975e-01 7.854076e-01 9.435193e-01 9.490902e-01
## [361] 9.552529e-01 9.300817e-01 9.619523e-01 9.393307e-01 7.887731e-01
## [366] 7.823917e-01 1.055310e+00 1.052892e+00 1.053898e+00 1.051999e+00
## [371] 1.053220e+00 1.052503e+00 1.049966e+00 1.052503e+00 1.034258e+00
## [376] 1.037583e+00 1.038171e+00 9.386731e-01 1.031717e+00 1.034332e+00
## [381] 9.173076e-01 1.033099e+00 1.049327e+00 1.036612e+00 1.036710e+00
## [386] 1.046421e+00 1.045693e+00 1.044425e+00 1.045837e+00 1.046120e+00
## [391] 1.046174e+00 1.045216e+00 1.035130e+00 1.036307e+00 1.035948e+00
## [396] 1.034948e+00 1.031632e+00 1.035082e+00 1.035932e+00 1.033961e+00
## [401] 1.036065e+00 1.037056e+00 1.036712e+00 1.514414e+00 9.513298e-01
## [406] 8.762408e-01 9.337309e-01 9.497006e-01 9.474243e-01 9.516308e-01
## [411] 9.430848e-01 7.824756e-01 7.700414e-01 1.055368e+00 1.052787e+00
## [416] 1.053864e+00 1.052272e+00 1.053524e+00 1.052429e+00 1.051034e+00
## [421] 1.050017e+00 9.387018e-01 9.209470e-01 1.033143e+00 1.049006e+00
## [426] 1.049746e+00
```

To test the assumption of independece, I conduct a Durbin-Watson test on the model. Because the D-W statistic is above 1, the assumption has been met, indicating the residuals are correlated.

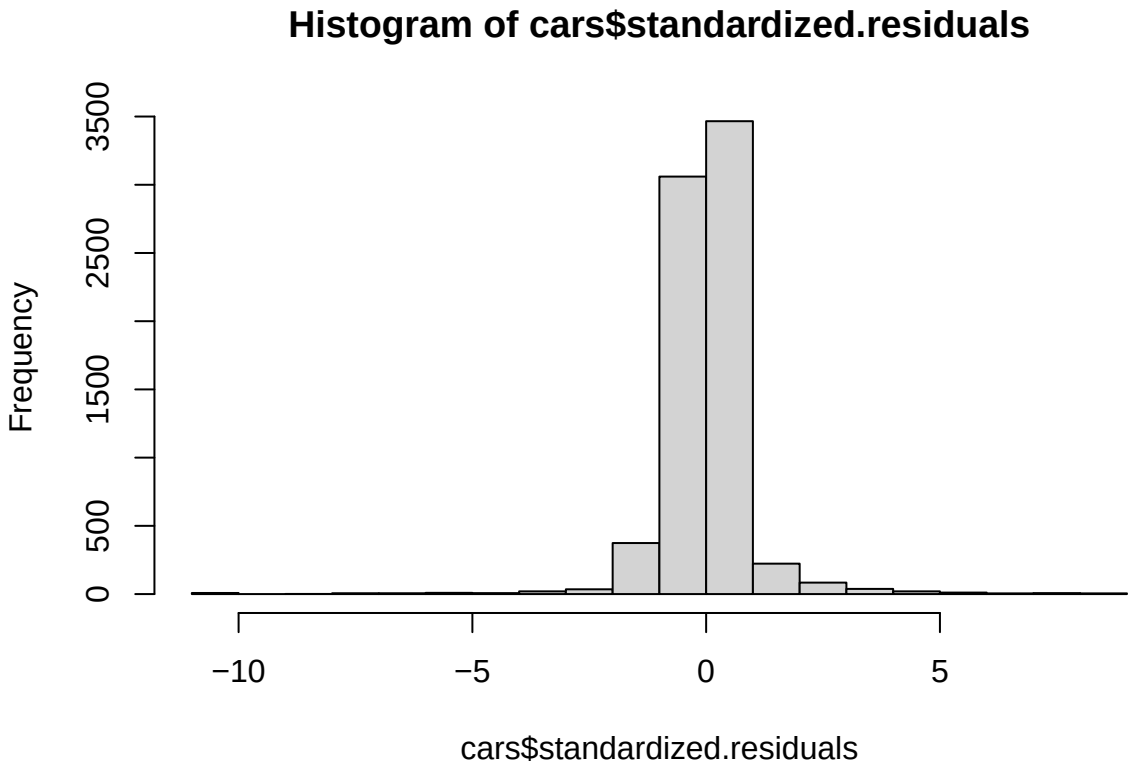
```
##  lag Autocorrelation D-W Statistic p-value
##    1      0.09631205      1.807353       0
##  Alternative hypothesis: rho != 0
```

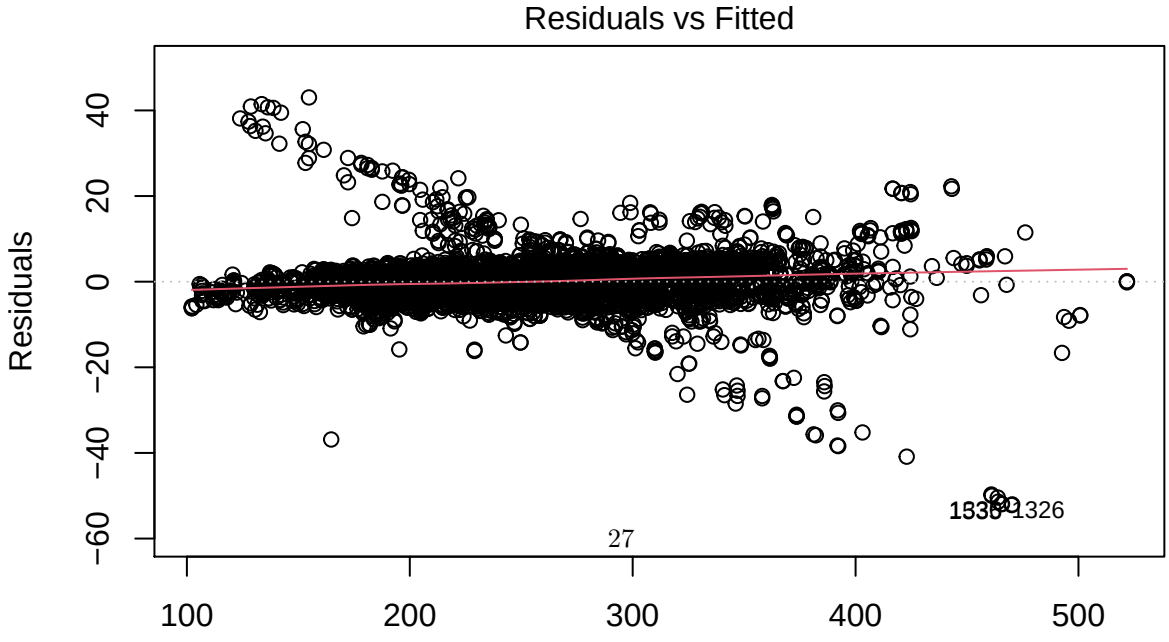
To determine if the model contains multicollinearity, I will construct a correlation matrix and calculate the determinant. If the determinant is greater than 0.00001, then multicollinearity is not a problem.

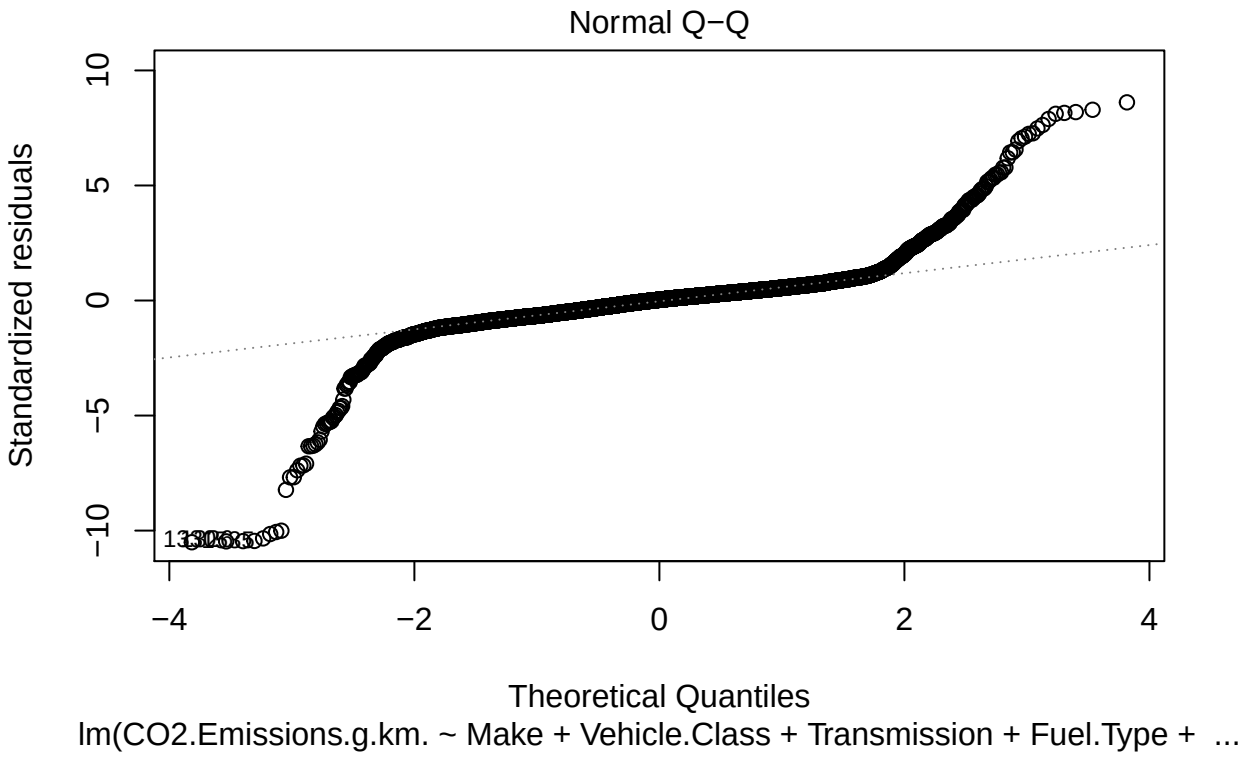
The determinant was found to be significantly smaller than this. This makes sense as many variables were highly linearly correlated with CO2 emissions, and thus, each other.

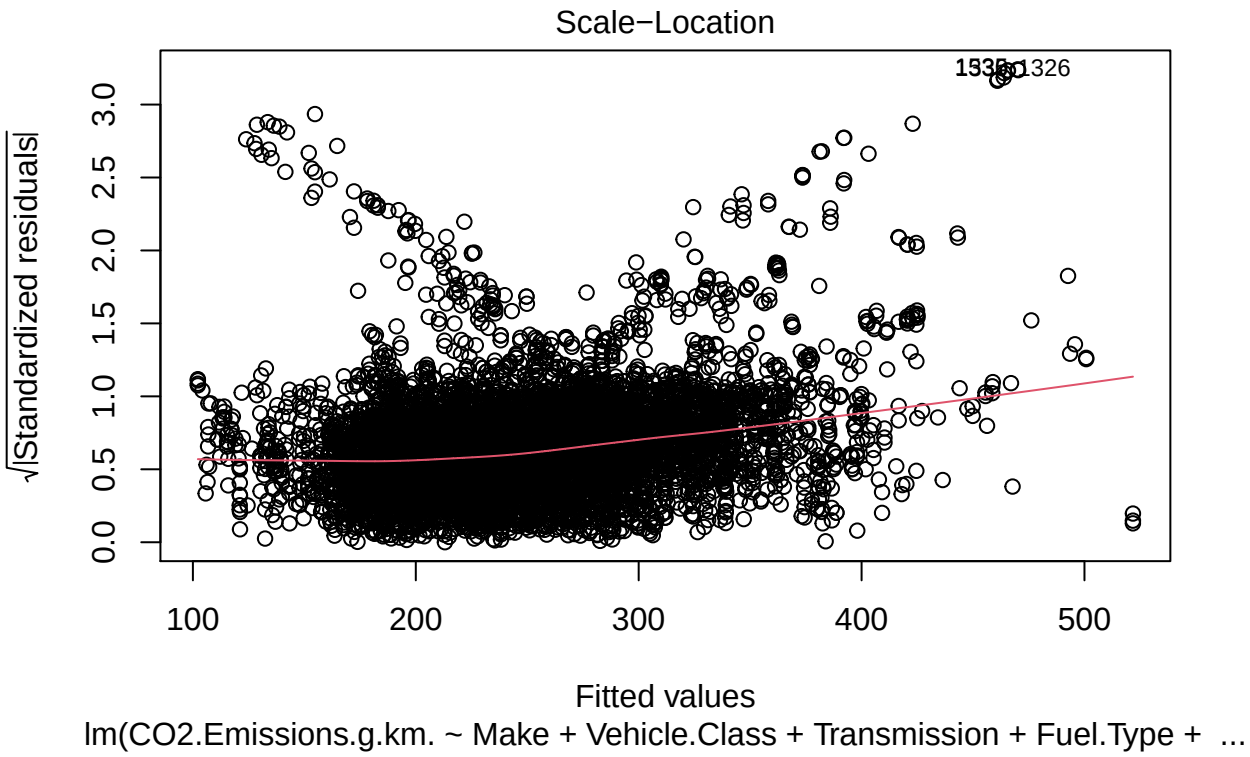
```
## [1] 4.153715e-09
```

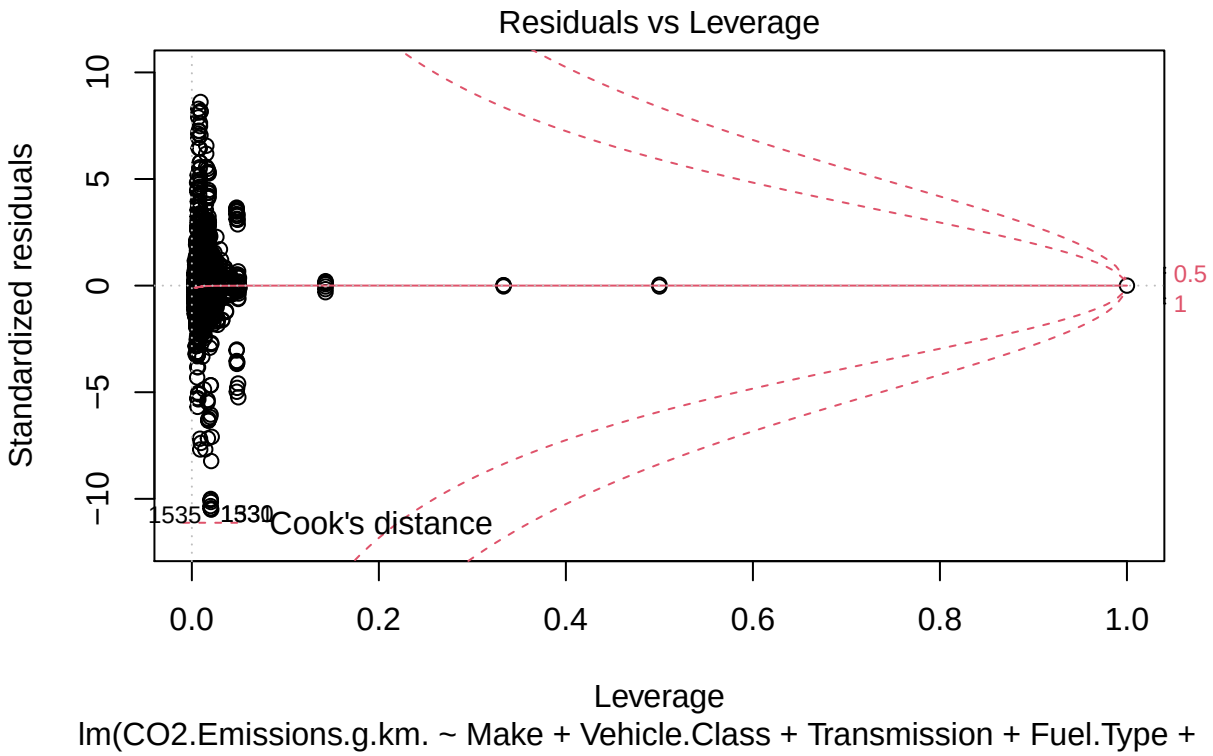
The histogram and Q-Q plot inform us of any deviations from normality seen in the residuals. Based on this, we can clearly see that the distribution of residuals is approximately normal. In addition, the plot of residuals vs fitted values appears to be not random. Residuals appear to remain fairly constant across the fitted values. Due to the non-random residual vs fitted value distribution, the normality of residuals, and the assumption of independence, this model has little bias.






Normal Q−Q
Standardized residuals
10
5
0
−5
−10
−4
−2
0
2
4
Theoretical Quantiles
lm(CO2.Emissions.g.km. ~ Make + Vehicle.Class + Transmission + Fuel.Type + ...


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
2.5
3.0
100
200
300
400
500
Fitted values
lm(CO2.Emissions.g.km. ~ Make + Vehicle.Class + Transmission + Fuel.Type + ...



CO2 emissions were highly linearly correlated with fuel consumption, so it is safe to assume that the same variables most influential on emissions will influence fuel consumption the most as well.