

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'

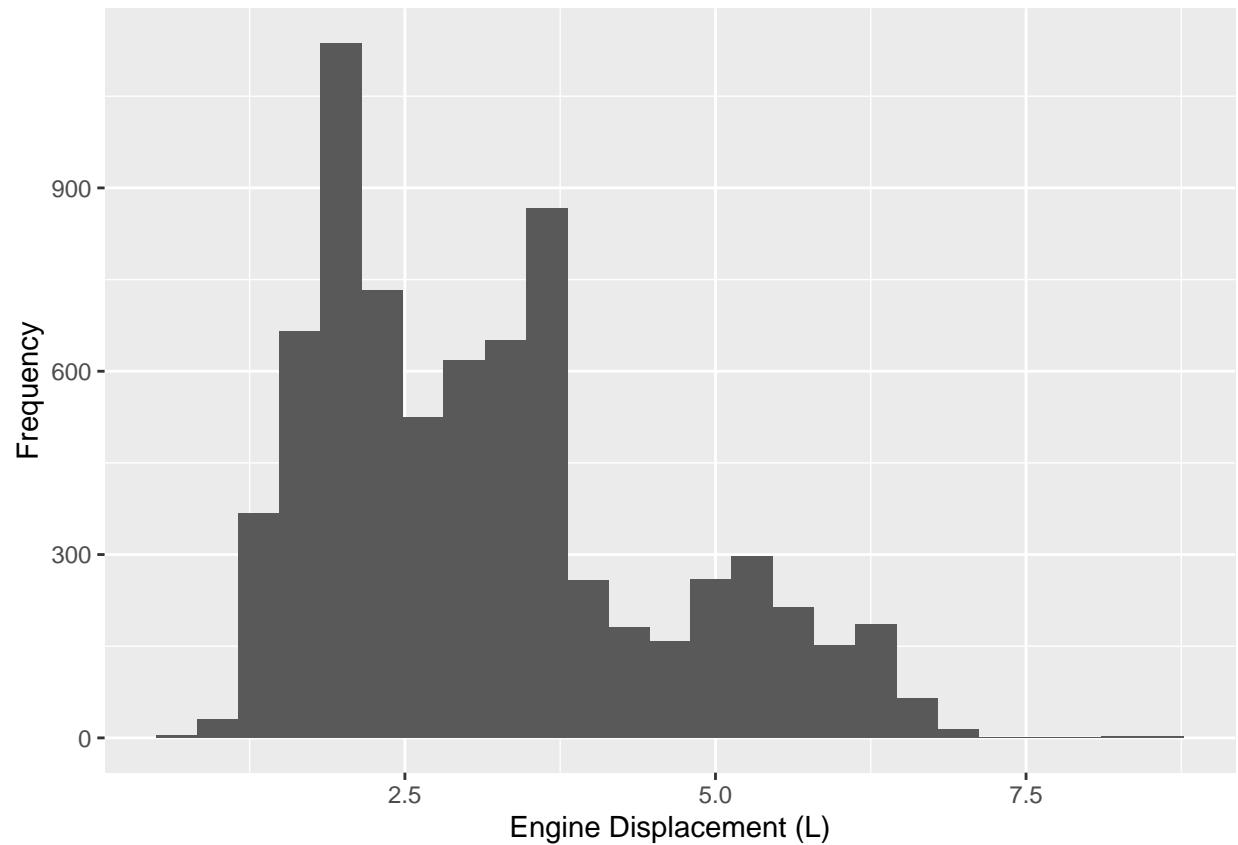
## The following object is masked from 'package:ggplot2':
## 
##     annotate
```

```

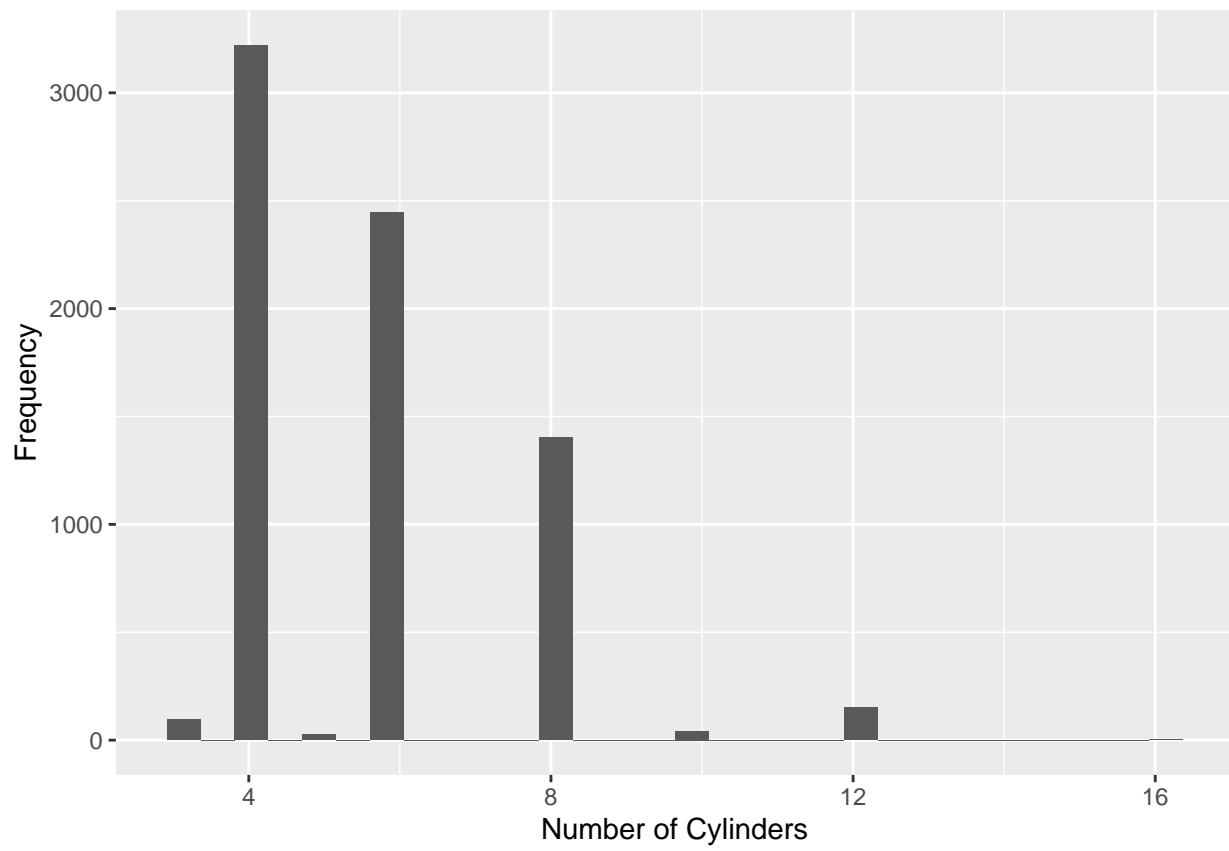
##          Make      Model      Vehicle.Class Engine.Size.L.
##  FORD       : 628 Length:7385     SUV - SMALL    :1217   Min.   :0.695
##  CHEVROLET : 588 Class  :character MID-SIZE      :1133   1st Qu.:2.035
##  BMW        : 527 Mode   :character COMPACT      :1022   Median  :2.936
##  MERCEDES-BENZ: 419                   SUV - STANDARD: 735   Mean    :3.161
##  PORSCHE    : 376                   FULL-SIZE      : 639   3rd Qu.:3.777
##  TOYOTA     : 330                   SUBCOMPACT    : 606   Max.    :8.639
##  (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.   : 96.1
##  1st Qu.:22.00                  1st Qu.:207.7
##  Median  :27.00                  Median :245.8
##  Mean    :27.48                  Mean   :250.6
##  3rd Qu.:32.00                  3rd Qu.:288.3
##  Max.    :69.00                  Max.   :522.2
##
## 'data.frame': 7385 obs. of 12 variables:
##   $ Make           : Factor w/ 42 levels "ACURA","ALFA ROMEO",...
##   $ Model          : chr "ILX" "ILX" "ILX HYBRID" "MDX 4WD" ...
##   $ Vehicle.Class : Factor w/ 16 levels "COMPACT","FULL-SIZE",...
##   $ Engine.Size.L. : num  1.97 2.44 1.25 3.54 3.66 ...
##   $ Cylinders      : int  4 4 4 6 6 6 6 6 6 4 ...
##   $ Transmission   : Factor w/ 5 levels "A","AM","AS",...
##   $ Fuel.Type       : Factor w/ 5 levels "D","E","N","X",...
##   $ 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 ...

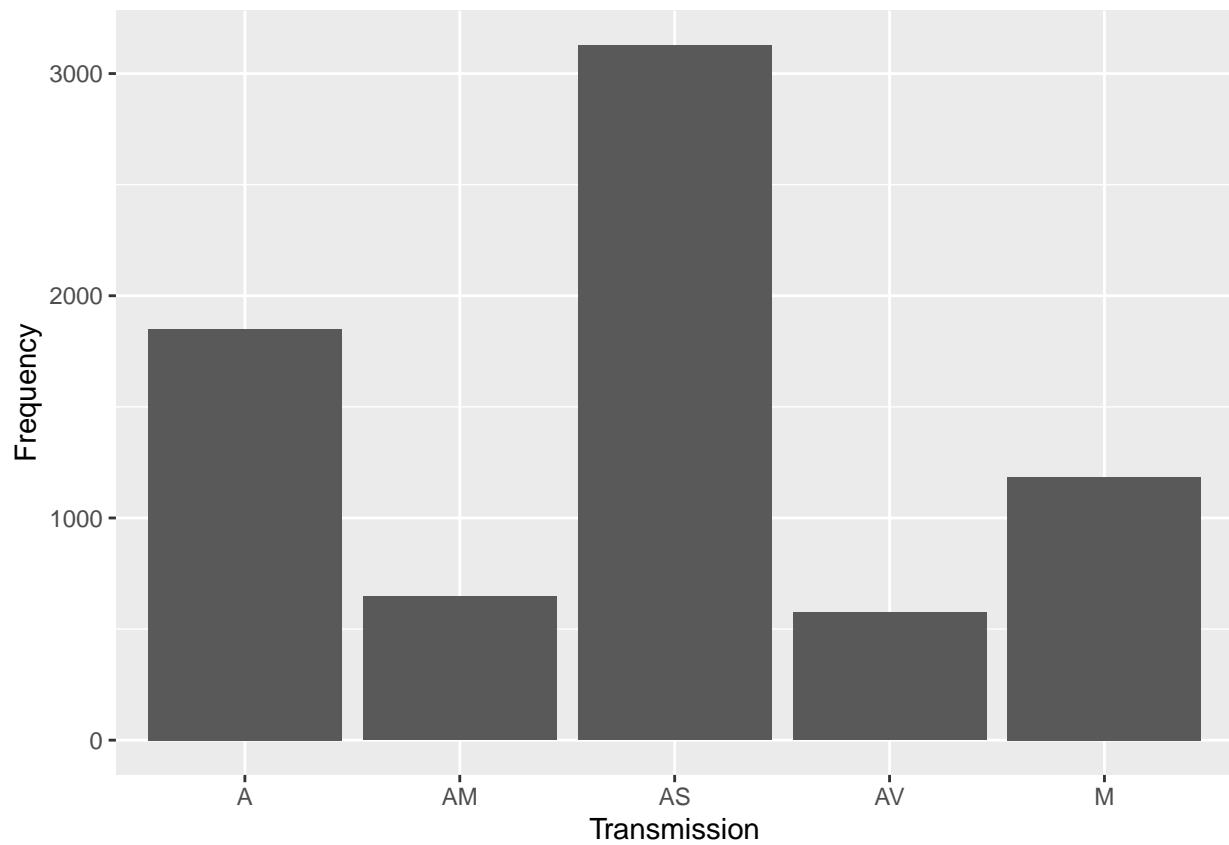
```

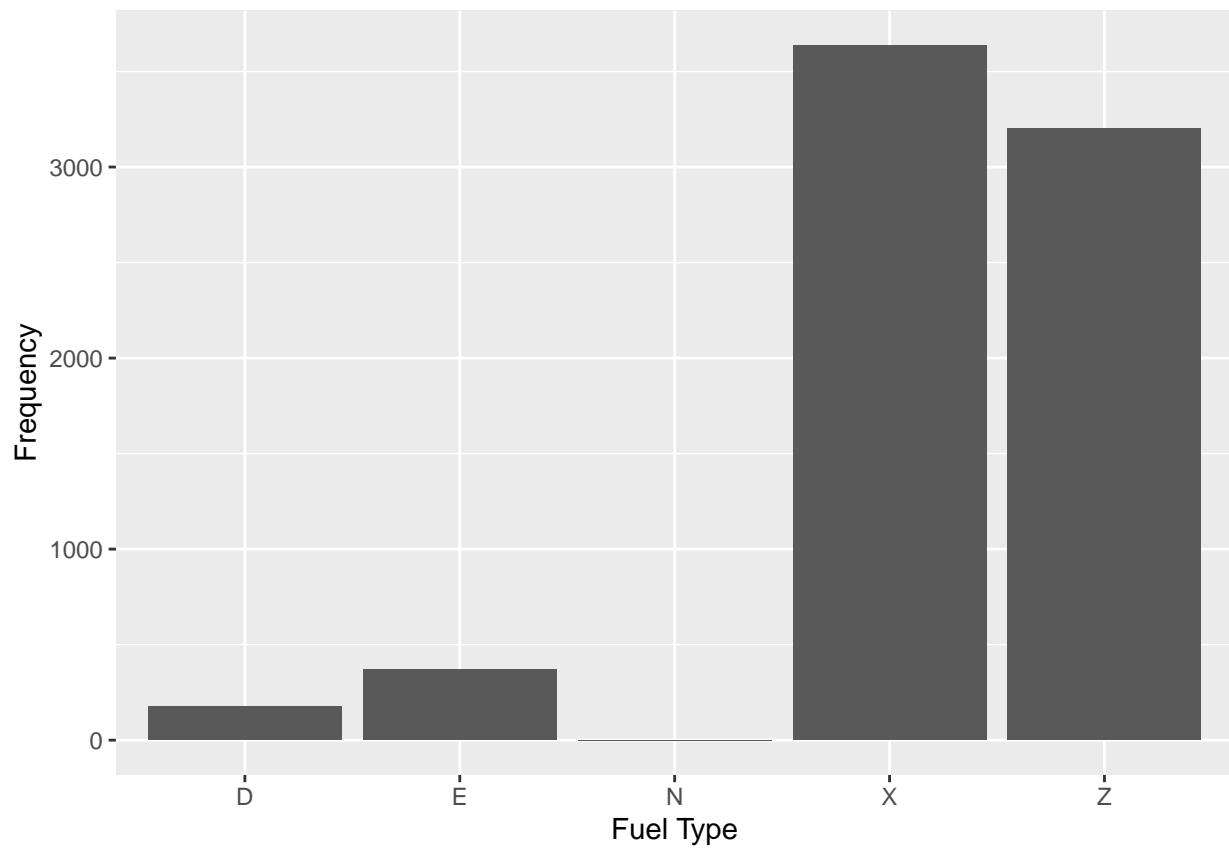
Histograms and bar charts of engine specifications to find the most common engine type



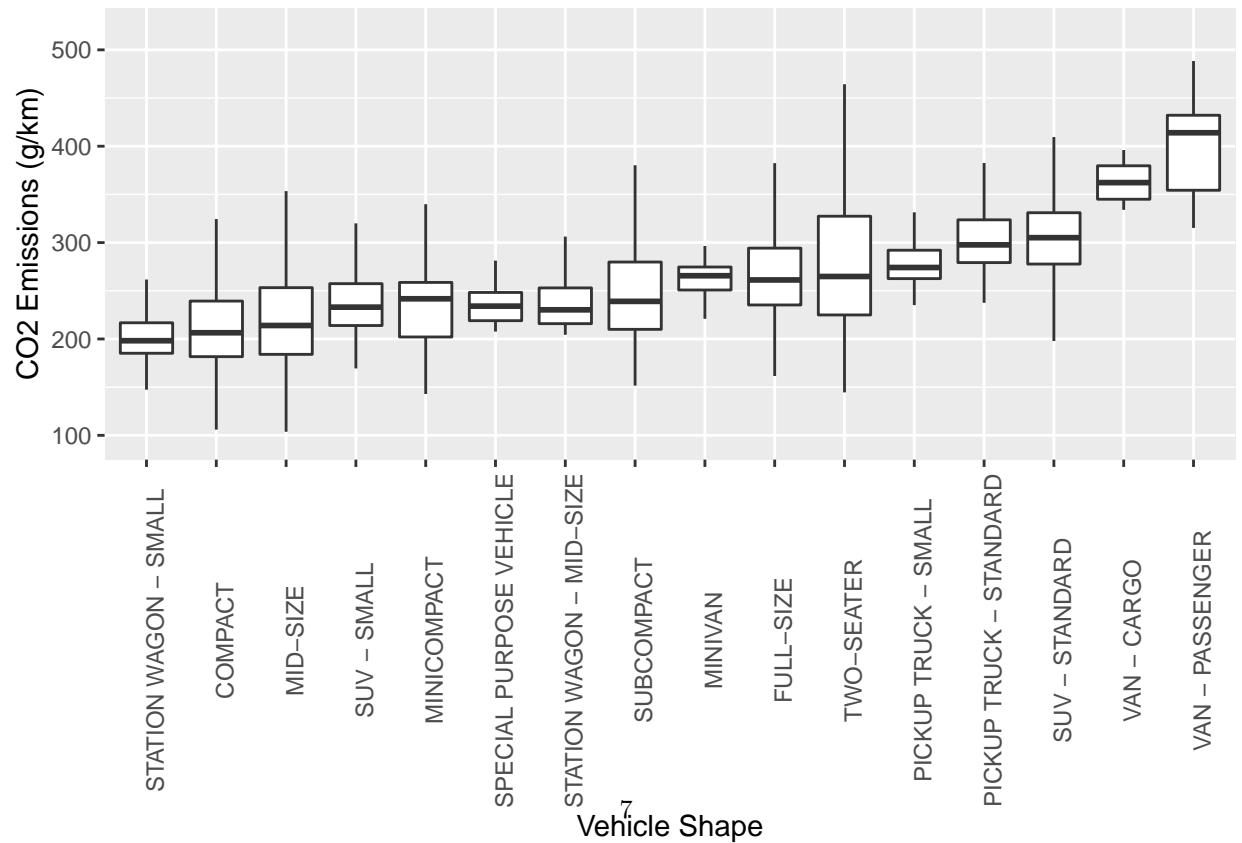
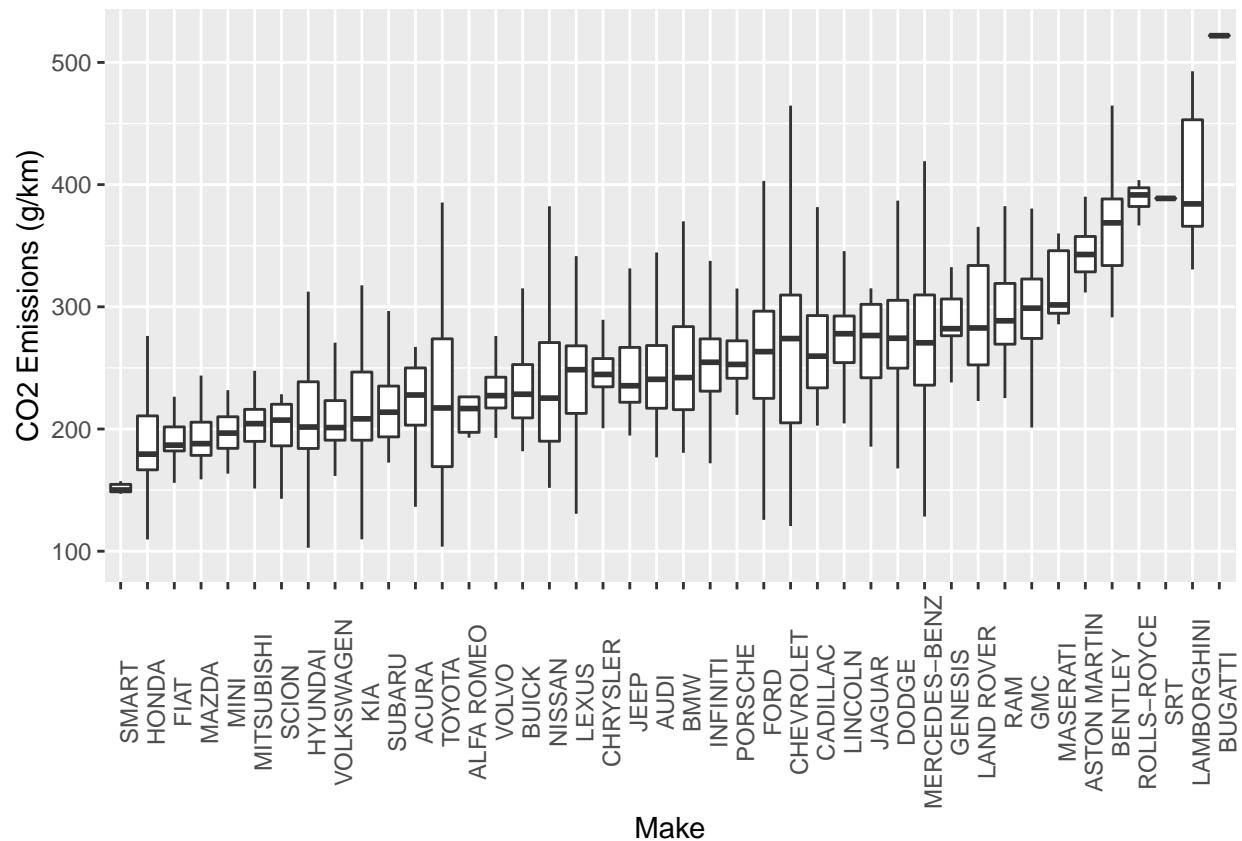
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

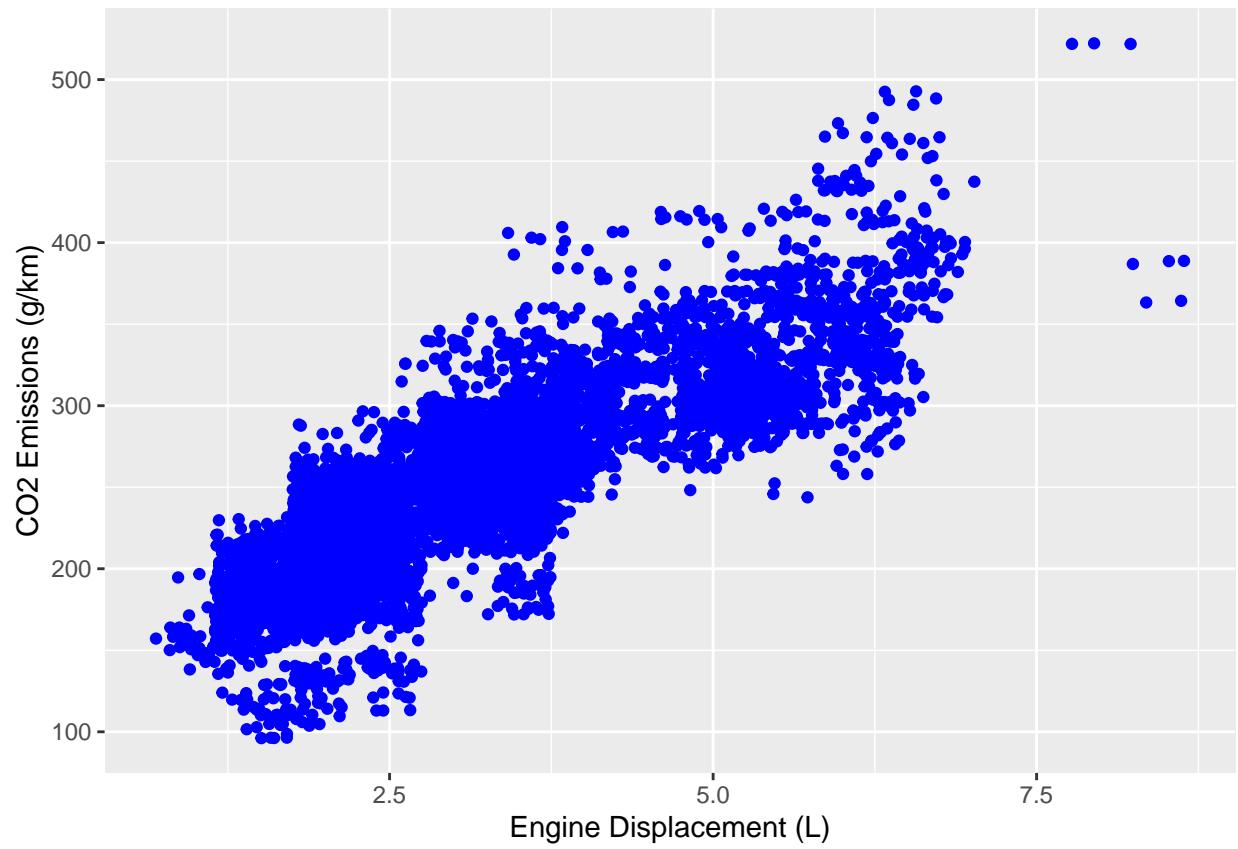


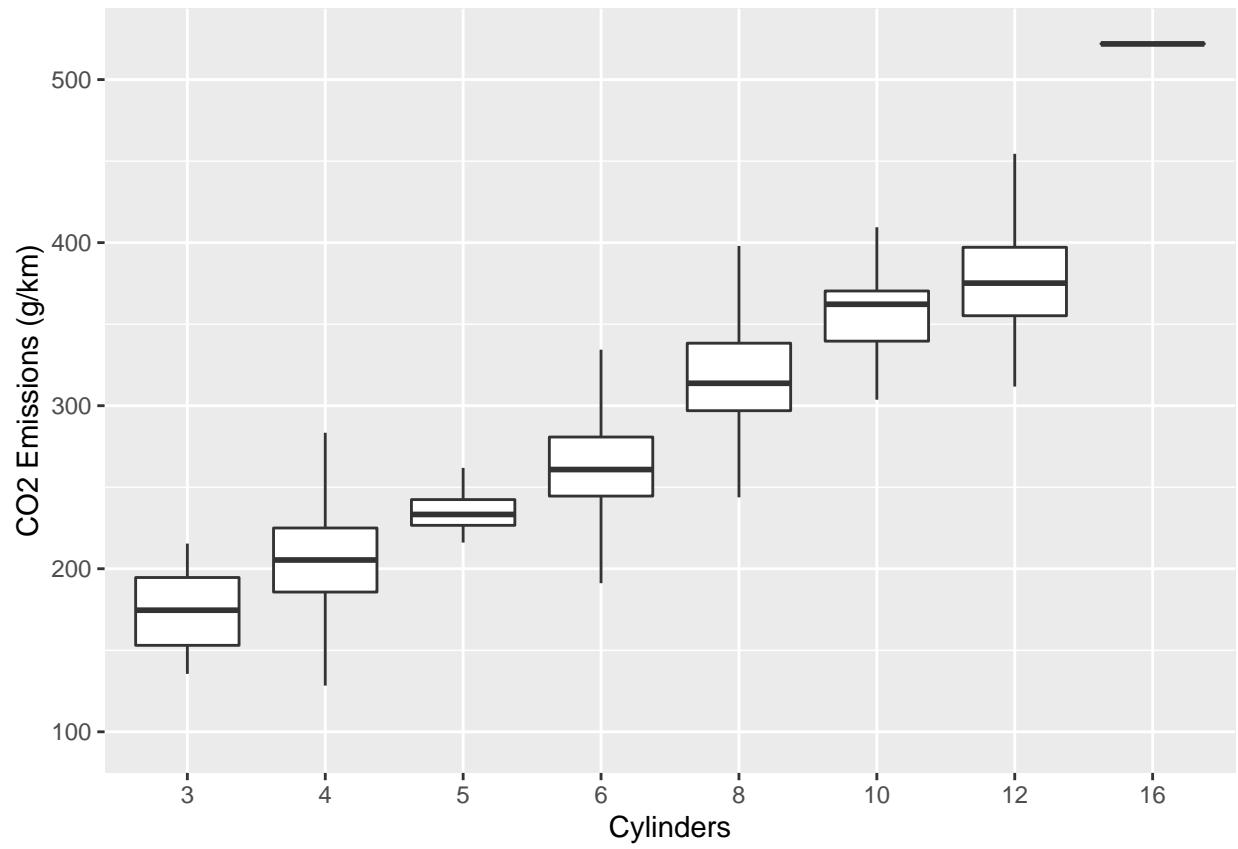


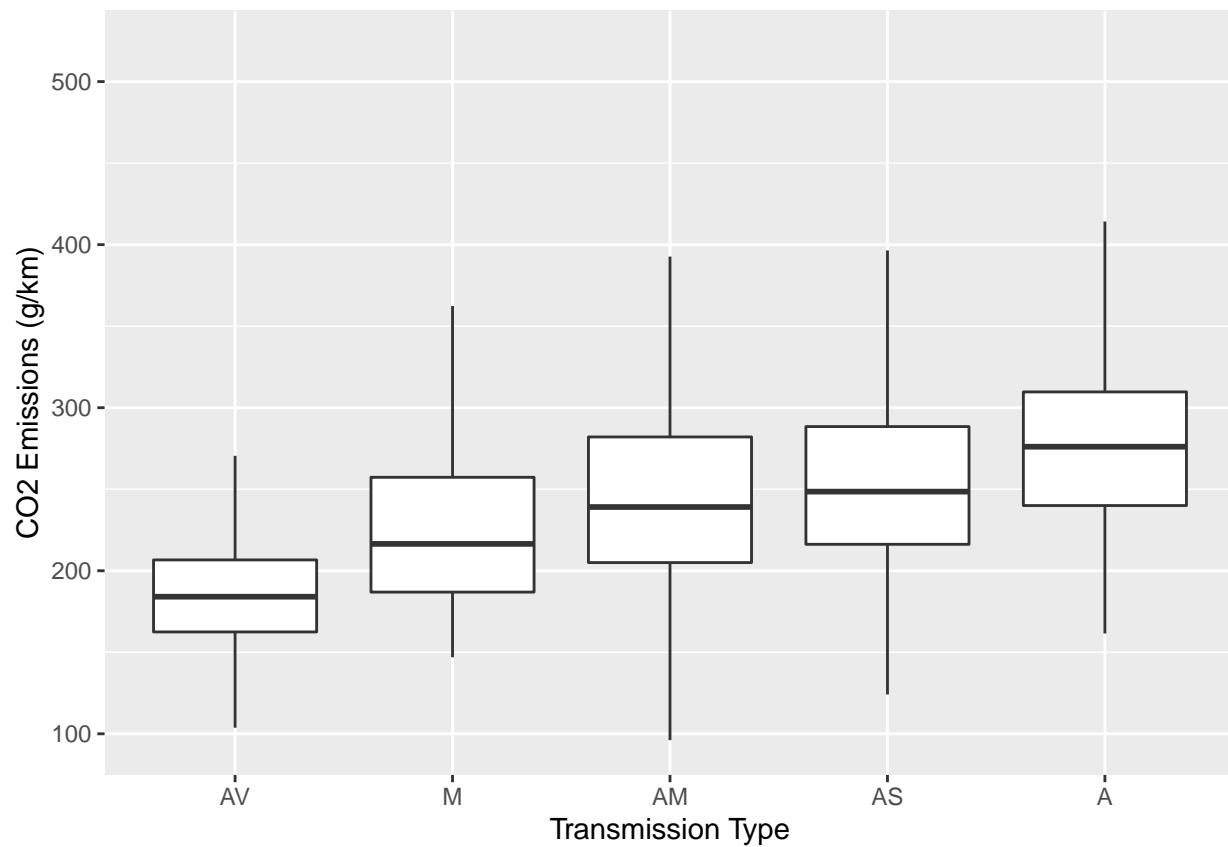


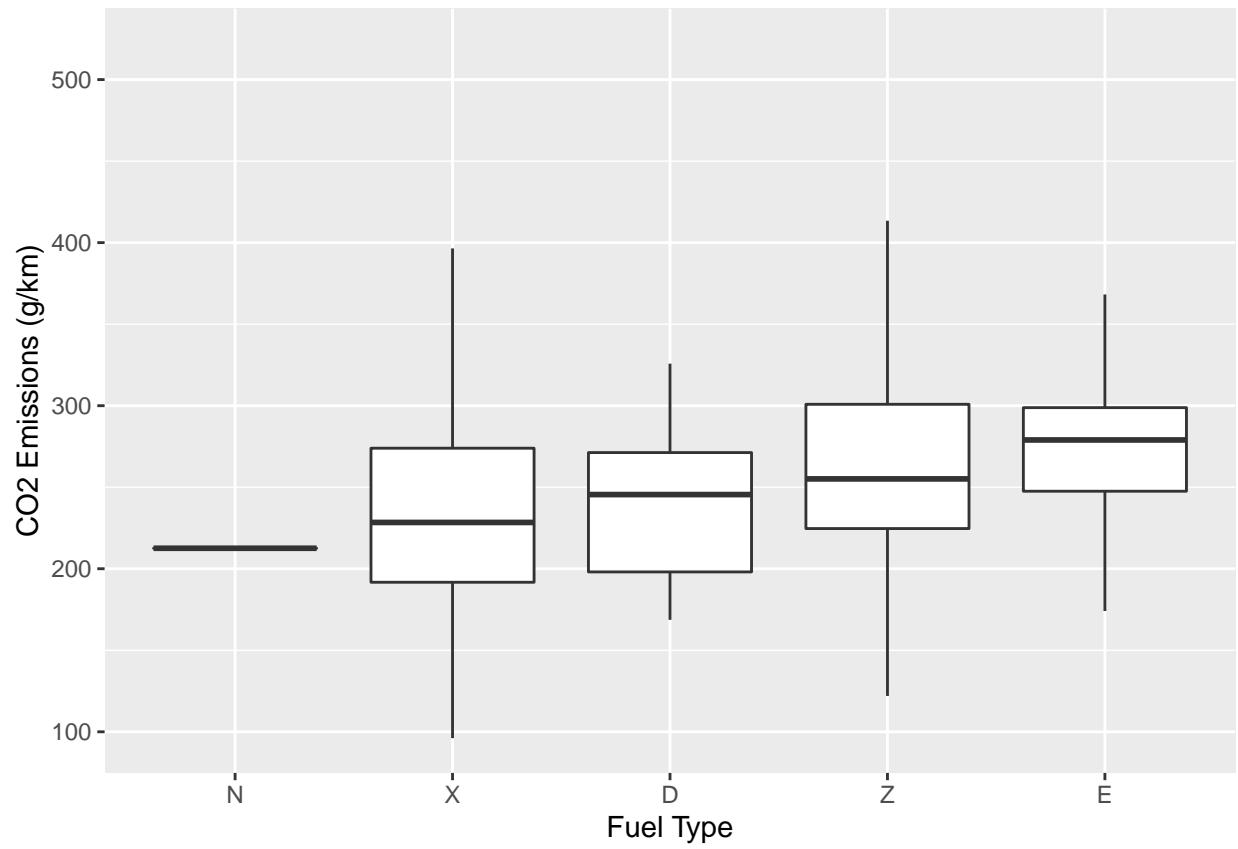
Scatter plots and box plots to find what the correlation is between each variable and CO2 emissions

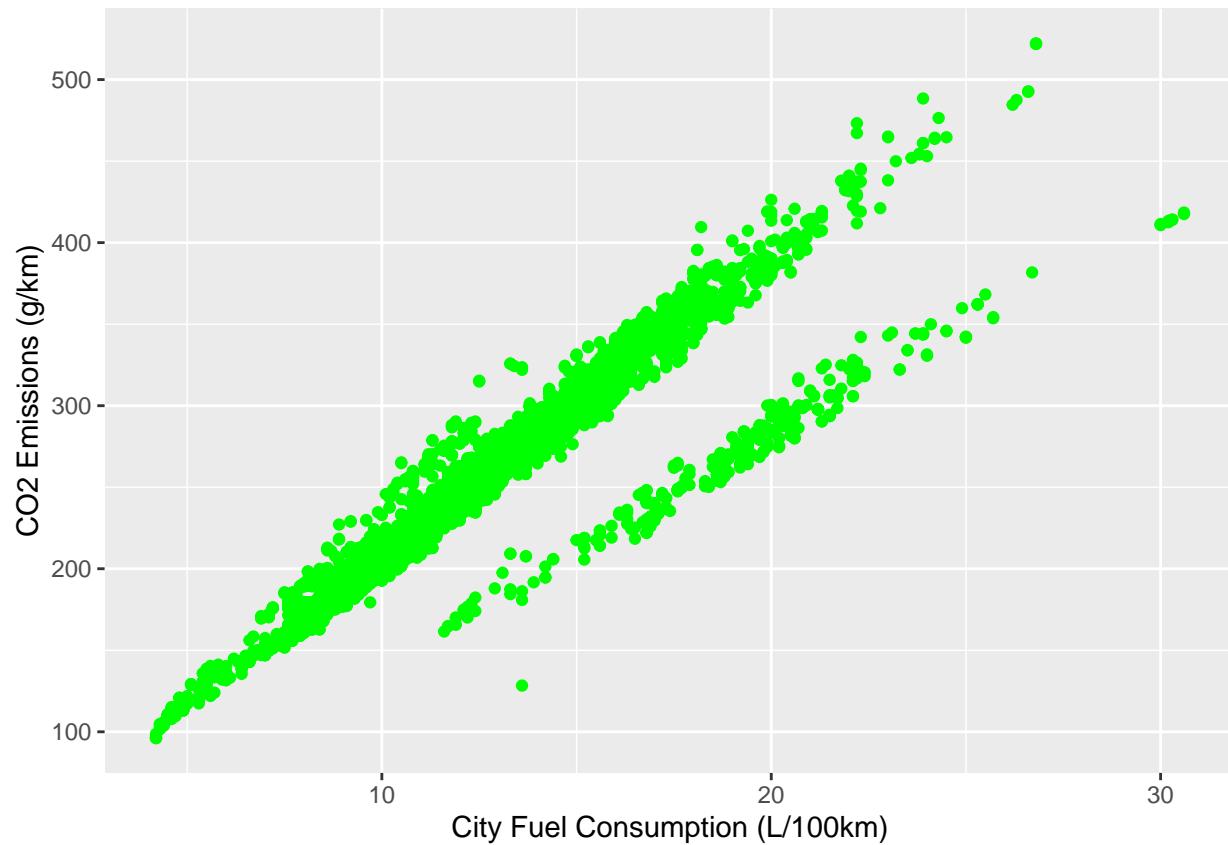


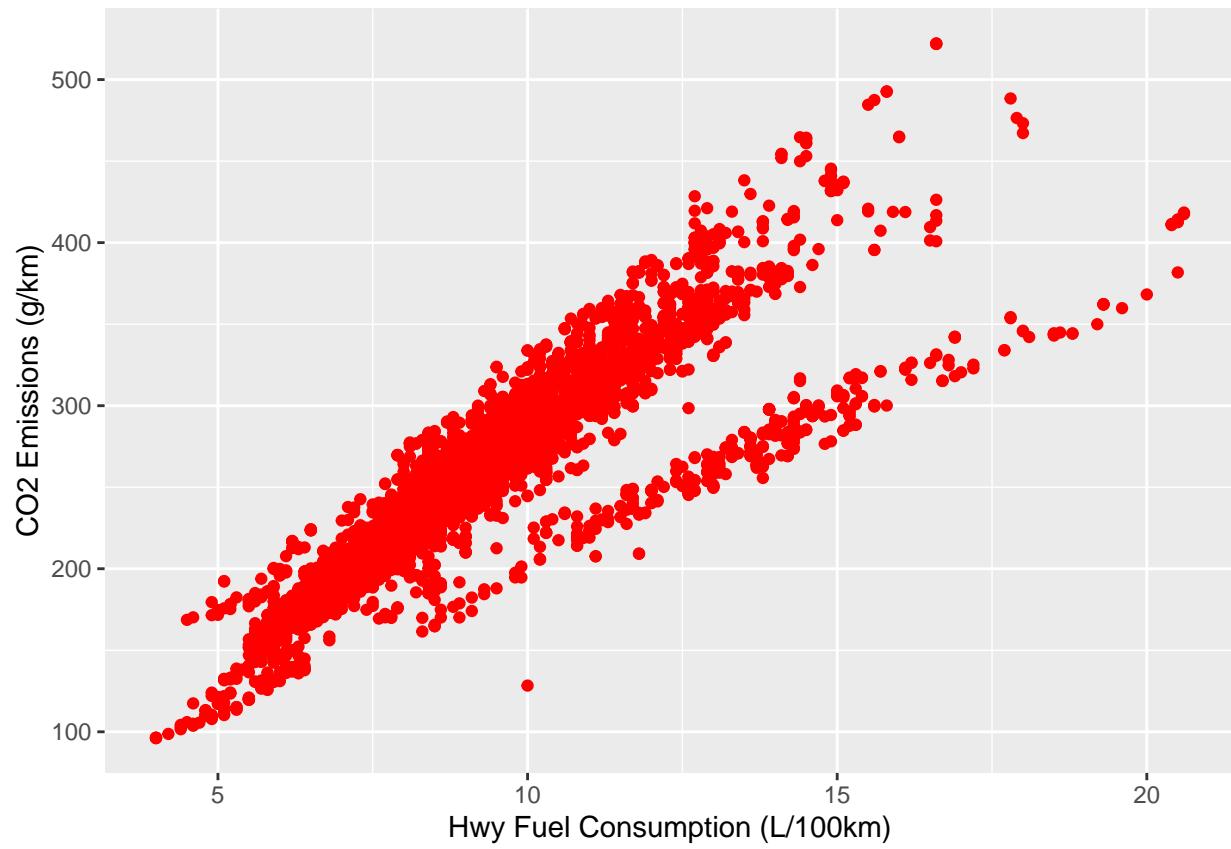


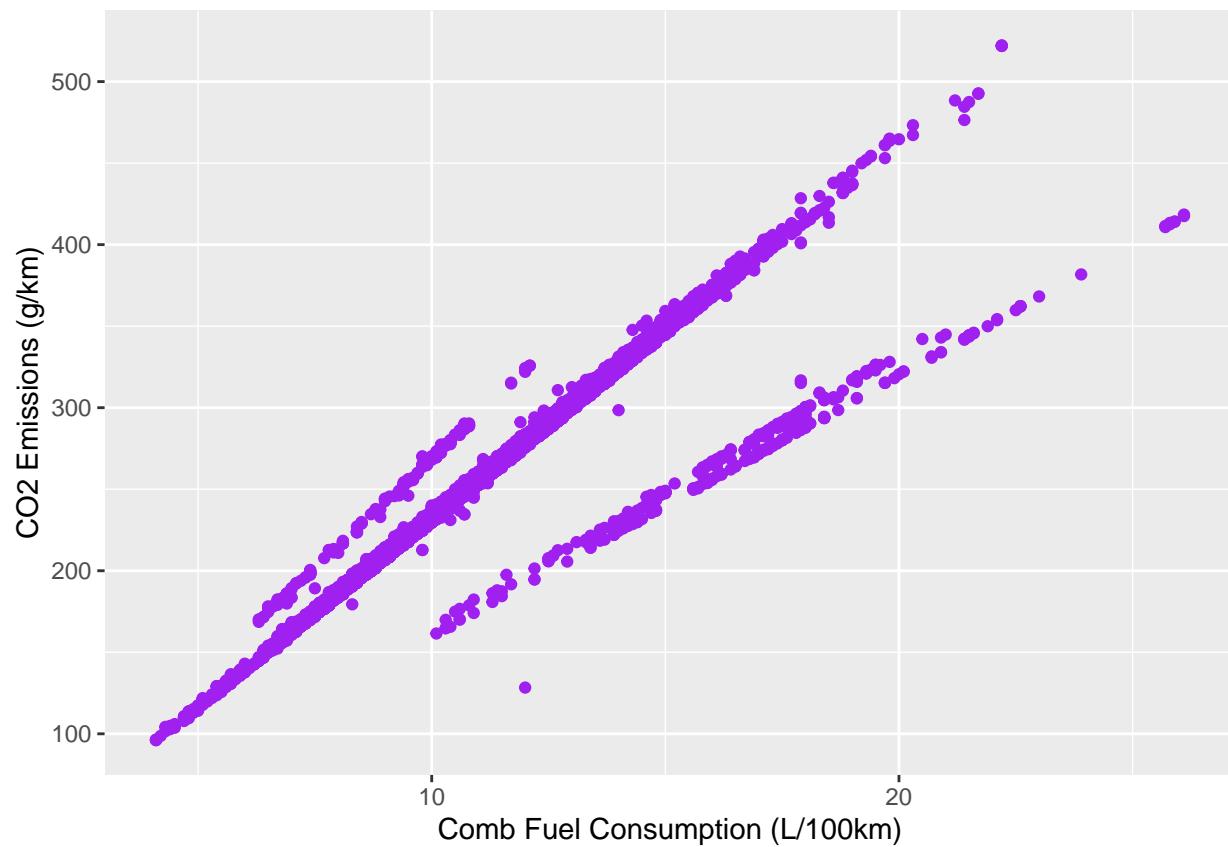






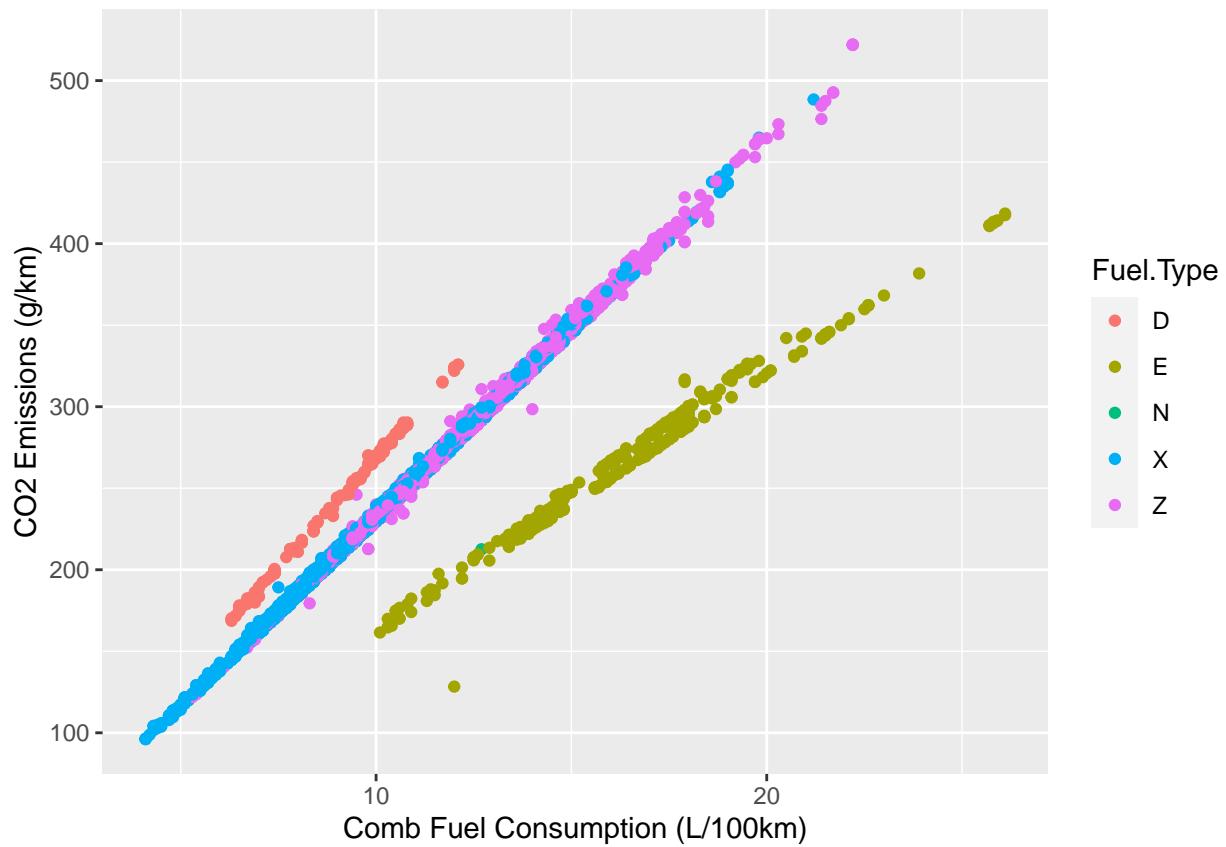






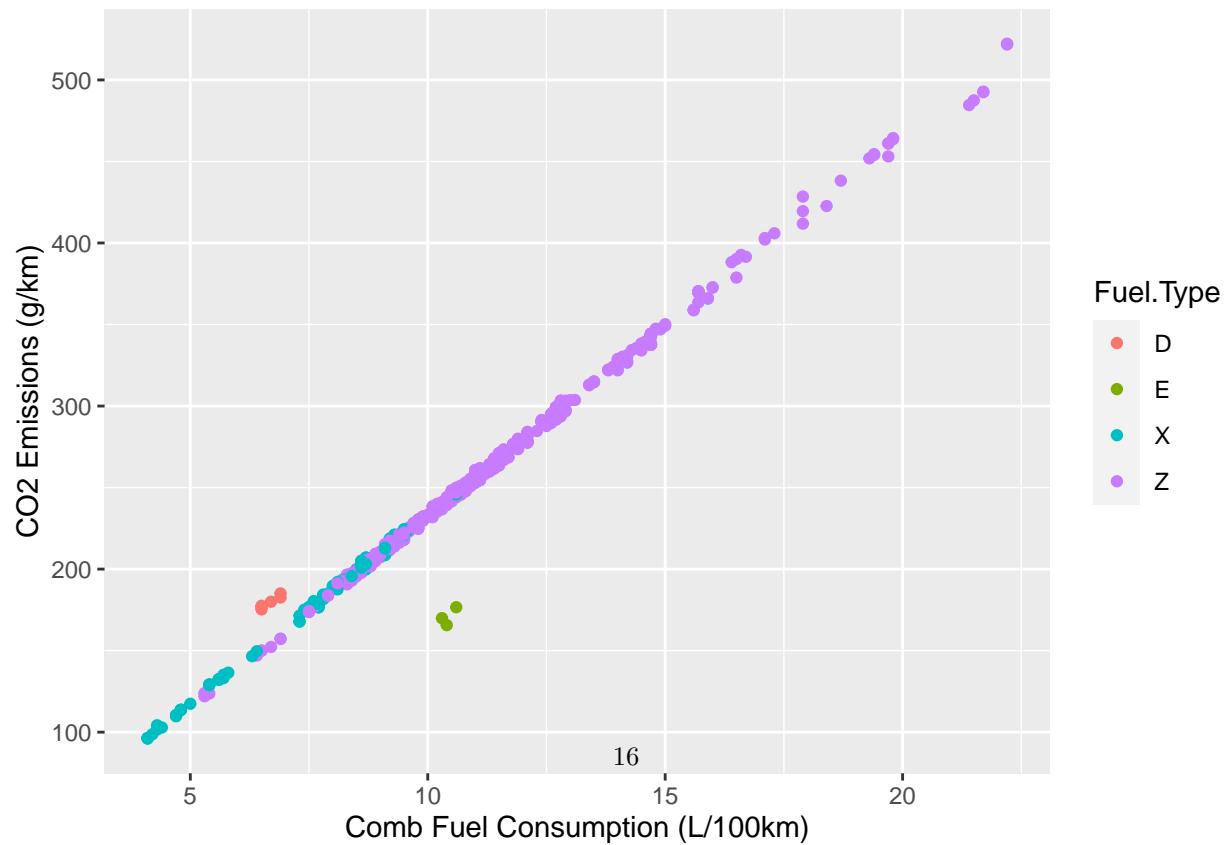
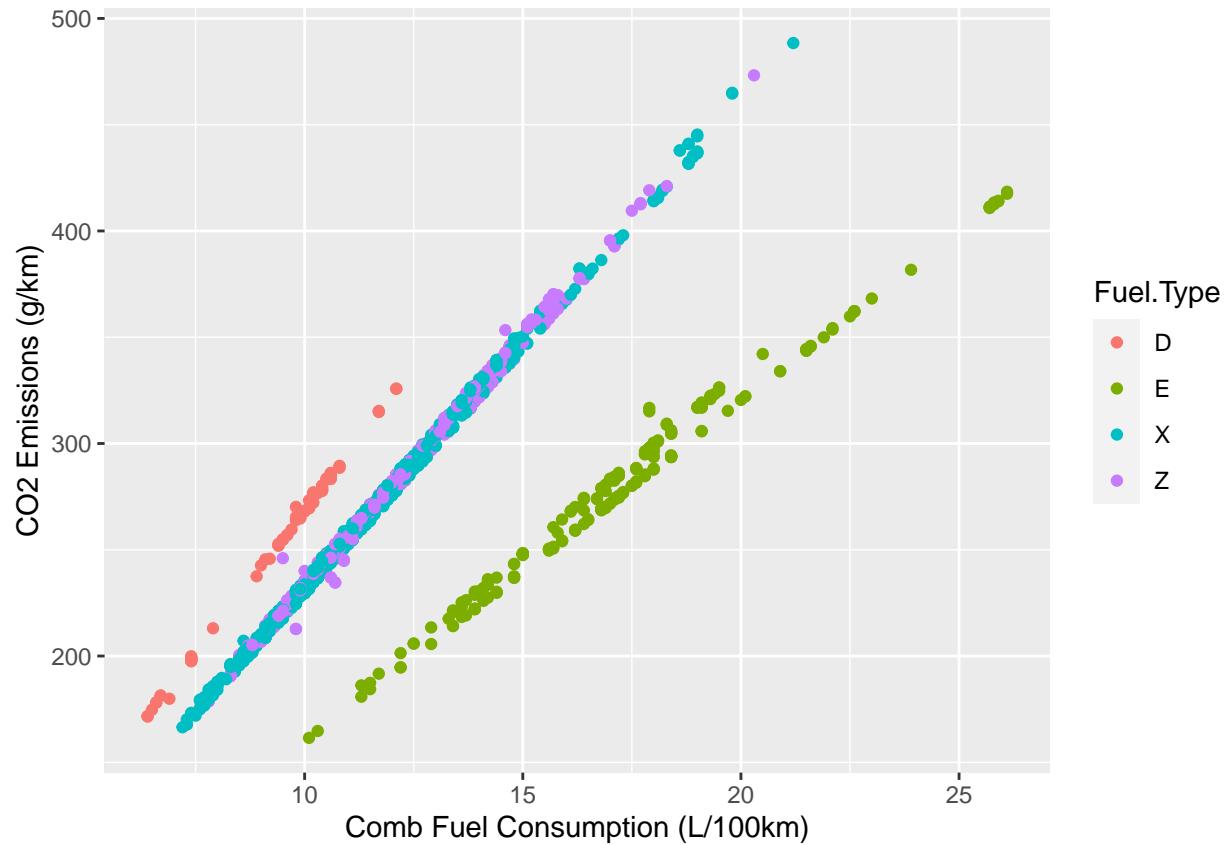
The emissions vs fuel consumption plot displayed multiple lines. Let's separate that data to see if we can see any interesting information.

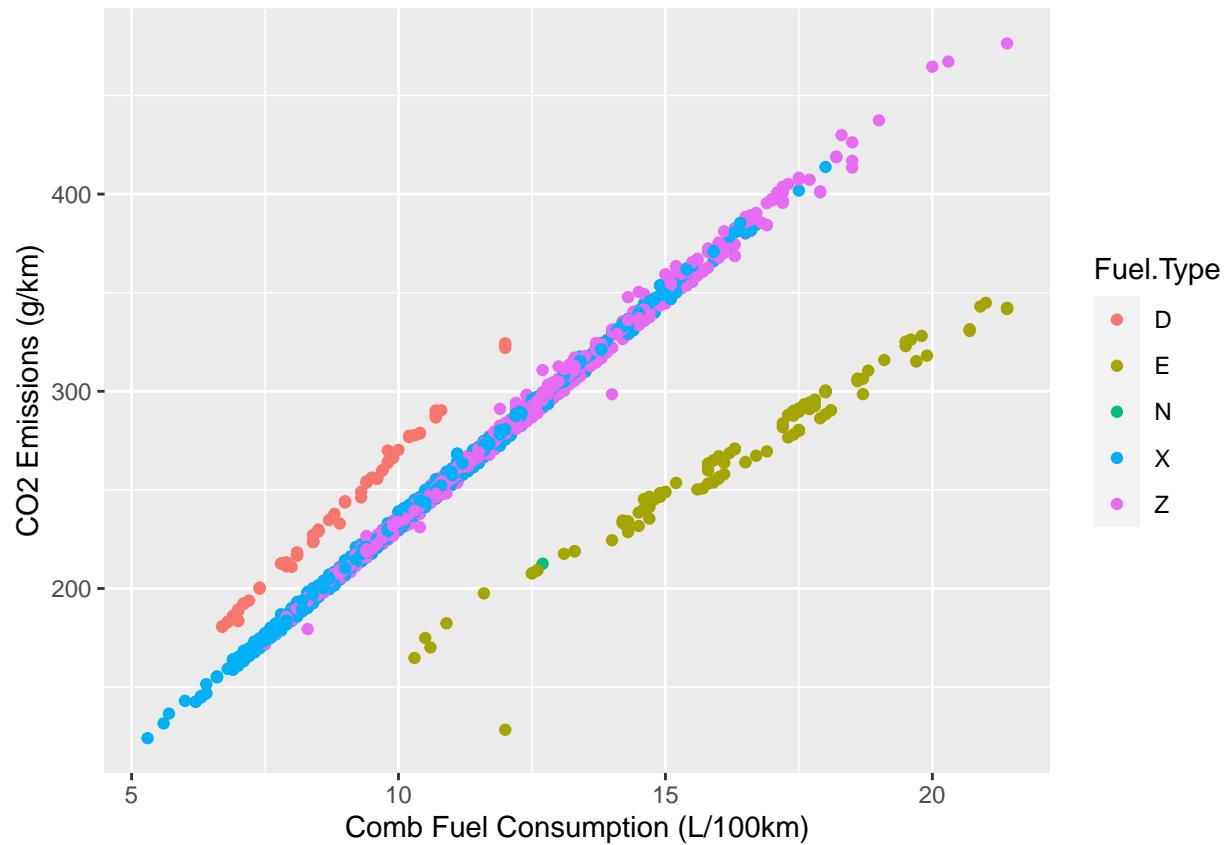
Starting with separating by fuel type:

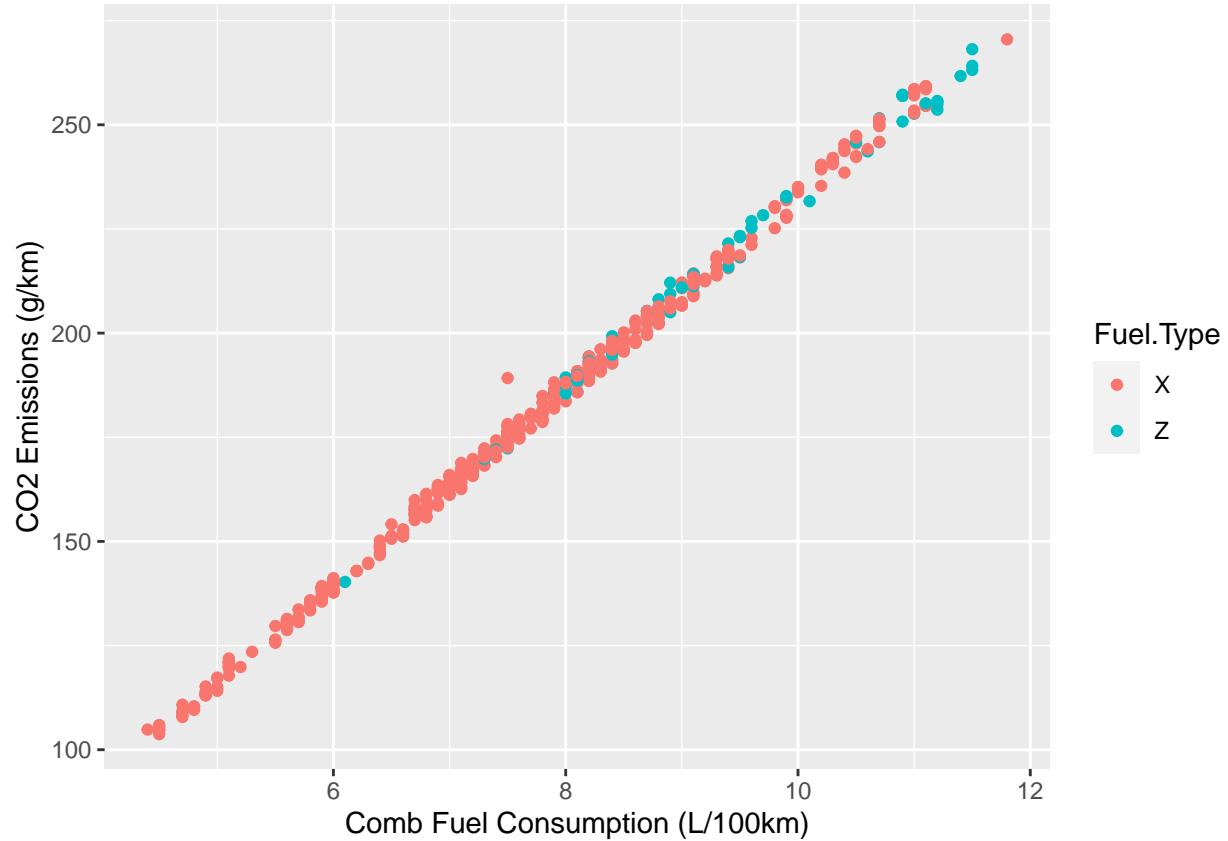


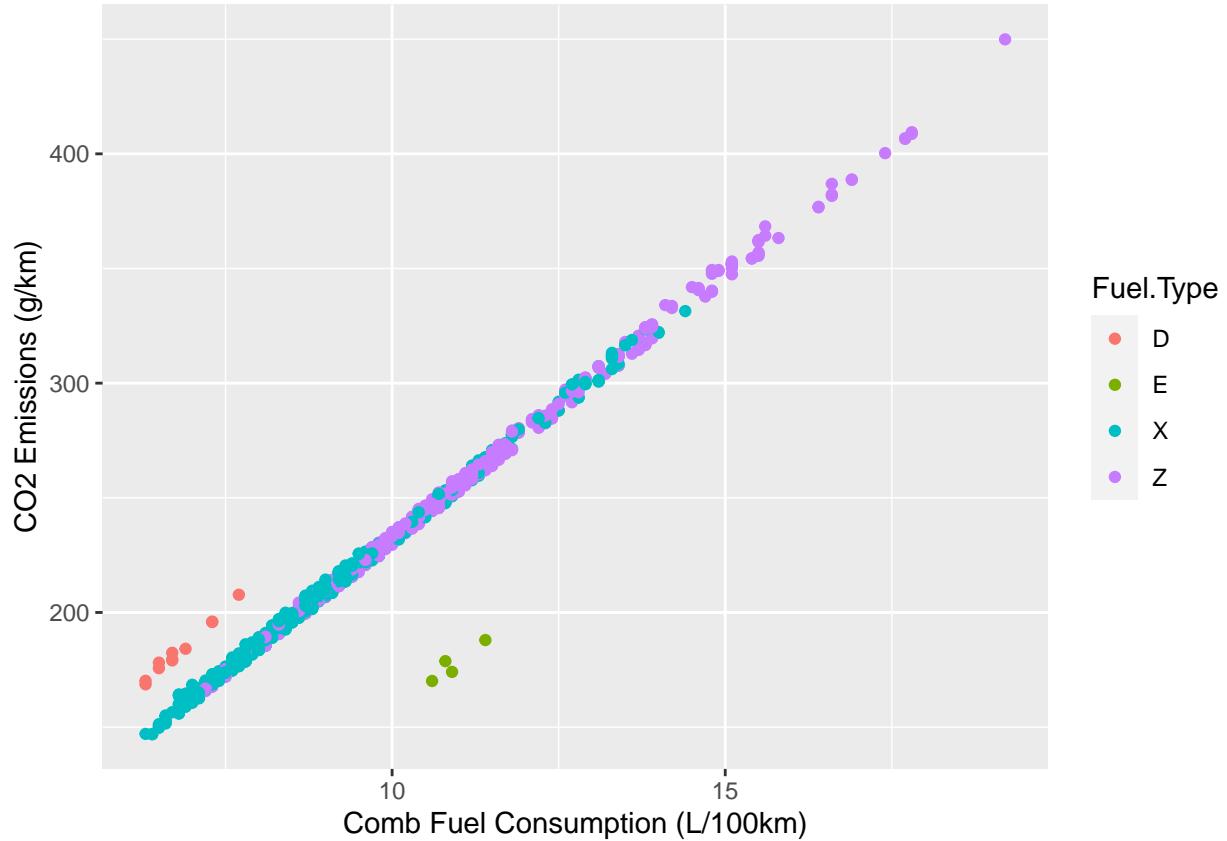
We can clearly see that each of the lines seen in the original plot is due to a different fuel type.

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









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 a significant difference between gas and ethanol CO2 emissions, but not gas and diesel 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 = -0.72904, df = 206.87, p-value = 0.4668  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -8.919348 4.103577  
## sample estimates:  
## mean of x mean of y  
## 235.1189 237.5268  
  
##  
## Welch Two Sample t-test  
##  
## data: cars_gas$CO2.Emissions.g.km. and cars_ethanol$CO2.Emissions.g.km.  
## t = -15.223, df = 487.83, p-value < 2.2e-16  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -45.15343 -34.82986  
## sample estimates:  
## mean of x mean of y  
## 235.1189 275.1105
```

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
## -52.619  -2.219    0.165   1.897  42.759
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)               3.456e+01  8.357e-01  41.350 < 2e-16
## MakeALFA ROMEO            2.295e+00  1.113e+00   2.061 0.039352
## MakeASTON MARTIN          3.002e+00  1.025e+00   2.930 0.003405
## MakeAUDI                  1.121e+00  6.828e-01   1.641 0.100825
## MakeBENTLEY                3.536e+00  1.014e+00   3.488 0.000489
## MakeBMW                   8.685e-01  6.484e-01   1.339 0.180497
## MakeBUGATTI                1.104e+01  3.057e+00   3.610 0.000308
## MakeBUICK                 2.614e+00  8.031e-01   3.255 0.001140
## MakeCADILLAC              2.031e+00  7.390e-01   2.748 0.006005
## MakeCHEVROLET             2.635e+00  6.786e-01   3.883 0.000104
## MakeCHRYSLER               7.430e+00  8.817e-01   8.427 < 2e-16
## MakeDODGE                  4.933e+00  7.433e-01   6.636 3.45e-11
## MakeFIAT                   2.158e+00  9.249e-01   2.333 0.019680
## MakeFORD                    4.577e+00  6.869e-01   6.663 2.87e-11
## MakeGENESIS                 5.552e+00  1.189e+00   4.668 3.10e-06
## MakeGMC                     2.442e+00  7.275e-01   3.357 0.000793
## MakeHONDA                  1.800e+00  7.396e-01   2.434 0.014957
## MakeHYUNDAI                 2.552e+00  7.261e-01   3.514 0.000444
## MakeINFINITI                1.625e+00  7.754e-01   2.096 0.036149
## MakeJAGUAR                  1.232e+00  7.464e-01   1.651 0.098830
## MakeJEEP                     4.370e+00  7.329e-01   5.962 2.60e-09
## MakeKIA                      2.884e+00  7.243e-01   3.982 6.90e-05
## MakeLAMBORGHINI             3.430e+00  1.097e+00   3.127 0.001770
## MakeLAND ROVER              -5.955e-01  8.307e-01  -0.717 0.473531
## MakeLEXUS                    7.668e-01  7.176e-01   1.069 0.285285
## MakeLINCOLN                 3.600e+00  8.382e-01   4.295 1.77e-05
## MakeMASERATI                4.134e+00  9.179e-01   4.503 6.79e-06
## MakeMAZDA                    1.260e+00  7.401e-01   1.703 0.088614
## MakeMERCEDES-BENZ           1.716e+00  6.637e-01   2.585 0.009759
## MakeMINI                     3.611e-01  7.170e-01   0.504 0.614558
## MakeMITSUBISHI              1.374e+00  8.324e-01   1.651 0.098835
## MakeNISSAN                  3.243e+00  7.194e-01   4.509 6.62e-06
## MakePORSCHE                 1.761e+00  6.953e-01   2.533 0.011341
## MakeRAM                      3.985e+00  8.809e-01   4.523 6.18e-06
## MakeROLLS-ROYCE              2.571e+00  1.007e+00   2.552 0.010734
## MakeSCION                    7.454e-01  1.255e+00   0.594 0.552478
## MakeSMART                    -3.907e+00  2.018e+00  -1.936 0.052923
## MakeSRT                      1.288e+00  3.631e+00   0.355 0.722901
## MakeSUBARU                  2.655e+00  7.760e-01   3.421 0.000627
## MakeTOYOTA                   1.679e+00  7.025e-01   2.390 0.016866
## MakeVOLKSWAGEN               1.140e+00  7.212e-01   1.581 0.113883
## MakeVOLVO                    9.191e-01  7.723e-01   1.190 0.234039
## Vehicle.ClassFULL-SIZE      7.246e-01  2.879e-01   2.517 0.011866
## Vehicle.ClassMID-SIZE      -2.459e-01  2.318e-01  -1.061 0.288922
## Vehicle.ClassMINICOMPACT   -5.884e-01  4.127e-01  -1.426 0.154010

```

## Vehicle.ClassMINIVAN	1.003e+00	6.195e-01	1.620	0.105356
## Vehicle.ClassPICKUP TRUCK - SMALL	4.891e+00	5.208e-01	9.391	< 2e-16
## Vehicle.ClassPICKUP TRUCK - STANDARD	2.277e+00	4.194e-01	5.428	5.87e-08
## Vehicle.ClassSPECIAL PURPOSE VEHICLE	3.665e+00	6.381e-01	5.744	9.63e-09
## Vehicle.ClassSTATION WAGON - MID-SIZE	7.765e-01	7.241e-01	1.072	0.283555
## Vehicle.ClassSTATION WAGON - SMALL	6.078e-02	3.689e-01	0.165	0.869124
## Vehicle.ClassSUBCOMPACT	-5.273e-03	2.738e-01	-0.019	0.984634
## Vehicle.ClassSUV - SMALL	1.497e+00	2.590e-01	5.779	7.82e-09
## Vehicle.ClassSUV - STANDARD	1.808e+00	3.325e-01	5.437	5.60e-08
## Vehicle.ClassTWO-SEATER	2.448e-01	3.362e-01	0.728	0.466494
## Vehicle.ClassVAN - CARGO	-1.435e+01	1.200e+00	-11.957	< 2e-16
## Vehicle.ClassVAN - PASSENGER	-8.410e+00	8.060e-01	-10.435	< 2e-16
## TransmissionAM	7.636e-01	3.144e-01	2.429	0.015169
## TransmissionAS	7.103e-01	2.158e-01	3.292	0.000999
## TransmissionAV	-1.214e+00	3.506e-01	-3.464	0.000535
## TransmissionM	-9.402e-03	2.433e-01	-0.039	0.969174
## Fuel.TypeE	-1.416e+02	6.120e-01	-231.386	< 2e-16
## Fuel.TypeN	-1.099e+02	5.044e+00	-21.786	< 2e-16
## Fuel.TypeX	-2.986e+01	4.237e-01	-70.477	< 2e-16
## Fuel.TypeZ	-2.804e+01	4.479e-01	-62.617	< 2e-16
## Cylinders	7.512e-01	1.157e-01	6.491	9.10e-11
## Engine.Size.L.	-1.936e-02	1.560e-01	-0.124	0.901202
## Fuel.Consumption.City..L.100.km.	6.522e+00	7.704e-01	8.466	< 2e-16
## Fuel.Consumption.Hwy..L.100.km.	4.861e+00	6.438e-01	7.550	4.87e-14
## Fuel.Consumption.Comb..L.100.km.	1.063e+01	1.395e+00	7.621	2.83e-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: 5.013 on 7315 degrees of freedom 
## Multiple R-squared:  0.9927, Adjusted R-squared:  0.9927 
## F-statistic: 1.447e+04 on 69 and 7315 DF,  p-value: < 2.2e-16 

## [1] 0.9927287

```

Calculating the RMSE of the model

```
## [1] 1.001918
```

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] 1.035096e+00 1.031452e+00 1.032325e+00 1.033702e+00 1.031836e+00
## [6] 1.033352e+00 1.031408e+00 1.033513e+00 1.033019e+00 1.031048e+00
## [11] 1.043849e+00 8.117494e-01 5.930457e-01 9.610861e-01 9.613728e-01
## [16] 9.442609e-01 9.379363e-01 9.458746e-01 9.377005e-01 9.372230e-01
## [21] 8.409622e-01 8.369198e-01 7.255915e-01 3.911598e-01 9.670087e-01
## [26] 3.645900e-01 9.688640e-01 9.209940e-01 9.154407e-01 9.151582e-01
## [31] 9.138126e-01 9.660789e-01 8.143566e-01 8.667242e-01 8.064292e-01
## [36] 6.408137e-01 5.307389e-01 6.134527e-01 7.974433e-01 7.829902e-01
## [41] 9.471791e-01 9.404145e-01 7.743168e-01 7.487716e-01 5.990261e-01
## [46] 6.104473e-01 6.169914e-01 5.902311e-01 9.544124e-01 9.654267e-01
## [51] 9.443422e-01 9.425778e-01 9.400516e-01 9.344959e-01 9.331062e-01
## [56] 8.450210e-01 8.317481e-01 7.242245e-01 3.840579e-01 9.653642e-01
## [61] 3.797815e-01 9.237441e-01 9.204229e-01 8.021455e-01 5.729225e-01
## [66] 9.182826e-01 9.261746e-01 1.032665e+00 1.031602e+00 1.033244e+00
## [71] 1.036165e+00 1.035721e+00 7.014408e-01 7.655894e-01 7.026264e-01
## [76] 6.271648e-01 9.359264e-01 9.195588e-01 1.038791e+00 1.057914e+00
## [81] 1.060188e+00 1.063633e+00 1.060307e+00 1.060272e+00 1.062845e+00
## [86] 1.061718e+00 1.062103e+00 1.061346e+00 1.177756e+00 1.177942e+00
## [91] 2.019236e+00 2.019236e+00 1.033026e+00 1.032832e+00 1.034151e+00
## [96] 1.032229e+00 1.032647e+00 1.033032e+00 1.034438e+00 1.033406e+00
## [101] 1.032483e+00 1.031214e+00 1.031150e+00 1.040059e+00 9.057034e-01
## [106] 9.447820e-01 3.637223e-01 9.682950e-01 3.593265e-01 9.611164e-01
## [111] 8.775908e-01 7.633567e-01 9.647540e-01 9.646706e-01 7.271865e-01
## [116] 9.321931e-01 8.482435e-01 6.378131e-01 6.378254e-01 6.760639e-01
## [121] 9.553438e-01 3.633451e-01 3.484085e-01 9.618008e-01 9.028604e-01
## [126] 8.786601e-01 8.994536e-01 9.388897e-01 8.397638e-01 1.031188e+00
## [131] 1.031658e+00 6.927931e-01 6.988320e-01 7.793416e-01 9.315388e-01
## [136] 9.676470e-01 1.038750e+00 1.056463e+00 1.062407e+00 1.063340e+00
## [141] 1.059751e+00 1.060523e+00 1.062809e+00 1.062528e+00 1.178186e+00
## [146] 1.177993e+00 1.031091e+00 1.046115e+00 1.031791e+00 1.036817e+00
## [151] 1.035017e+00 1.035427e+00 1.035289e+00 1.031363e+00 1.034883e+00
## [156] 1.033408e+00 1.034101e+00 1.032615e+00 1.034266e+00 1.031647e+00
## [161] 9.625809e-01 8.035539e-01 8.377999e-01 8.749485e-01 8.590455e-01
## [166] 9.208125e-01 5.349301e+13 6.734287e-01 8.870402e-01 9.413122e-01
```

```

## [171] 9.286417e-01 1.031337e+00 1.032448e+00 8.939275e-01 9.512765e-01
## [176] 6.769557e-01 1.031805e+00 9.495992e-01 9.574862e-01 5.298709e-01
## [181] 5.267340e-01 5.497306e-01 9.503206e-01 9.477901e-01 9.574835e-01
## [186] 9.289529e-01 9.414268e-01 9.676258e-01 9.680065e-01 8.727536e-01
## [191] 8.513983e-01 8.137881e-01 7.883962e-01 8.495184e-01 7.346020e-01
## [196] 1.033617e+00 1.036162e+00 1.034898e+00 1.032881e+00 1.033723e+00
## [201] 1.032006e+00 9.221785e-01 9.610379e-01 5.028965e-01 1.031460e+00
## [206] 9.483804e-01 7.840261e-01 1.031838e+00 1.034173e+00 1.062038e+00
## [211] 1.059790e+00 1.059951e+00 1.059732e+00 1.055907e+00 1.058284e+00
## [216] 1.177924e+00 1.177899e+00 1.033658e+00 1.046829e+00 1.045514e+00
## [221] 1.045320e+00 1.044755e+00 1.039902e+00 1.035940e+00 1.036627e+00
## [226] 1.034113e+00 1.033869e+00 1.033546e+00 1.034073e+00 1.033020e+00
## [231] 1.035011e+00 1.032749e+00 1.033935e+00 1.034062e+00 8.208496e-01
## [236] 8.766718e-01 8.650245e-01 8.713489e-01 9.238728e-01 9.207576e-01
## [241] 7.187374e-01 9.060985e-01 9.528267e-01 1.031534e+00 9.600250e-01
## [246] 8.285444e-01 9.540559e-01 9.660593e-01 9.455684e-01 5.444553e-01
## [251] 5.470131e-01 6.183575e-01 9.391308e-01 9.373496e-01 9.609939e-01
## [256] 9.399575e-01 9.319795e-01 9.652352e-01 1.053430e+00 1.052116e+00
## [261] 1.050245e+00 1.052471e+00 8.602441e-01 8.805599e-01 8.996875e-01
## [266] 8.186088e-01 7.587140e-01 1.031830e+00 1.035842e+00 1.034907e+00
## [271] 1.033951e+00 1.031295e+00 9.159403e-01 9.569163e-01 9.213662e-01
## [276] 8.102806e-01 6.110978e-01 1.031588e+00 1.032632e+00 1.032649e+00
## [281] 1.032097e+00 1.177356e+00 1.034081e+00 1.036829e+00 1.036704e+00
## [286] 1.036639e+00 1.047036e+00 1.046801e+00 1.045241e+00 1.044208e+00
## [291] 1.041239e+00 1.046030e+00 1.043535e+00 1.035543e+00 1.036031e+00
## [296] 1.035281e+00 1.033779e+00 1.033276e+00 1.033678e+00 1.033405e+00
## [301] 1.032786e+00 1.032751e+00 1.514541e+00 9.382200e-01 9.187252e-01
## [306] 9.658909e-01 8.723490e-01 9.404444e-01 1.032023e+00 9.519711e-01
## [311] 9.098656e-01 9.571912e-01 8.341839e-01 9.464250e-01 9.515126e-01
## [316] 9.641064e-01 9.485865e-01 9.553221e-01 9.379716e-01 9.547735e-01
## [321] 1.052034e+00 1.052369e+00 1.052438e+00 1.049753e+00 1.052434e+00
## [326] 8.426075e-01 8.903159e-01 7.576686e-01 1.035482e+00 1.036013e+00
## [331] 1.035882e+00 1.036146e+00 1.034341e+00 1.034395e+00 1.032072e+00
## [336] 1.031252e+00 1.032922e+00 1.033392e+00 1.034418e+00 1.036486e+00
## [341] 1.036117e+00 1.036670e+00 1.046096e+00 1.046099e+00 1.045661e+00
## [346] 1.044597e+00 1.041228e+00 1.045560e+00 1.045935e+00 1.043795e+00
## [351] 1.035513e+00 1.036390e+00 1.036329e+00 1.036011e+00 1.033515e+00
## [356] 1.034782e+00 1.514509e+00 9.512055e-01 8.718245e-01 9.457826e-01
## [361] 1.032012e+00 9.168056e-01 9.646957e-01 7.956025e-01 9.520003e-01
## [366] 9.522268e-01 9.492158e-01 9.375788e-01 9.590529e-01 9.468031e-01
## [371] 7.857441e-01 7.812311e-01 1.055412e+00 1.053897e+00 1.053785e+00
## [376] 1.052067e+00 1.052719e+00 1.052556e+00 1.050413e+00 1.052191e+00
## [381] 1.031096e+00 1.034318e+00 1.037594e+00 1.037770e+00 9.300542e-01
## [386] 1.031683e+00 1.034430e+00 9.162446e-01 1.033174e+00 1.049506e+00
## [391] 1.036360e+00 1.036803e+00 1.046575e+00 1.045555e+00 1.045236e+00
## [396] 1.045853e+00 1.046088e+00 1.046149e+00 1.046061e+00 1.034580e+00
## [401] 1.036332e+00 1.035886e+00 1.034857e+00 1.031832e+00 1.035793e+00
## [406] 1.036029e+00 1.032325e+00 1.036126e+00 1.035969e+00 1.036476e+00
## [411] 1.514377e+00 9.509865e-01 8.703021e-01 9.316020e-01 9.550149e-01
## [416] 9.480555e-01 9.524038e-01 9.421385e-01 7.865832e-01 7.810676e-01
## [421] 1.055617e+00 1.052390e+00 1.053583e+00 1.052239e+00 1.052347e+00
## [426] 1.052558e+00 1.051635e+00 1.049814e+00 9.293827e-01 9.162919e-01
## [431] 1.032002e+00 1.049184e+00 1.049695e+00

```

To test the assumption of independence, 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      0.09690374    1.806167      0
## Alternative hypothesis: rho != 0
```

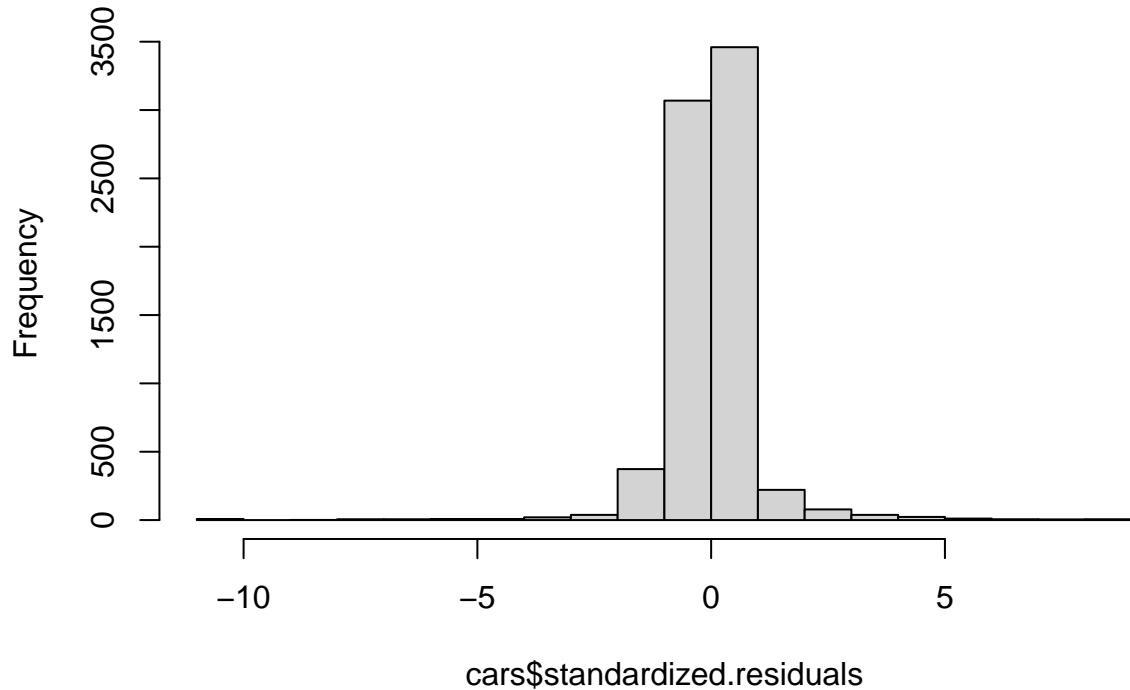
To determine if the model contains multicollinearity, I will construct a correlation matrix and calculate the determinant. If the determinant is greater than 0.00001, then multicollinearity is not a problem.

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.

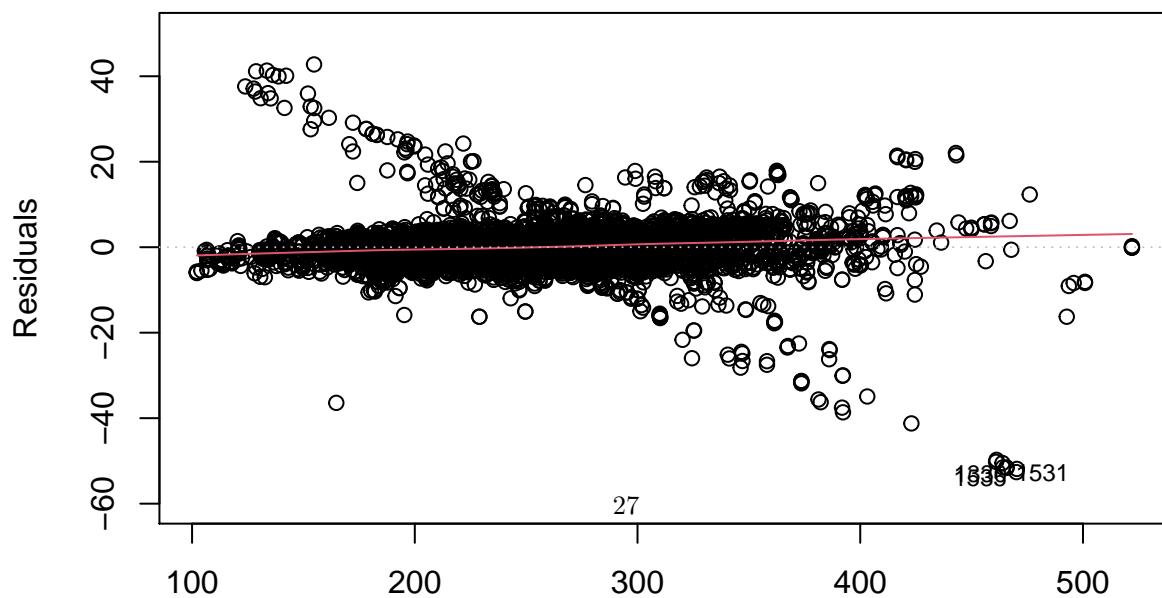
```
## [1] 4.102824e-09
```

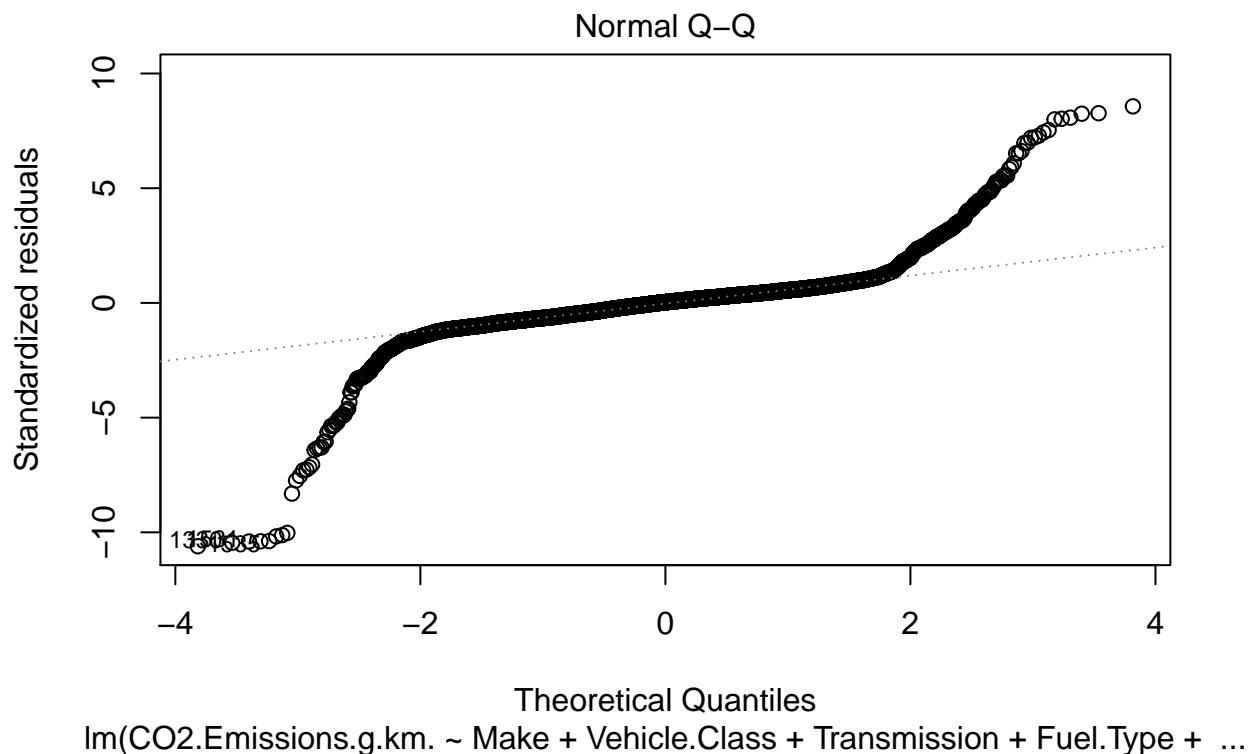
The histogram and Q-Q plot inform us of any deviations from normality seen in the residuals. Based on this, we can clearly see that the distribution of residuals is approximately normal. In addition, the plot of residuals vs fitted values appears to be not random. Residuals appear to remain fairly constant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

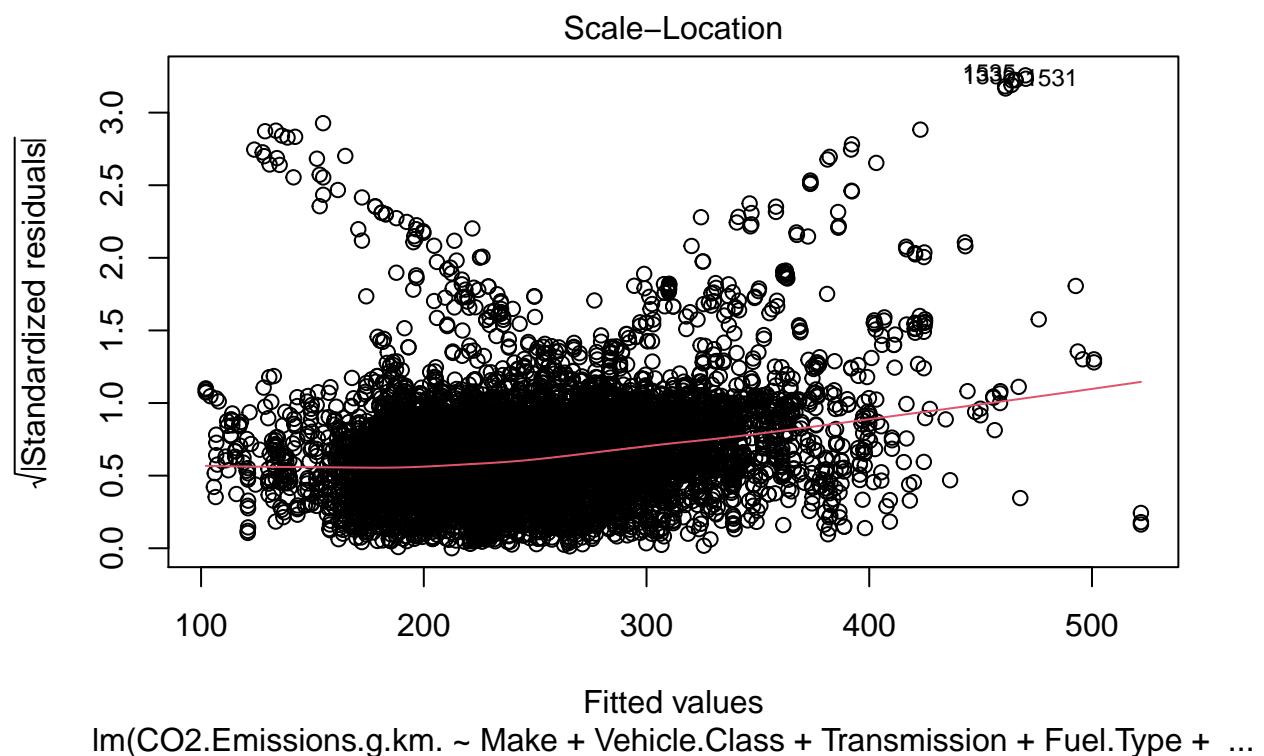
Histogram of cars\$standardized.residuals

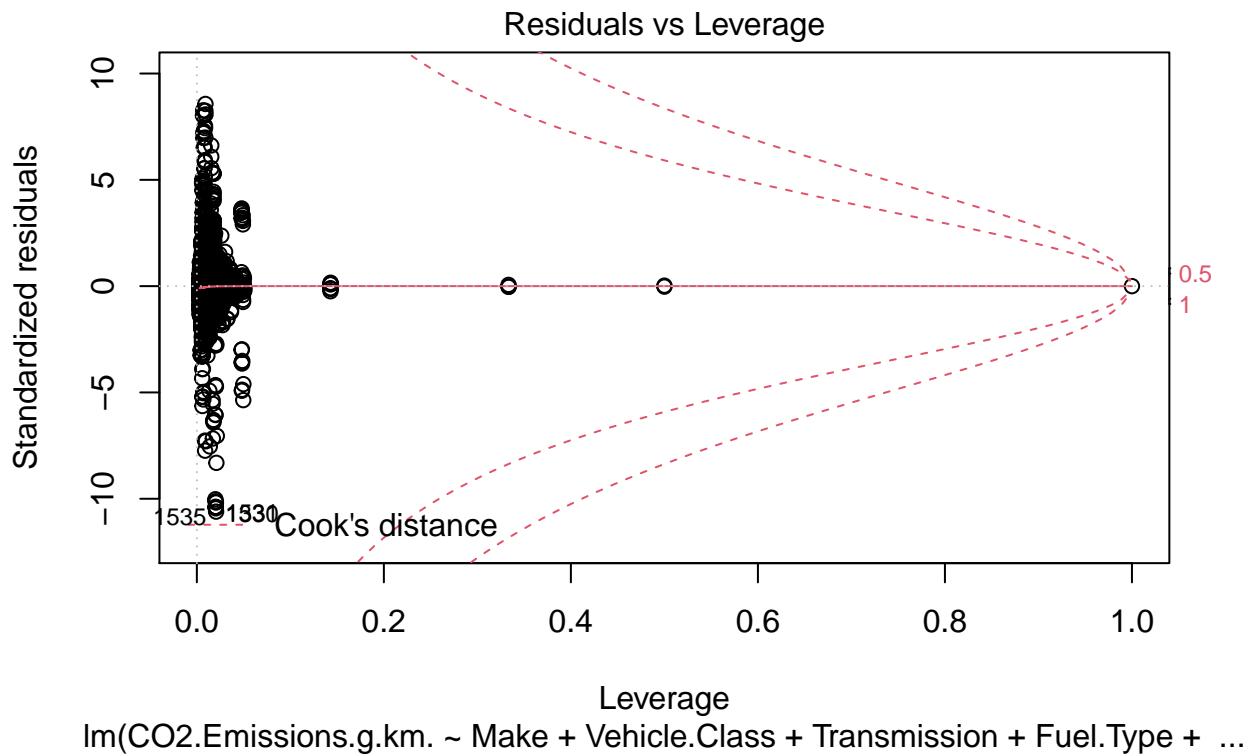


Residuals vs Fitted









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.