Has modern day car technology truly evolved?

Lakshay Badlani

11/27/2017

Introduction

Modern day car manufacturers strive to make their cars as fuel efficient as possible and thus do their best in order to ensure that the number of miles to the gallon their products achieve is maximized. In most cases, many car manufacturers make a vehicles MPG rating a key point in advertisements. There are many factors that influence a cars fuel efficiency rating however according to most engineers the three key influences that affect a cars MPG statistic is the displacement of its engine, the total weight of the car, and the number of gears its engine can cycle through[1]. Each one of these factors seem to influence the rating equally so there is no dominating factor. In addition to this it is claimed that in today's generation the miles to the gallon for both types of transmission, automatic and manual respectively, are roughly equal[2] and thus one can draw the conclusion that both forms of transmission are equally impacted by all three of the aforementioned key factors. Whilst this may be true for modern day cars, I would like to investigate if these statements held true for cars built in the late 1970's and later. The motivation to do this lies in the fact that I am interested if the same issues plagued past manufacturers as it was during the late twentieth century that the environmental impact of cars was beginning to be realised and the price of gas was slowly creeping up[3]. If any parallels can be drawn it would speak volumes about how technology has revolutionized.

In order to carry out this investigation I will be using the mtcars dataset that is included with R. It is a medium sized dataset containing information on a variety of car models made in 1974 including their fuel efficiency. The data was taken from Motor Trends US magazine and is considered to be reliable. I will split the dataset into two smaller subsets, one referring to cars that were of automatic transmission and the other referring to cars of manual transmission. This will allow me to isolate whether or not transmission had any impact on the cars fuel efficiency. In addition we will calculate simple correlations of our three key factors and also look at the most relevant multivariable correlations to determine which variables really impacted the MPG rating of the cars.

Data Check

Lets first take a quick glance at the variables and attributes our dataset includes by running a simple summary call:

```
## We don't have to preload the dataset since its inbuilt.
summary(mtcars)
```

```
##       mpg             cyl             disp             hp       
##  Min.   :10.40   Min.   :4.000   Min.   : 71.1   Min.   : 52.0  
##  1st Qu.:15.43   1st Qu.:4.000   1st Qu.:120.8   1st Qu.: 96.5  
##  Median :19.20   Median :6.000   Median :196.3   Median :123.0  
##  Mean   :20.09   Mean   :6.188   Mean   :230.7   Mean   :146.7  
##  3rd Qu.:22.80   3rd Qu.:8.000   3rd Qu.:326.0   3rd Qu.:180.0  
##  Max.   :33.90   Max.   :8.000   Max.   :472.0   Max.   :335.0  
##       drat             wt             qsec             vs        
##  Min.   :2.760   Min.   :1.513   Min.   :14.50   Min.   :0.0000  
##  1st Qu.:3.080   1st Qu.:2.581   1st Qu.:16.89   1st Qu.:0.0000  
##  Median :3.695   Median :3.325   Median :17.71   Median :0.0000  
##  Mean   :3.597   Mean   :3.217   Mean   :17.85   Mean   :0.4375  
##  3rd Qu.:3.920   3rd Qu.:3.610   3rd Qu.:18.90   3rd Qu.:1.0000  
##  Max.   :4.930   Max.   :5.424   Max.   :22.90   Max.   :1.0000  
##        am              gear            carb      
##  Min.   :0.0000   Min.   :3.000   Min.   :1.000  
##  1st Qu.:0.0000   1st Qu.:3.000   1st Qu.:2.000  
##  Median :0.0000   Median :4.000   Median :2.000  
##  Mean   :0.4062   Mean   :3.688   Mean   :2.812  
##  3rd Qu.:1.0000   3rd Qu.:4.000   3rd Qu.:4.000  
##  Max.   :1.0000   Max.   :5.000   Max.   :8.000  
```

Since the column names are clearly abbreviated and the units of each could be useful we can utilise data dictionary for this dataset[4]. It states that the columns included in the data set are as follows:

1. mpg = Miles per (US) gallon
2. cyl = Number of cylinders
3. disp = Displacement (cubic inches)
4. hp = Gross horsepower
5. drat = Rear axle ratio
6. wt = Weight (measured in 1000 lbs)
7. qsec = 1/4 mile time
8. vs = V shaped engine or straight (0 = V engine, 1 = straight)
9. am = Transmission (0 = automatic, 1 = manual)
10. gear = Number of forward gears
11. carb = Number of carburetors

As we can see from the list above, all our necessary attributes (namely transmission, mpg, weight, displacement and gears) are contained within the dataset and thus we can carry out our investigation.

One last check if there are any bad or missing records in the dataset:

```
# Complete.cases is a function that returns a list of true or false
# values depending on if there are NA/Null values in a row. The negation
# highlights all the missing cases by converting any incomplete rows to
# true values.
mtcars[!complete.cases(mtcars),]
```

```
## [1] mpg  cyl  disp hp   drat wt   qsec vs   am   gear carb
## <0 rows> (or 0-length row.names)
```

Since no output was returned by the preceding function call we know our dataset is complete.

Data Preparation

The first step in analyzing our dataset is splitting it into two subgroups: automatic and manual transmission. In order to do this observe the follow code:

```
# Here we check the condition that the am column equals 0 (ie automatic)
# and pick out the rows for which the condition holds true
automatic = mtcars[mtcars$am == 0,]
head(automatic)
```

```
##                    mpg cyl  disp  hp drat    wt  qsec vs am gear carb
## Hornet 4 Drive    21.4   6 258.0 110 3.08 3.215 19.44  1  0    3    1
## Hornet Sportabout 18.7   8 360.0 175 3.15 3.440 17.02  0  0    3    2
## Valiant           18.1   6 225.0 105 2.76 3.460 20.22  1  0    3    1
## Duster 360        14.3   8 360.0 245 3.21 3.570 15.84  0  0    3    4
## Merc 240D         24.4   4 146.7  62 3.69 3.190 20.00  1  0    4    2
## Merc 230          22.8   4 140.8  95 3.92 3.150 22.90  1  0    4    2
```

```
# Here we check the condition that the am column equals 0 (ie manual)
# and pick out the rows for which the condition holds true
manual = mtcars[mtcars$am == 1,]
head(manual)
```

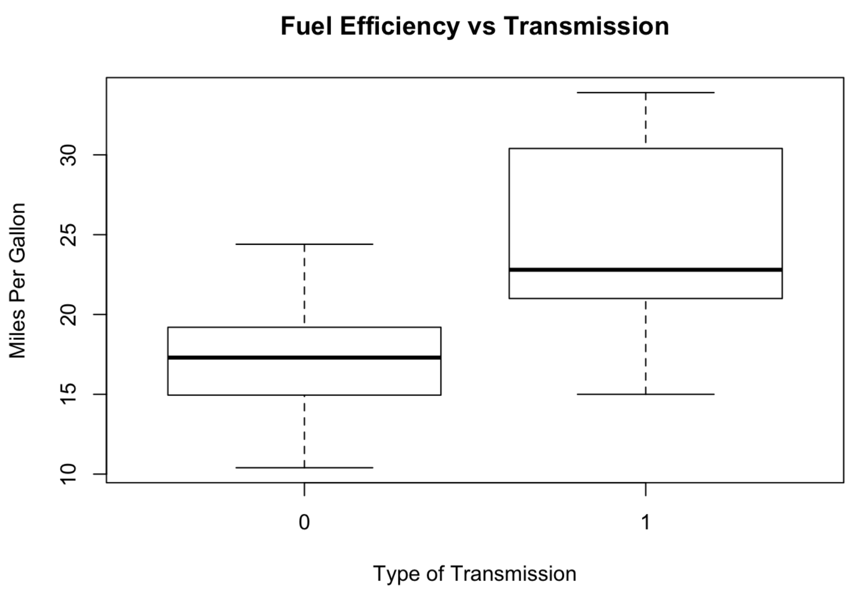
```
##                 mpg cyl  disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4      21.0   6 160.0 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag  21.0   6 160.0 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710     22.8   4 108.0  93 3.85 2.320 18.61  1  1    4    1
## Fiat 128       32.4   4  78.7  66 4.08 2.200 19.47  1  1    4    1
## Honda Civic    30.4   4  75.7  52 4.93 1.615 18.52  1  1    4    2
## Toyota Corolla 33.9   4  71.1  65 4.22 1.835 19.90  1  1    4    1
```

Now that we have reduced our data frame into two specific subgroups as necessary we can press forward with our analysis.

Basic Analysis

First let's take a step back and look at the whole mtcars dataset again. Since nowadays there's is very little to distinguish cars with different transmissions in terms of MPG we should check and see if that was always the case. A simple of method of comparing the miles per gallon for the different types of transmission for cars in 1974 is to create a boxplot. This can be done using the inbuilt boxplot method as follows:

```
#Create boxplot using standard method as shown in class
boxplot(mpg~am,data=mtcars, main="Fuel Efficiency vs Transmission",
      xlab="Type of Transmission", ylab="Miles Per Gallon")
```



Remembering that a 0 signifies automatic and 1 signifies manual transmission it is clear to see that in older manual cars had a distinct edge over those that ran with automatic transmission. There is actually quite a large difference between the two, with the median MPG of the manual transmission well above the upper quartile of the MPG for automatic cars. However, one thing to note is that the MPG for manual cars was seemingly more variant than automatic since the interquartile range for automatic cars is much smaller than their counterparts. Just from this we can see how far technology has improved in the sense that whilst the absolute values for MPG have not changed so much, the efficiency of automatic cars has drastically increased.

Correlation between key factors and MPG

Since we have seen that the transmission clearly had a significant impact on the fuel efficiency of the car, we will utilize our two smaller subsets that we created previously in order to remove the transmission influence and investigate how impactful the key influences we discussed at the beginning of this post affected the different types of transmission.

In order to begin this investigation we will first look at the strength of the linear relationship between the key influences and the MPG variable. This will give us an indication of the strength of the association between the two and thus allow us to glean some insight into how impactful each variable is.

To do this we will look at the correlation matrix of each variable against one another to see if we can isolate any significant relationships. To do this we just have to run the following code for each subset:

```
# Computes the correlation matrix for the automatic car dataset
# for each variable as shown in class
correlation_auto = cor(automatic)
```

```
## Warning in cor(automatic): the standard deviation is zero
```

```
correlation_auto
```

```
##             mpg        cyl       disp         hp       drat         wt
## mpg   1.0000000 -0.7959989 -0.7926335 -0.8315065  0.4682566 -0.7676554
## cyl  -0.7959989  1.0000000  0.8294544  0.8454881 -0.6301768  0.6036585
## disp -0.7926335  0.8294544  1.0000000  0.8343294 -0.6135817  0.8190264
## hp   -0.8315065  0.8454881  0.8343294  1.0000000 -0.3426359  0.6797596
## drat  0.4682566 -0.6301768 -0.6135817 -0.3426359  1.0000000 -0.4011936
## wt   -0.7676554  0.6036585  0.8190264  0.6797596 -0.4011936  1.0000000
## qsec  0.6571077 -0.8563202 -0.6697201 -0.8040275  0.3543292 -0.3707429
## vs    0.7358827 -0.9166985 -0.8209662 -0.8460032  0.5674290 -0.5799984
## am           NA         NA         NA         NA         NA         NA
## gear  0.5400502 -0.6688689 -0.6486873 -0.5857106  0.7792264 -0.3165919
## carb -0.6564196  0.4620689  0.4367031  0.6794894  0.1903169  0.6412858
##            qsec         vs am       gear       carb
## mpg   0.6571077  0.7358827 NA  0.5400502 -0.6564196
## cyl  -0.8563202 -0.9166985 NA -0.6688689  0.4620689
## disp -0.6697201 -0.8209662 NA -0.6486873  0.4367031
## hp   -0.8040275 -0.8460032 NA -0.5857106  0.6794894
## drat  0.3543292  0.5674290 NA  0.7792264  0.1903169
## wt   -0.3707429 -0.5799984 NA -0.3165919  0.6412858
## qsec  1.0000000  0.7993329 NA  0.5579752 -0.4948623
## vs    0.7993329  1.0000000 NA  0.6761234 -0.4063322
## am           NA         NA  1         NA         NA
## gear  0.5579752  0.6761234 NA  1.0000000  0.1217161
## carb -0.4948623 -0.4063322 NA  0.1217161  1.0000000
```

If we want to isolate the correlations of just the MPG variable against all the others since in order to attain a more concise set of results we can just run:

```
# This takes the first 11 rows of the first column.
correlation_auto[1:11, 1]
```

```
##        mpg        cyl       disp         hp       drat         wt
##  1.0000000 -0.7959989 -0.7926335 -0.8315065  0.4682566 -0.7676554
##       qsec         vs         am       gear       carb
##  0.6571077  0.7358827         NA  0.5400502 -0.6564196
```

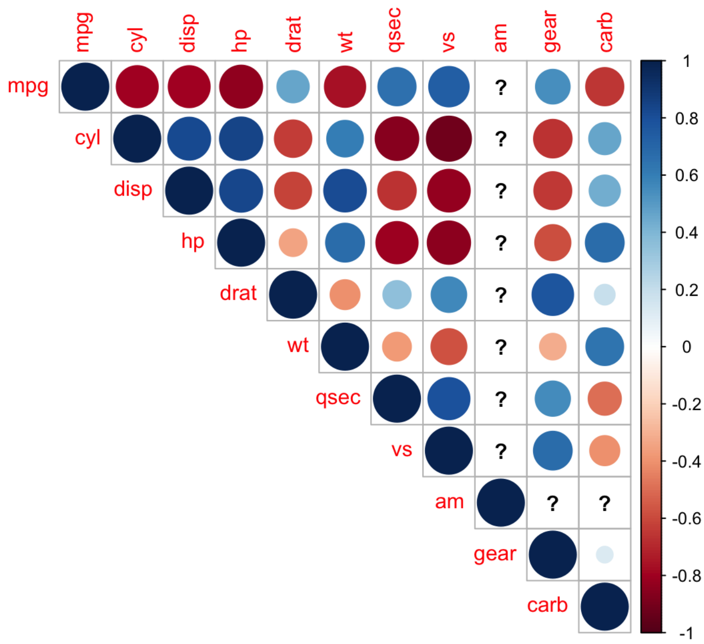
Given the untidiness of the numbers sometimes it is better to visualize the data instead of looking at just the numbers. I say this because it is often easier to see how different the values are instead of just computing the differences. In order to do this we will take advantage of the "corrplot" package[5]. First run the command "install.packages("corrplot")" in your console and then run the following code:

```
# Loads the appropriate library
library("corrplot")
```

```
## Warning: package 'corrplot' was built under R version 3.4.2
```

```
## corrplot 0.84 loaded
```

```
# Function call to help visualise data
corrplot(correlation_auto, type = "upper")
```



Now we can clearly see which attributes are most significantly correlated with MPG. Let us just analyze our three main components first. As you can see displacement and weight are both heavily negatively correlated (given by the red color) with respect to MPG which makes intuitive sense and match the aforementioned research[6]. In addition, we can also see that weight and displacement have a heavy positive correlation in between themselves and thus can draw the conclusion that they go hand in hand. However contrary to our expectations the number of gears is actually positively correlated with MPG which is quite interesting since the more gears an engine needs usually means its heavier. Moreover, there is a significant negative correlation between displacement and the number of gears alongside a smaller negative correlation with weight which again is quite confusing. These anomalies might be due to the fact that older generation automatic cars had very few gears so it did not really impact the efficiency of the car[7].

We can perform these same computations on the manual cars dataset we have:

```
# Computes the correlation matrix for the manual car dataset
# for each variable as shown in class
correlation_man = cor(manual)
```

```
## Warning in cor(manual): the standard deviation is zero
```

```
correlation_man
```

```
##             mpg        cyl       disp         hp       drat         wt
## mpg   1.0000000 -0.8259983 -0.8348954 -0.8006683  0.4700605 -0.9089148
## cyl  -0.8259983  1.0000000  0.9408836  0.9004347 -0.4157897  0.8470235
## disp -0.8348954  0.9408836  1.0000000  0.9240353 -0.3045196  0.8308084
## hp   -0.8006683  0.9004347  0.9240353  1.0000000 -0.4971463  0.8145279
## drat  0.4700605 -0.4157897 -0.3045196 -0.4971463  1.0000000 -0.4999675
## wt   -0.9089148  0.8470235  0.8308084  0.8145279 -0.4999675  1.0000000
## qsec  0.8022104 -0.8588975 -0.8448990 -0.8494566  0.4050637 -0.6791201
## vs    0.7254418 -0.7798436 -0.6926936 -0.6188664  0.3043988 -0.6973598
## am           NA         NA         NA         NA         NA         NA
## gear -0.4019587  0.4892461  0.5562408  0.6730025 -0.3028749  0.2955416
## carb -0.7737834  0.8644802  0.7322192  0.8492580 -0.5328569  0.7818783
##            qsec         vs am       gear       carb
## mpg   0.8022104  0.7254418 NA -0.4019587 -0.7737834
## cyl  -0.8588975 -0.7798436 NA  0.4892461  0.8644802
## disp -0.8448990 -0.6926936 NA  0.5562408  0.7322192
## hp   -0.8494566 -0.6188664 NA  0.6730025  0.8492580
## drat  0.4050637  0.3043988 NA -0.3028749 -0.5328569
## wt   -0.6791201 -0.6973598 NA  0.2955416  0.7818783
## qsec  1.0000000  0.8404934 NA -0.7896317 -0.8462017
## vs    0.8404934  1.0000000 NA -0.5367450 -0.7714333
## am           NA         NA  1         NA         NA
## gear -0.7896317 -0.5367450 NA  1.0000000  0.5579887
## carb -0.8462017 -0.7714333 NA  0.5579887  1.0000000
```

Isolating the main variable, MPG, we want to look at:

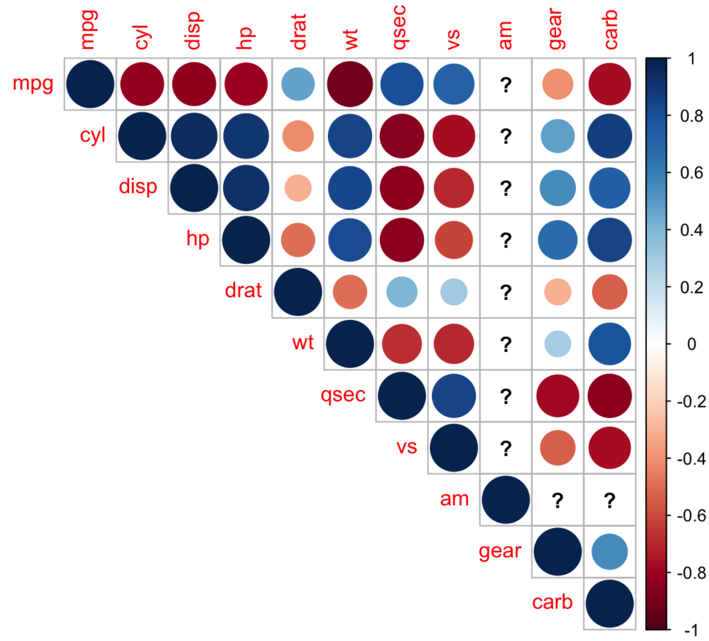
```
# This takes the first 11 rows of the first column.
correlation_man[1:11, 1]
```

```
##        mpg        cyl       disp         hp       drat         wt
##  1.0000000 -0.8259983 -0.8348954 -0.8006683  0.4700605 -0.9089148
##       qsec         vs         am       gear       carb
##  0.8022104  0.7254418         NA -0.4019587 -0.7737834
```

Lastly we can utilize corrplot again to give us a visualization of the different correlations[8]:

```
# Loads the appropriate library
library("corrplot")

# Function call to help visualize data
corrplot(correlation_man, type = "upper")
```



Once again let us start of our analysis by looking at the three highlighted features. With respect manual cars, we have some interesting results: weight seems to have a larger impact than displacement on manual cars and seems to affect manual cars more than automatic cars in general. In addition gears seem to have a negative albeit small impact in this case suggesting that manual cars must have had more gears during this time period than automatic cars and thus could be considered to be more advanced. Moreover, both weight and displacement seem to have positive correlations with the number of gears consolidating the previously stated theory.

Overall it seems like individually weight and displacement have a more significant impact on MPG for both types of transmission and the number of gears seemingly had opposite effects on the two different types. Since the effect of the number of gears in absolute terms was quite low in comparison to the other two one can assume that the displacement and weight of manual cars was just a fair bit smaller than those of their automatic counterparts. I say this because it would explain the large distinction between the MPG values for the two different types of transmissions as we saw previously. In addition, given the low impact of the number of gears on both types of cars I don't believe that cars made in this generation actually were affected by the number of gears too much (unlike modern day cars). There could be several reasons for this but I think the simplest one would be that there just weren't as many. Moreover, it seems like they were more affected by things like the number of cylinders and carburetors.

Multiple Correlation

We can extend our analysis further by utilizing what is known as a multiple correlation coefficient. Essentially what this tells us is the strength and direction of the association between two variables (x and y for example) and a dependent variable z.[9] To fit it to our case an example of x and y would be weight and displacement and MPG would be our z value. With this coefficient then we can can find the pair of variables that together have the highest correlation with MPG. This will give us a better idea on which attributes of a car actually affect the fuel efficiency since in the real world there is no way to isolate each of these components since they have to work in tandem.

Side Note: Something to keep in mind is that this coefficient only showcases the magnitude of the relationship between the independent variables and the one dependent variable[10]. This is because it can never be negative if you look at the value so we dont know the nature of the relationship just the strength of it.

The formula and function

The formula for the multiple correlation coefficient (pictured below)[11] is quite complex and thus instead of typing out code to calculate it each time, we will pass in the appropriate columns of data corresponding to different attributes and the column of data corresponding to MPG.

$$R = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{yx_1} \cdot r_{yx_2} \cdot r_{x_1x_2}}{1 - r_{x_1x_2}^2}}$$

where

r_{yx_1} = correlation coefficient for y and x_1

r_{yx_2} = correlation coefficient for y and x_2

$r_{x_1x_2}$ = correlation coefficient for x_1 and x_2

In the picture above x1 and x2 would be equivalent to weight and displacement and y would be equivalent to MPG for example.

The following code defines a function that mirrors the formula above:

```
multi_cor = function(x1, x2, y){

  # Create appropriate variables as shown in the formula
  r_x1_y = cor(x1, y)
  r_x2_y = cor(x2, y)
  r_x1_x2 = cor(x1, x2)

  # Compute the numerator as shown in the formula
  numerator = r_x1_y^2 + r_x2_y^2 - 2*r_x1_y*r_x2_y*r_x1_x2

  # Compute the denominator as shown in the formula
  denominator = 1 - r_x1_x2^2

  # Return the square root of the numerator divider by the denominator
  # as shown in the formula
  return ((numerator/denominator)^0.5)
}
```

Finding the highest correlated pair of attributes

The following code will return the highest correlated pair and their given correlation to the MPG variable for automatic cars:

```
# Variable to keep track of max correlation
tracker = 0

# Variable that keeps track of pair of variables that
# attained the highest correlation
highest_pair = c(0, 0)

# Double for loop to attain all the possible pairs of variables
# We start at 2 since the first column indicates MPG values
for (i in 2: 11){
  for (j in 2:11){

    # Call to function that computes the multiple correlation coefficient
    # We suppress warnings here because sometimes the standard deviation
    # is 0 since we pass in the same variable twice and R highlights this as bad.
    correlation = suppressWarnings(multi_cor(automatic[, i], automatic[,j], automatic$mpg))

    # Some columns are binary values and hence result in a NA coefficient
    # so we want to skip to the next iteration as we don't need to care about
    # this result
    if (is.na(correlation)){
      next()
    }

    # Like before if we use the same variable twice we sometimes get a coefficient
    # of 1 which is meaningless so we want to skip to the next iteration.
    if (correlation == 1){
      next()
    }

    # This if statement checks to see if we computed a new max correlation.
    # If we did we update all the necessary values.
    if (correlation > tracker){
      tracker = correlation
      highest_pair = c(i, j)
    }
  }
}

# Here we just iterate through both numbers in the highest
# pair vector.
for (i in 1:2){

  attribute_num = highest_pair[i]

  # This attains the column name from the number we stored
  # and prints it out
  print(names(automatic)[attribute_num])
}
```

```
## [1] "gear"
## [1] "carb"
```

```
# Prints out the max correlation we found
print(tracker)
```

```
## [1] 0.9060908
```

We can do the same for the manual cars but we just have to change the dataframe we are looking at. Thus, if we want to check the highest correlated pair for manual cars we can simply run the same code again but with a small change:

```r
# Variable to keep track of max correlation
tracker = 0

# Variable that keeps track of pair of variables that
# attained the highest correlation
highest_pair = c(0, 0)

# Double for loop to attain all the possible pairs of variables
# We start at 2 since the first column indicates MPG values
for (i in 2: 11){
  for (j in 2:11){

    # Call to function that computes the multiple correlation coefficient
    # We suppress warnings here because sometimes the standard deviation
    # is 0 since we pass in the same variable twice and R highlights this as bad.
    correlation = suppressWarnings(multi_cor(manual[, i], manual[,j], manual$mpg))

    # Some columns are binary values and hence result in a NA coefficient
    # so we want to skip to the next iteration as we don't need to care about
    # this result
    if (is.na(correlation)){
      next()
    }

    # Like before if we use the same variable twice we sometimes get a coefficient
    # of 1 which is meaningless so we want to skip to the next iteration.
    if (correlation == 1){
      next()
    }

    # This if statement checks to see if we computed a new max correlation.
    # If we did we update all the necessary values.
    if (correlation > tracker){
      tracker = correlation
      highest_pair = c(i, j)
    }
  }
}

# Here we just iterate through both numbers in the highest
# pair vector.
for (i in 1:2){

  attribute_num = highest_pair[i]

  # This attains the column name from the number we stored
  # and prints it out
  print(names(automatic)[attribute_num])
}
```

```
## [1] "wt"
## [1] "qsec"
```

```r
# Prints out the max correlation we found
print(tracker)
```

```
## [1] 0.9431923
```

Thus, from these results we can see that the most impactful pair of attributes in terms of fuel efficiency for automatic cars in 1974 is the number of gears and carburetors. This is quite surprising since individually both of these variables have quite a low correlation with MPG especially the number of gears. Moreover, the very high multiple correlation coefficient implies a very strong linear relationship between the variables and MPG. As a result, this significant relationship might have actually kickstarted the technological innovation we see in todays cars in terms of number of gears. Some modern day cars have up to 10 gears in comparison to the measly 4 most older cars have![12] In addition notice that the weight of the car and its displacement do not make an appearance which is quite interesting given their high individual associations.

In the case of manual cars we see that the most impactful pair of attributes in terms of fuel efficiency is the weight of the cars and the cars quarter mile time. The latter variable is quite surprising since it is not an attribute that comes to mind when thinking about tuning a cars fuel efficiency. I personally believe that it appears here because of how closely linked the weight of a car and its quarter mile time[13]. As a result, I believe that this indicates that the weight of the car is the variable that most significantly impacts the cars MPG rating. This makes intuitive sense and is in line with the modern day line of thinking.

Take Home Message

After performing two different types of analysis, univariate and multivariate, there are quite a few conclusions we can draw. First of all, looking at the simple boxplots we created we can see that technology has progressed a fair amount in terms of closing the gap between the difference in fuel efficiency between automatic and manual transmission since in 1974 there was a very clear distinction. Moreover, from our single variable coefficient analysis we saw that both weight and displacement were heavily negatively correlated with MPG for both types of transmission thus implying that were large factors in determining fuel efficiency as they are today. However, the number of gears did not meet the expectation set by the research cited in the introduction. This is because not only was its impact quite negligible in comparison to the aforementioned variables but it was also positively correlated with MPG for automatic cars! Thus, if we stopped our analysis there we would say that number of gears affecting the fuel efficiency of cars was not a problem that plagued older manufacturers as they do today.

However, our multivariate correlation showcased different relationships that we weren’t expecting. It showed us that for automatic cars the pairing of number of gears and carburetors seemed to present the strongest association with MPG. On the other hand, weight and consequently the cars quarter mile time were the pair with the strongest relationship with MPG. We can interpret these results in different ways. Clearly our third key variable of displacement is not present in these “important” variables but this does not mean its unimportant since we only looked at the pairing that caused the max correlation. Moreover, the difference in pairings for the two types of transmission could indicate the different purposes of the two types: for example, manual cars were often used for racing back then and thus engine size (and consequently displacement) were minimized. Lastly if we take a combination of our findings from the univariate and multivariate analysis we can argue that the three factors highlighted at the beginning of this post were in-fact the biggest influences on fuel efficiency as they all had a strong relationship with MPG in their own respective manners. Thus, we can claim that whilst technology might have improved the same problems that plagued old cars in regards to their efficiency has not been solved by today’s engineers.

Moreover, if we just look at the absolute value of the miles per gallon these old cars attained we see that they are quite similar to the average number of miles attained by modern cars. This is puzzling because one would think given the amount of innovation in the past half century we would achieved better results but this similarity can be explained by the increase in sophistication of our vehicles. Nowadays we have more gadgets and complex electronic systems in our car that make them more comfortable, safe etc but also less efficient. Thus, perhaps we need to find a better balance to truly raise the average fuel efficiency of modern day cars.

References

1. “Top 10 Factors Contributing To Fuel Economy.” - Car Maintenance and Car Repairs, DriverSide, www.driverside.com/auto-library/top_10_factors_contributing_to_fuel_economy-317.↩
2. “Manual or automatic, which is more likely to use less fuel?” Daily Nation, 18 Feb. 2015, www.nation.co.ke/lifestyle/travel/Manual-or-automatic-which-is–more-likely-to-use-less-fuel/1950822-2626598-ei078q/index.html.↩
3. Melosi, Martin V. “Automobile and the Environment in American History: Auto Emissions and Air Pollution.” Autolife, www.autolife.umd.umich.edu/Environment/E_Overview/E_Overview4.htm.↩
4. “R: Motor Trend Car Road Tests.” R Documentation, stat.ethz.ch/R-manual/R-devel/library/datasets/html/mtcars.html.↩
5. Taiyun Wei and Viliam Simko (2017). R package “corrplot”: Visualization of a Correlation Matrix (Version 0.84). Available from https://github.com/taiyun/corrplot↩
6. “Top 10 Factors Contributing To Fuel Economy.” - Car Maintenance and Car Repairs, DriverSide, www.driverside.com/auto-library/top_10_factors_contributing_to_fuel_economy-317.↩
7. “Automatic transmission.” Wikipedia, Wikimedia Foundation, 10 Nov. 2017, en.wikipedia.org/wiki/Automatic_transmission#History.↩
8. An Introduction to corrplot Package. CRAN Project, cran.r-project.org/web/packages/corrplot/vignettes/corrplot-intro.html. (cited since I learnt how to used package from here.)↩
9. Montgomery, Jayson. “Multiple Correlation and Regression.” SlidePlayer, 30 July 2017, slideplayer.com/slide/8688492/.↩
10. Multiple Correlation Coefficient. Middle Tennessee State University, mtweb.mtsu.edu/stats/regression/level3/multicorrel/multicorrcoef.htm.↩
11. Montgomery, Jayson. “Multiple Correlation and Regression.” SlidePlayer, 30 July 2017, slideplayer.com/slide/8688492/.↩
12. Haj-Assaad, Sami. “Is Having a Transmission With More Gears Really Better? » AutoGuide.Com News.” AutoGuide, 30 Mar. 2017, www.autoguide.com/auto-news/2017/03/is-having-a-transmission-with-more-gears-really-better-.html.↩
13. “How to Compare Weight Savings to Horsepower Gains.” Tuner University, 30 Mar. 2012, www.tuneruniversity.com/blog/2012/03/power-to-weight-ratio/.↩