Effect or transmission on MPG

Executive Summary

- Overall, this analysis suggests that the mpg for **manual trannsmission is better than that of the automatic transmission**.

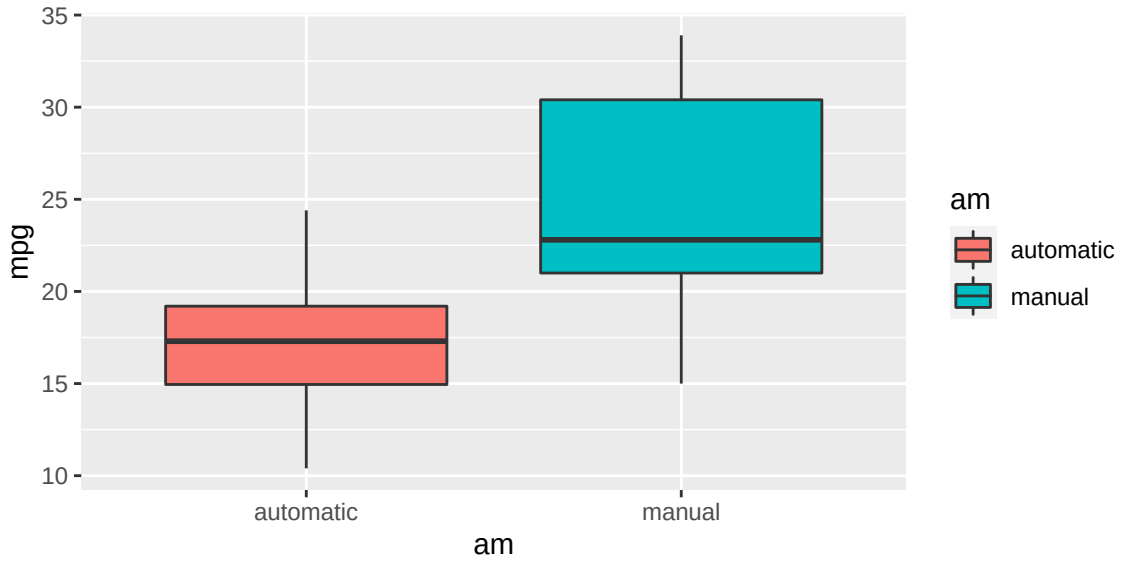
- It suggests that keeping the other covariates constant, the mean mpg for manual transmission is greater than that of automatc transmission by *2.1592707* with the p-value of *0.1440531*.

- This behavior can also be shown using the two sided t test as shown below.

```
## [1] -11.280194  -3.209684
## attr(,"conf.level")
## [1] 0.95
```

The confidence interval lies entirely below zero suggesting that mpg for manual transmission is better than that of the automatic transmission.

Basic Exploratory Analysis



This plot gives a rough estimate that the mpg for manual is better than that of automatic transmission.
Let's dig into it and try to fit a linear regression model which may give a better picture of it.
We also find that:
- cyl,gear and carb are discrete variables.
- vs and am are factor variables.
- mpg,disp,hp,drat,wt and qsec are continuous variables.

Model Selection

```
##           Mazda RX4       Mazda RX4 Wag          Datsun 710      Hornet 4 Drive
##           0.3025065           0.2902207           0.2388171           0.2277394
##   Hornet Sportabout             Valiant          Duster 360           Merc 240D
##           0.1995118           0.2822841           0.3259181           0.3302312
##            Merc 230            Merc 280            Merc 280C          Merc 450SE
##           0.7422870           0.4293261           0.3748573           0.3032810
##          Merc 450SL         Merc 450SLC  Cadillac Fleetwood Lincoln Continental
```

```
##           0.1921148           0.2236587           0.3744802           0.3090439
##   Chrysler Imperial            Fiat 128         Honda Civic      Toyota Corolla
##           0.3066962           0.1789510           0.5119321           0.2328717
##       Toyota Corona    Dodge Challenger         AMC Javelin          Camaro Z28
##           0.4334135           0.2180100           0.1744450           0.4080732
##    Pontiac Firebird           Fiat X1-9       Porsche 914-2        Lotus Europa
##           0.2053054           0.1421645           0.6232257           0.4310982
##      Ford Pantera L        Ferrari Dino       Maserati Bora          Volvo 142E
##           0.6632516           0.3910191           0.6427573           0.2905077
```

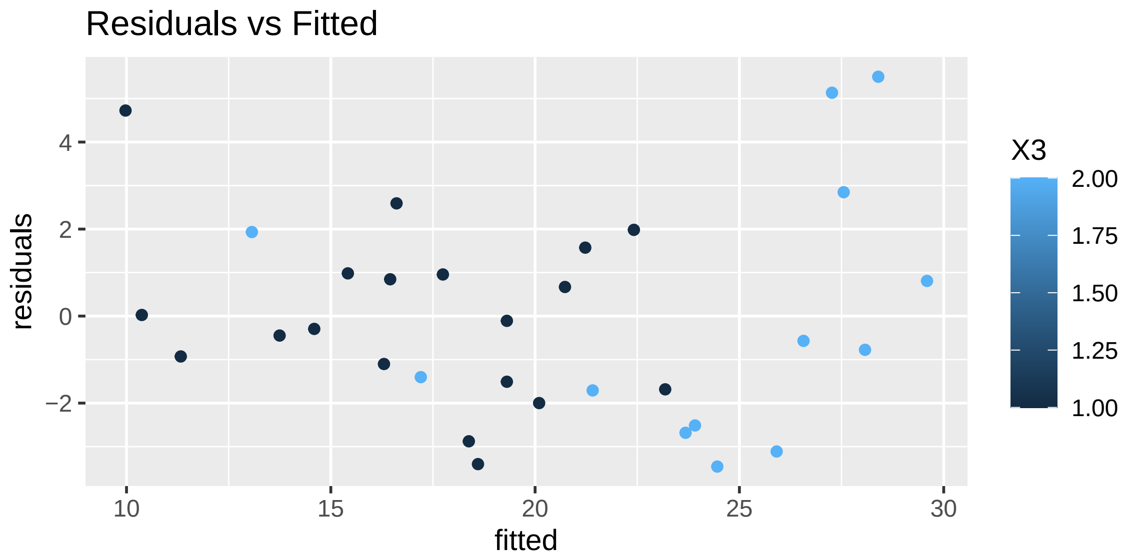
The hatvalues suggests that there is no such potential influencer in our data set that we may need to look upon.

Let's fit the model with different parameters depending upon their correlation with the mpg and see which one fits the best.

```
## Analysis of Variance Table
##
## Model 1: mpg ~ disp + am
## Model 2: mpg ~ disp + am + wt
## Model 3: mpg ~ disp + am + wt + hp
## Model 4: mpg ~ disp + am + wt + hp + drat
## Model 5: mpg ~ disp + am + wt + hp + drat + qsec
##   Res.Df    RSS Df Sum of Sq       F   Pr(>F)
## 1     29 300.28
## 2     28 246.56  1    53.725  8.9486 0.006164 **
## 3     27 179.91  1    66.649 11.1012 0.002685 **
## 4     26 175.67  1     4.240  0.7062 0.408668
## 5     25 150.09  1    25.574  4.2598 0.049551 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

As we can see that model 2 performs better than the first model and the model 3 perfoms even better than the model 2 with the F value of apprx. 0.0026.
After that the F value increases which suggests that the later models are not good.
This analayis of variance suggests that the third model is best among all of them.

Analysing



There is no such significant pattern in the residuals and the fitted values as such.

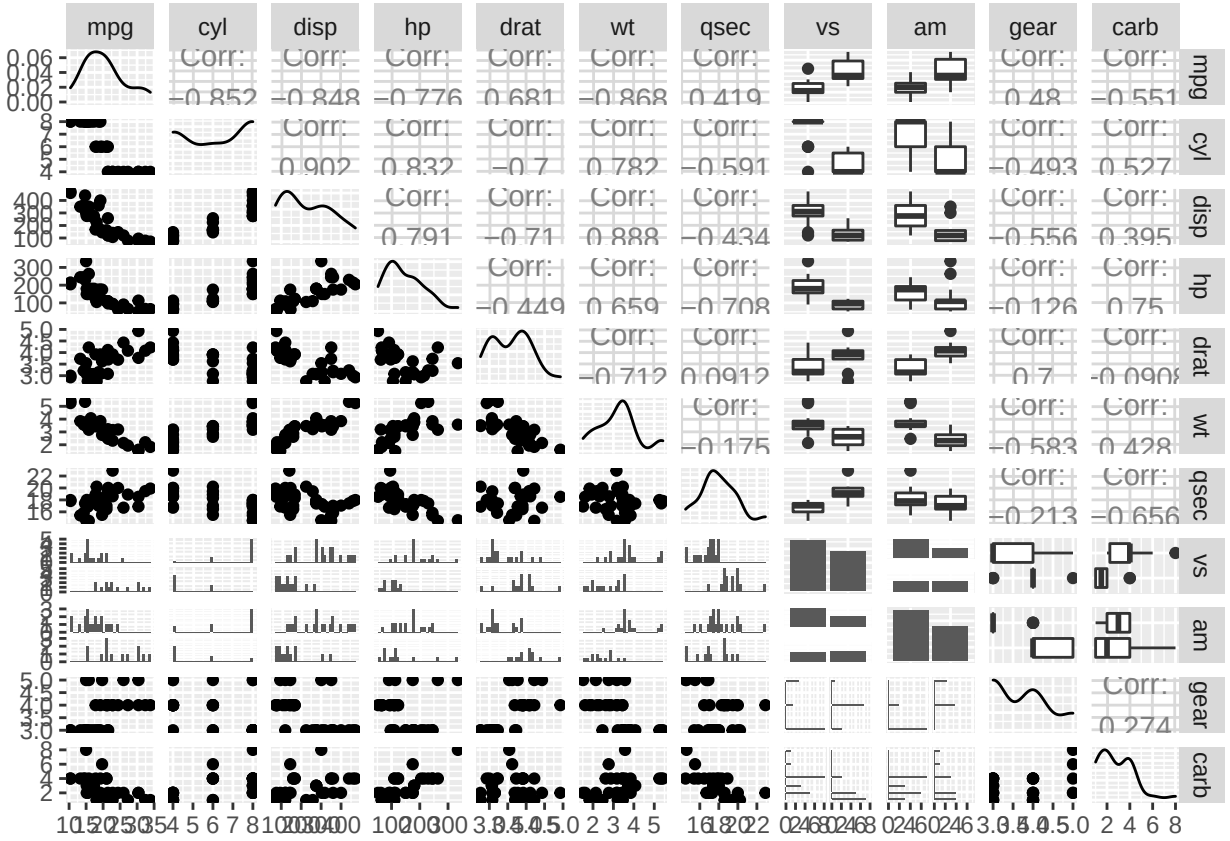
```
##                 Estimate Std. Error    t value     Pr(>|t|)
## (Intercept) 34.209443370 2.82282610 12.1188632 1.979953e-12
## disp         0.002489354 0.01037681  0.2398959 8.122229e-01
## ammanual     2.159270737 1.43517565  1.5045341 1.440531e-01
## wt          -3.046747000 1.15711931 -2.6330448 1.382936e-02
## hp          -0.039323213 0.01243358 -3.1626624 3.842032e-03
```

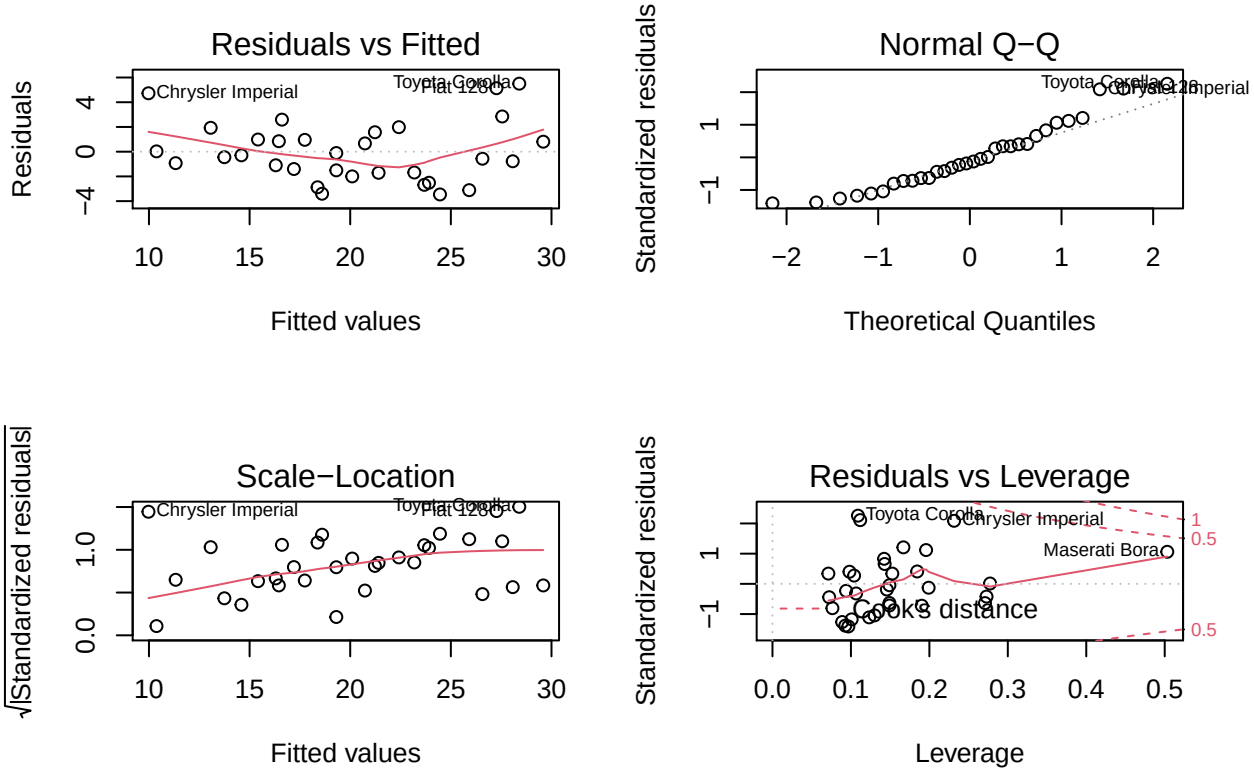
The coefficient corresponding to "ammanual" suggests that keeping the other covariates constant, the **mean mpg for manual transmission is greater than that of automatc transmission** by 2.1592707 with the p-value of 0.1440531.
Also the coefficient corresponding to displ suggests that there is a very *minor increase in the mpg with displacement* keeping other covariates constant.
The coefficient corresponding to weight suggests that *wt has a negative impact on the the mpg* keeping others constant.
Also, there is a very *minor decrease in the mpg with hp* keeping other covariates constant.

Appendix Figures

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

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




Residuals vs Fitted
Residuals
Fitted values
Chrysler Imperial
Toyota Corolla
Fiat 128
Normal Q−Q
Standardized residuals
Theoretical Quantiles
Toyota Corolla
Chrysler Imperial
Scale−Location
√|Standardized residuals|
Fitted values
Chrysler Imperial
Toyota Corolla
Fiat 128
Residuals vs Leverage
Standardized residuals
Leverage
Toyota Corolla
Chrysler Imperial
Maserati Bora
Cook's distance
1
0.5
0.5