project

Dehshini

2023-12-28

Regression Model Course Project

Executive summary:

This analysis was done to answer the following:

"Is an automatic or manual transmission better for MPG"?
"Quantify the MPG difference between automatic and manual transmissions"

There was a total of 19 automatic transmission cars and 13 manual transmission cars. The average mpg for automatic transmission cars was 17.15mpg while that for manual cars was higher by 7.25 at 24.39mpg. After controlling for factors that may increase fuel consumption such as car weight, number of cylinders and gross horsepower of engines, there was no significant difference in mpg between automatic and manual transmission cars. The difference in mpg between manual and automatic was 1.8 (95% CI -1.1, 4.7)

Conclusion: there is a 1.8mpg difference between manual and automatic transmission cars after controlling for car weight, gross horsepower and number of cylinders of engine. Both automatic and manual transmission have similar mpg.

```
##                    mpg cyl disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
## Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
## Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
## Valiant           18.1   6  225 105 2.76 3.460 20.22  1  0    3    1
```

```
##                 mpg cyl  disp  hp drat    wt qsec vs am gear carb
## Porsche 914-2  26.0   4 120.3  91 4.43 2.140 16.7  0  1    5    2
## Lotus Europa   30.4   4  95.1 113 3.77 1.513 16.9  1  1    5    2
## Ford Pantera L 15.8   8 351.0 264 4.22 3.170 14.5  0  1    5    4
## Ferrari Dino   19.7   6 145.0 175 3.62 2.770 15.5  0  1    5    6
## Maserati Bora  15.0   8 301.0 335 3.54 3.570 14.6  0  1    5    8
## Volvo 142E     21.4   4 121.0 109 4.11 2.780 18.6  1  1    4    2
```

```
##       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
```

```
## # A tibble: 2 x 3
##      am     n mean_mpg
##   <dbl> <int>    <dbl>
## 1     0    19     17.1
## 2     1    13     24.4
```

mpg seems to be associated with transmission with the automatic transmission cars having less mpg compared to manual transmission cars.

```
##
## Call:
## lm(formula = mpg ~ factor(am), data = mt)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.3923 -3.0923 -0.2974  3.2439  9.5077
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   17.147      1.125  15.247 1.13e-15 ***
## factor(am)1    7.245      1.764   4.106 0.000285 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.902 on 30 degrees of freedom
## Multiple R-squared:  0.3598, Adjusted R-squared:  0.3385
## F-statistic: 16.86 on 1 and 30 DF,  p-value: 0.000285
```

```
##                2.5 %   97.5 %
## (Intercept) 14.85062 19.44411
## factor(am)1  3.64151 10.84837
```

the average mpg of automatic transmission cars is 17.147 (95% confidence interval 14.8 - 19.4). the mean mpg of manual transmission vehicles is 7.245 more than automatic transmission vehicles(95% CI 3.6 - 10.8).

this relationship may be confounded by other variables such as weight, horsepower, number of cylinders. Do a matrix graph to explore the associations.

control for other variables

```
##
## Call:
## lm(formula = mpg ~ factor(am) + factor(cyl) + wt + hp, data = mt)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.9387 -1.2560 -0.4013  1.1253  5.0513
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  33.70832    2.60489  12.940 7.73e-13 ***
## factor(am)1   1.80921    1.39630   1.296  0.20646
## factor(cyl)6 -3.03134    1.40728  -2.154  0.04068 *
## factor(cyl)8 -2.16368    2.28425  -0.947  0.35225
## wt           -2.49683    0.88559  -2.819  0.00908 **
## hp           -0.03211    0.01369  -2.345  0.02693 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.41 on 26 degrees of freedom
## Multiple R-squared:  0.8659, Adjusted R-squared:  0.8401
## F-statistic: 33.57 on 5 and 26 DF,  p-value: 1.506e-10
```

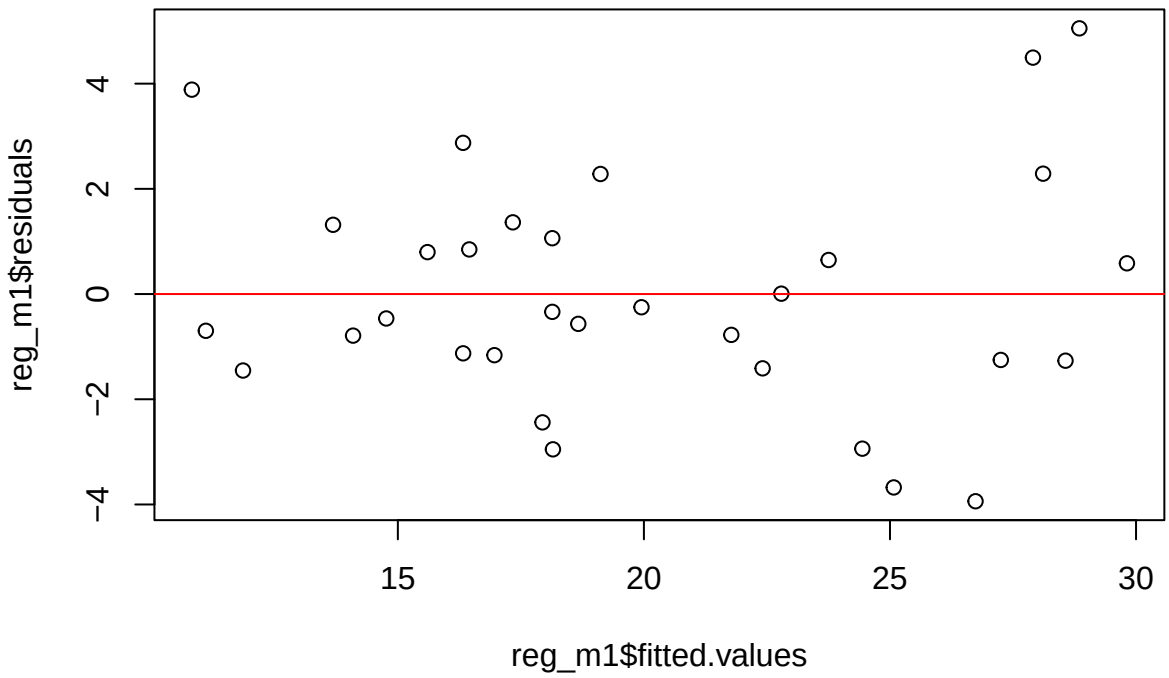
```
##                    2.5 %       97.5 %
## (Intercept)  28.35390366 39.062744138
## factor(am)1  -1.06093363  4.679356394
## factor(cyl)6 -5.92405718 -0.138631806
## factor(cyl)8 -6.85902199  2.531671342
## wt           -4.31718120 -0.676477640
## hp           -0.06025492 -0.003963941
```

after controlling for weight of car, number of cylinders, and gross horsepower, the mpg does not significantly differ by transmission type. the average mpg for automatic transmission cars is 33.7 (95% CI 28.4 - 39.1). the average for manual transmission is 1.8 more than that of automatic (95% CI -1.1 - 4.7)

compare the two models

the multivariate model seems better as the residual sum of squares is significantly lower.

check model assumptions.



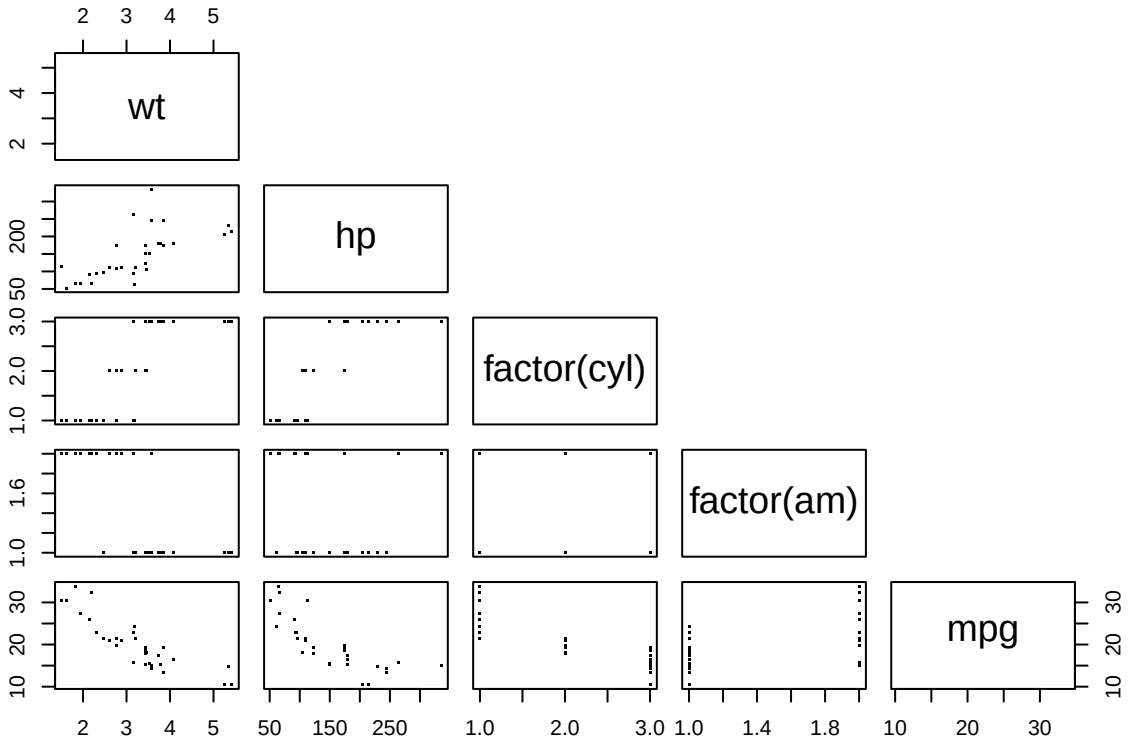
residuals look random with no clear pattern. Linear, somewhat equal variance except for the higher fitted values where variance is higher.

normal distribution mostly, except for the tail ends.

perform diagnostic, dfits

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.