Course Project

Pat Mitoraj

December 17, 2017

Course Project

Executive Summary

Examining data from the mtcars dataset, there does not appear to be a statistically significant relationship between transmission type and a car's average miles per gallon. While manual transmission cars have a higher miles per gallon on average, once car weight and engine cylinders are accounted for there is not a statistical difference between automatic and manual transmission cars.

Exploratory Analysis

Exploratory analysis was done to investigate the relationships between several of the variables in the dataset. First, the relationship between transmission type and miles per gallon was plotted.

Figure 1 of the appendix shows that manual cars tend to have a higher miles per gallon. However, this difference could be accounted for by other variables. To explore this possibility, the researcher next plotted transmission type vs. car weight to see if there was a relationship.

Figure 2 showed that manual cars do tend to have a lower weight than automatic transmission cars, meaning that weight may be influencing the relationship between transmission and mpg.

The other relationship explored was that between transmission type and engine cylinders

FIGURE 3 shows that the majority of automatic cars were 8 cylinders, while the majority of manual cars were only 4 cylinder engines. Again, this is a relationship that could be influencing the relationship between transmission and miles per gallon.

Model Selection

The initial model used was a linear model with transmission predicting miles per gallon. This model was used as a baseline to understand how, without accounting for any other variables, transmission and mpg were related.

```
model<-lm(mtcars$mpg~mtcars$am)
```

A second model was also used, which was again a linear regression. However, this model also accounted for car weight and engine cylinders, both of which were shown to have a possible relationship with transmission type in the exploratory data analysis.

```
model2<-lm(mtcars$mpg~mtcars$am+mtcars$wt+mtcars$cyl)
```

Results

The results of the first model showed a strongly signifcant relationship between miles per gallon and transmission type

```
##
## Call:
## lm(formula = mtcars$mpg ~ mtcars$am)
##
## 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 ***
## mtcars$am      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
```

These model results show that a manual transmission car was predicted to have 7.245 more miles per gallon than an automatic transmission car. Furthermore with a p value of less than .01, the relationship was significant at the 1% significance level. The plot of the predicted versus actual values (Figure 4) shows that the predicted values fell near the center of the distribution for both automatic and manual transmission cars.

The results for the second model show a very different result.

```
##
## Call:
## lm(formula = mtcars$mpg ~ mtcars$am + mtcars$wt + mtcars$cyl)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.1735 -1.5340 -0.5386  1.5864  6.0812
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  39.4179     2.6415  14.923 7.42e-15 ***
## mtcars$am     0.1765     1.3045   0.135  0.89334
## mtcars$wt    -3.1251     0.9109  -3.431  0.00189 **
## mtcars$cyl   -1.5102     0.4223  -3.576  0.00129 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.612 on 28 degrees of freedom
## Multiple R-squared:  0.8303, Adjusted R-squared:  0.8122
## F-statistic: 45.68 on 3 and 28 DF,  p-value: 6.51e-11
```

Once car weight and engine cylinders are held constant, the coefficient for transmsission drops dramatically (from 7.245 to .18) and is no longer statistically significant. In contrast, both car weight and engine cylinders both appear to have negative relationships with miles per gallon and are signifcant at the 1% significance level. Additionally, this model is a much better predictor of miles per gallon as evidenced by its r-squared value. The first model had an r-squared of .36, while model 2 had an r-squared of .83.

The results of model 2 are confirmed by the plot of predicted vs. actual results and the residuals (Figure 5). The am plot has larger residuals, as evidence by the higher proportion of solid colored blue and red dots, whereas the cyl and wt graphs are better fits with smaller residuals.

Conclusion

In summary, the results of this analysis show that there is not a statistically significant difference between the miles per gallon of automatic cars versus manual transmission cars. Instead, engine cylinders or car weight are much more important factors in identifying which car will get you more miles per gallon.

Appendix

FIGURE 1

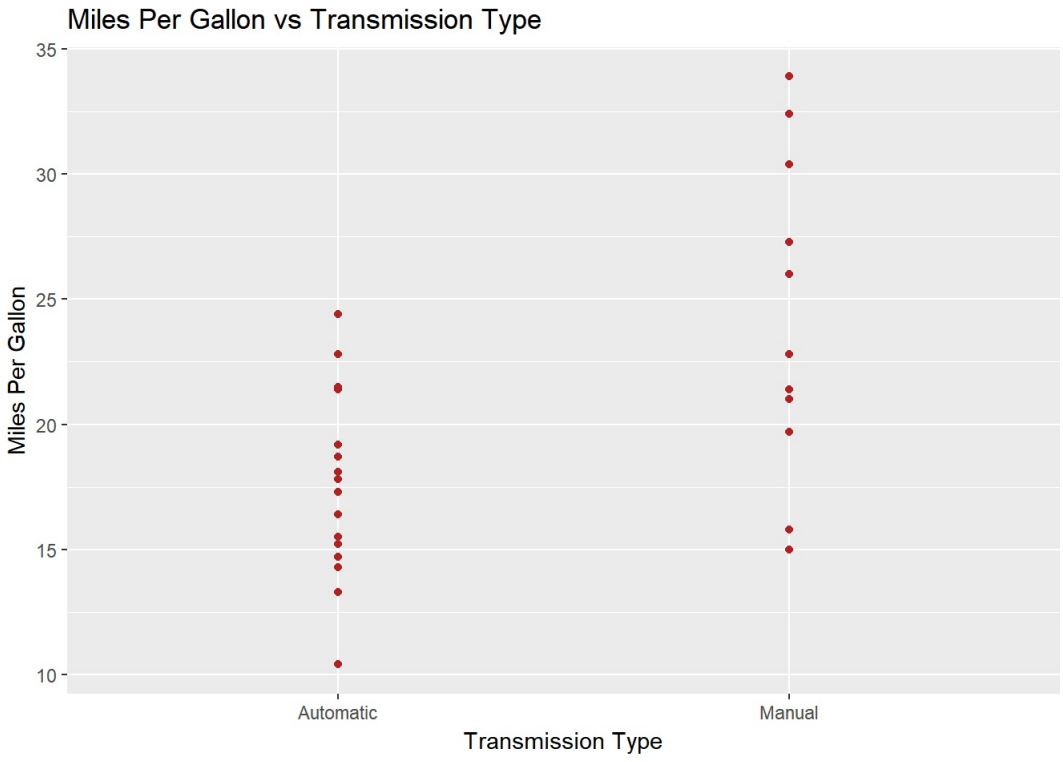


FIGURE 2

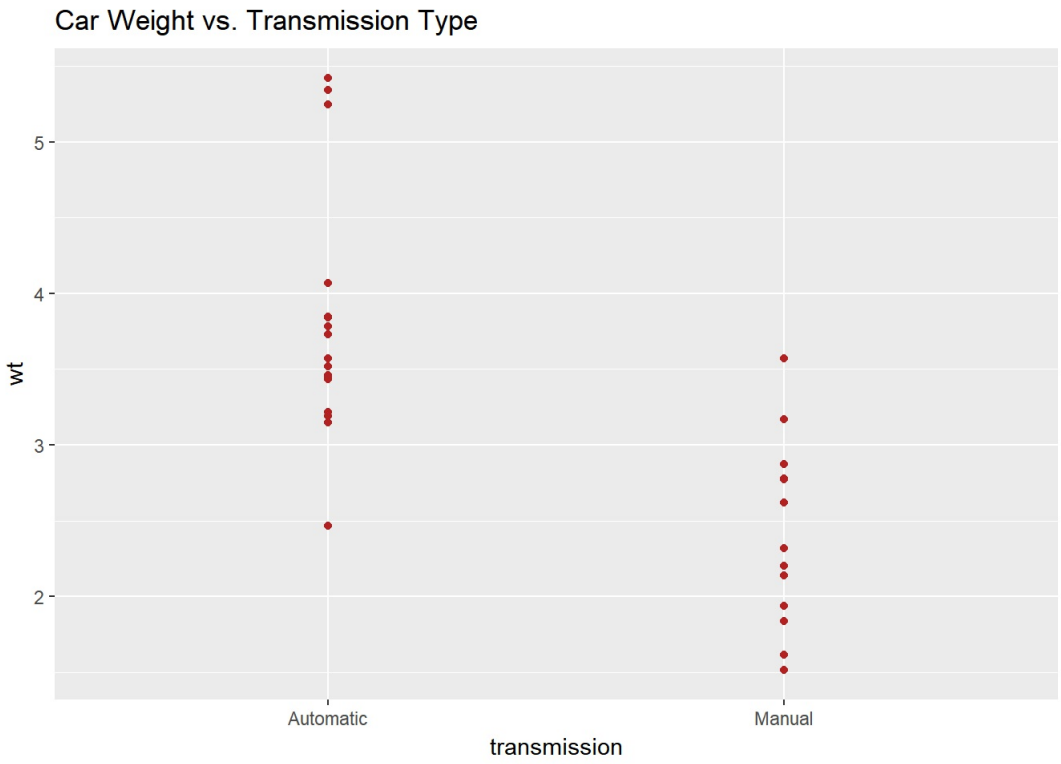


FIGURE 3

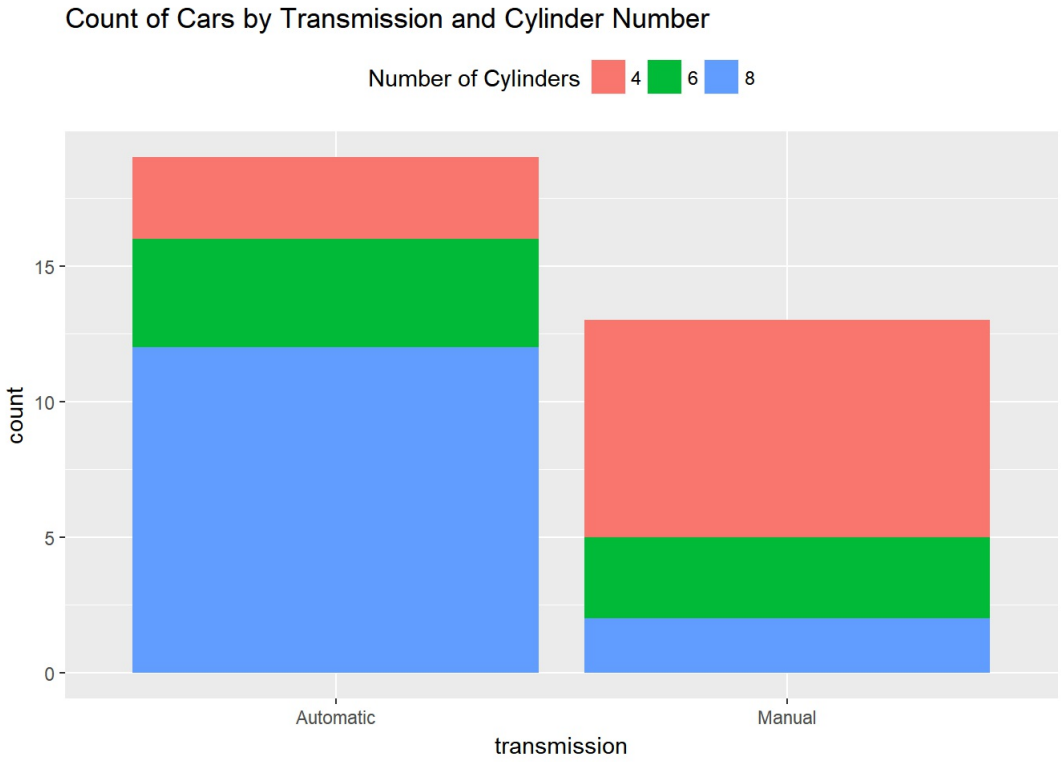


FIGURE 4

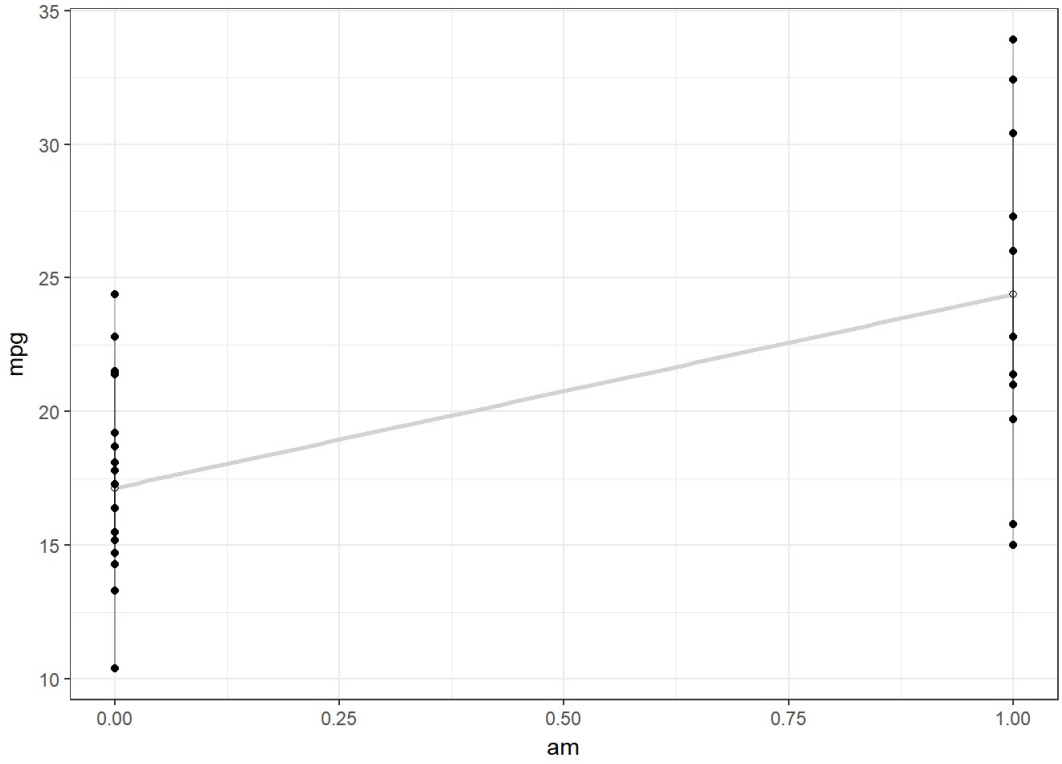
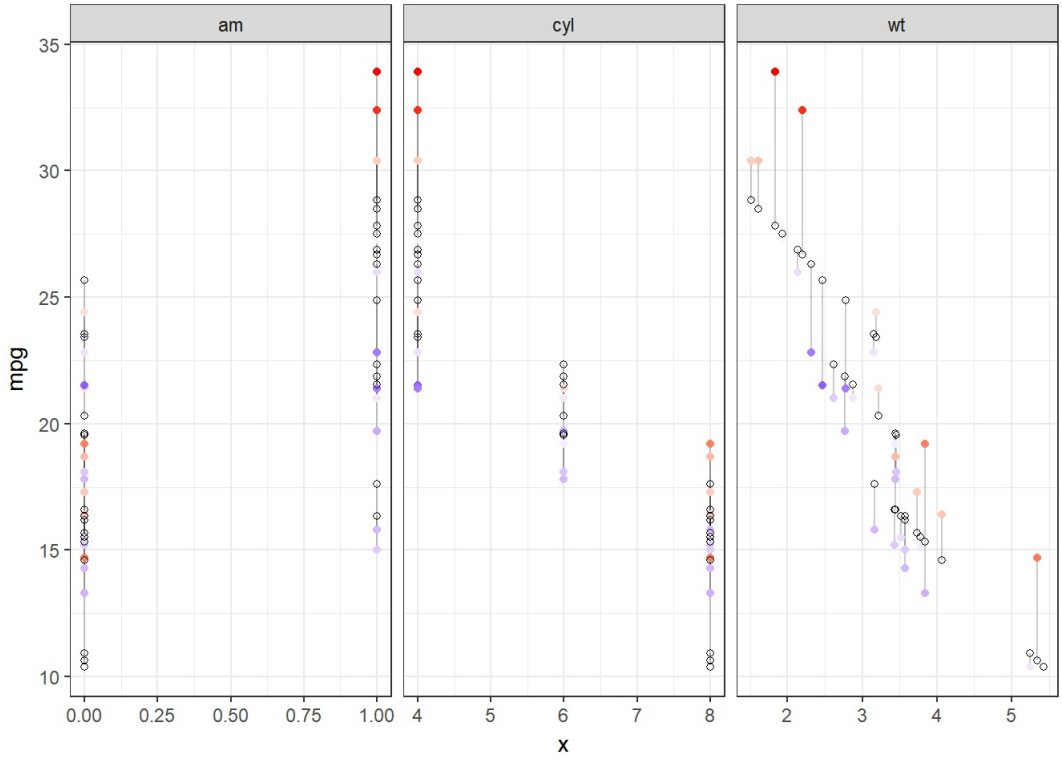

mpg
35
30
25
20
15
10
0.00
0.25
0.50
0.75
1.00
am

FIGURE 5


am
cyl
wt
mpg
35
30
25
20
15
10
0.00
0.25
0.50
0.75
1.00
4
5
6
7
8
2
3
4
5
x