Motor Trend MPG Analysis: Manual vs. Automatic

Brad Tweardy

Sunday, April 19, 2015

Executive Summary

In this project we use a dataset, mtcars, to analyze a selection of 30 cars (observations) and 11 variables (mpg, cyl, disp, hp, drat, wt, qsec, vs, am, gear, carb). The mission of this report is to use Statistics and Regression Models through Exploratory Data Analysis to answer the following two questions:

1. Is an automatic or manual transmission better for MPG?
2. Quantifying, how different is the MPG between automatic and manual transmissions?

Overall, our results have shown a better result for miles per gallon in manual transmission automobiles versus automatic transmission automobiles. Whether we use just the transmission type or all variables as predictors, we can see evidence in this in the report.

Analysis

First, we'll set a few variables to factors then we will conduct a linear model (begin_model) that uses all ten variables as predictors of mpg (see Figure 2 for scatter plot of the results). We use the step function in this model to determine our "best model" (best_model). The step function selects the best variables (cyl, hp, am, wt) in predicting mpg by the Akaike information criterion that implements both forward selection and backward elimination and eliminates variables that are not useful in predicting mpg. We'll then review a summary of our best model.

```
data(mtcars)
mtcars$cyl <- factor(mtcars$cyl)
mtcars$vs <- factor(mtcars$vs)
mtcars$gear <- factor(mtcars$gear)
mtcars$carb <- factor(mtcars$carb)
mtcars$am <- factor(mtcars$am, labels = c("Automatic", "Manual"))
begin_model <- lm(mpg ~ ., data = mtcars)
best_model <- step(begin_model, direction = "both")
```

```
summary(best_model)
```

```
##
## Call:
## lm(formula = mpg ~ cyl + hp + wt + am, data = mtcars)
##
## 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 ***
## cyl6        -3.03134    1.40728  -2.154  0.04068 *
## cyl8        -2.16368    2.28425  -0.947  0.35225
## hp          -0.03211    0.01369  -2.345  0.02693 *
## wt          -2.49683    0.88559  -2.819  0.00908 **
## amManual     1.80921    1.39630   1.296  0.20646
## ---
## 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
```

From the summary we see that R squared is .84, which tells us that 84% of the variablity is described with our best model.

If we create a basic model (simple) using only transmission type (am) as a predictor for mpg, we can see that manual transmissions have an obvious advantage versus automatic transmissions with regards to their effect on a automobile's MPG (*See Figure 2*).

```
simple <- lm(mpg ~ am, data = mtcars)
```

Now, we can create a comparison model and compare the basic model to our prior two models (best_model, begin_model). In comparison we can see that the p-value for the begin model is very high so we cannot reject the null hypothesis that the additional variables that it includes from our begin model do not contribute. Conversely, we can reject the null hypothesis in the best model because the p-value is very low.

```
anova(simple, best_model, begin_model)
```

```
## Analysis of Variance Table
##
## Model 1: mpg ~ am
## Model 2: mpg ~ cyl + hp + wt + am
## Model 3: mpg ~ cyl + disp + hp + drat + wt + qsec + vs + am + gear + carb
##   Res.Df    RSS Df Sum of Sq       F    Pr(>F)
## 1     30 720.90
## 2     26 151.03  4    569.87 17.7489 1.476e-05 ***
## 3     15 120.40 11     30.62  0.3468    0.9588
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Conclusion

In conclusion, we can answer our two questions with the conducted analysis.

1. We can say that a manual transmission is better for MPG in automobiles when we adjust for all variables.
2. If we are to keep all variables constant, except for the transmission type, we can say that we see an increase of about 1.8 MPG from automatic to manual transmission.

Appendix

Figure 1

```
pairs(mtcars)
```

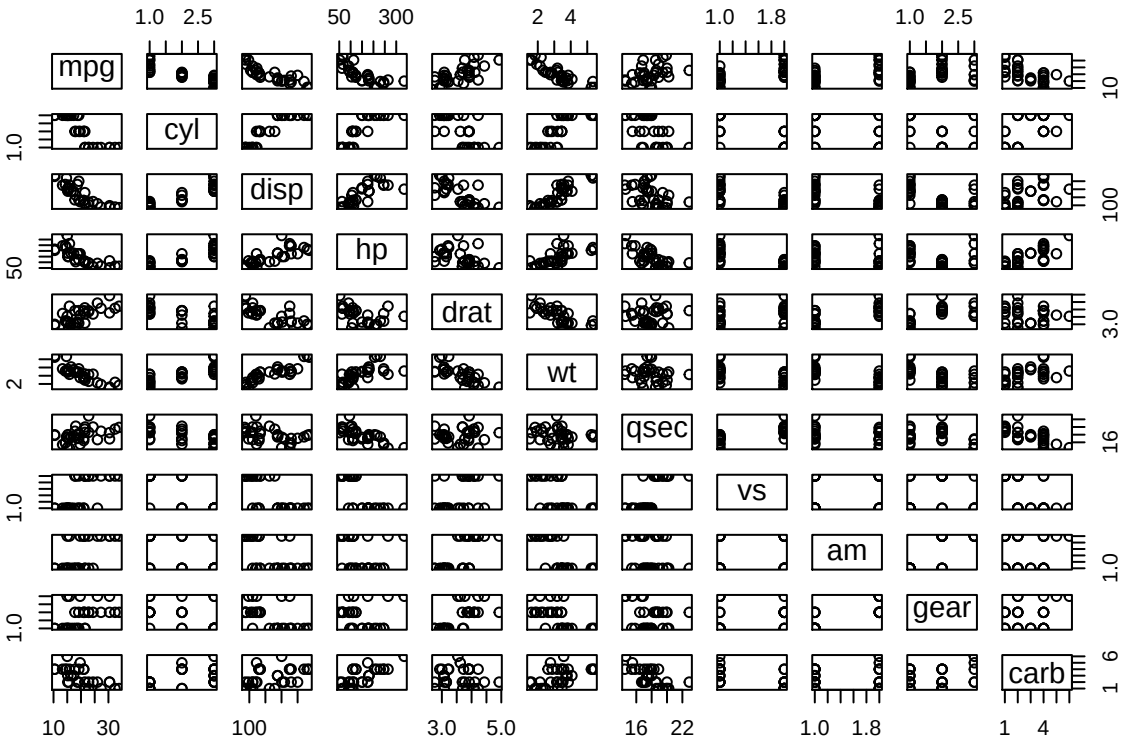
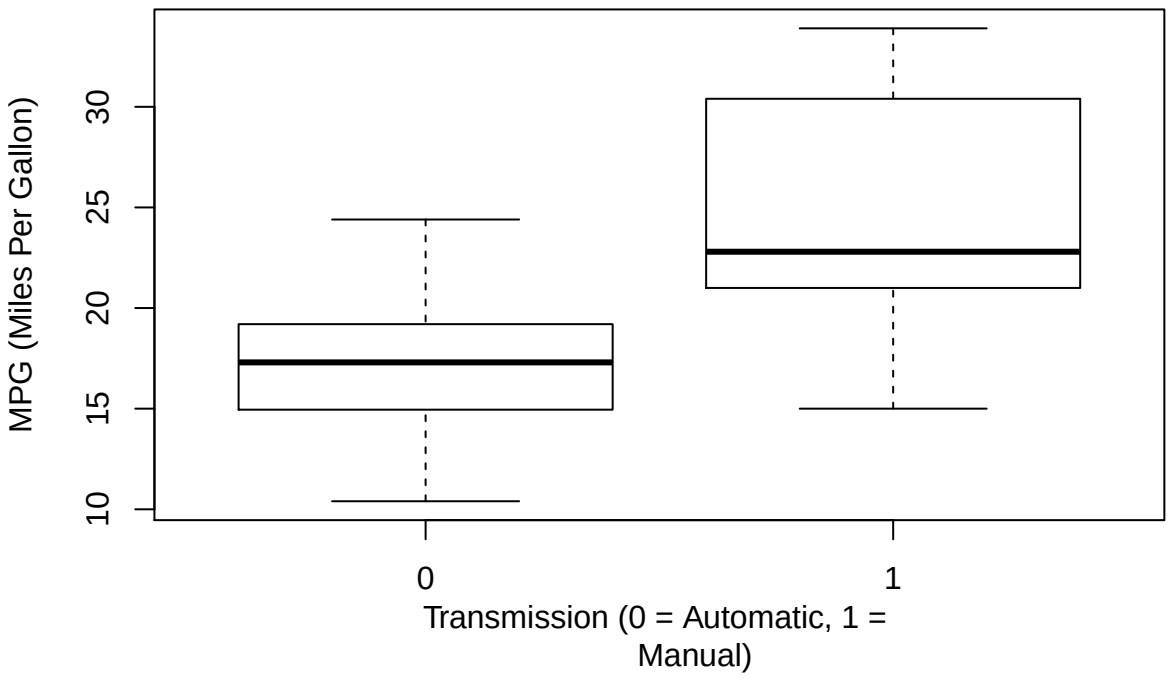


Figure 2

```
library(datasets)
data(mtcars)
boxplot(mpg ~ am, data = mtcars, xlab = "Transmission (0 = Automatic, 1 =
        Manual)", ylab = "MPG (Miles Per Gallon)")
```


MPG (Miles Per Gallon)
10
15
20
25
30
0
1
Transmission (0 = Automatic, 1 = Manual)