Motor Trend - Impact of Automatic vs Manual Transmission on MPG Analysis

Song Ray Yang

Saturday, May 23, 2015

Executive Summary

Motor Trend magazine used the dataset "mtcars"" from 1974 which comprises fuel consumption and performance for 32 automobiles to analyze the impact of different transmission types on miles per gallon (MPG). The following analysis proves that manual transmission outperforms automatic transmission by 7.24 MPG when all other variables are held constant.

Exploratory Data Analysis

Before we build a model, let us explore the dataset first.

```
library(datasets)
data(mtcars)
summary(mtcars$mpg)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   10.40   15.42   19.20   20.09   22.80   33.90
```

```
mtcars$am <- factor(mtcars$am, labels=c("automatic","manual"))
```

Please see Figure 1 in the Appendix for a boxplot by type of transmission. Figure 1 shows that manual transmission has higher mean compared to automatic transmission. To prove this is indeed the case, we run a regression analysis.

Regression Analysis

```
simplemodel <- lm(mpg ~ am, mtcars)
summary(simplemodel)
```

```
##
## Call:
## lm(formula = mpg ~ am, data = mtcars)
##
## 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 ***
## ammanual       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
```

From the coefficients we can see that manual transmission performs better as measured in MPG on average 17.147 + 7.245 = 24.392 MPG as compared to 17.147 MPG for automatic transmission. The Adjusted R-squared says that 34% of the variability is explained by this model.

However, transmission type may not be the only significant variable that determines MPG. Let us run a multivariable model taking into account weight, number of cyclinders, and gross horsepower.

```
bettermodel <- lm(mpg ~ am + wt + cyl + hp, mtcars)
summary(bettermodel)
```

```
##
## Call:
## lm(formula = mpg ~ am + wt + cyl + hp, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.4765 -1.8471 -0.5544  1.2758  5.6608
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 36.14654    3.10478  11.642 4.94e-12 ***
## ammanual     1.47805    1.44115   1.026   0.3142
## wt          -2.60648    0.91984  -2.834   0.0086 **
## cyl         -0.74516    0.58279  -1.279   0.2119
## hp          -0.02495    0.01365  -1.828   0.0786 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.509 on 27 degrees of freedom
## Multiple R-squared:  0.849,  Adjusted R-squared:  0.8267
## F-statistic: 37.96 on 4 and 27 DF,  p-value: 1.025e-10
```

The second model is better as the Adjusted R-squared says that 83% of the variabilty is explained by this model. We also see weight as the variable with the most impact (ie lowest p-value).

Residual Plots and Diagnostics

In order to compare the models, we also look at the residual plots for both models in Figure 2 and Figure 3. They confirm that while a simple model can show significance of transmission variable, it is not the best model. The better model would be a better predictor of MPG.

Conclusion

While simple model shows that manual transmission is better than automatic but the low adjusted R-squared of 34% isn't convincing as for using transmission type as primary predictor. The better model with the higher adjusted R-squared of 83% show that weight has more effect on MPG.

Appendix

Figure 1

```
boxplot(mpg ~ am, mtcars, ylab = "miles per gallon", xlab = "transmission type")
```

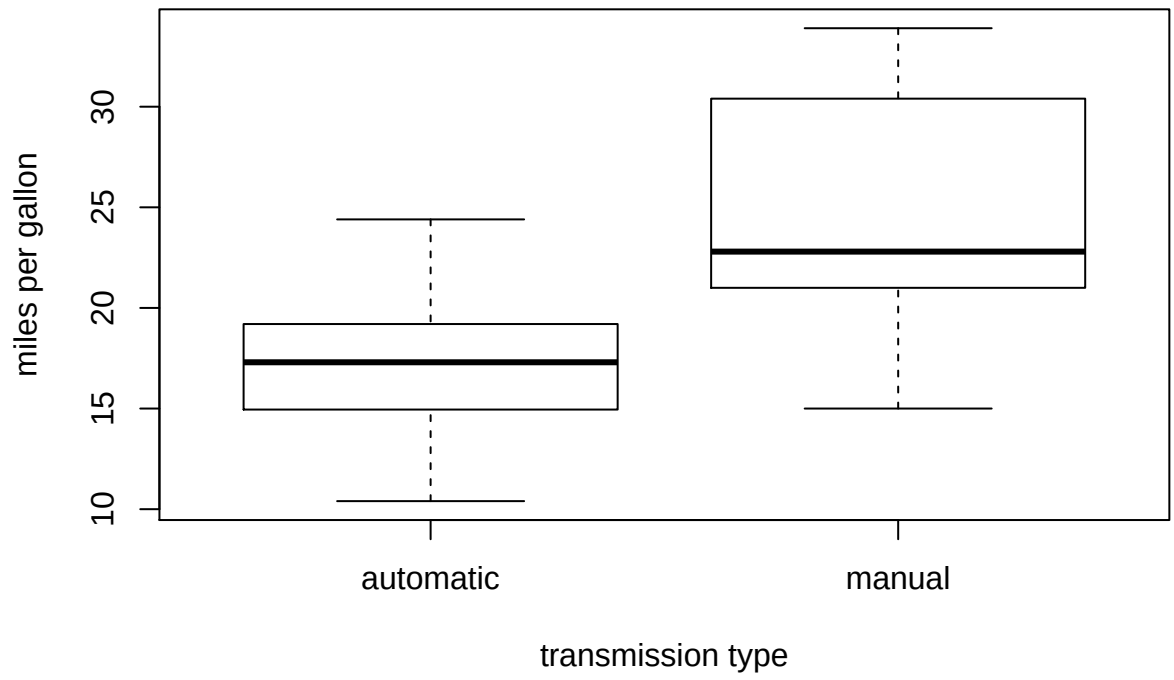


Figure 2

```
par(mfrow=c(2,2))
plot(simplemodel)
```

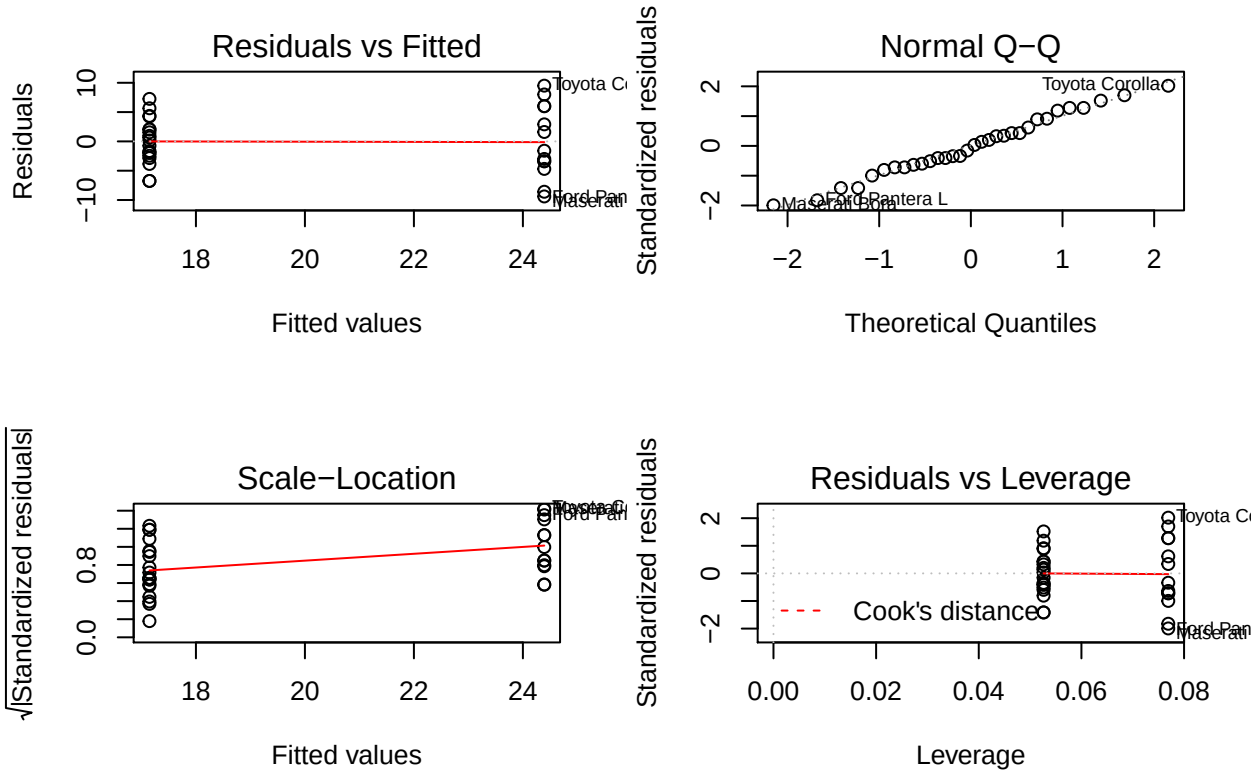
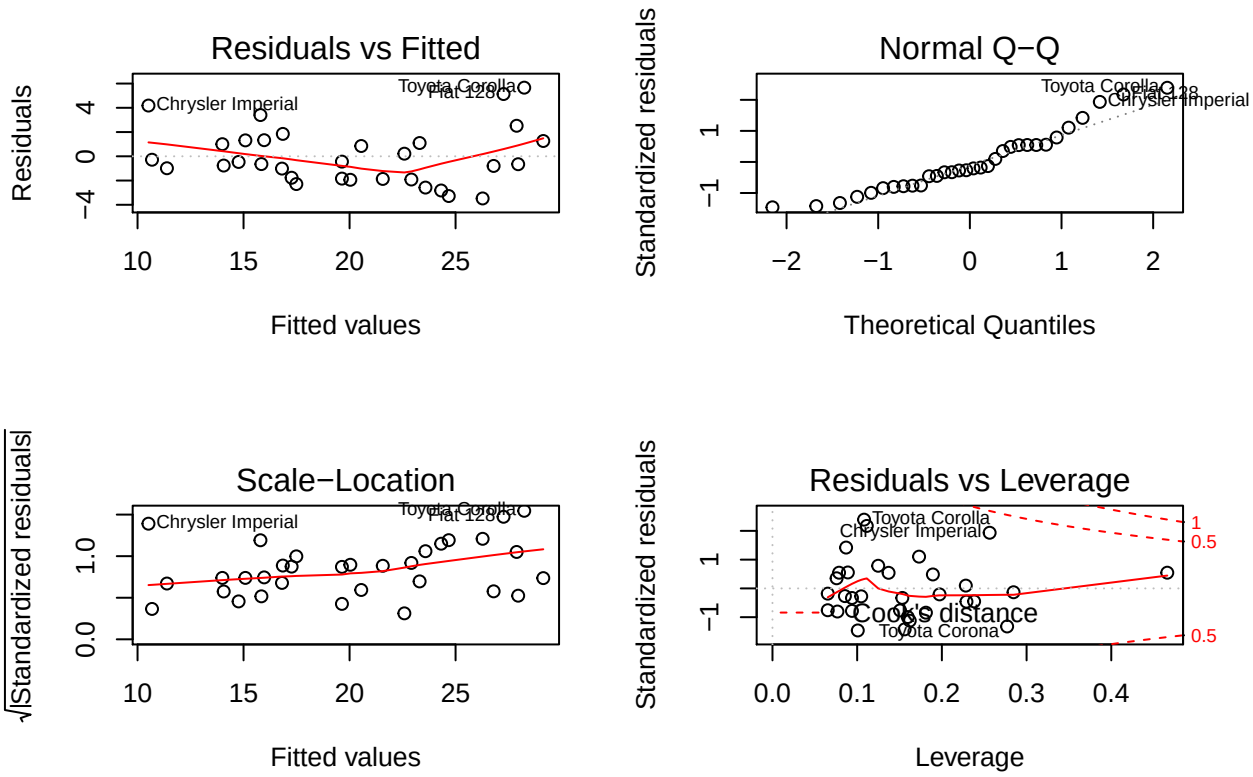


Figure 3

```
par(mfrow=c(2,2))
plot(bettermodel)
```


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