Are Manual Transmissions Better Consumers?

Exclusive Summary

The `mtcars` dataset is investigated. Relationship between variables in the dataset and `MPG`, the outcome, is explored. At the first look, cars with manual transmission make, on average, **7.24** more miles per gallon than cars with automatic transmission. However, there are confounding variables. These confounding variables are alsa highly correlated to each other. A selection based on adjusted r squares of fitted models gave the best observable as common logarithm of the `disp` variable, the displacement in cubic inches. First, a model with two lines, same slope revealed only a difference of **0.326** miles per gallon and p value is **0.79**. Second, a model with two lines different slope obtained. This model shows **2.72** miles per gallon difference for 100 cu. in. displacement and **-2.27** difference for 350 cu. in. displacements. Both models, we **fail to reject** the null hypothesis that there is no difference between manual transmission and automatic transmission.

The echo of R code is disabled to have a compact report. A copy of R Markdown file that produced this document can be foun at the github repository .

Exploratory Data Analyses

Average MPG by Transmission Type

Average `MPG` may be obtained simply by fitting a model with factor variable transmission as observable:

```
## (Intercept)   famManual
##       17.15        7.24
```

Therefore, cars with manual transmission make, on average, 7.24 more miles per gallon than cars with automatic transmission. One can say that manual transmissions are better than automatic ones considering `MPG`. Although there is a variability in `MPG` by transmission type, one cannot conclude that the cause of this variability is transmission type. There may be confounding variables. Boxplots in Figure 1 shows `MPG` distributions by transmission types.

Other Factors Correlated with MPG

After investigating `mtcars` data, correlation between MPG and observables wt, weight; disp, displacement; and, hp, horse power are found to be high. Correlation between observables are high also, hence, a multi variable model would not be suitable. Models with one observable are compared below:

```
##           -  log10
## wt   0.7446 0.8038
## disp 0.7090 0.8169
## hp   0.5892 0.7111
```

The model with the observable `log10(disp)` has the highest adjusted r squared.

Regression Models

The common logarithm of displacement, `log10(disp)`, is chosen as observation variable and the transmission type is dummy factor variable. Two adjusted regression models are observed Using these variables: (1) Model with two lines, same slope, and, (2) Model with two lines, different slope.

Model with Two Lines Same Slope

```
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept)  68.1335      5.877 11.5929 2.086e-12
## x           -20.9897      2.407 -8.7210 1.337e-09
## famManual     0.3255      1.232  0.2641 7.936e-01
```

Cars with manual transmissions make, only, 0.326 more miles per gallon on average than cars with automatic transmission. There is no significant difference between manual transmission and automatic transmission.

The plot of the model is shown in Figure 3.

Model with Two Lines Different Slope

```
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept)    57.22      7.810   7.326 5.613e-08
## x             -16.50      3.206  -5.145 1.868e-05
## famManual      21.08     10.426   2.022 5.284e-02
## x:famManual    -9.18      4.582  -2.003 5.491e-02
```

```
##               Manual Automatic Difference
## 100 (cu. in.)  26.95     24.23       2.72
## 350 (cu. in.)  12.98     15.25      -2.27
```

The difference of miles per gallon varies with common logarithm of displacements. Cars with 100 cu. in. displacement and manual transmissions make **2.72** *more* miles per gallon on average than cars with 100 cu. in. displacement and automatic transmission. On the other hand, cars with 350 cu. in. displacement and manual transmissions make **2.27** *less* miles per gallon on average than cars with 350 cu. in. displacement and automatic transmission. However, the p-values of coefficients are greater than 0.05, and, are not significant.

The plot of the model is shown in Figure 3.

Diagnostics

Figure 4 shows diagnostics for regression model two lines with same slope. Conditions of regression are good considering plots.

Figure 5 shows diagnostics for regression model two lines with different slope. Similary, diagnostics plots are in good conditions..

Appendix

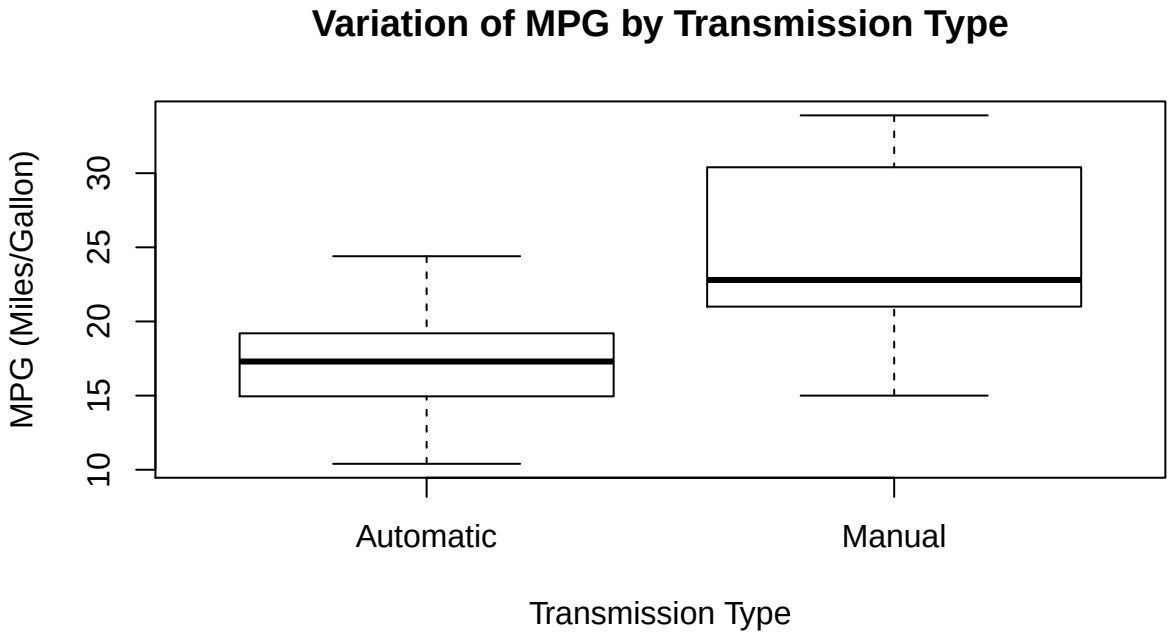

Figure 1. Variation of MPG by Transmission Type

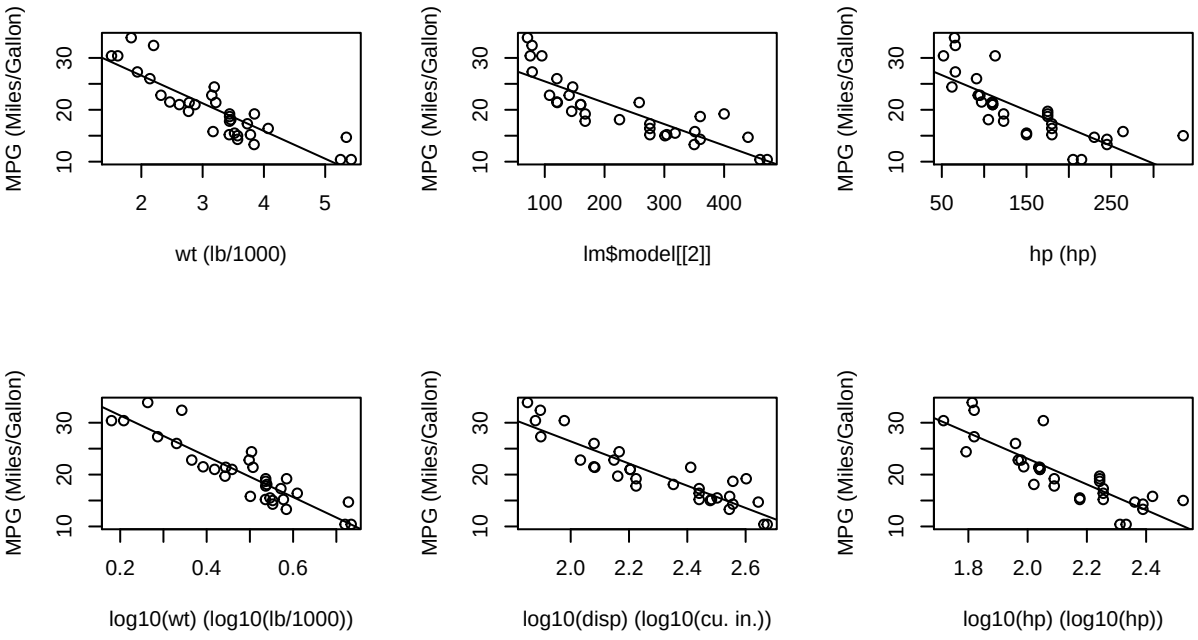

Figure 2. Comparison of Models with Potentially Confounding Observables

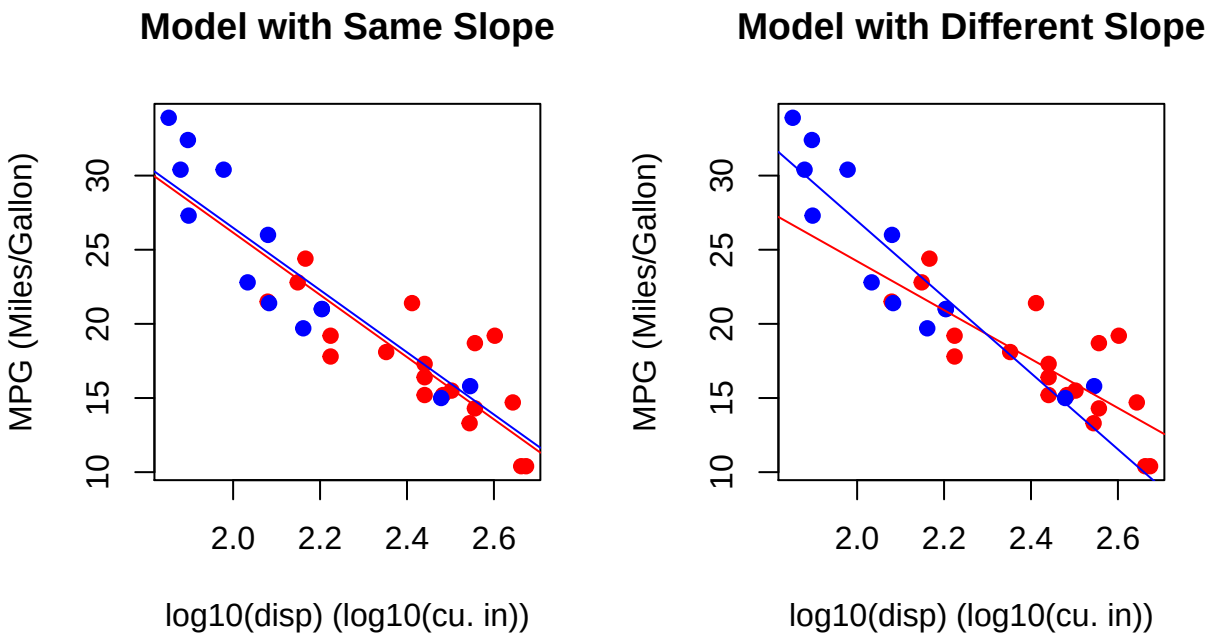

Figure 3. Comparison of Models with Same Slope and Different Slope

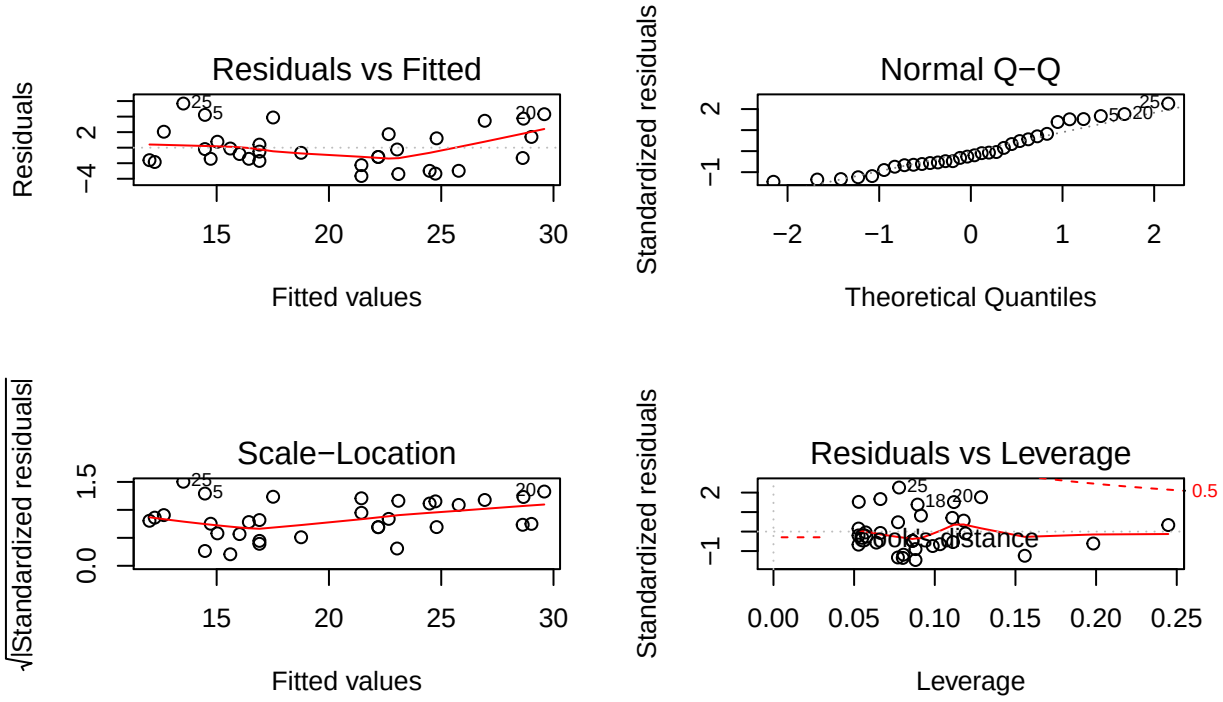

Figure 4. Diagnostics for Model with Same Slope

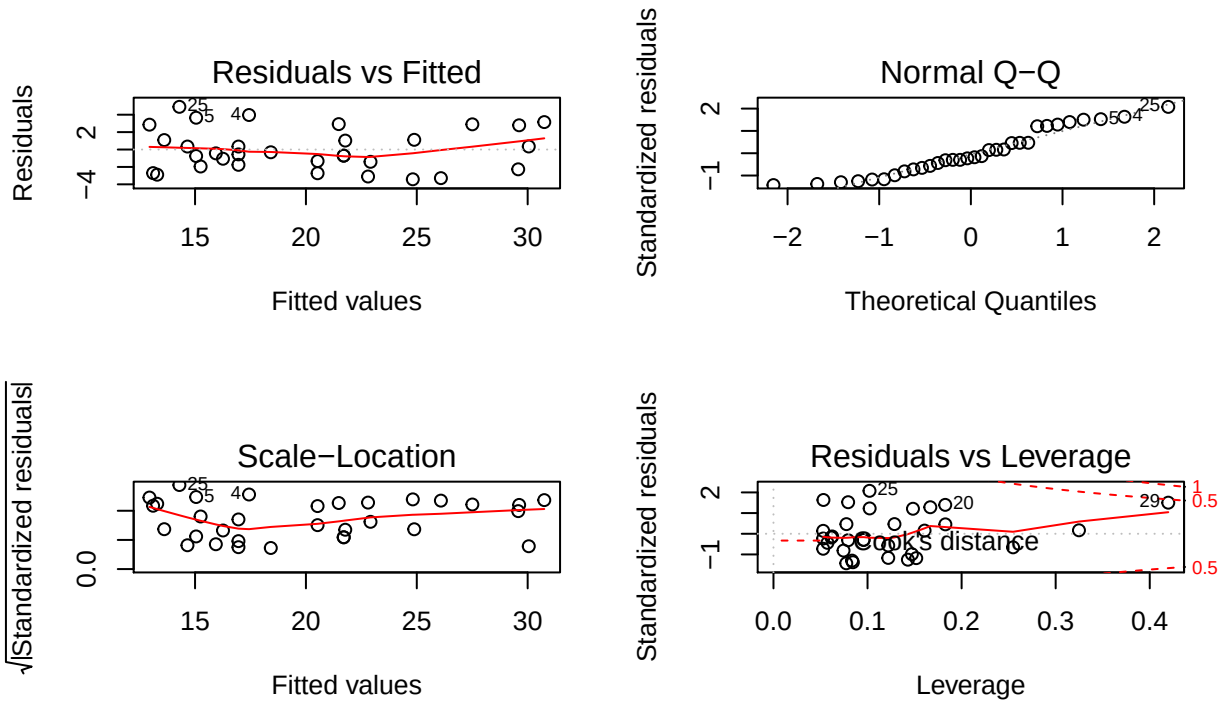

Figure 5. Diagnostics for Model with Different Slope