

Regression Models: Motor Trend Cars Analysis

Mark Culp

April 10, 2017

Executive Summary

This report examines Motor Trend car road tests extracted from the 1974 Motor Trend US magazine. These road tests were conducted on 1973-1974 car models. The tests examined 10 aspects of automobile design and performance for 32 different models. We focused here on the miles per gallon (mpg) performance of automatic versus manual transmissions.

Appendix

```
# Load data set
library(datasets)
data("mtcars")
```

Exploratory Analysis

```
# Examine columns and data types
str(mtcars)
```

```
## 'data.frame':   32 obs. of  11 variables:
## $ mpg : num  21 21 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 ...
## $ cyl : num   6  6  4  6  8  6  8  4  4  6 ...
## $ disp: num  160 160 108 258 360 ...
## $ hp  : num  110 110 93 110 175 105 245 62 95 123 ...
## $ drat: num   3.9 3.9 3.85 3.08 3.15 2.76 3.21 3.69 3.92 3.92 ...
## $ wt  : num   2.62 2.88 2.32 3.21 3.44 ...
## $ qsec: num  16.5 17 18.6 19.4 17 ...
## $ vs  : num   0  0  1  1  0  1  0  1  1  1 ...
## $ am  : num   1  1  1  0  0  0  0  0  0  0 ...
## $ gear: num   4  4  4  3  3  3  3  4  4  4 ...
## $ carb: num   4  4  1  1  2  1  4  2  2  4 ...
```

```
# Range of MPG values
range(mtcars$mpg)
```

```
## [1] 10.4 33.9
```

```
# Car transmissions:
# -----
# 0 = automatic
# 1 = manual
table(mtcars$am)
```

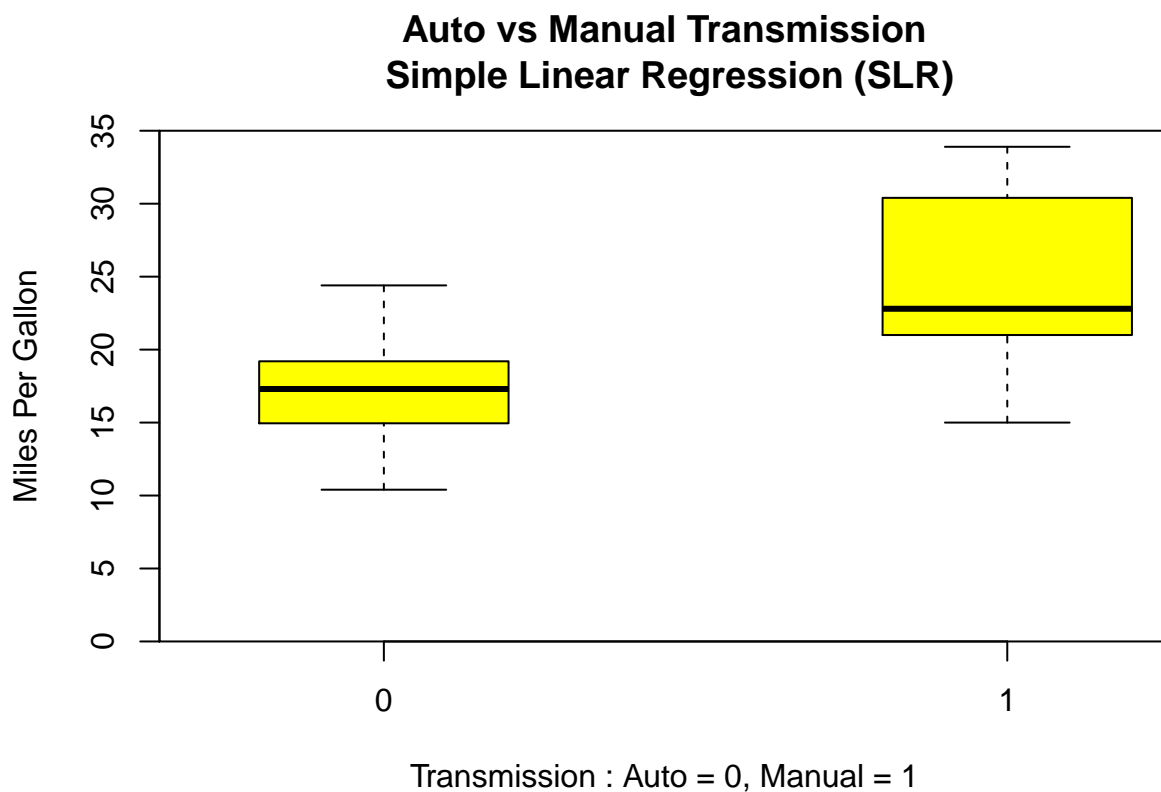
```
##
##  0  1
## 19 13
```

Fitting Multiple Models, Strategy

```
library(knitr)

fit1 <- lm(mpg ~ am, data = mtcars)

boxplot(mpg ~ am, data = mtcars,
        boxwex = 0.4, at = 0:1 - 0.2,
        main = "Auto vs Manual Transmission\n Simple Linear Regression (SLR)",
        col = "yellow",
        xlab = "Transmission : Auto = 0, Manual = 1",
        ylab = "Miles Per Gallon",
        xlim = c(-0.5,1), ylim = c(0,35), yaxs = "i")
```



```
regressor <- c("cyl", "disp", "hp", "wt", "am", "gear", "carb")
correlate <- c(round(cor(mtcars$mpg, mtcars$cyl), 3),
              round(cor(mtcars$mpg, mtcars$disp), 3),
              round(cor(mtcars$mpg, mtcars$hp), 3),
              round(cor(mtcars$mpg, mtcars$wt), 3),
              round(cor(mtcars$mpg, mtcars$am), 3),
              round(cor(mtcars$mpg, mtcars$gear), 3),
              round(cor(mtcars$mpg, mtcars$carb), 3)
              )

mtCor <- data.frame(regressor, correlate)
```

```
kable(t(mtCor[,2]),caption = "Potential Regressors", col.names = t(mtCor[,1]))
```

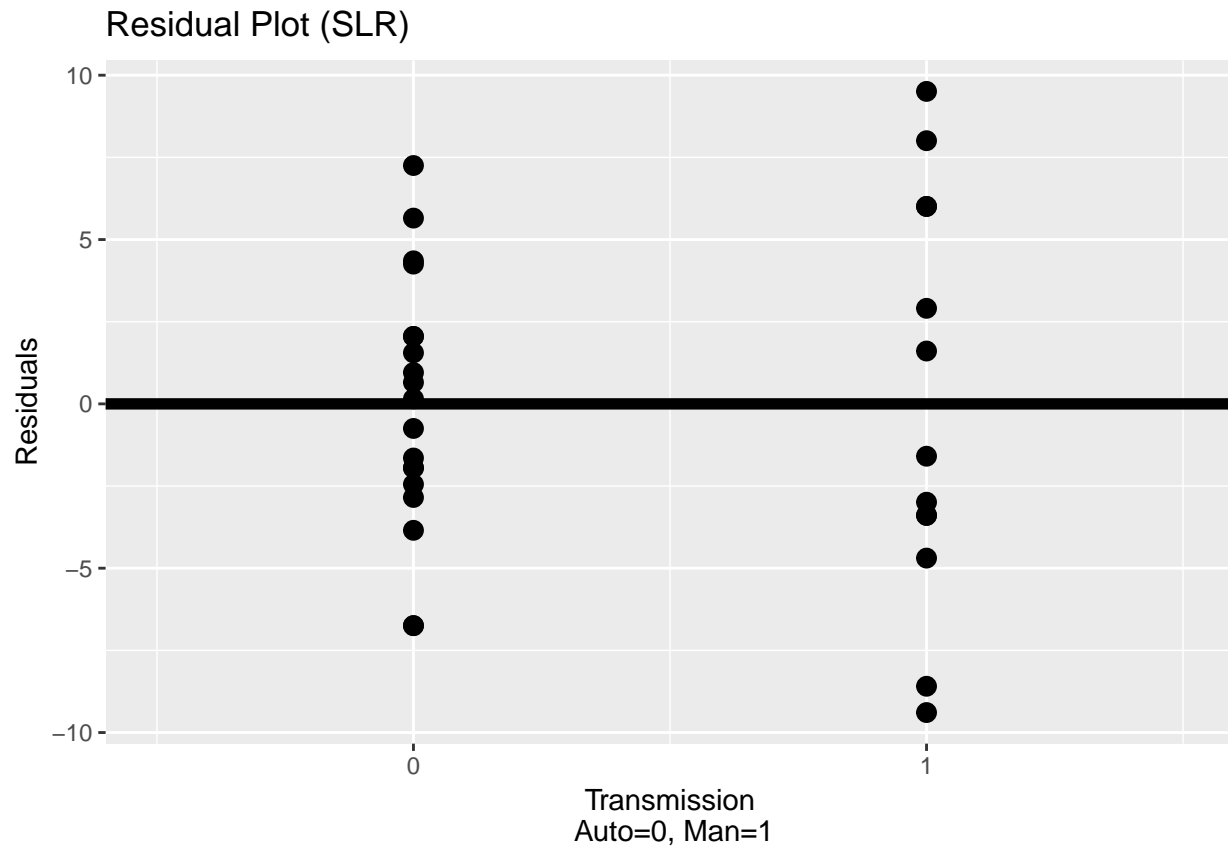
Table 1: Potential Regressors

cyl	disp	hp	wt	am	gear	carb
-0.852	-0.848	-0.776	-0.868	0.6	0.48	-0.551

Residual Plot and Diagnostics

```
library(ggplot2)

g <- ggplot(data.frame(x = mtcars$am, y = resid(fit1)), aes(x=x, y=y))
g <- g + ggtitle("Residual Plot (SLR)")
g <- g + xlab("Transmission\n Auto=0, Man=1")
g <- g + ylab("Residuals")
g <- g + geom_point(size = 3, colour = "black")
g <- g + geom_hline(yintercept = 0, size = 2)
g <- g + scale_x_continuous(breaks = 0:1, limits = c(-0.5,1.5))
g
```



Conclusions : Quantified Uncertainty

Cars with manual transmissions have better gas mileage than cars with automatic transmissions. This can be quantified by ...