

# Regression Models Project 1

*Dann Hekman*

*Saturday, March 07, 2015*

## Executive Summary

There is a common believe that cars with a manual transmission are more fuel-efficient than automatics. Using 1974 [data](#) from *Motor Trends* we can analyze this claim using regression analysis. Based on the analysis presented below, although manual transmission cars get on average 7.24 more miles per gallon than automatic cars, we can confidently say that this is because automatic transmission vehicles weigh more, and therefore have lower gas mileage. We do not believe that manual transmission vehicles have a better gas mileage, once you account for the weight of the vehicle.

## Research Question

We start our analysis by examining the claim that manual transmission cars have better gas mileage than automatic cars. We can load the data and do some basic analysis in R on this dataset.

```
data(mtcars)
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 ...
```

```
summary(mtcars)
```

```
##           mpg           cyl           disp           hp
## Min.      :10.40   Min.      :4.000   Min.      : 71.1   Min.      : 52.0
## 1st Qu.:15.43   1st Qu.:4.000   1st Qu.:120.8   1st Qu.: 96.5
## Median :19.20   Median :6.000   Median :196.3   Median :123.0
## Mean      :20.09   Mean      :6.188   Mean      :230.7   Mean      :146.7
## 3rd Qu.:22.80   3rd Qu.:8.000   3rd Qu.:326.0   3rd Qu.:180.0
## Max.      :33.90   Max.      :8.000   Max.      :472.0   Max.      :335.0
##           drat           wt           qsec           vs
## Min.      :2.760   Min.      :1.513   Min.      :14.50   Min.      :0.0000
## 1st Qu.:3.080   1st Qu.:2.581   1st Qu.:16.89   1st Qu.:0.0000
## Median :3.695   Median :3.325   Median :17.71   Median :0.0000
## Mean      :3.597   Mean      :3.217   Mean      :17.85   Mean      :0.4375
## 3rd Qu.:3.920   3rd Qu.:3.610   3rd Qu.:18.90   3rd Qu.:1.0000
```

```
## Max.      :4.930   Max.      :5.424   Max.      :22.90   Max.      :1.0000
##          am          gear          carb
## Min.      :0.0000   Min.      :3.000   Min.      :1.000
## 1st Qu.:0.0000   1st Qu.:3.000   1st Qu.:2.000
## Median :0.0000   Median :4.000   Median :2.000
## Mean    :0.4062   Mean    :3.688   Mean    :2.812
## 3rd Qu.:1.0000   3rd Qu.:4.000   3rd Qu.:4.000
## Max.    :1.0000   Max.    :5.000   Max.    :8.000
```

Specifically, our variables of interest are mpg and am, which is coded to 0 for automatic transmissions, and 1 for manual transmission. We can regress

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

where  $X_i$  is the *as* variable in the mtcars dataset. Simple Least Squares regression will give the average mpg for automatics ( $X = 0$ ) and  $\beta_1$  will give the added average miles per gallon of a manual transmission. Looking at just the transmission, we can see that automatic transmission vehicles have an average of 17.15 miles per gallon, and manuals get an additional 7.24 miles per gallon on average (24.39 total).

```
model=lm(mpg ~ am, data=mtcars)
summary(model)$coef
```

```
##              Estimate Std. Error  t value    Pr(>|t|)
## (Intercept) 17.147368   1.124603 15.247492 1.133983e-15
## am          7.244939   1.764422  4.106127 2.850207e-04
```

## Correction for Car Weight

Before jumping to the conclusion that manual transmissions are more fuel efficient, we should try to control for other variables in the dataset. Specifically, let's look at the number of cylinders, horsepower, and vehicle weight.

$$Y_i = \beta_0 + \beta_1 X_i + \sum \beta_j X_{ij} + \epsilon_i$$

where  $\beta_1$  is the coefficient for the dummy variable *as* and the sum of  $\beta_j$  and  $X_j$  includes the *cyl*, *hp*, and *wt* variables from the data frame.

```
model=lm(mpg ~ am + cyl + hp + wt, data=mtcars)
summary(model)$coef
```

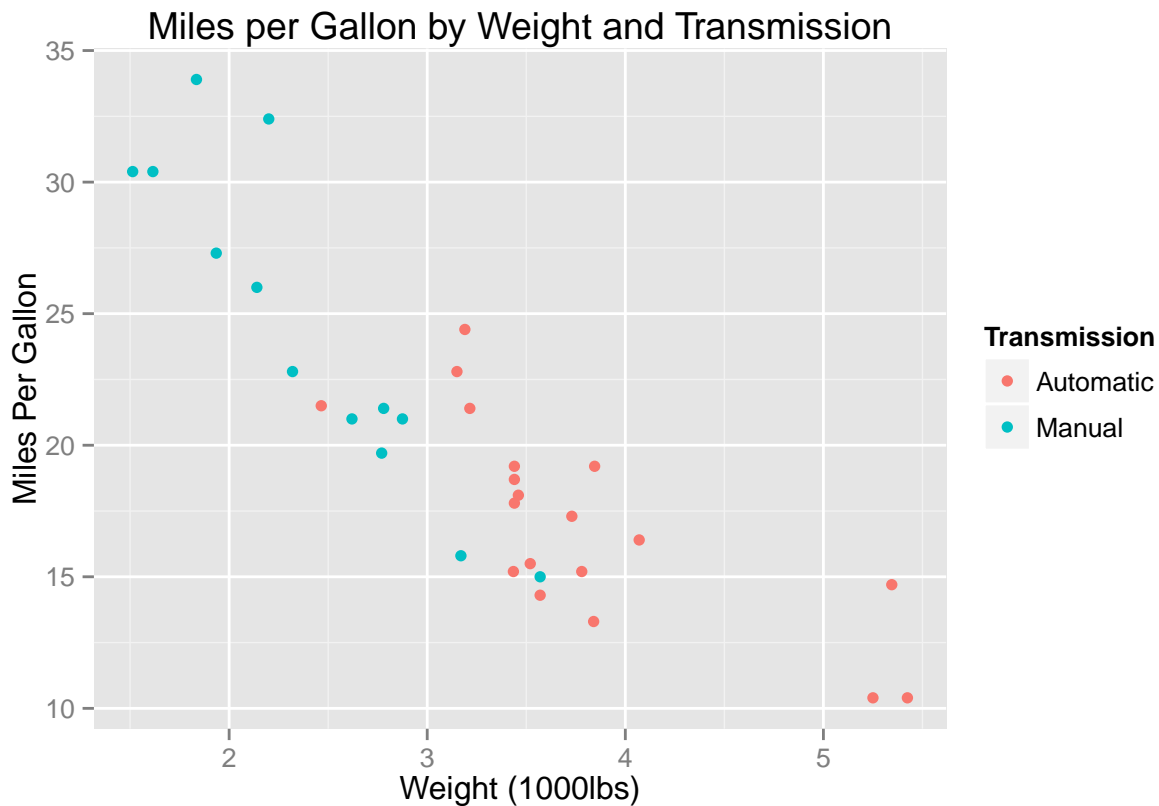
```
##              Estimate Std. Error  t value    Pr(>|t|)
## (Intercept) 36.14653575 3.10478079 11.642218 4.944804e-12
## am          1.47804771 1.44114927  1.025603 3.141799e-01
## cyl        -0.74515702 0.58278741 -1.278609 2.119166e-01
## hp         -0.02495106 0.01364614 -1.828433 7.855337e-02
## wt         -2.60648071 0.91983749 -2.833632 8.603218e-03
```

Controlling for weight, manual vs. automatic ceases to be a statistically significant predictor of gas mileage

## Conclusions

Despite the fact that manual transmission vehicles do, on average, have a better gas mileage, when controlling for the number of cylinders, horsepower, and most importantly vehicle weight, this difference ceases to be substantive or significant. This spurious relationship between mpg, transmission, is very well illustrated by the following graph.

```
library(ggplot2)
q<-qplot(x=wt,y=mpg,data=mtcars,colour=as.factor(am))
q<-q + scale_color_discrete(name="Transmission",labels=c("Automatic","Manual"))
q<-q + ggtitle("Miles per Gallon by Weight and Transmission")
q<-q + xlab("Weight (1000lbs)") + ylab("Miles Per Gallon")
q
```



So we can see clearly that automatics tend to be both heavier and have a lower gas mileage, but comparably weighted manual transmission vehicles tend to have comparable gas mileage.