Vehicle Transmission Type Effects on Gas Mileage

Joe Phaneuf

02/10/2015

Executive Summary

This report uses Motor Trend Magazine data from 1974 to determine if there is a measurable difference in gas mileage (mpg) between manual and automatic transmission cars. After fitting a linear model to account for vehicle weight, there was not enough evidence to conclude that manual or automatic transmissions achieve better gas mileage.

Exploratory Data Analysis

This dataset contains miles-per-gallon, number of cylinders, engine displacement, rear axle ratio, weight, qsec, V-or-Straight engine, transmission type (*am*), number of forward gears, and number of carburetors. We'd expect weight to be one of the largest predictors of gas-mileage, so to get our bearings, let's view a plot of gas-mileage against weight, colored by transmission type (0=automatic, 1 = manual)

```
fmpgvwt <- getNewFigureNumber(figures)
figures<-c(figures,list(ggplot(mtcars,aes(x=wt,y=mpg,colour=factor(am)))+
        geom_point()+scale_colour_brewer(palette="Set1")+
        geom_smooth(method="lm",fill=NA)+xlab("weight (x1000lbs)")+
        ggtitle(paste0(fmpgvwt,"MPG vs. Weight, Colored by Transmission Type"))))
```

See Fig 1. in the appendix.

Model Selection

The simplest possible model for this problem is fitting gas-mileage on transmission type:

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

However, the R^2 for this model is quite low, indicating that transmission type alone is not sufficient to predict gas mileage. Fig 1. showed that weight is a very strong predictor for gas mileage, and the varying slopes for transmission type indicate an interaction term:

```
mpgModel <- lm(mpg~wt*am,data=mtcars)
summary(mpgModel)
```

```
##
## Call:
## lm(formula = mpg ~ wt * am, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6004 -1.5446 -0.5325  0.9012  6.0909
##
## Coefficients:
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  31.4161     3.0201  10.402 4.00e-11 ***
## wt           -3.7859     0.7856  -4.819 4.55e-05 ***
## am           14.8784     4.2640   3.489  0.00162 **
## wt:am        -5.2984     1.4447  -3.667  0.00102 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.591 on 28 degrees of freedom
## Multiple R-squared:  0.833,  Adjusted R-squared:  0.8151
## F-statistic: 46.57 on 3 and 28 DF,  p-value: 5.209e-11
```

```
fresq <- getNewFigureNumber(figures)
figures <- c(figures,list(qplot(sample=residuals(mpgModel),stat="qq",
            main=paste0(fresq,"Qunatile-Quantile Plot of mpgModel Residuals"))))
```

All of the terms in this model are significant ($p<0.05$), and the R^2 indicates that these terms explain most of the gas-mileage variability. Introducing other variables in the dataset into this model has little effect on increasing the model fit, so this model will be used for the final analysis.

Analysis

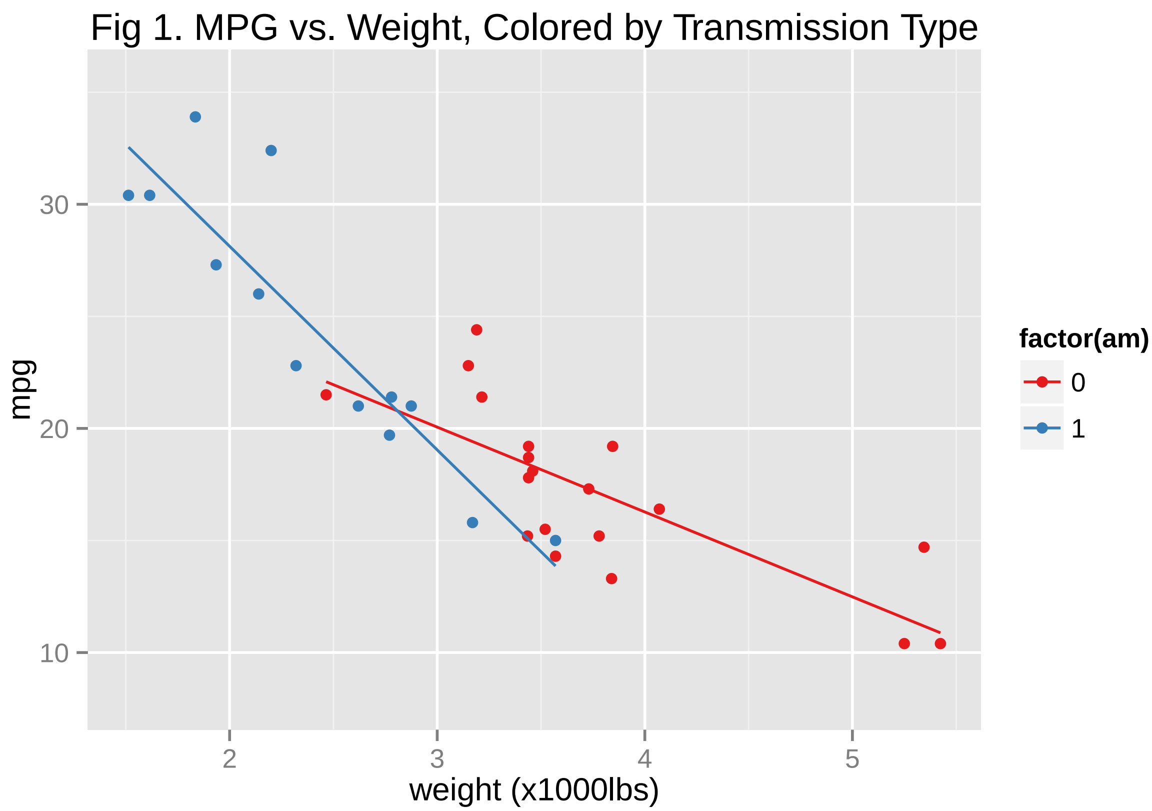
```
#get coefficients for 95% confidence interval:
sumCoef <- summary(mpgModel)$coefficients
amSlope95Int <- sumCoef[3,1] + c(-1,0,1) * qt(.975, df = mpgModel$df) * sumCoef[3, 2]
amWtSlope95Int <-sumCoef[4,1] + c(-1,0,1) * qt(.975, df = mpgModel$df) * sumCoef[4, 2]
mockCars<- data.frame(wt=0:7)
pI <- function(wt,i){amSlope95Int[i]+wt*amWtSlope95Int[i]} #predictInterval
mockCars<-cbind(mockCars,lwr=pI(mockCars$wt,1),mpg=pI(mockCars$wt,2),
                upr=pI(mockCars$wt,3))
fmpgdiff <- getNewFigureNumber(figures)
figures<-c(figures,list(ggplot(mockCars,aes(x=wt,y=mpg,colour="mpg"))+
        geom_line()+geom_line(aes(x=wt,y=lwr,colour="lower"))+
        geom_line(aes(x=wt,y=upr,colour="upper"))+
        xlab("weight (x1000lbs)")+
        ggtitle(paste0(fmpgdiff,"Mpg Increase for Manual vs. Weight: 95% Interval"))))
```

For model diagnostics, refer to figure Fig 2. for a quantile-quantile plot of the model's residuals. The qq plot exhibits some non-linearity, indicating some skew in the distribution of the residuals.

The interaction term indicates that the mileage difference due to transmission type is dependant on weight. Using the model coefficients, we can say with 95% confidence that gas mileage for manual tranmission cars falls within these two bounding lines (functions of weight) 6.1439279 -8.2576928 x weight(x1000lbs) and 23.6129171 -2.3390281 x weight(x1000lbs) . See Fig 3. for a plot of manual transmission mpg difference as a function of weight, including 95% confidence bounds (upper and lower). The 95% interval contains zero for almost the entire range of vehicle weight, so we **cannot claim a gas mileage difference between automatic and manual tranmission vehicles**, after controlling for vehicle weight. Note that there is a small weight window where the 95% interval does not contain zero, however the vehicles in this dataset in that weight window are exclusively manual transmissions (see Fig 1.).

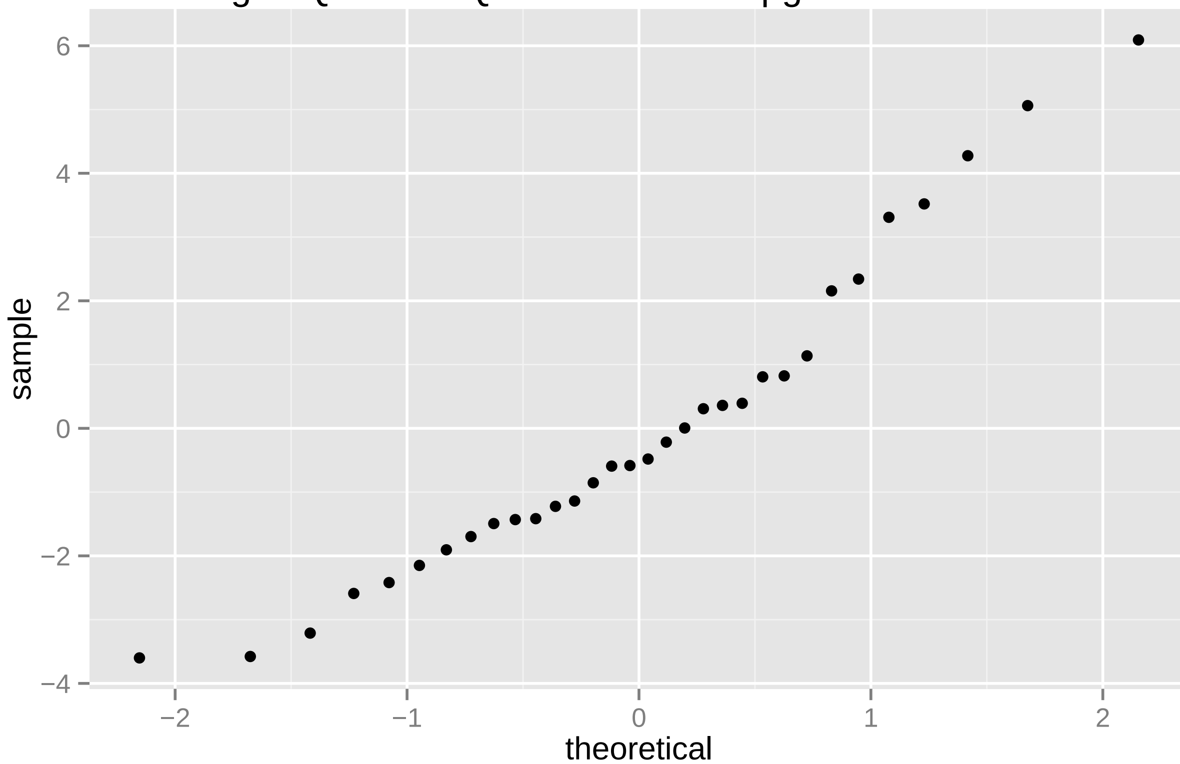
Appendix

```r
for(f in figures){
  print(f)
  print("")
}
```



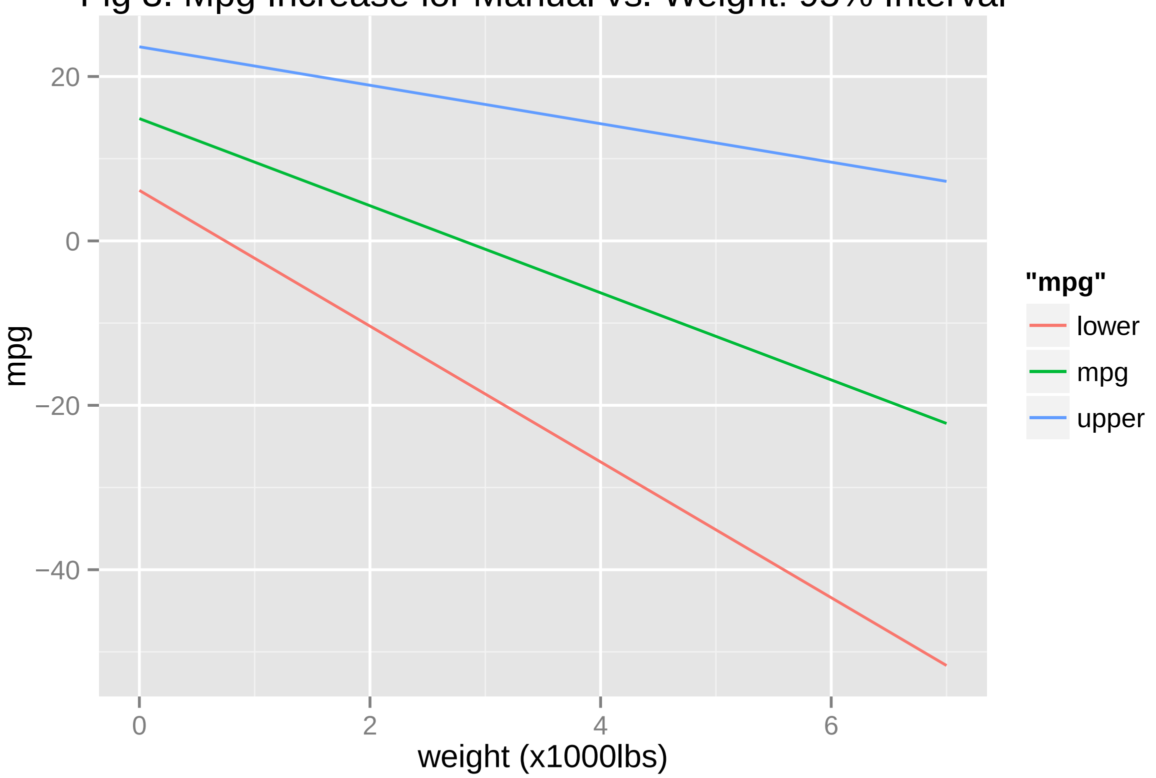

```
## [1] ""
```

Fig 2. Qunatile−Quantile Plot of mpgModel Residuals




```
## [1] ""
```

Fig 3. Mpg Increase for Manual vs. Weight: 95% Interval



```
## [1] ""
```