Influence from Transmission Type on MPG

Shang Ju

March 12, 2017

Executive Summary

Mpg (miles-per-gallon) for automatic transmission cars tends to be lower than that of manual transmission cars. The average difference is 5.277 mpg.

This result is robust by controlling effect from horsepower.

However, the causal effect should be drawn with caution because other factors that are highly correlated to transmission type such as number of gears, the rear axle ratio, and weight, could affect the accuracy of the conclusion.

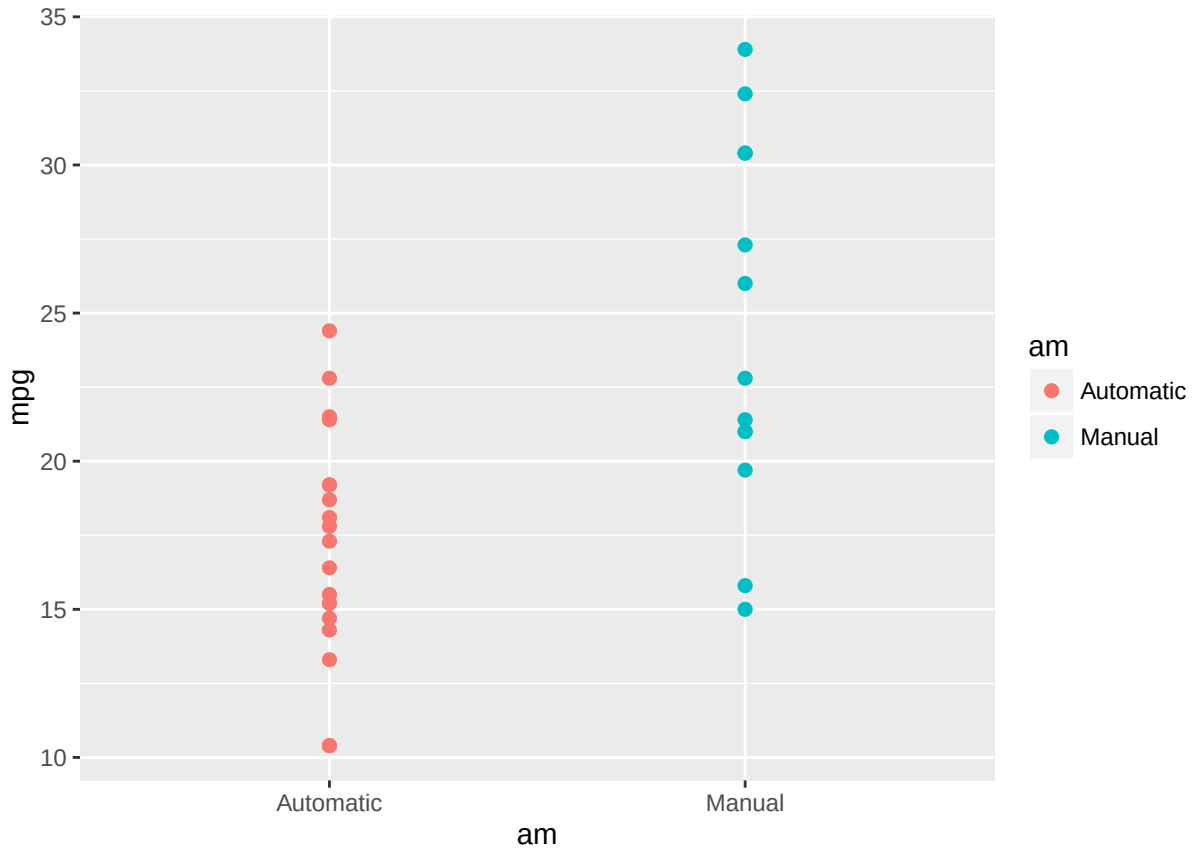
Background

Motor Trend, a magazine about the automobile industry, wants to know: 1. Is an automatic or manual transmission better for MPG? 2. Quantify the MPG difference between automatic and manual transmissions.

Dataset: mtcars in R

Exploratory Data Analyses

```
data(mtcars)
library(ggplot2)
mtcars$am <- factor(mtcars$am,levels = c(0,1), labels=c('Automatic','Manual'))
ggplot(data=mtcars, aes(x=am, y=mpg, color = am)) +
    geom_point(size = 2)
```



From the graph we see that automatic transmission cars seem to have lower miles-per-gallon than manual transmission cars.

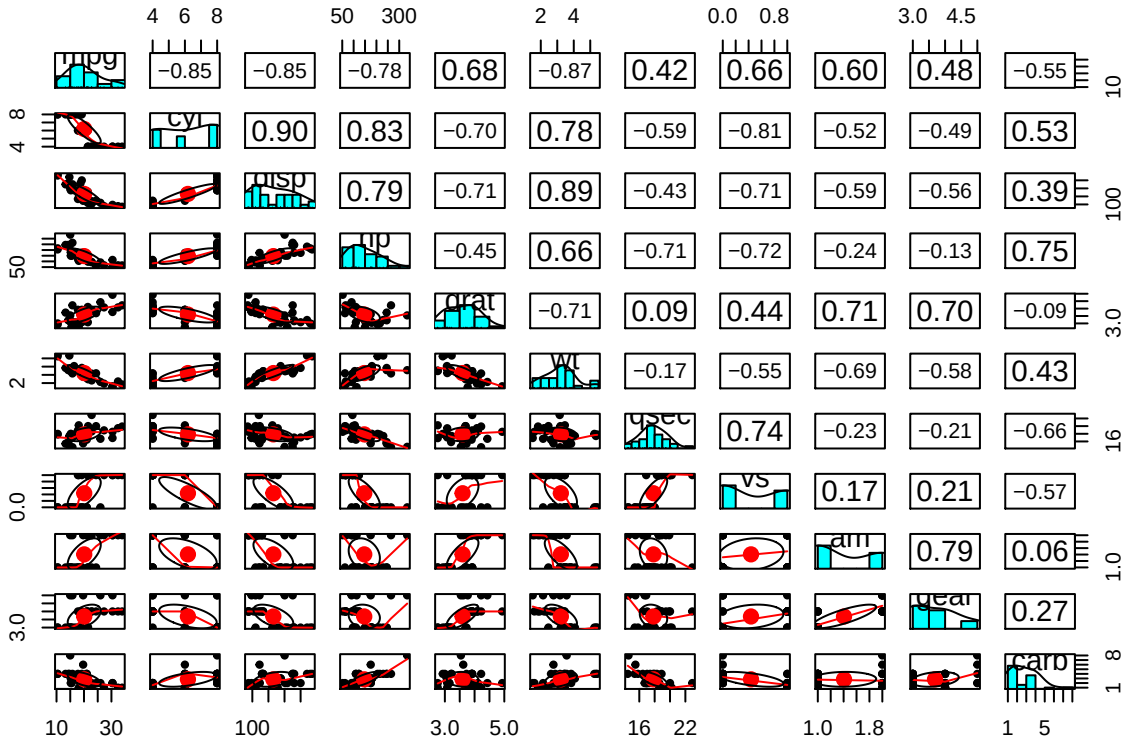
What about correlations between different factors?

```
library(psych)
```

```
##
## Attaching package: 'psych'

## The following objects are masked from 'package:ggplot2':
##
##     %+%, alpha
```

```
pairs.panels(mtcars)
```



}

From the correlation chart we can see that both mpg and transmission type are correlated to some other variables. So we need to control the influence from other variables before drawing conclusions between mpg and transmission type.

Are the questions answerable?

From the correlation chart we see that transmission type is highly correlated with number of gears, the rear axle ratio, and weight. So the questions might not be answered with absolute confidence until we collect more data with less correlation between these variables.

For this task, we'll do our best to discover insights with existing data. But we should keep in mind of the caveat of this analysis.

Use regression models to answer the questions

Model 1: mpg ~ am

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

Model 2: mpg ~ .

```
fit2 <- lm(mpg~.,data=mtcars)
anova(fit1,fit2)
```

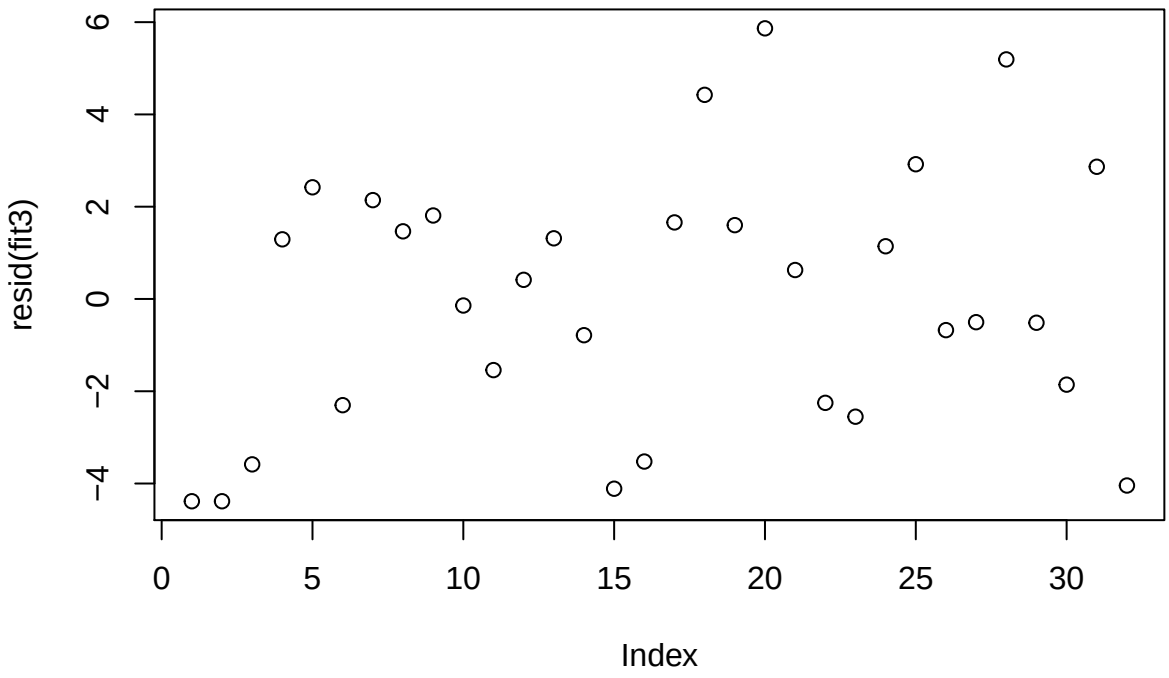
```
## Analysis of Variance Table
##
## Model 1: mpg ~ am
## Model 2: mpg ~ cyl + disp + hp + drat + wt + qsec + vs + am + gear + carb
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1     30 720.90
## 2     21 147.49  9     573.4 9.0711 1.779e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model 3: mpg ~ am + hp

```
fit3 <- lm(mpg~ am + hp,data=mtcars)
anova(fit2,fit3)
```

```
## Analysis of Variance Table
##
## Model 1: mpg ~ cyl + disp + hp + drat + wt + qsec + vs + am + gear + carb
## Model 2: mpg ~ am + hp
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     21 147.49
## 2     29 245.44 -8   -97.945 1.7432 0.1466
```

```
plot(resid(fit3))
```



Anova test result suggests that model 2 & 3 are not significantly different. Plus, this is no strong multicolinearily in model 3. Residual plot looks normal.

Conclusion

Based on result from model 3, mpg of automatic transmission cars tends to be lower than that of manual transmission cars. With horsepower hold constant, mpg of automatic transmission cars is 5.3 lower than that of manual transmission cars.

However, the causal effect should be concluded with caution given that other factors, which are highly correlated to transmission type: number of gears, the rear axle ratio, and weight, could affect the accuracy of the conclusion.