Trans Type vs. MPG

Phillip Chaffee

September 24, 2015

Abstract

This study attempts to relate a vehicles transmission type, automatic or manual, to the vehicles mpg using the mtcars data. The method used will be linear regression. It will evaluate which transmission type causes greater MPG, and will attempt to quantify the MPG difference between automatic and manual transmissions.

Methods

First a regression line is fit to the data. See Fig. 1. AM at 0 marks an automatic transmission. AM at 1 marks a manual transmission.

The graph indicates that only observing the two variables in question seems to suggest an increase in mpg in manual transmission cars. The average automatic car seems to get 17.147 mpg, and manual cars see a 7.245 mpg increase from an automatic on average.

There are two foreseeable problems with this fit. First, there is an obvious lack of data between the two groups of points. Second, it is very possible that the difference in mpg between automatic and manual cars is explainable by their difference in weight.

Not much can be done for the lack of data. Let's look at the difference in weight between automatic and manual transmission cars. Fig 2. The boxplot shows that automatic transmission cars are much heavier than manual cars. It is much more likely that this is what is affecting the difference in MPG.

Let's look at a new fit accounting for the difference in weight. Fig. 3. It is apparent in figure three that weight has much more of an effect on MPG than transmission type.

Take a look at the affect of weight on mpg, split by transmission type, and a residual plot of the initial fit. Fig. 4 and Fig. 5.

Results

The data does not support any significant relation between transmission type and MPG. There is too much missing data between the two groups of plots, and the change is MPG is explained by weight, leaving no significant effect caused by transmission type. There is no inference to be performed since the questions cannot be answered.

Appendix

Fig. 1

```
fit <- lm(mpg ~ am, data=mtcars)
summary(fit)$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
```

```
library(ggplot2)
g <- ggplot(mtcars, aes(am, mpg)) + geom_point() + geom_smooth(method="lm")
g
```

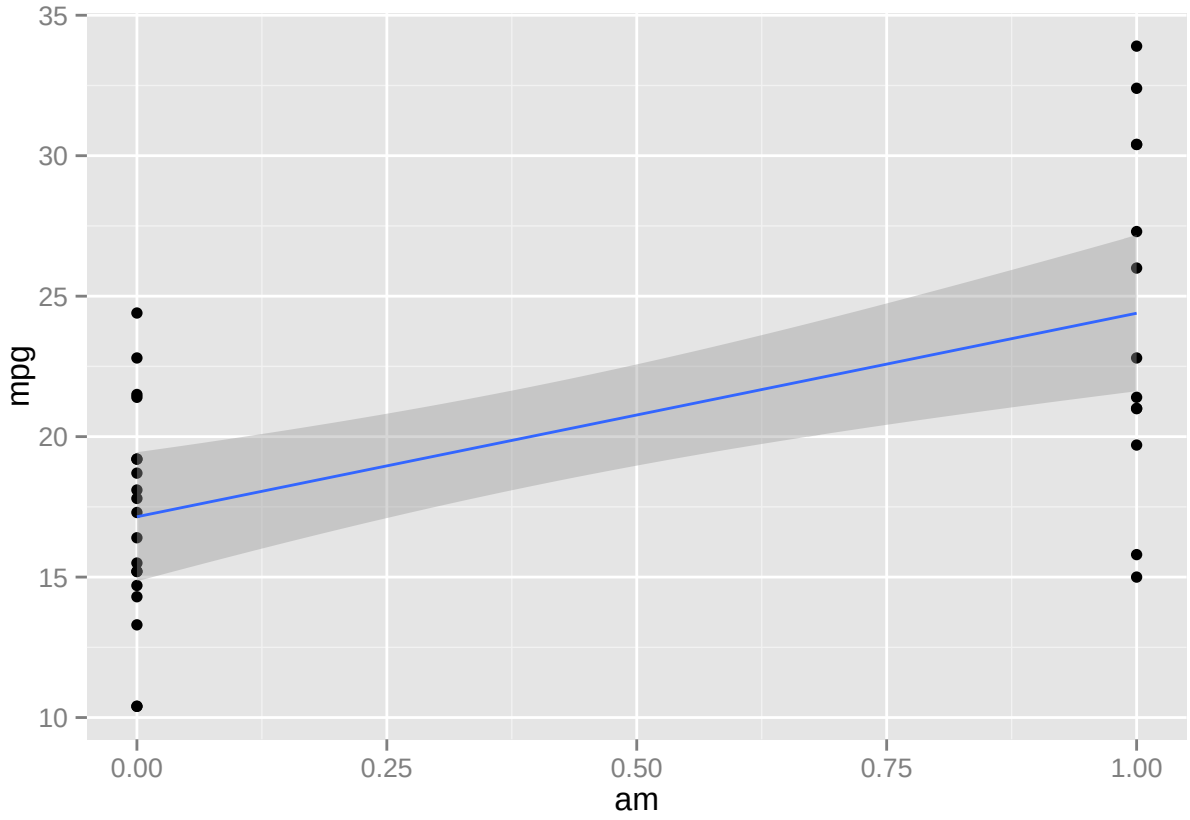


Fig. 2

```
library(ggplot2)
g <- ggplot(mtcars, aes(as.factor(am), wt)) + geom_boxplot()
g
```

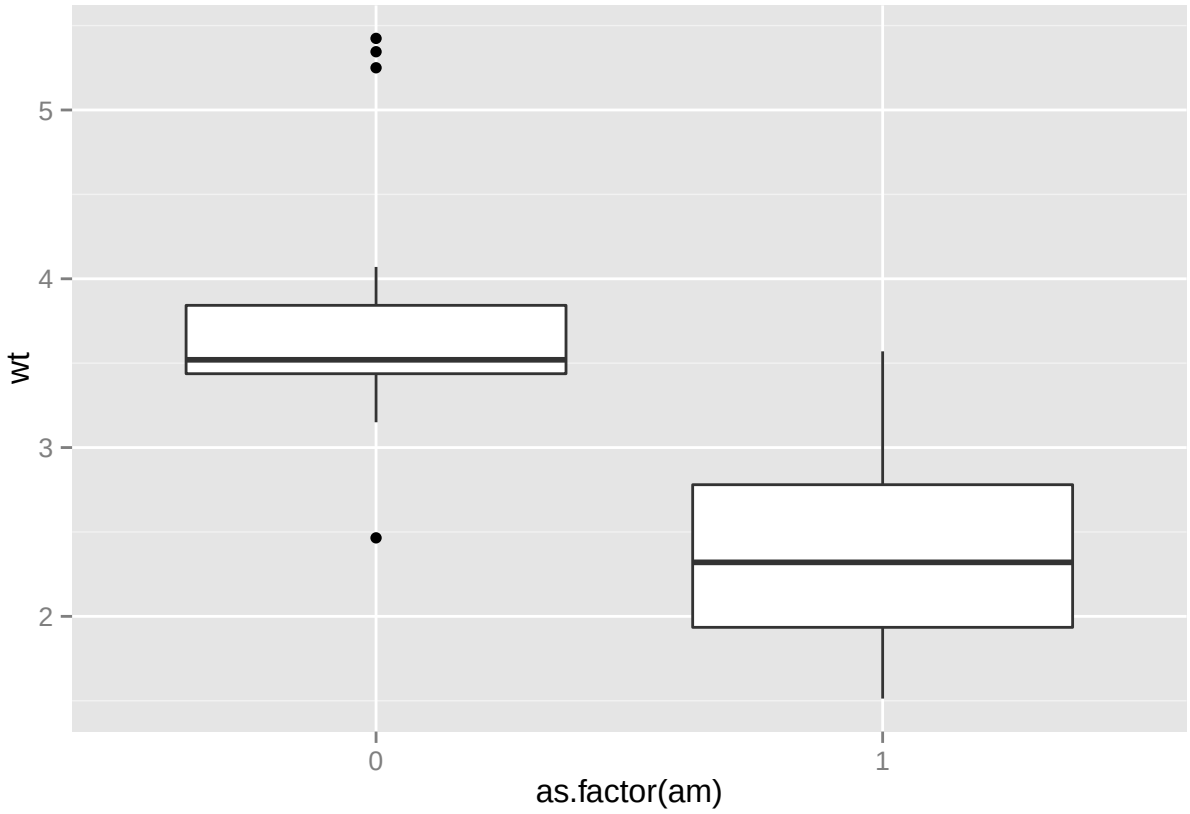


Fig. 3

```
fit <- lm(mpg ~ wt + am, data=mtcars)
summary(fit)$coef
```

```
##                Estimate Std. Error     t value     Pr(>|t|)
## (Intercept) 37.32155131  3.0546385 12.21799285 5.843477e-13
## wt          -5.35281145  0.7882438 -6.79080719 1.867415e-07
## am          -0.02361522  1.5456453 -0.01527855 9.879146e-01
```

Fig. 4

```
library(ggplot2)
g <- ggplot(mtcars, aes(wt, mpg)) + geom_point() + facet_grid(. ~ am) + geom_smooth(method="lm")
g
```

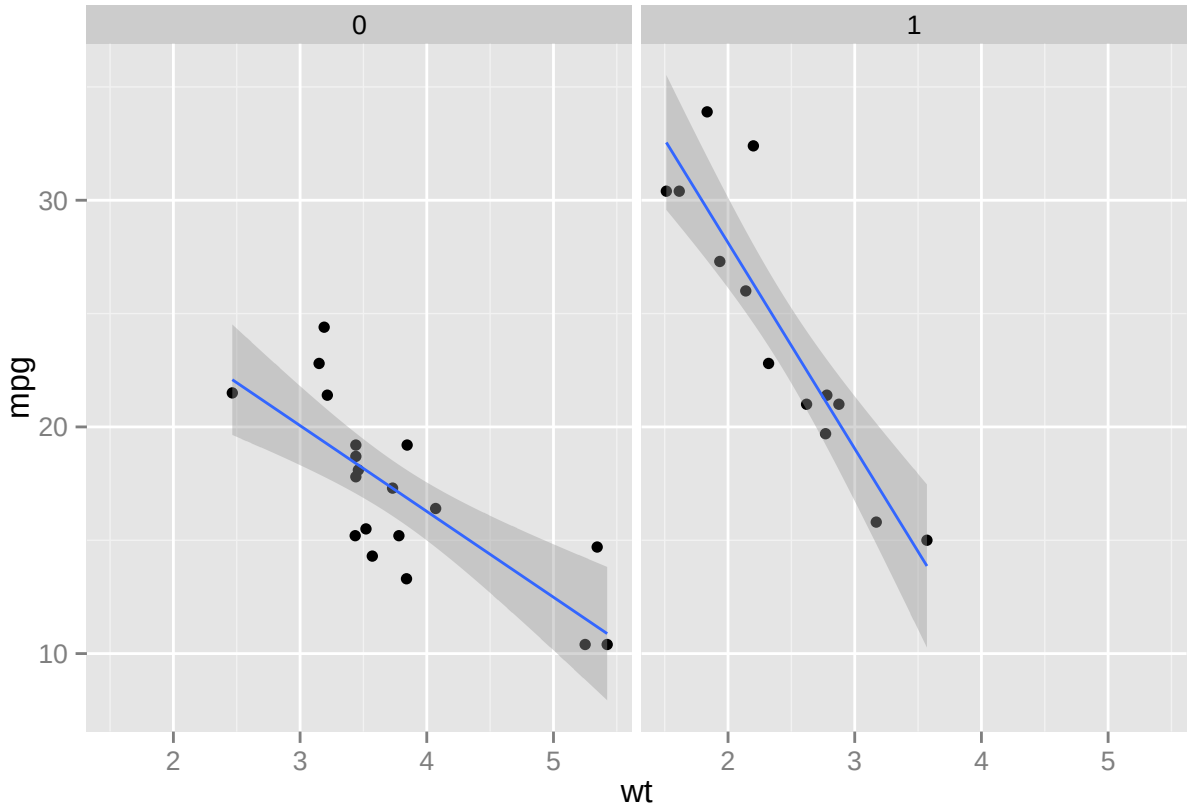
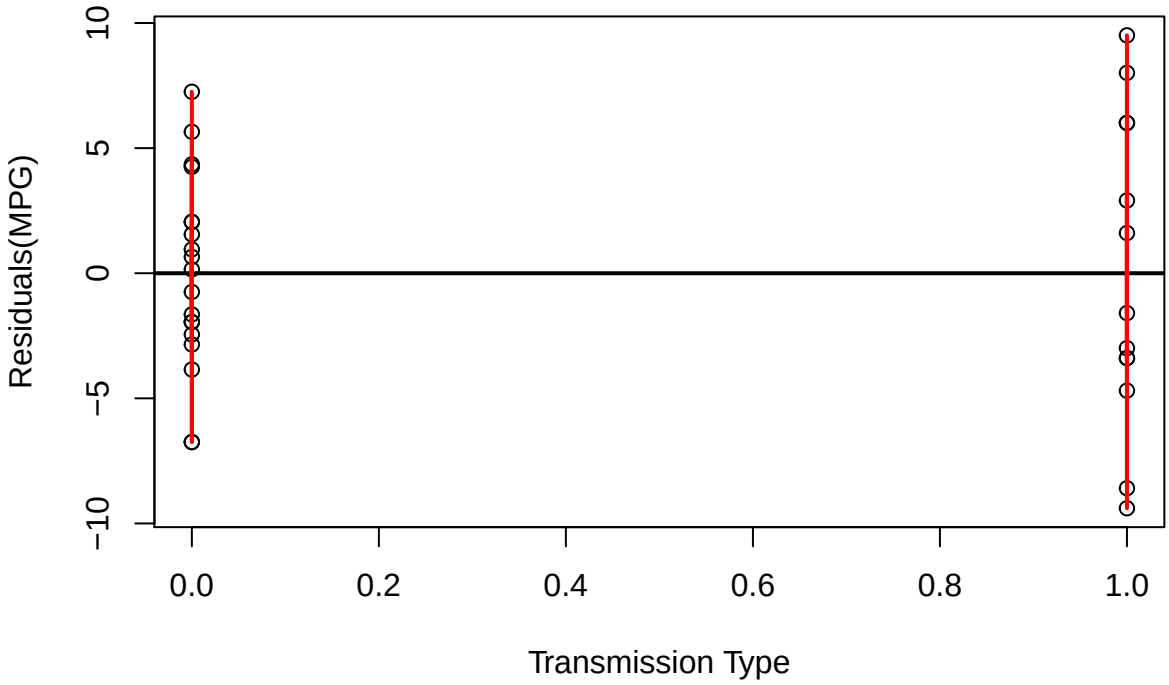


Fig. 5

```
data(mtcars)
fit <- lm(mpg ~ am, data=mtcars)
e <- resid(fit)
x <- mtcars$am
plot(x, e, xlab = "Transmission Type", ylab = "Residuals(MPG)")
abline(h=0, lwd=2)
for(i in 1:length(x))
  lines(c(x[i], x[i]), c(e[i], 0), col="red", lwd=2)
```


Residuals(MPG)
10
5
0
−5
−10
0.0
0.2
0.4
0.6
0.8
1.0
Transmission Type