

Fuel efficiency analysis

Saurabh Yadav

Executive summary

In this we assume that we work for a motor trend magazine and we have to look at the dataset of the mtcars and explore the relationship between the few variables and the mpg variable. In this we have to specially give the answers for the following questions

1. "Is an automatic or manual transmission better for MPG"
2. "Quantify the MPG difference between automatic and manual transmissions"

In our analysis we have shown that the manual transmission is better for lighter weight and less displacements.

Loading the data

```
##  
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':  
##  
##     filter, lag
```

```
## The following objects are masked from 'package:base':  
##  
##     intersect, setdiff, setequal, union
```

Adding a column to identify the auto and manual transmission.

```
mtcars$mode<-ifelse(mtcars$am==0,"automatic","manual")
```

Exploratory analysis

1. First taking the simply mean of the mpg of transmissions

```
sumry<-group_by(mtcars,mode)  
meandata<-summarise(sumry,mean(mpg))  
meandata<-data.frame(meandata)  
colnames(meandata)<-c("mode","mean")  
meandata
```

```
##           mode      mean  
## 1 automatic 17.14737  
## 2   manual  24.39231
```

Conclusion from the exploratory analysis

1. By taking the mean we can analyse that the manual transmission is better than the automatic. But to be sure we have to do some regression analysis to check if the manual is always better or it depends on some factors of the data.

Regerssion analysis

T test for the variable mpg

```
t.test(mtcars$mpg[mtcars$am==0],mtcars$mpg[mtcars$am==1],paired = FALSE)
```

```
##
## Welch Two Sample t-test
##
## data: mtcars$mpg[mtcars$am == 0] and mtcars$mpg[mtcars$am == 1]
## t = -3.7671, df = 18.332, p-value = 0.001374
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -11.280194 -3.209684
## sample estimates:
## mean of x mean of y
## 17.14737 24.39231
```

Conclusion: The confidence interval clearly shows that mpg for manual transmission is better than the automamtic because the interval is negative.

Fitting the different models

1.Fitting the model considering all the variables

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

```
##
## Call:
## lm(formula = mpg ~ ., data = mtcars)
##
## Coefficients:
## (Intercept)      cyl      disp      hp      drat
## 12.30337      -0.11144      0.01334     -0.02148      0.78711
##      wt      qsec      vs      am      gear
## -3.71530      0.82104      0.31776      2.52023      0.65541
##      carb  modemanual
## -0.19942           NA
```

Conclusion: The model including all the variables does not show a good estimation for transmission. So we are including the differnt variables with interaction factor am(0=automatic,1=manual).

2. Fitting the model considering the wt variable

```
fit2<-lm(mpg~wt*factor(mode),data = mtcars)
fit2$coefficients
```

```
##           (Intercept)           wt  factor(mode)manual
##           31.416055        -3.785908           14.878423
```

```
## wt:factor(mode)manual
## -5.298360
```

Conclusion:

1. For automatic transmission the intercept is 31.4161 and for every 1000 lbs increase there is 3.7859 mpg decrease.
 2. For manual transmission the intercept is 46.28 and for every 1000 lbs increase there is 9.07 mpg decrease mpg.
 3. But according to Plot 2 in the appendix the manual transmission is better for the light weight cars.
- ### 3. Fitting the model considering the disp variable

```
fit3<-lm(mpg~disp*factor(mode),data = mtcars)
fit3$coefficients
```

```
##          (Intercept)          disp  factor(mode)manual
##          25.15706407         -0.02758360           7.70907298
## disp:factor(mode)manual
##          -0.03145482
```

Conclusion:

1. The intercept is 25.1570 for automatic transmission and for increase in displacement the decrease in mpg is 0.027584 and for manual transmission the intercept is 32.85 and for each increase in mpg the decrease is 0.059039.
 2. But according to the plot 3 in appendix the manual transmission gives higher mpg for the displacement less than 245 (cu.in.)
- ### 4. Fitting the model considering the hp variable

```
fit4<-lm(mpg~hp*factor(mode),data = mtcars)
fit4$coefficients
```

```
##          (Intercept)          hp  factor(mode)manual
##          26.6248478696         -0.0591369818           5.2176533777
## hp:factor(mode)manual
##          0.0004028907
```

Conclusion:

1. For automatic transmission the intercept is 26.62 and for each increase in horsepower there is 0.0591 decrease in mpg.
 2. For manual transmission the intercept is 31.83 and for each increase in horsepower there is .058 decrease in mpg.
 3. So the manual transmission is better when we include the hp variable.
- ### 5. Fitting the model considering the cyl variable

```
fit5<-lm(mpg~cyl*factor(mode),data = mtcars)
fit5$coefficients
```

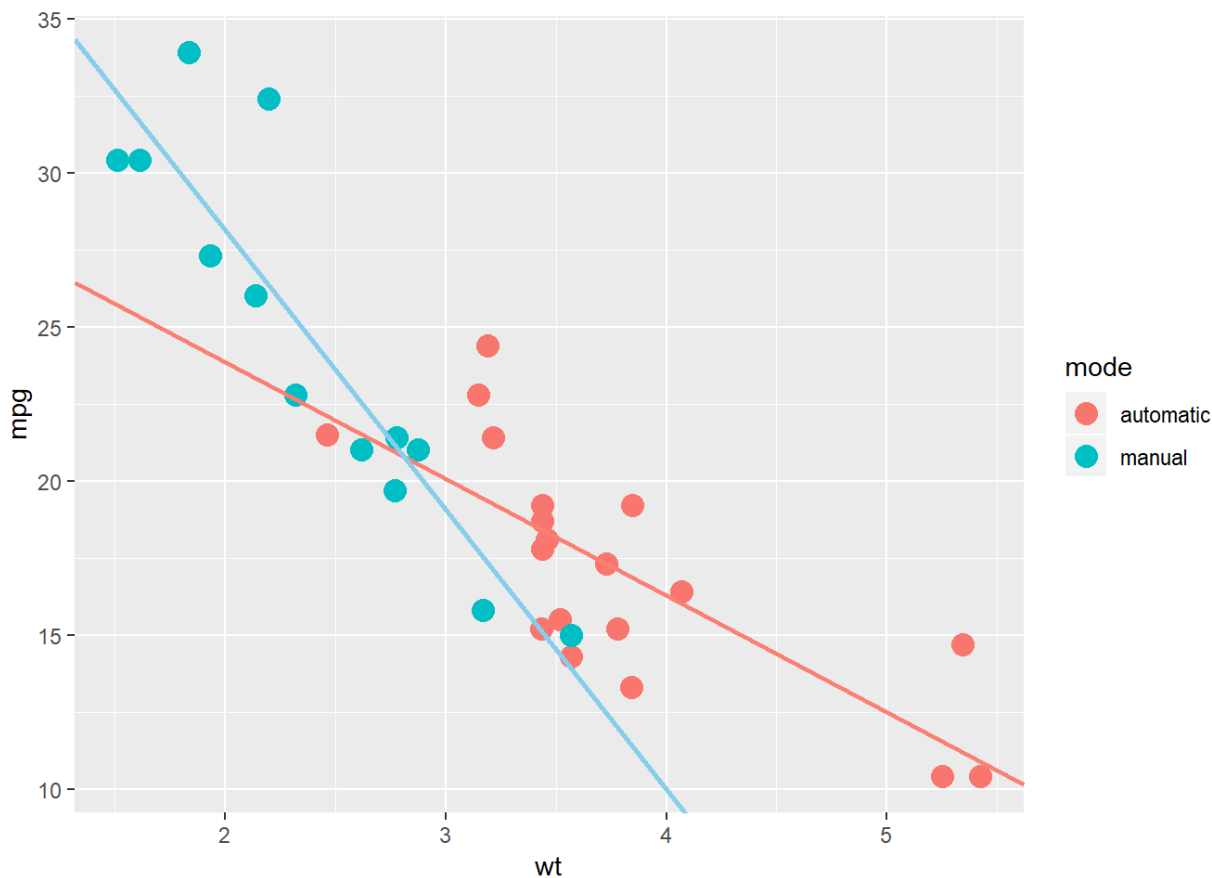
```
##          (Intercept)          cyl  factor(mode)manual
##          30.873529         -1.975735           10.175407
## cyl:factor(mode)manual
##          -1.305116
```

Conclusion: For auto transmission there is 1.9757 decrease in mpg when cylinder increases and manual transmission there is 2.28 decrease in mpg when cylinder increases. So the manual transmission is better.

Appendix

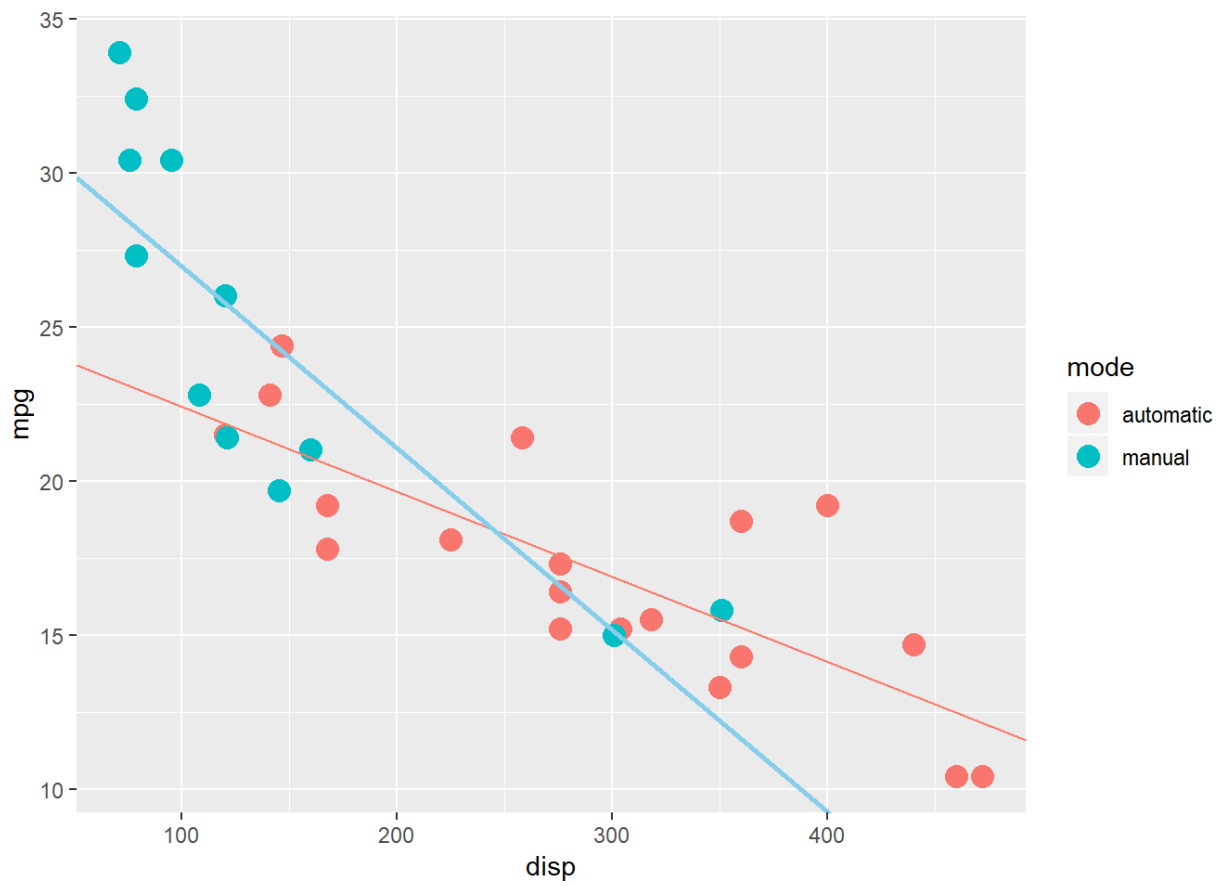
Plot 1 (fitting the model with variable wt)

```
ggplot(data = mtcars, aes(x=wt, y=mpg, color=mode))+geom_point(size=4)+geom_abline(slope = -3.786, intercept = 31.416, color="salmon", size=1)+geom_abline(slope = -9.08, intercept = 46.28, color="skyblue", size=1)
```



Plot 2 (fitting the model with variable disp)

```
par(mfrow = c(2, 2))
ggplot(data = mtcars, aes(x=disp, y=mpg, color=mode))+geom_point(size=4)+geom_abline(slope = -0.027584, intercept = 25.157, color="salmon")+geom_abline(slope = -0.0590, intercept = 32.85, color="skyblue", size=1)
```



Plot 3 (Residuals and diagnostics)

```
par(mfrow = c(2, 2))  
plot(fit1)
```

