Motor Trend Project

Dhiraj Deshmukh

October 3, 2017

Summary

This report explores the relationship between manual and automatic transmission on the fuel efficiency. For this analysis, a dataset from Motor Trend US Magazine is used. Following questions are answered through the analysis.

1. Is an automatic or manual transmission better for MPG?
2. How different is the MPG between automatic and manual transmissions?

The result obtained for the data shows that the manual transmission presents higher average MPG value than the automatic models.

Load Data

```
library(ggplot2)
data(mtcars)
head(mtcars)
```

```
##                   mpg cyl disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
## Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
## Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
## Valiant           18.1   6  225 105 2.76 3.460 20.22  1  0    3    1
```

Exploratory Analysis

```
mtcars$cyl  <- factor(mtcars$cyl)
mtcars$vs   <- factor(mtcars$vs)
mtcars$gear <- factor(mtcars$gear)
mtcars$carb <- factor(mtcars$carb)
```

```
mtcars$am   <- factor(mtcars$am,labels=c("Automatic","Manual"))
```

Appendix - Plot 1 below shows that Automatic Transmisson has lower MPG than the Manual transmission.
Regression Analysis

```
aggregate(mpg~am, data = mtcars, mean)
```

```
##          am      mpg
## 1 Automatic 17.14737
## 2    Manual 24.39231
```

Lets do the t test now

```
automatic <- mtcars[mtcars$am == "Automatic",]
manual <- mtcars[mtcars$am == "Manual",]
t.test(automatic$mpg, manual$mpg)
```

```
##
##  Welch Two Sample t-test
##
## data:  automatic$mpg and manual$mpg
## 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
```

Since the p-value is 0.001374, this is very significant difference. Lets quantify this now

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

```
##
## Call:
## lm(formula = mpg ~ am, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
```

```
## -9.3923 -3.0923 -0.2974  3.2439  9.5077
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   17.147      1.125  15.247 1.13e-15 ***
## amManual       7.245      1.764   4.106 0.000285 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.902 on 30 degrees of freedom
## Multiple R-squared:  0.3598, Adjusted R-squared:  0.3385
## F-statistic: 16.86 on 1 and 30 DF,  p-value: 0.000285
```

This shows that automatic mpg is around 17.1 and manual is around 7.2. It also shows that the adjusted R square value is only around .338 and multiple R squared value is around .36, which mean only 36% of the variance can be explained. We will show now the mulivariable regresion model.

Multivariable Regression Model

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

```
##
## Call:
## lm(formula = mpg ~ ., data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.4506 -1.6044 -0.1196  1.2193  4.6271
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 12.30337   18.71788   0.657   0.5181
## cyl         -0.11144    1.04502  -0.107   0.9161
## disp         0.01334    0.01786   0.747   0.4635
```

```
## hp          -0.02148    0.02177  -0.987   0.3350
## drat         0.78711    1.63537   0.481   0.6353
## wt          -3.71530    1.89441  -1.961   0.0633 .
## qsec         0.82104    0.73084   1.123   0.2739
## vs           0.31776    2.10451   0.151   0.8814
## am           2.52023    2.05665   1.225   0.2340
## gear         0.65541    1.49326   0.439   0.6652
## carb        -0.19942    0.82875  -0.241   0.8122
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.65 on 21 degrees of freedom
## Multiple R-squared:  0.869,  Adjusted R-squared:  0.8066
## F-statistic: 13.93 on 10 and 21 DF,  p-value: 3.793e-07
```

Correlation

```
cor(mtcars)[1,]
```

```
##        mpg        cyl       disp         hp       drat         wt
##  1.0000000 -0.8521620 -0.8475514 -0.7761684  0.6811719 -0.8676594
##       qsec         vs         am       gear       carb
##  0.4186840  0.6640389  0.5998324  0.4802848 -0.5509251
```

This model explains 86.9% of the variance and thus variables cyl, disp and hp shows strong correlation. Use these variables in final regression model

```
final <- lm(mpg ~ wt+hp+disp+cyl+am, data = mtcars)
summary(final)
```

```
##
## Call:
## lm(formula = mpg ~ wt + hp + disp + cyl + am, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.5952 -1.5864 -0.7157  1.2821  5.5725
##
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 38.20280    3.66910  10.412 9.08e-11 ***
## wt          -3.30262    1.13364  -2.913  0.00726 **
## hp          -0.02796    0.01392  -2.008  0.05510 .
## disp         0.01226    0.01171   1.047  0.30472
## cyl         -1.10638    0.67636  -1.636  0.11393
## am           1.55649    1.44054   1.080  0.28984
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.505 on 26 degrees of freedom
## Multiple R-squared:  0.8551, Adjusted R-squared:  0.8273
## F-statistic:  30.7 on 5 and 26 DF,  p-value: 4.029e-10
```

Thus we can say that the difference between manual and automatic transmission is 1.55 MPG.

Appendix

Plot 1 Boxplot of MPG

```
boxplot(mpg ~ am, data = mtcars, ylab = "MPG", xlab = "Transmission Type")
```

Plot 2 Pairs Plot

```
pairs(mpg ~ ., data = mtcars)
```

Plot 3 Residual Plot

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

So multivariable regression model shows that multiple R square value is higher at .855, where 85.5% of the regression variance can be explained by chosen variables. It concludes that wt and cyl are confouding variables in relation to am and mpg and manual transmission cars on an average have around 1.55 mpg more that that of automatic cars.