Course7Project-YR

Yonatan Rafael

==========

Executive Summary

This report gagues whether automatic or maual transmission is better for MPG, along with quantifying the MPG difference between automatic and manual transmissions. The report uses the mtcars data set in R. below you will see an exploratory analysis, a t.test to test significance of the difference, and an attempt to model the data.

Exploratory Data Analysis

Take a look at how the data is set up. Rename the “am” data for presentation purposes.

```
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
```

```
mtcars$am2[mtcars$am == 0] <- "Auto"
mtcars$am2[mtcars$am == 1] <- "Manual"
```

Run a boxplot comparing automatic transmission (x = 0) versus manual transmission (x = 1). The plot shows that the manual transmission looks to be faster than automatic transmission. Saved in Appexi

Testing our Hypothesis

Based on the chart above, we set up a t-test.

- Null hypothesis: Manual MPG equals Automatic MPG
- Alternate hypothesis: Manual MPG does not equal Automatic MPG

```
t.test(mtcars$mpg ~ mtcars$am2, conf.level = .95)
```

```
##
##  Welch Two Sample t-test
##
## data:  mtcars$mpg by mtcars$am2
## 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 in group Auto mean in group Manual
##            17.14737             24.39231
```

The results state that with 95% confidence, we reject the null hypothesis that the manual MPG and automatic MPG are the same. And based on our exploratory analysis it is clear that manual transmission MPG > automatic transmission MPG. This does however assume that transmission type is the only variable in the analysis.

Modeling

The first model includes all the variables. We’ll look at the coefficients to determine which we’d like to include for future models. The results show that wt, am, and qsec are most relevant, based on having the lowest p-values.

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

```
##
## Call:
## lm(formula = mpg ~ ., data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.4506 -1.6044 -0.1196  1.2193  4.6271
##
## Coefficients: (1 not defined because of singularities)
##             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
## am2Manual         NA         NA      NA       NA
## ---
## 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
```

Our next model is focused on the above three variables (wt, am, qsec). This model shows that all three variables are signficant with 95% confdience interval. Additionally the R^2 value of .85 means that 85% of variance is explained in this model. The second model determines that switching from automatic to manual transmission increases MPG by 2.93 when holding qsec and wt constant.

```
model2 <- lm(mpg ~ factor(am2) + wt + qsec, data = mtcars)
summary(model2)
```

```
##
## Call:
## lm(formula = mpg ~ factor(am2) + wt + qsec, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.4811 -1.5555 -0.7257  1.4110  4.6610
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         9.6178     6.9596   1.382 0.177915
## factor(am2)Manual   2.9358     1.4109   2.081 0.046716 *
## wt                 -3.9165     0.7112  -5.507 6.95e-06 ***
## qsec                1.2259     0.2887   4.247 0.000216 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.459 on 28 degrees of freedom
## Multiple R-squared:  0.8497, Adjusted R-squared:  0.8336
## F-statistic: 52.75 on 3 and 28 DF,  p-value: 1.21e-11
```

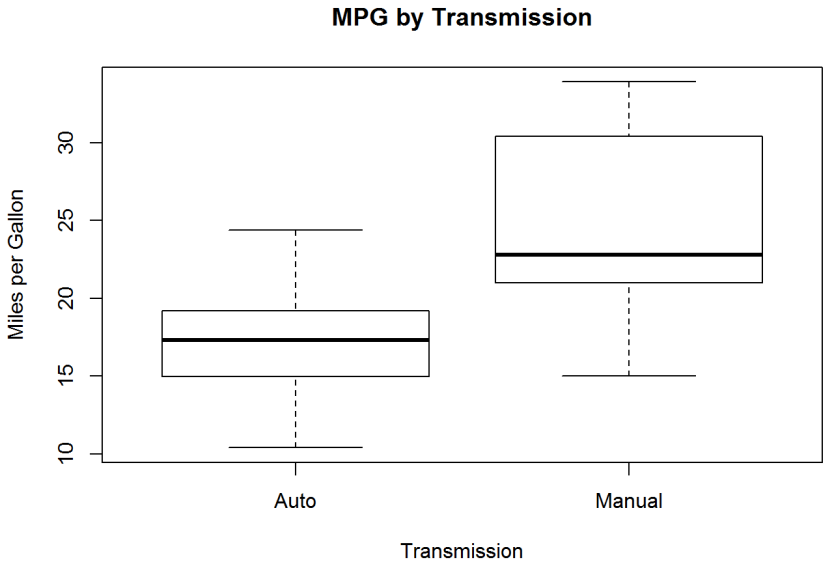
Conclusion

- Manual transmission is better for MPG.
- We need to factor in wt and qsec when evaluating effect on MPG.
- When wt and qsec are held constant, the difference in MPG between Manual and Automatic transmission is 2.93.

APPENDIX.

Boxplot to gauge difference between manual and automatic transmission as it pertains to MPG.

```
boxplot(mtcars$mpg ~ factor(mtcars$am2), xlab = "Transmission", ylab =  "Miles per Gallon")
title(main = "MPG by Transmission")
```



ANOVA test to gauge the significance of the variables in our final model.

```
anova(model2)
```

```
## Analysis of Variance Table
##
## Response: mpg
##             Df Sum Sq Mean Sq F value    Pr(>F)
## factor(am2)  1 405.15  405.15  67.012 6.542e-09 ***
## wt           1 442.58  442.58  73.203 2.673e-09 ***
## qsec         1 109.03  109.03  18.034 0.0002162 ***
## Residuals   28 169.29    6.05
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual Check and Normality Check.

```
par(mfrow=c(2,2))
plot(model2)
```