Regression Analysis Course Project

Vicki Cocco

### Motor Trend 1970's Automobile Transmission Impact on MPG

### Executive Summary

#### Based on the given data and input from a Systems Design Engineer from the

#### automobile industry, this report will:

#### 1) Determine if a manual or automatic transmission was better for miles per gallon fuel usage.

#### 2) Quantify the MPG difference between manual and automatic transmissions.

The intended audience for this report are people interested in the interaction of transmission type and MPG and who have a working knowledge of common statistics and basic linear regression techniques. The chosen regression model will have a small p-value, indicating that there is little evidence against the model. It will have a approximate normal distribution of residuals indicating ????????????, and each predictor(????) will be relevant indicate by the p-value of their coeficients(???????).

Basis of analysis: In order to determine the transmission type impact on MPG, the industry expert recommended focusing on the data specific to transmissions such as the type and the number of forward gears. He also advised that the drivers should be assumed to be of the same ability, inexperienced drivers will affect the MPG of manual transmissions. That information was extracted from the given data and a portion of it is listed in Table 1.

Looking at Table 2, both the small p-value and a close normal distribution of residuals are good indicators that manual transmissions obtain 7.245 MPG over automatic. But, this is not the full story.

Image 3 shows the number of gears for each transmission type over the cars tested. Manual transmissions have data for 4 and 5 gears and automatic transmissions have data for 3 and 4 gears. Given this unique property of each type of transmission, the success of the relationship between transmission type and MPG may be influenced by the number of gears.

Image 2 compared side-by-side with Image 1, shows the the MPG for both types of transmissions across the common 4th gear. Looking at Table 2, although there is still an MPG advantage of 5.225 for manual transmissions the data analysis p-value of 0.108 does not have strong support for the absolute conclusion that Manual transmissions have an MPG advantage across common gears.

Conversely, Table 3 shows the results of the regression model between MPG and Transmission type combined with the number of gears. The large p-values of each of the coefficients suggests that in this combination, neither the transmission type nor any number of gears is influential to the model.

Table 4 shows the coefficients for the linear regression between MPG versus transmission type and number of gears in combination. Table 5 shows the coeficients for the linear regression between MPG versus number of gears alone. Since automatic transmissions are the only type with 3 gears, the intercept coeficient is the same in each regression model. With the regression model including the transmission, none of the other coeficients have strong p-values. However, with the regression model based on gears alone, the coeficients have small p-values. ??????? need to explain residuals

### Conclusions:

Since automatic transmissions are the only transitions type with three gears, the intercept for

### Appendix

#### Table 1

|  |  |  |
| --- | --- | --- |
| Miles Per Gallon | Number of Forward Gears | Transmission Type |
| 21.0 | 4 | Manual |
| 21.0 | 4 | Manual |
| 22.8 | 4 | Manual |
| 21.4 | 3 | Automatic |
| 18.7 | 3 | Automatic |
| 18.1 | 3 | Automatic |

#### Table 2

##   
## Call:  
## lm(formula = mpg ~ Transmission, data = dtCars)  
##   
## 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 \*\*\*  
## TransmissionManual 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

#### Table 3

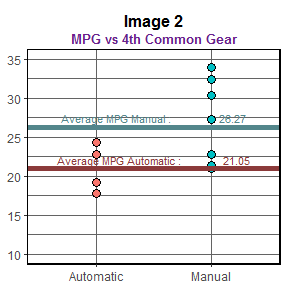
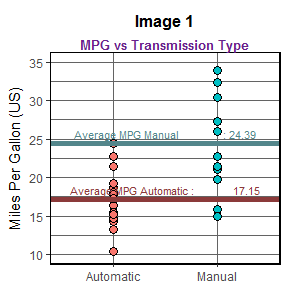
##   
## Call:  
## lm(formula = mpg ~ Transmission, data = dtCars, subset = (gears ==   
## "4"))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -5.2750 -3.8250 -0.4125 3.5437 7.6250   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 21.050 2.416 8.713 5.53e-06 \*\*\*  
## TransmissionManual 5.225 2.959 1.766 0.108   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.832 on 10 degrees of freedom  
## Multiple R-squared: 0.2377, Adjusted R-squared: 0.1615   
## F-statistic: 3.118 on 1 and 10 DF, p-value: 0.1079

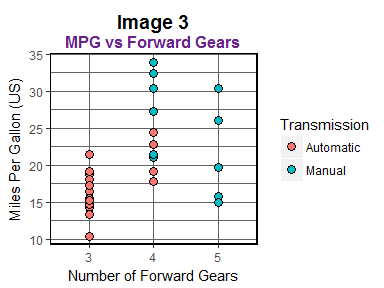
#### Table 4

##   
## Call:  
## lm(formula = mpg ~ Transmission + factor(gears), data = dtCars)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -6.3800 -3.3062 -0.7567 3.1575 9.0200   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 16.10667 1.16497 13.826 4.92e-14 \*\*\*  
## TransmissionManual 5.22500 2.76296 1.891 0.0690 .   
## factor(gears)4 4.94333 2.53899 1.947 0.0616 .   
## factor(gears)5 0.04833 3.61422 0.013 0.9894   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.512 on 28 degrees of freedom  
## Multiple R-squared: 0.4938, Adjusted R-squared: 0.4396   
## F-statistic: 9.105 on 3 and 28 DF, p-value: 0.0002282

#### Table 5

##   
## Call:  
## lm(formula = mpg ~ factor(gears), data = dtCars)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -6.7333 -3.2333 -0.9067 2.8483 9.3667   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 16.107 1.216 13.250 7.87e-14 \*\*\*  
## factor(gears)4 8.427 1.823 4.621 7.26e-05 \*\*\*  
## factor(gears)5 5.273 2.431 2.169 0.0384 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.708 on 29 degrees of freedom  
## Multiple R-squared: 0.4292, Adjusted R-squared: 0.3898   
## F-statistic: 10.9 on 2 and 29 DF, p-value: 0.0002948





### Images 4a and 4b

