Motor Trends

Ian Castillo Rosales

December 21, 2014

Summary

In this number of Motor Trend, we'll analyse the *relationship* between a set of variables from the mtcars data set and miles per gallon (MPG) to answer some interesting questions, some of them were sent us from our readers. We choose only two:

- What transmission is better for MPG? Manual of Automatic?
  Manual transmission delivers 7.245 MPG more than Automatic transmission
- What is the MPG difference between automatic and manual transmissions?
  We use a multivariate regression model that incorporated number of cylinders, displacement, horsepower, and weight to find manual transmissions to be 0.1412 MPG better than automatic transmissions with a goodness of fit reaching 86%.

The next analysis support our conclutions.

Load and process data

The data used in this analysis was obtained from the data set **mtcars**. The data was extracted from the 1974 Motor Trend US magazine, and comprises fuel consumption and 10 aspects of automobile design and performance for 32 automobiles (1973–74 models).

We can explore some of the structure of the data:

```
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
```

The variables are intepreted in the next way:

```
mpg = Miles/(US) gallon
cyl = Number of cylinders
disp = Displacement (cu.in.)
hp = Gross horsepower
drat = Rear axle ratio
wt = Weight (lb/1000)
qsec = 1/4 mile time
vs = V/S
am = Transmission (0 = automatic, 1 = manual)
gear = Number of forward gears
carb = Number of carburetors
```

Exploratory Analysis

A first approach to answer the question “What transmission is better for MPG? Manual of Automatic?” is look into the mean of the MPG (Miles Per Gallon) over all the observations and compare the mean with the mean of the automatic transmission (AT) and the mean of the manual transmission (MT).

An easy view is show in the figure 1 on the Appendix

```
mpg_auto <- mtcars$mpg[mtcars$am == 0]
mpg_manu <- mtcars$mpg[mtcars$am == 1]

summary(mtcars$mpg)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    10.4    15.4    19.2    20.1    22.8    33.9
```

```
mean(mpg_auto)
```

```
## [1] 17.15
```

```
mean(mpg_manu)
```

```
## [1] 24.39
```

To identify the best variables for analyze the difference between the transmission types through a multivariable regression model we can see the relation in the variables from the data with the MPG variable. We choose the variables with higher correlation for our model.

```
mpg_allcor <- abs(cor(mtcars)[1, -1])
vars <- names(mtcars)[-1]

vars[quantile(mpg_allcor, 0.75) < mpg_allcor]
```

```
## [1] "cyl"  "disp" "wt"
```

We can see more clearly the relation in the figure 2

Fit the models

Model 1

We can first, fit a lineal model wihout any variable.

```
model1 <- lm(mpg ~ factor(am), data = mtcars)
summary(model1)
```

```
##
## Call:
## lm(formula = mpg ~ factor(am), data = mtcars)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -9.392 -3.092 -0.297  3.244  9.508
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    17.15       1.12   15.25  1.1e-15 ***
## factor(am)1     7.24       1.76    4.11  0.00029 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.9 on 30 degrees of freedom
## Multiple R-squared:  0.36,   Adjusted R-squared:  0.338
## F-statistic: 16.9 on 1 and 30 DF,  p-value: 0.000285
```

From the model 1 summary we can conclude that

- (Intercept) 17.147 is the mean MPG for Automatic Transmission .
- factor(am)1 7.245 tells us that 24.392 (17.147+7.245) is the mean MPG for Manual Transmission.

Manual transmission delivers 7.245 MPG more than Automatic transmission

Model 2

The next model analyze the situation with all the variables.

```
model2 <- lm(mpg ~ ., data = mtcars)
summary(model2)
```

```
##
## Call:
## lm(formula = mpg ~ ., data = mtcars)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -3.45  -1.60  -0.12   1.22   4.63
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  12.3034    18.7179    0.66    0.518
## cyl          -0.1114     1.0450   -0.11    0.916
## disp          0.0133     0.0179    0.75    0.463
## hp           -0.0215     0.0218   -0.99    0.335
## drat          0.7871     1.6354    0.48    0.635
## wt           -3.7153     1.8944   -1.96    0.063 .
## qsec          0.8210     0.7308    1.12    0.274
## vs            0.3178     2.1045    0.15    0.881
## am            2.5202     2.0567    1.23    0.234
## gear          0.6554     1.4933    0.44    0.665
## carb         -0.1994     0.8288   -0.24    0.812
## ---
## 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:  0.807
## F-statistic: 13.9 on 10 and 21 DF,  p-value: 3.79e-07
```

From the model 2 summary, we can conclude that am 2.5202 is the increase in MPG as a result of switching to Manual. The result still remane the same.

Manual transmission delivers 2.5202 MPG more than Automatic transmission.

Model 3

The las model analyze the model with the variables that we choose.

```
model3 <- lm(mpg ~ am + wt + disp + factor(cyl), data = mtcars)
summary(model3)
```

```
##
## Call:
## lm(formula = mpg ~ am + wt + disp + factor(cyl), data = mtcars)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -4.503 -1.283 -0.482  1.495  5.789
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  33.81607    2.91427   11.60  8.8e-12 ***
## am            0.14121    1.32675    0.11   0.9161
## wt           -3.24918    1.24910   -2.60   0.0151 *
## disp          0.00163    0.01376    0.12   0.9065
## factor(cyl)6 -4.30478    1.49236   -2.88   0.0078 **
## factor(cyl)8 -6.31841    2.64766   -2.39   0.0246 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.65 on 26 degrees of freedom
## Multiple R-squared:  0.838,  Adjusted R-squared:  0.806
## F-statistic: 26.8 on 5 and 26 DF,  p-value: 1.73e-09
```

From the model summary, we can conclude that am 0.14121 is the increase in MPG as a result of switching to Manual, keeping other factors unchanged.

Manual transmission delivers 0.14121 MPG more than Automatic transmission.

To choose the best model we can realize an ANOVA analysis.

```
anova(model1, model2, model3)
```

```
## Analysis of Variance Table
##
## Model 1: mpg ~ factor(am)
## Model 2: mpg ~ cyl + disp + hp + drat + wt + qsec + vs + am + gear + carb
## Model 3: mpg ~ am + wt + disp + factor(cyl)
##   Res.Df RSS Df Sum of Sq    F  Pr(>F)
## 1     30 721
## 2     21 147  9       573 9.07 1.8e-05 ***
## 3     26 183 -5       -35 1.01    0.44
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The best model with the lowest p-value is the Model 2 with and 5% of uncertainty with a R square (goodness of fit) of

```
summary(model2)$r.square
```

```
## [1] 0.869
```

Appendix

Figure 1

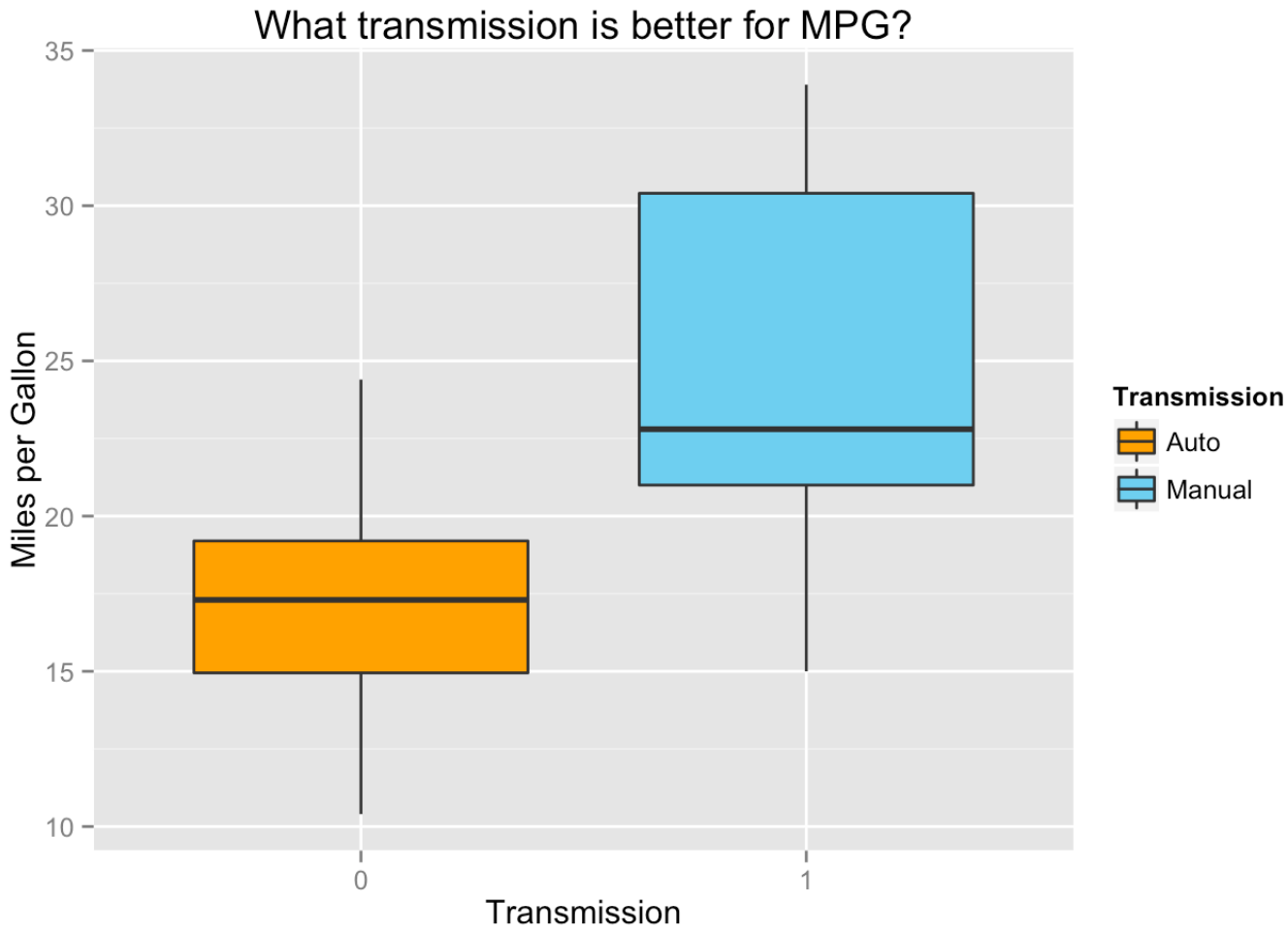

Figure 2

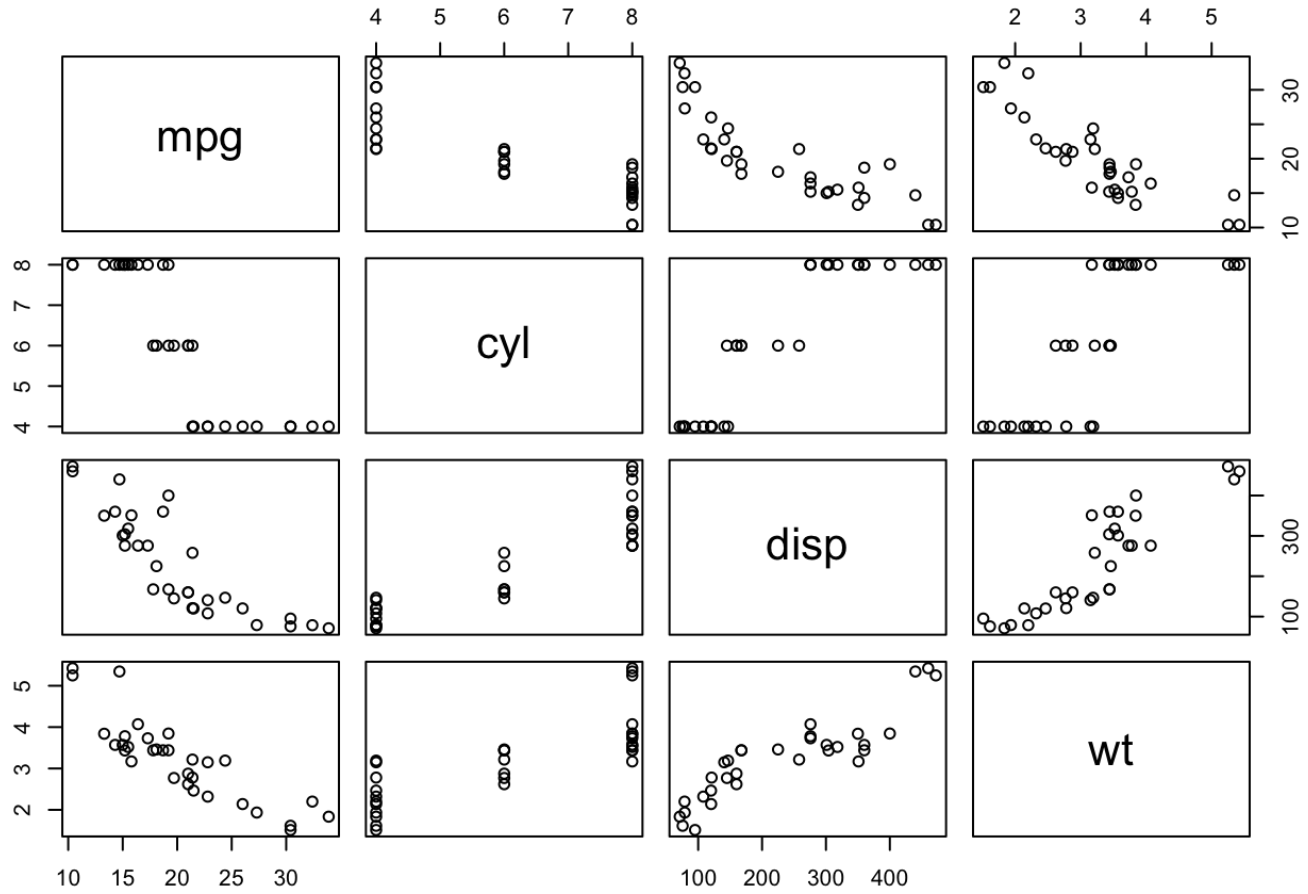

Figure 3: Residuals from the model 2