Which transmission type gets better MPG and by how much?

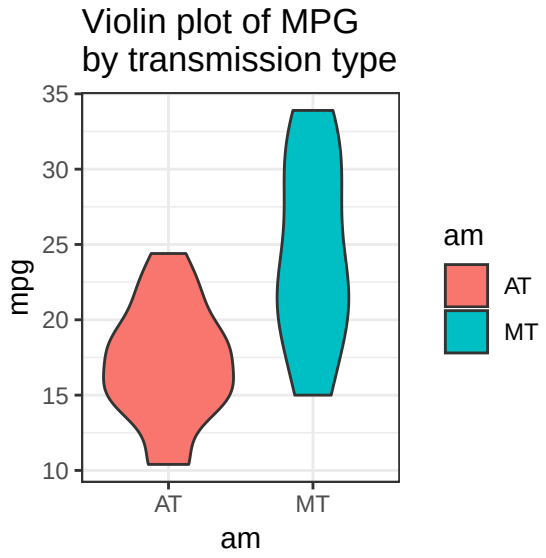
Kirill Setdekov

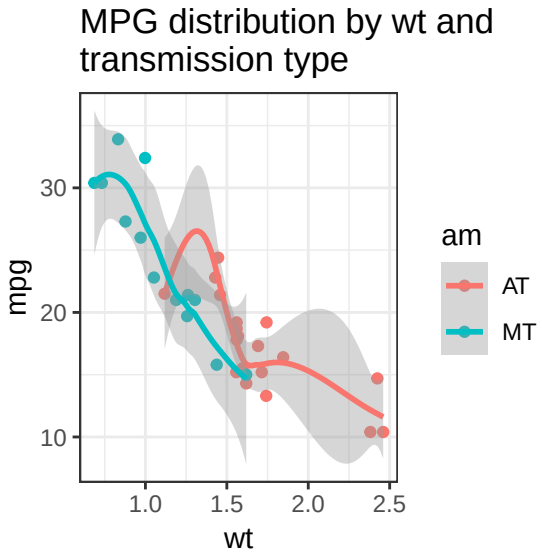
September 11, 2019

```
mtcars <- mtcars %>%
    mutate(
        wt = wt * 0.45359237,
        am = factor(am, levels = c(0, 1), labels = c("AT", "MT")),
        cylsize = disp / cyl
    ) # convert weight to metric t and create cylinder size variable
```

Summary

According to the analysis, manual transmission vehicles have 7.245 better MPG compared to automatic transmission ones. However, adjusting the MPG by other variables, transmission type does not play a statistically significant role in MPG, having other factors constant. This suggests that not manual transmission leads to better MPG, but lighter cars with less cylinders and smaller cylinders.





We can see that on average manual cars get better MPG, however if wee look at MPG vs weight for both transmission types, we see that: 1. on general MT cars are lighter; 2. there is a significant overlap of confidence intervals for the region of 1 to 1,6 t weights.

Regression analysis

I performed several sets of regression.

Basis mpg vs transmission type

Results are significant, residuals look normal and no outliers.

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

```
##
## 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 ***
## amMT           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
```

All original factors included

None of the coefficients are significant, but we find, that cyl, displ, wt are individually contributing to a good model (also, cylsize, not shown).

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

```
##
## Call:
## lm(formula = mpg ~ . - cylsize, 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          -8.19084    4.17647  -1.961   0.0633 .
## qsec         0.82104    0.73084   1.123   0.2739
## vs           0.31776    2.10451   0.151   0.8814
## amMT         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
```

```
anova(mt)
```

```
## Analysis of Variance Table
##
## Response: mpg
##           Df Sum Sq Mean Sq  F value    Pr(>F)
## cyl        1 817.71  817.71 116.4245 5.034e-10 ***
## disp       1  37.59   37.59   5.3526  0.030911 *
## hp         1   9.37    9.37   1.3342  0.261031
## drat       1  16.47   16.47   2.3446  0.140644
## wt         1  77.48   77.48  11.0309  0.003244 **
## qsec       1   3.95    3.95   0.5623  0.461656
## vs         1   0.13    0.13   0.0185  0.893173
## am         1  14.47   14.47   2.0608  0.165858
## gear       1   0.97    0.97   0.1384  0.713653
## carb       1   0.41    0.41   0.0579  0.812179
## Residuals 21 147.49    7.02
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Best model search

Model mt3 is the best found one that includes and does not loose according to anova compared to a model with all factors. There are also no significant issues with residuals, however they fail the normality test. However, exluding transmission type `am` in a model mt3 is also possible, which again suggests that given other factors, transmission type does not affect MPG.

```
mt3 <- lm(mpg~disp+cyl+wt+cylsize+am, data = mtcars)
summary(mt3)
```

```
##
## Call:
## lm(formula = mpg ~ disp + cyl + wt + cylsize + am, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.3767 -1.5719 -0.2125  1.6865  3.6178
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 62.76208    7.66267   8.191 1.13e-08 ***
## disp         0.10207    0.03205   3.184 0.003744 **
## cyl         -4.77298    1.09670  -4.352 0.000186 ***
## wt          -6.71411    2.30406  -2.914 0.007244 **
## cylsize     -0.75086    0.24022  -3.126 0.004328 **
## amMT        -1.44739    1.25402  -1.154 0.258917
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.295 on 26 degrees of freedom
## Multiple R-squared:  0.8784, Adjusted R-squared:  0.855
## F-statistic: 37.55 on 5 and 26 DF,  p-value: 4.308e-11
```

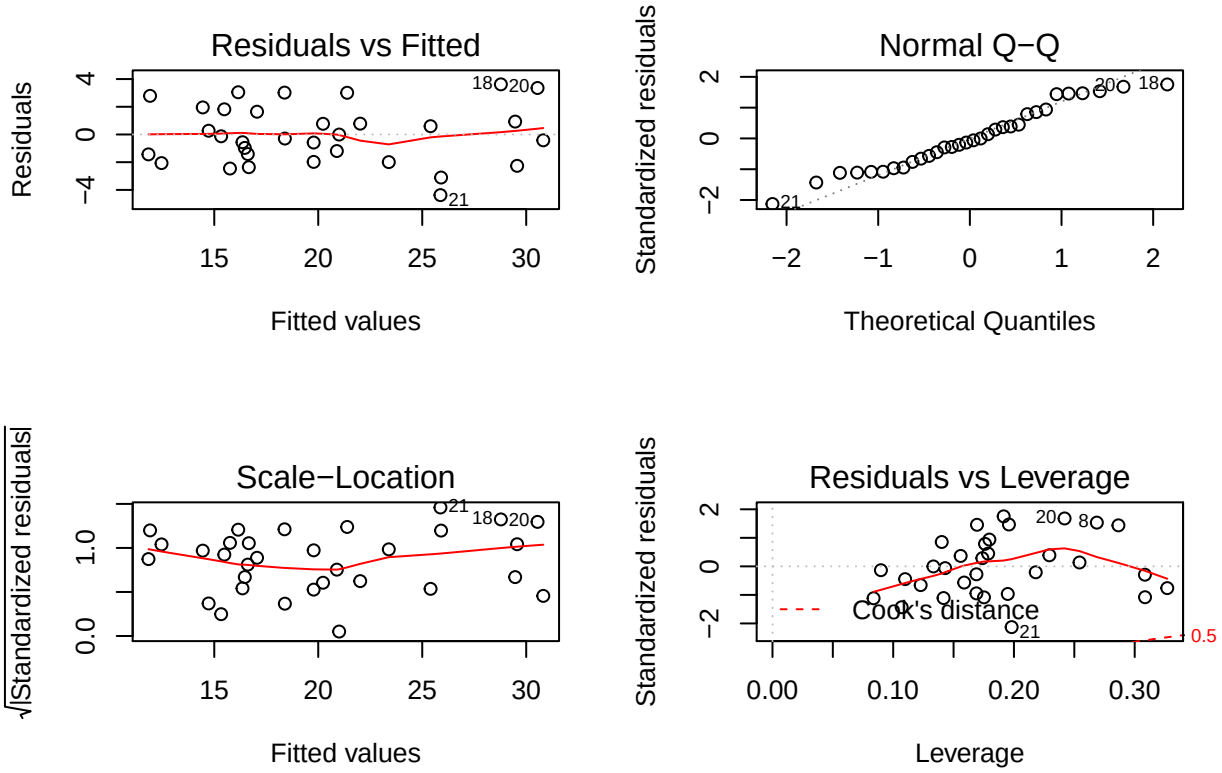
```r
anova(mt3,mt)
```

```
## Analysis of Variance Table
##
## Model 1: mpg ~ disp + cyl + wt + cylsize + am
## Model 2: mpg ~ (cyl + disp + hp + drat + wt + qsec + vs + am + gear +
##     carb + cylsize) - cylsize
##   Res.Df    RSS Df Sum of Sq F Pr(>F)
## 1     26 136.96
## 2     21 147.49  5   -10.536
```

```r
shapiro.test(mt3$residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  mt3$residuals
## W = 0.9669, p-value = 0.4184
```

```r
par(mfrow = c(2, 2))
plot(mt3)
```

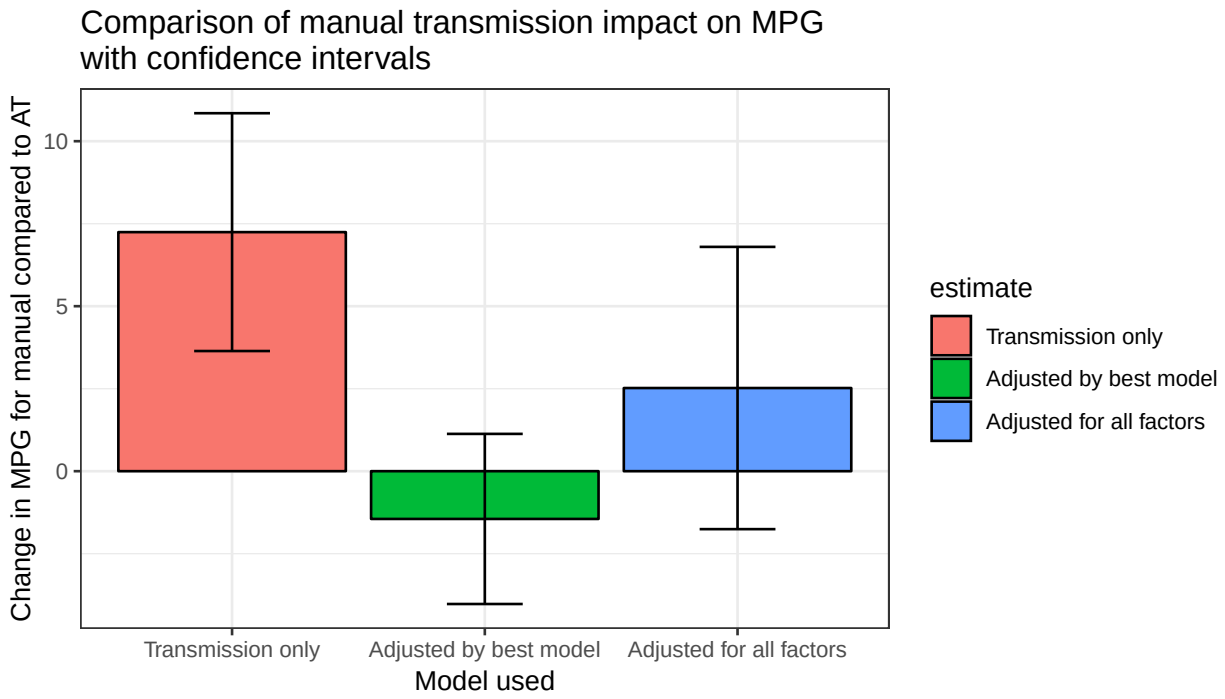



```r
mt4 <- lm(mpg~disp+cyl+wt+cylsize, data = mtcars)
anova(mt3,mt4)
```

```
## Analysis of Variance Table
##
## Model 1: mpg ~ disp + cyl + wt + cylsize + am
```

```
## Model 2: mpg ~ disp + cyl + wt + cylsize
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     26 136.96
## 2     27 143.98 -1   -7.0174 1.3322 0.2589
```

Numeric results and comparisons



By itsetlf manual transmission yields `7.245` higher MPG compared to an automatic transmission

Adjusted for all factors - manual transmissions in a car gives it 2.52 better MPG compared to an automatic transmission. However, this result is statistically insignificant and the best predictors of MPG are cylinder count, displacement, cylinder size and weight. Adjusting for the factors in the best model, manual transmission provides -1.447 lower MPG compared to an automatic transmission.

In both cases effect of manual vs auto transmission, given all other factors equal is statistically indistinguishible from 0, while other factors remain significant.