

Developing Data Products - Week 4 - Assignment

Sandeep Acharyya

7/30/2019

Executive Summary

There is a set of variables related to miles per gallon (MPG) (outcome).

Here the difference of the MPG between automatic and manual transmissions is focused, the `mtcars` dataset in car package is taken as the data source, and techniques about regression models is used to solve the following two questions:

- “Is an automatic or manual transmission better for MPG”
- “Quantifying how different is the MPG between automatic and manual transmissions?”

Data Processing

Load and Test Data

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).

```
library(datasets)
mpgData <- with(mtcars, data.frame(mpg, am))
mpgData$am <- factor(mpgData$am, labels = c("Automatic", "Manual"))
#summary(mpgData)
```

Process Data

Q1: Is an automatic or manual transmission better for MPG?

```
summary(mpgData[mpgData$am == "Automatic",])
```

```
##           mpg           am
##  Min.      :10.40   Automatic:19
##  1st Qu.:14.95   Manual    : 0
##  Median :17.30
##  Mean    :17.15
##  3rd Qu.:19.20
##  Max.    :24.40
```

```
summary(mpgData[mpgData$am == "Manual",])
```

```
##      mpg      am
## Min.   :15.00   Automatic: 0
## 1st Qu.:21.00   Manual   :13
## Median :22.80
## Mean   :24.39
## 3rd Qu.:30.40
## Max.   :33.90
```

Q2: Quantifying how different is the MPG between automatic and manual transmissions?

```
fit <- lm(mpg ~ as.integer(am), data=mpgData)
summary(fit)
```

```
##
## Call:
## lm(formula = mpg ~ as.integer(am), data = mpgData)
##
## 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)      9.902       2.628   3.768 0.000720 ***
## as.integer(am)    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
```

Results

So, manual transmission is better than automatic for MPG, which increased by 7.2449393.

Appendix

Fig 1. MPG between automatic and manual transmissions

```
par(mfrow=c(1,2))
with(mpgData,{
  boxplot(mpg ~ am,
    ylab = "miles per gallon (MPG)")
  plot(mpg ~ as.integer(am),
    xlab = "Automatic (1) or Manual(2)",
    ylab = "miles per gallon (MPG)")
})
```

```
abline(fit, col=2)
})
```

