

The Effect of Automobile transmission types on Fuel

Efficiency

1.Summary

In this analysis, the relationship between automobile transmission types and fuel efficiency was explored. Particularly, MPG(miles per gallon) was used as a measure of fuel efficiency in both automatic and manual transmission. After exploring the automobile data from the 1974 Motor Trend US magazine, a further multivariate regression model was generated to quantify the relationship between the analysis objects.

2.Data

Before analyzing the data, variables on observation from the data set were stated as follow:

- 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

mpg and am were the two main variable analyzed in this report, along with cyl, hp, wt.

3.Exploratory Data Analysis

Excluded outlier of the data set variables

In this section, used R code below to find out out-liner in box-plots and removed them from the data set. From the box-plot, out-liner in hp, wt,qsec and carb can be found. Moving further to identify the out-liners.

```
boxplot(mtcars$hp)
identify(rep(1, length(mtcars$hp)), mtcars$hp, labels = seq_along(mtcars$hp))
```

Identified hp[31],wt[15],wt[16],wt[17],qsec[9],carb[31] as out-liners. Remove the them from the data.

```
mtcars<-mtcars[-c(31,15,16,17,9),]
```

Explor the basic relationship between Automobile transmission types and Fuel Efficiency

The average MPG of automatic transmission is 17.4461538. And the average MPG of manual transmission is 23.78. The cooperation between transmission type and MPG is 0.6274936. At this level of relationship analysis, transmission type and MPG shows a strong correlation and manual transmission have MPG **6.3338462**mpg than auto transmission.

4.Regression Analysis

To quantify the relationship between transmission type and MPG, regression analysis was conducted.

One Variable Regression

Generated a regression model between MPG and transmission type

```
model1<-summary(lm(mtcars$mpg~ mtcars$am))
```

From the result, this regression model has Adjusted R-squared 40% with coefficient of transmission type 7.342, meaning if the car is manual transmission, MPG will be **7.342mpg** more.

However, from the residual plot in the appendix, this model showed an obvious pattern of residual. That means a multivariate regression model was needed.

Discover the relationship between MPG and other variables from the data set

Before building the multivariate regression model, a pair plot was generated to discover the potential relationship between MPG and other variables.(Refer to pair plot in appendix)

From the pair plot, variable wt,disp, hp and drat showed potential regression relationships with mpg. Therefore, a a multivariate regression model with mpg to those variables was generated.

Begin with variables wt,disp, hp and drat, several regression model were tried, and the following model was the model had significant "am" coefficient and a maximum Adjust R-squared.

```
model2<-summary(lm(mtcars$mpg~ mtcars$am + mtcars$drat+ mtcars$qsec +  
mtcars$hp))  
model2
```

The residual plot of this regress was generated in the appendix. No pattern was found in this plot. The regression model was accepted.

To further avoid multidisciplinary problem, a correlation matrix of these variable was created in the appendix. From the matrix, three pairs of variables had correlation more than 60%. Since this report only cared about the relationship

between transmission type and MPG, variable "drat" was removed to improve the model(It had correlation more than 60% with "am").

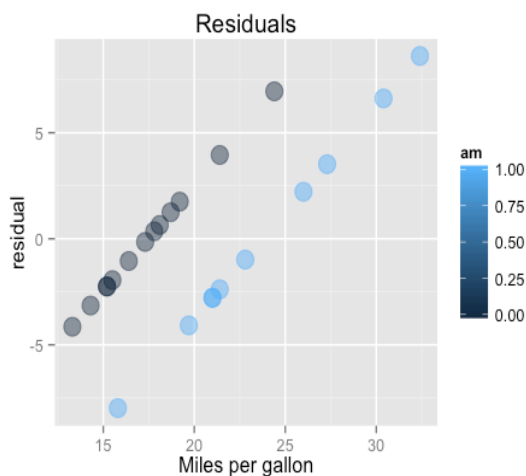
```
model13<-summary(lm(mtcars$mpg~ mtcars$am + mtcars$qsec + mtcars$hp))
model13
```

5.Conclusion

Finally, a model had Adjusted R-squared **74.2%** with only significant coefficient of transmission type **5.63511**. The writer can conclude that manual transmission is better for MPG with 5.64 mpg more than automatic.

Appendix

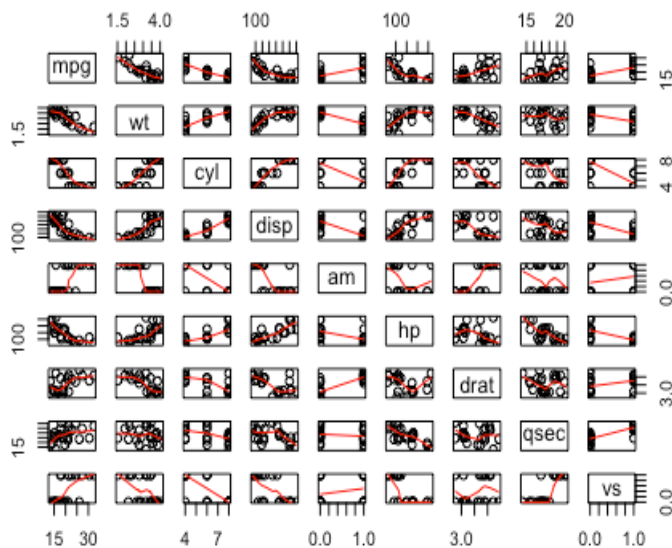
Residual plot for One Variable Regression



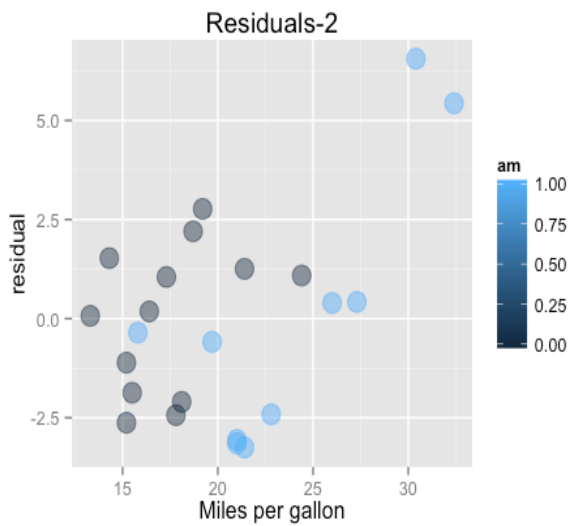
Pair plot for One Variable Regression

```
pairs(data=mtcars, mpg ~ wt + cyl + disp + am + hp+ drat+ qsec+ vs,
      panel=panel.smooth, main="Relationships between MPG and other varia
ble")
```

Relationships between MPG and other variable



Residual plot for multivariate Variable Regression



Correlation matrix

```
##          mtcars.am mtcars.drat mtcars.qsec
## mtcars.am      1.0000000    0.7965562   -0.1685339
## mtcars.drat    0.7965562    1.0000000   -0.1222086
## mtcars.qsec   -0.1685339   -0.1222086    1.0000000
```