MPG and Transmission Types

The `mtcars` dataset has vehicle performance and characteristics for a wide variety of vehicles. We are initially concerned with the role that transmission type, either automatic or manual, can have on a vehicles fuel consumption, specifically its Miles per Gallon(MPG). Let's take a look at our dataset below and note from `?mtcars` we can see that the transmission type is indicated in the `am` variable where `0` indicates automatics and `1`, manual.

```
##                    mpg        am
## Mazda RX4         21.0    Manual
## Mazda RX4 Wag     21.0    Manual
## Datsun 710        22.8    Manual
## Hornet 4 Drive    21.4 Automatic
## Hornet Sportabout 18.7 Automatic
## Valiant           18.1 Automatic
```

Looking at *Figure.1* and *Figure.2*, we can see that there is some sort of trend that automatic transmission vehicles have lower MPGs than manual transmission. Now, let's test this hypothesis with a simple Student-T test to see if this claim is accurate and if so, how confident can we be in this claim.

Manual or Automatic?

Let's test our hypothesis that the average MPG for a manual transmission vehicle is greater than the average MPG for an automatic transmission vehicle. We are going to use 95% confidence independent group T-test.

```
aut <- df$mpg[which(df$am == "Automatic")]
man <- df$mpg[which(df$am == "Manual")]
t.test(man, aut)$conf
```

```
## [1]  3.209684 11.280194
## attr(,"conf.level")
## [1] 0.95
```

Looking at the confidence interval we see that it does not include 0 which means that we reject our null hypothesis that the two means are equal. Instead, we can confidently say, that with our criterion, manual transmissions vehicles have higher MPG ratings than their automatic counterpart.

How much though?

Now that we know that manual transmissions vehicles have better fuel efficiency, MPG. The natural question is to ask, by how much? We can do this by assigning transmission type to a dummy variable in a linear regression model. In our model, we will assign `automatic` to `0` and `manual` to `1`.

```
fit <- lm(mpg ~ am, data = df)
fit$coef
```

```
## (Intercept)    amManual
##   17.147368    7.244939
```

The above output tells us that averge MPG for an automatic transmission is 17.15. The second coefficient tells us the additional MPG earned for a manual transmission vehicle, specifically that manual transmission accounts for an extra 7.24 which comes to 24.39.

We can further analyze our model by looking at the residual plot to really highlight any potential patterns or trends that our model did not account for. Looking at the residual plot in *Figure.3*, the scattering of points looks fairly random, equally spaced above and below 0 with no trends of heteroskedascity.

Inference

Now lets take a closer look at our parameters and try to determine how confident we can be in the exact effect of transmission type on MPG.

```r
std.am <- coef(summary(fit))[1,2]
std.man <- coef(summary(fit))[2,2]

ci.am <- fit$coef[1] + c(-1,1)*qt(.975, fit$df - 2)*std.am
ci.man <- fit$coef[2] + c(-1,1)*qt(.975, fit$df - 2)*std.man

inf.df <- cbind(data.frame(c(fit$coef[1], fit$coef[2]),rbind(ci.am, ci.man)))
names(inf.df) <- c("Parameter", "Lower CI", "Upper CI")
rownames(inf.df) <- c("Automatic", "Manual")
inf.df
```

```
##           Parameter  Lower CI Upper CI
## Automatic 17.147368 14.843725 19.45101
## Manual     7.244939  3.630685 10.85919
```

Summary

So again, our linear regression model tells us that, with our criterion, we can be confident that the Manual transmission types have a positive effect on MPG and that the magnitude of that effect is somewhere between 3.5 and 11.

Appendix

Figure. 1

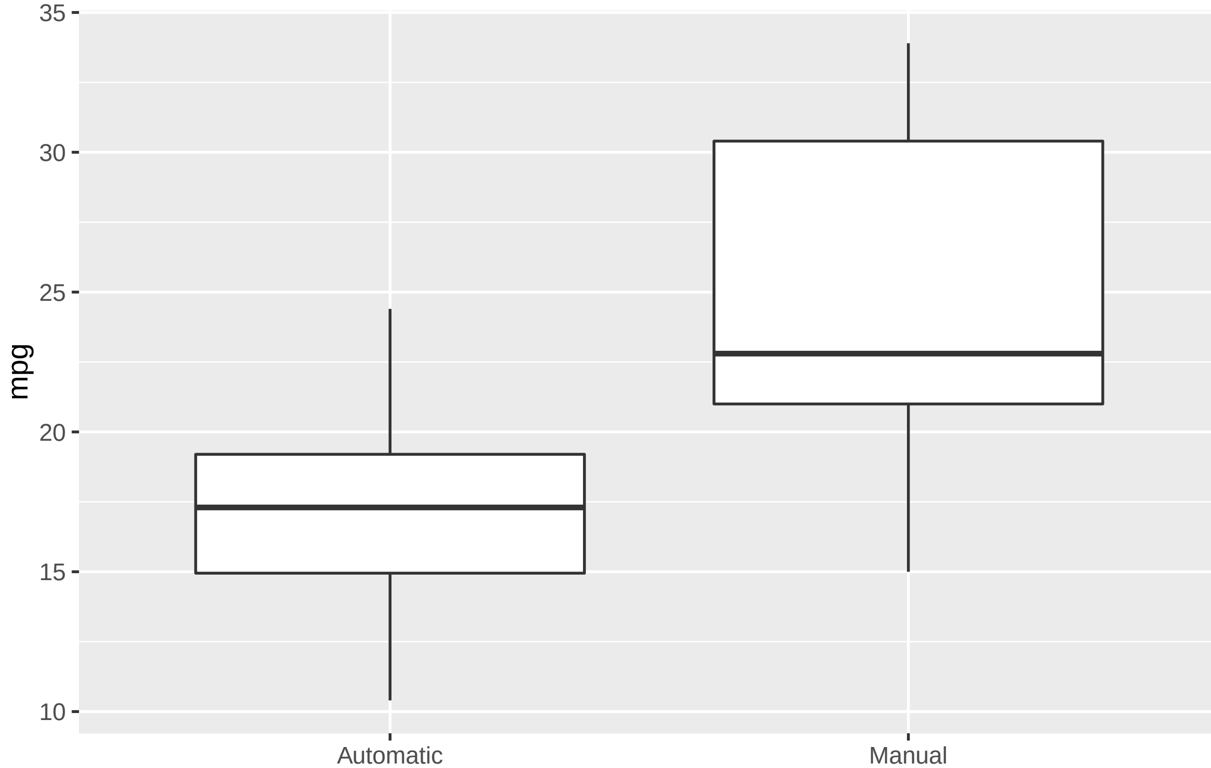
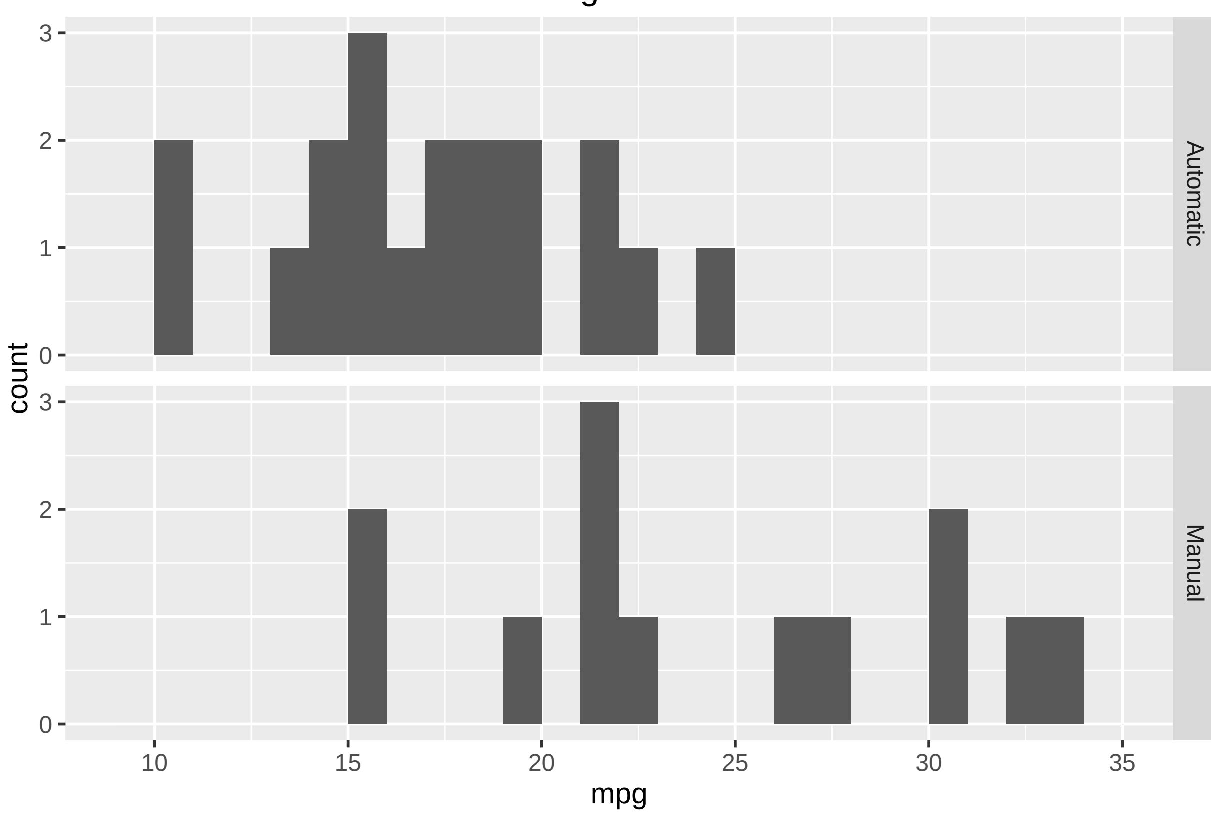
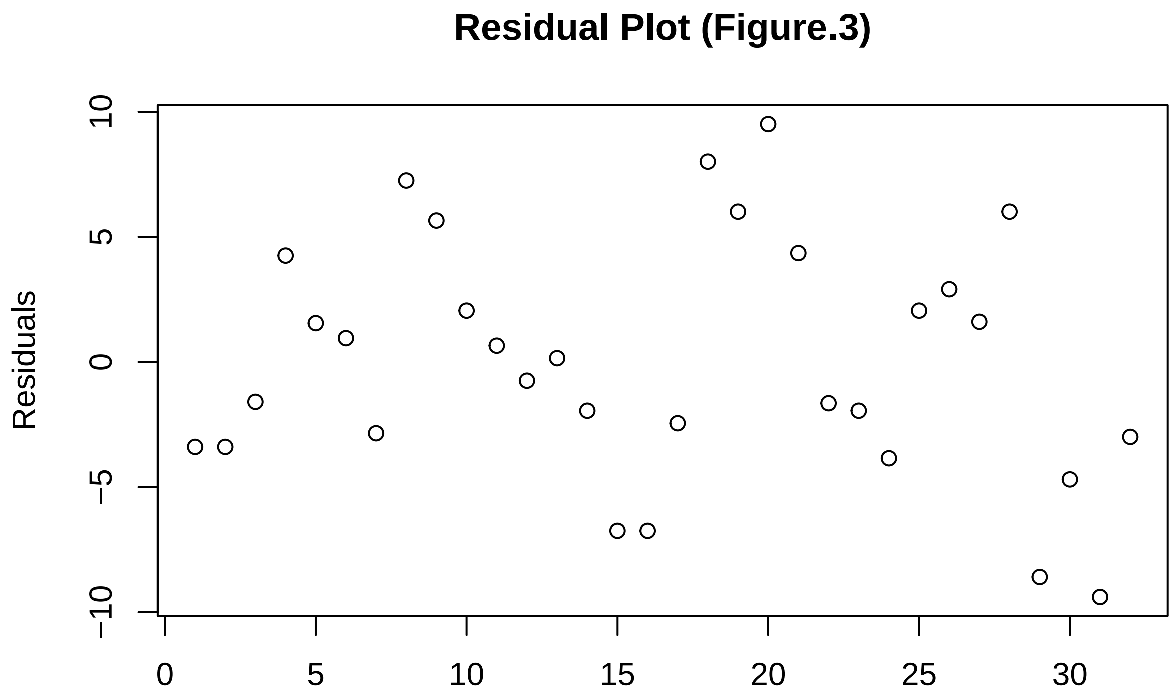


Figure. 2




Residual Plot (Figure.3)
Residuals
10
5
0
−5
−10
0
5
10
15
20
25
30