DATA605_Discussion 11

Load Data

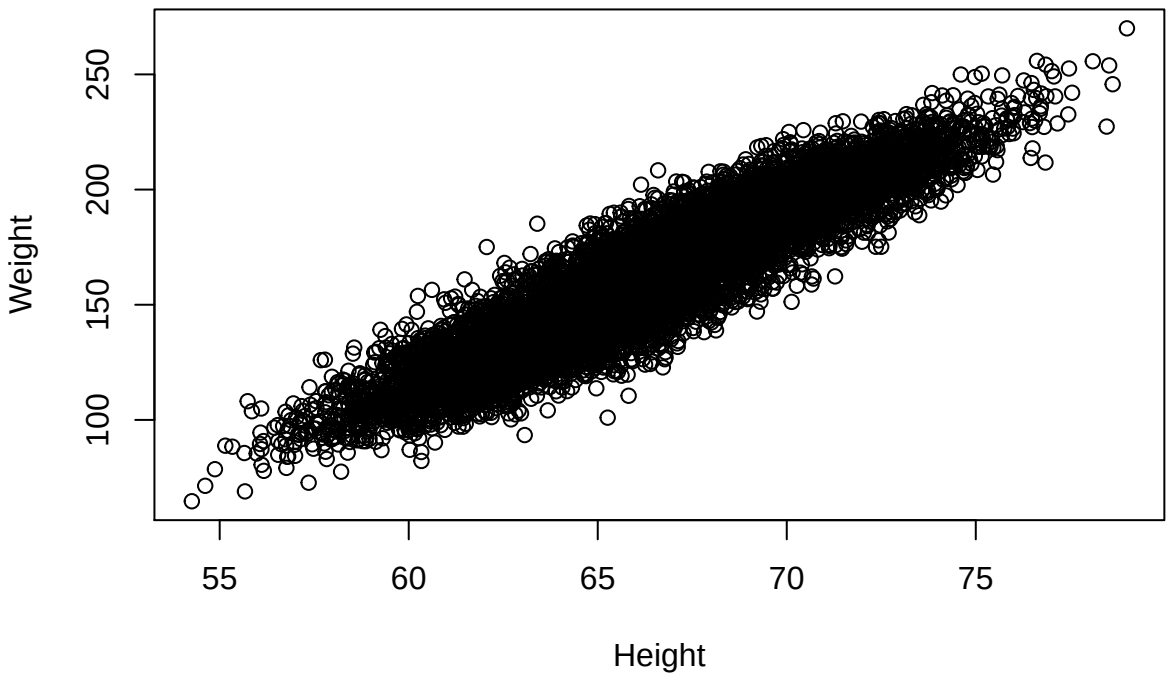
The dataset is from kaggle:https://www.kaggle.com/mustafaali96/weight-height

```
h_w <- read.csv("https://raw.githubusercontent.com/irene908/DATA605/main/h_w.csv")
head(h_w)
```

```
##     Height   Weight
## 1 73.84702 241.8936
## 2 68.78190 162.3105
## 3 74.11011 212.7409
## 4 71.73098 220.0425
## 5 69.88180 206.3498
## 6 67.25302 152.2122
```

Plot

```
plot(h_w)
```



Linear Model

```
L <- lm(Weight ~ Height, h_w)
summary(L)
```

```
##
## Call:
## lm(formula = Weight ~ Height, data = h_w)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -51.934  -8.236  -0.119   8.260  46.844
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -350.73719    2.11149  -166.1   <2e-16 ***
## Height         7.71729    0.03176   243.0   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.22 on 9998 degrees of freedom
## Multiple R-squared:  0.8552, Adjusted R-squared:  0.8552
## F-statistic: 5.904e+04 on 1 and 9998 DF,  p-value: < 2.2e-16
```

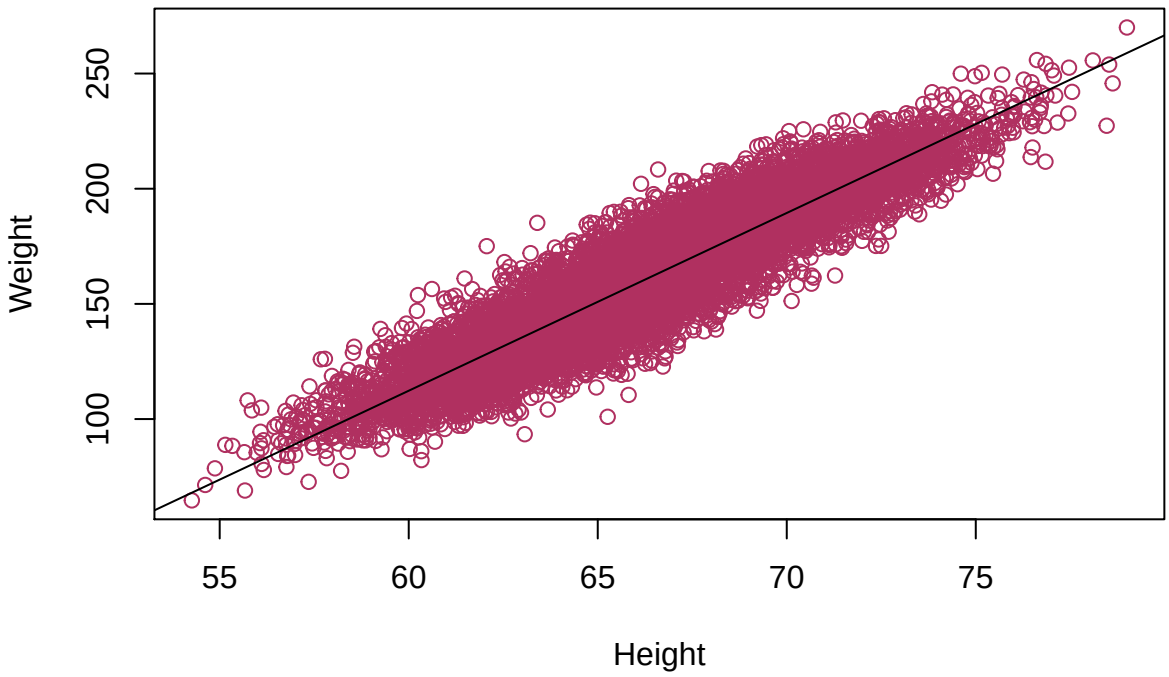
intercept is -350.73719 and *slope* is 7.71729

So the One-Factor Regression Model is:

$Weight = -350.73719 + 7.71729 * Height$

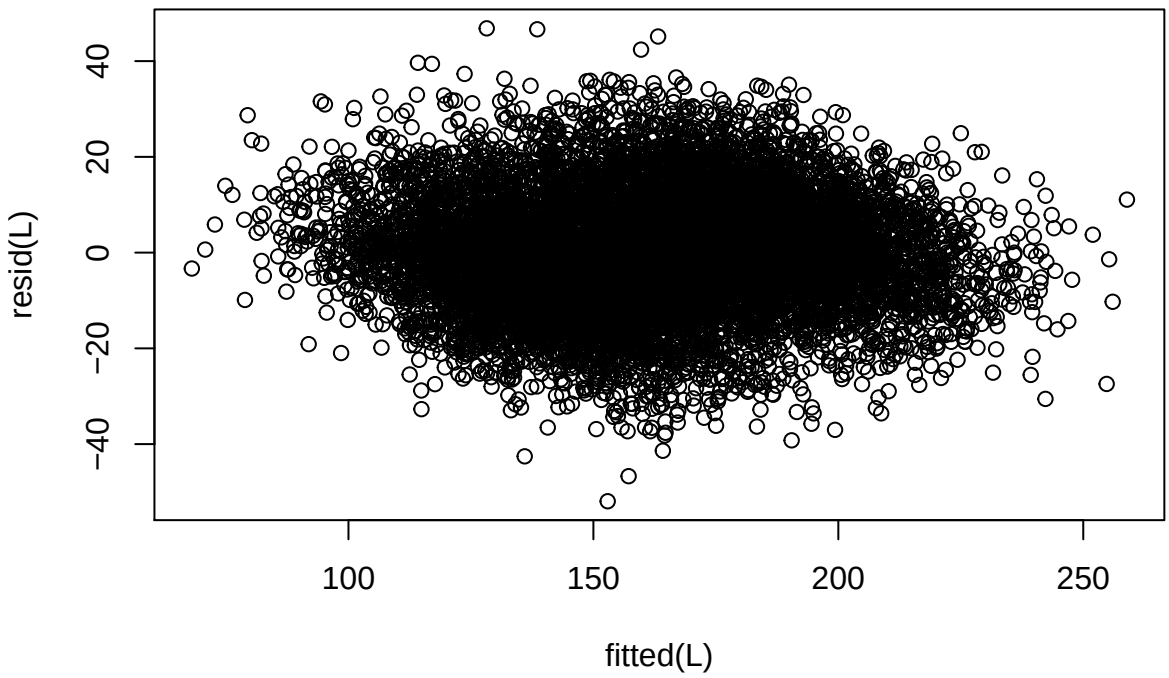
Plot with linear model

```
plot(h_w,col="maroon")
abline(L)
```



Residual Plot

```
plot(fitted(L), resid(L))
```

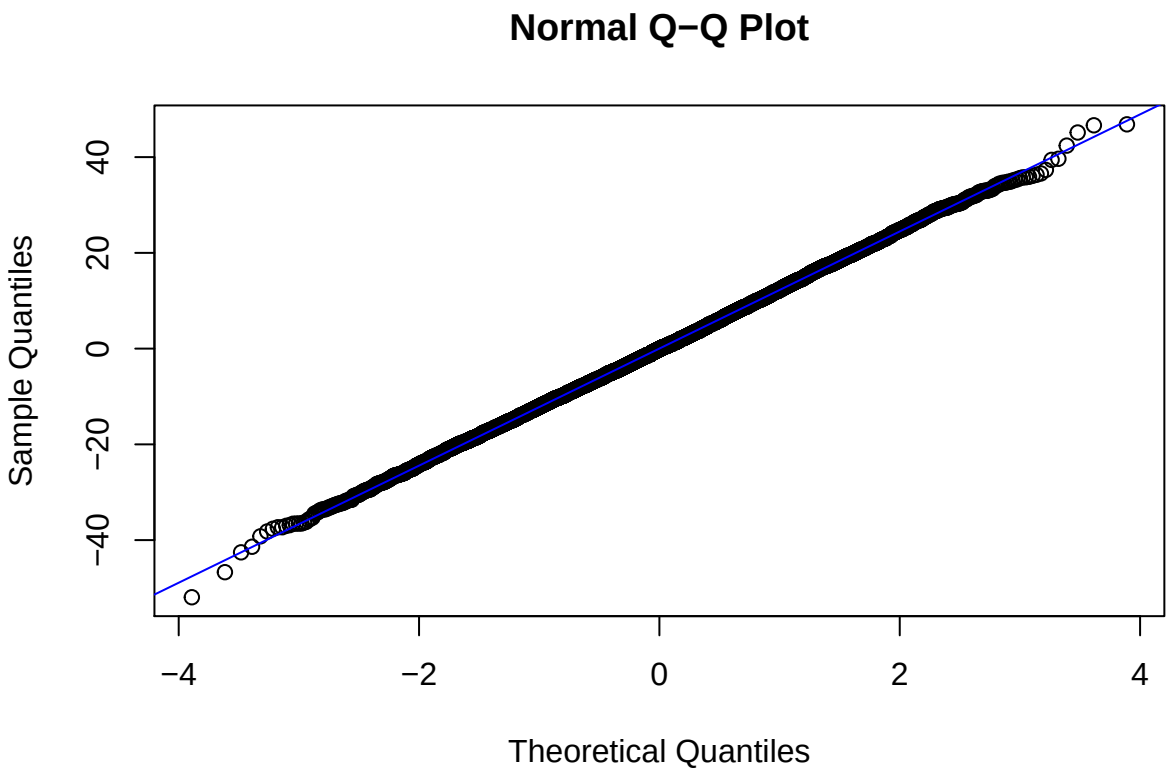


The above plot shows that around zero the residuals are uniformly distributed.

Q-Q Plot

The below Q-Q plot also show that the residuals are normally distributed.

```
qqnorm(resid(L))
qqline(resid(L),col="blue")
```


Normal Q−Q Plot
Sample Quantiles
Theoretical Quantiles
−40
−20
0
20
40
−4
−2
0
2
4