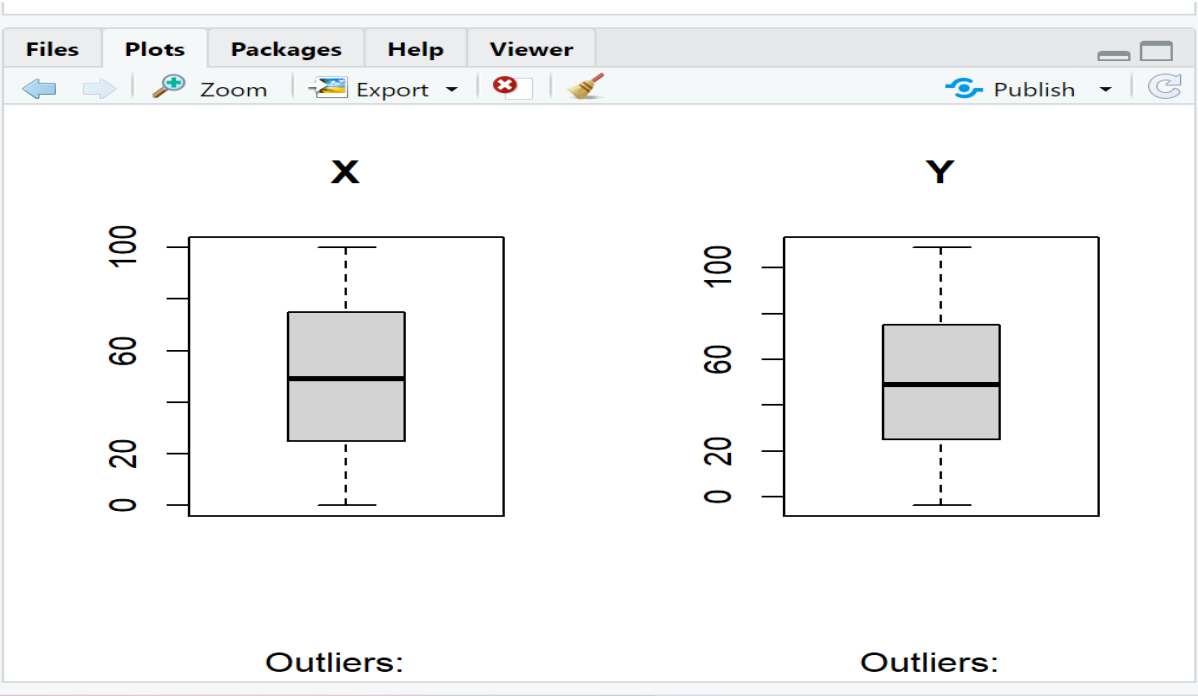
ADA
Assignment:-2

Name:-k.gokul
19BCD7006

```
> if(numberOfNA > 0){
+     cat('Number of missing values found: ', numberOfNA)
+     cat('\nRemoving missing values...')
+     mydata = mydata[complete.cases(mydata), ]
+ }
Number of missing values found:  1
Removing missing values...
```



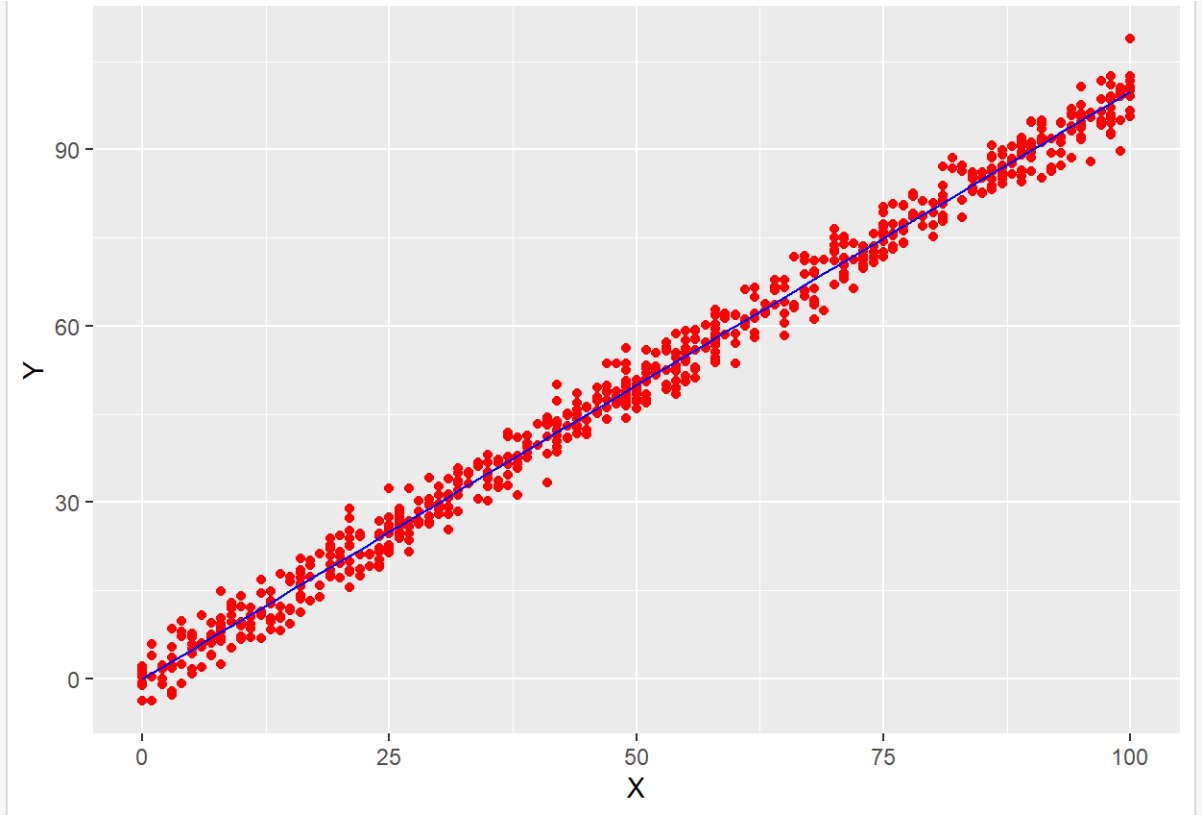

```
> boxplot(mydata$x, main='X', sub=paste('Outliers: ', boxplot.stats(mydata$x)$out))
> boxplot(mydata$y, main='Y', sub=paste('Outliers: ', boxplot.stats(mydata$y)$out))
```

Console Terminal Jobs

R 4.1.3 · ~/

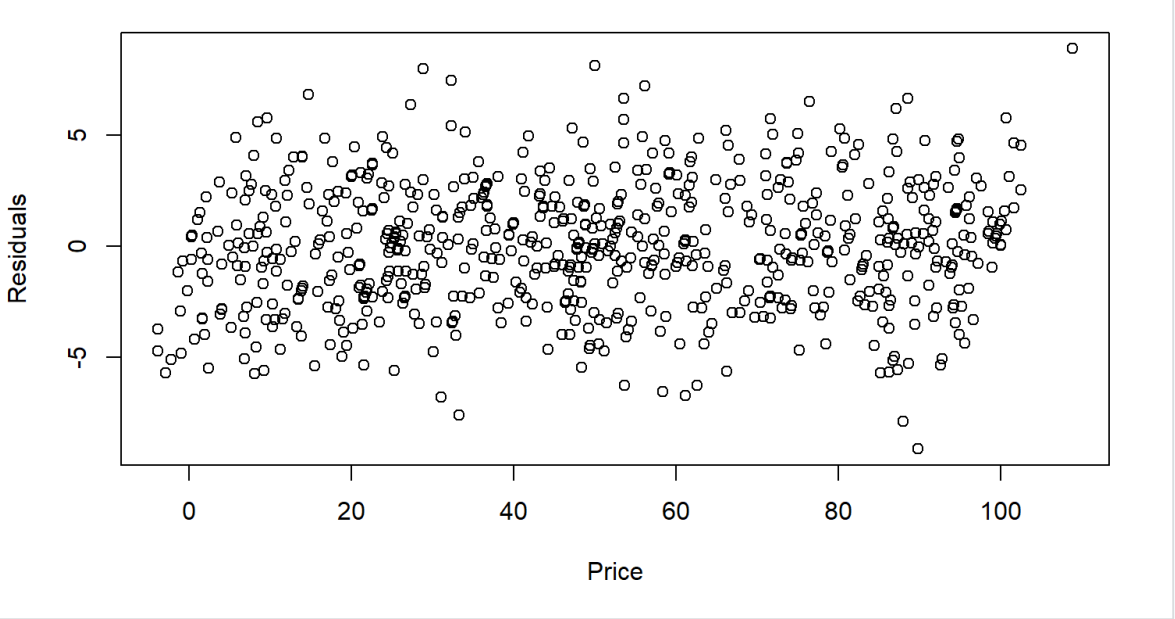
```
> regressor = lm(formula = y ~.,data = mydata)
> summary(regressor)

Call:
lm(formula = y ~ ., data = mydata)

Residuals:
    Min      1Q  Median
-9.1523 -2.0179  0.0325
     3Q     Max
 1.8573  8.9132

Coefficients:
             Estimate
(Intercept) -0.107265
x            1.000656
            Std. Error
(Intercept)   0.212170
x             0.003672
            t value Pr(>|t|)
(Intercept)  -0.506    0.613
x           272.510   <2e-16

(Intercept)
x           ***
---
Signif. codes:
  0 ‘***’ 0.001 ‘**’ 0.01
  ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.809 on 697 degrees of freedom
Multiple R-squared:  0.9907,    Adjusted R-squared:  0.9907
F-statistic: 7.426e+04 on 1 and 697 DF,  p-value: < 2.2e-16
```

```
> ggplot() +
+     geom_point(aes(x = mydata$x, y = mydata$y),
+                colour = 'red') +
+     geom_line(aes(x = mydata$x, y = predict(regressor, newdata = mydata)),
+               colour = 'blue') +
+     ggtitle('X vs Y (Training set)') +
+     xlab('X') +
+     ylab('Y')
> |
```




```
> mean (apply(compare, 1, min)/apply(compare, 1, max))
[1] -Inf
> mean(0.9,0.9,0.9,0.9)
[1] 0.9
>
```

```
> plot(mydata$y, resid(regressor),
+      ylab="Residuals", xlab="Price",
+      main="Residual plot,19bce7358")
> mean(regressor$residuals)
[1] -2.207445e-17
>
```


Residuals
5
0
-5
0
20
40
60
80
100
Price