3) Emp\_data -> Build a prediction model for Churn\_out\_rate

**Ans :**

**R Code :**

## Simple Linear Regression

########## Emp Data Set #########

emp <- read.csv('D:\\Data Science\\Excelr\\Assignments\\Assignment\\Simple Linear Regression\\emp\_data.csv')

CR <- emp$Churn\_out\_rate

SH <- emp$Salary\_hike

boxplot(CR, col="dodgerblue4")

boxplot(SH,col="dodgerblue4")

plot(SH,CR,main="Scatter Plot",

col="Dodgerblue4",

col.main="Dodgerblue4",

col.lab="Dodgerblue4",

xlab="Churn Out Rate",

ylab="Salary Hike", pch=20)

reg.model<-lm(CR~SH, data=emp)

summary(reg.model)

plot(SH,CR,main="Line of Best Fit",col="Dodgerblue4", col.main="Dodgerblue4")

abline(reg.model, col="red")

res <- signif(residuals(reg.model))

pre <- predict(reg.model)

segments(SH, CR, SH, pre)

predict(reg.model,newdata= data.frame(SH=c(2000,2200)))

**Results :**

> reg.model<-lm(CR~SH, data=emp)

> summary(reg.model)

Call:

lm(formula = CR ~ SH, data = emp)

Residuals:

Min 1Q Median 3Q Max

-3.804 -3.059 -1.819 2.430 8.072

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 244.36491 27.35194 8.934 1.96e-05 \*\*\*

SH -0.10154 0.01618 -6.277 0.000239 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.469 on 8 degrees of freedom

Multiple R-squared: 0.8312, Adjusted R-squared: 0.8101

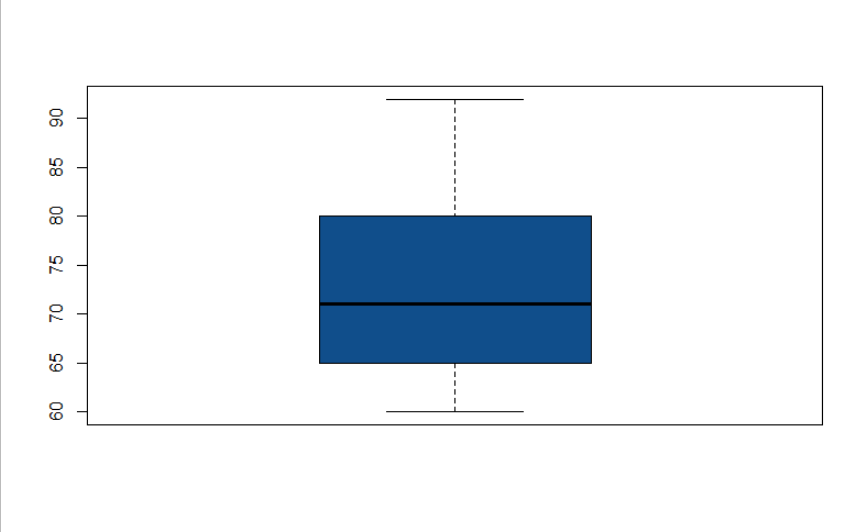
F-statistic: 39.4 on 1 and 8 DF, p-value: 0.0002386

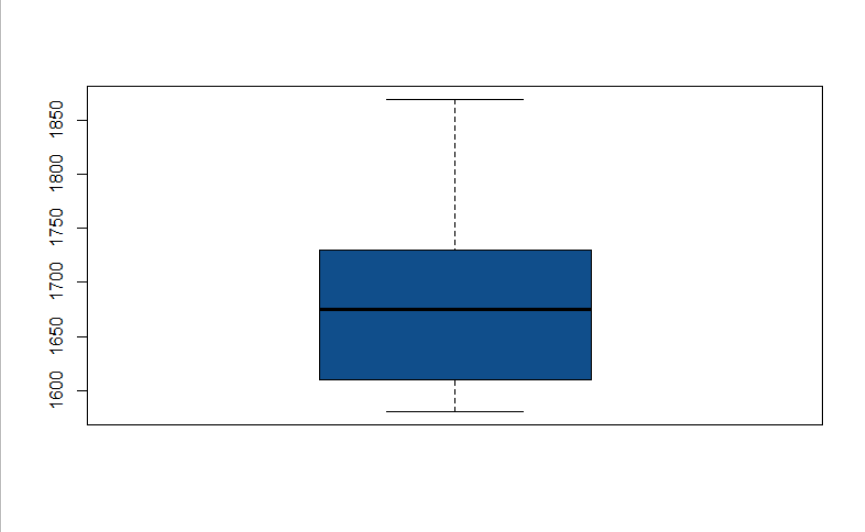
> predict(reg.model,newdata= data.frame(SH=c(2000,2200)))

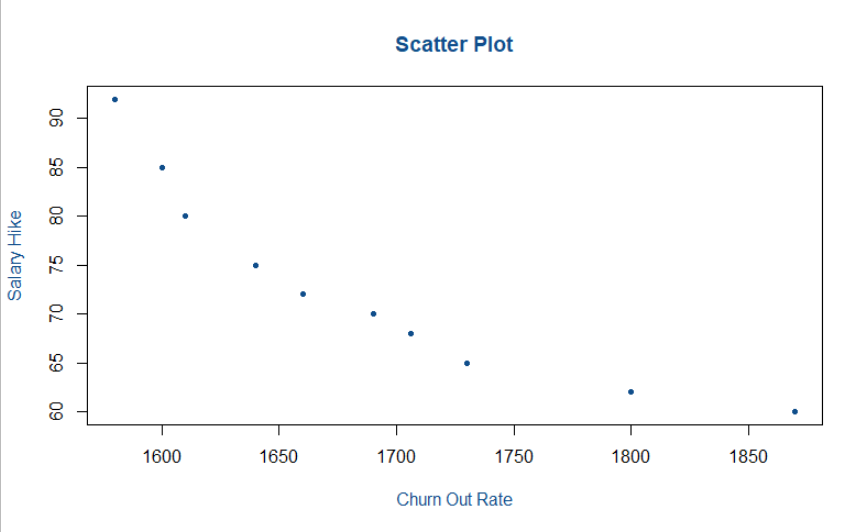
1 2

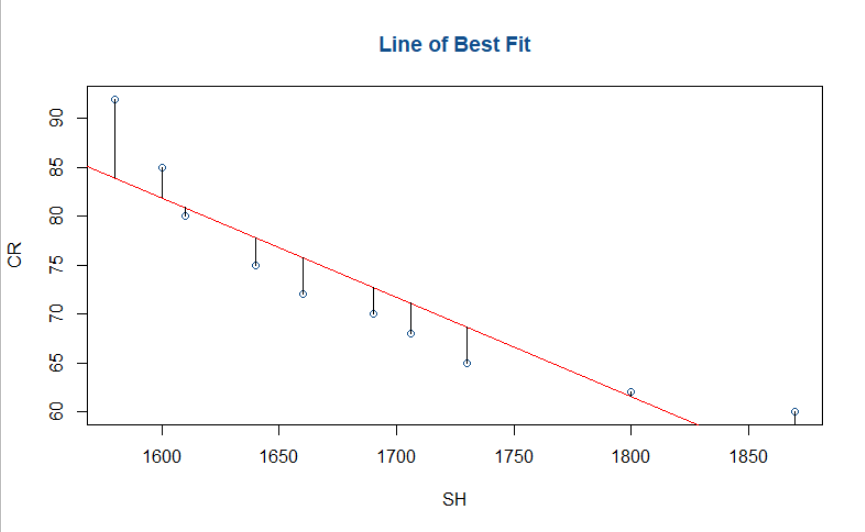
41.27962 20.97109

**Plots :**









**Inference :**

Getting good R squared value of 0.8312.