Experiment-3.2

Student Name: Nabha Varshney **UID: 20BCS4995**

Branch: CSE **Section/Group: 20BCS-DM-704 (A)**

Semester: 6th **Date of Performance: 5thMay 2023**

Subject Name: Data Mining **Subject Code: 20CSP- 351**

Aim – To perform the regression analysis using R programming.

Objective-

- Represent the reading of file using R studio
- Displaying the graph using linear regression.
- Demonstration of regression analysis by linear regression.

Script and Output-

Regression Analysis Code-

```
# Generate random IQ values with mean = 30 and sd =2
IQ <- rnorm(40, 30, 2)

# Sorting IQ level in ascending order
IQ <- sort(IQ)
IQ

# Generate vector with pass and fail values of 40 students
result <- c(0, 0, 0, 1, 0, 0, 0, 0, 0, 1,
            1, 0, 0, 0, 1, 1, 0, 0, 1, 0,
            0, 0, 1, 0, 0, 1, 1, 0, 1, 1,
            1, 1, 1, 0, 1, 1, 1, 1, 0, 1)

# Data Frame
df <- as.data.frame(cbind(IQ, result))

# Print data frame
print(df)
```

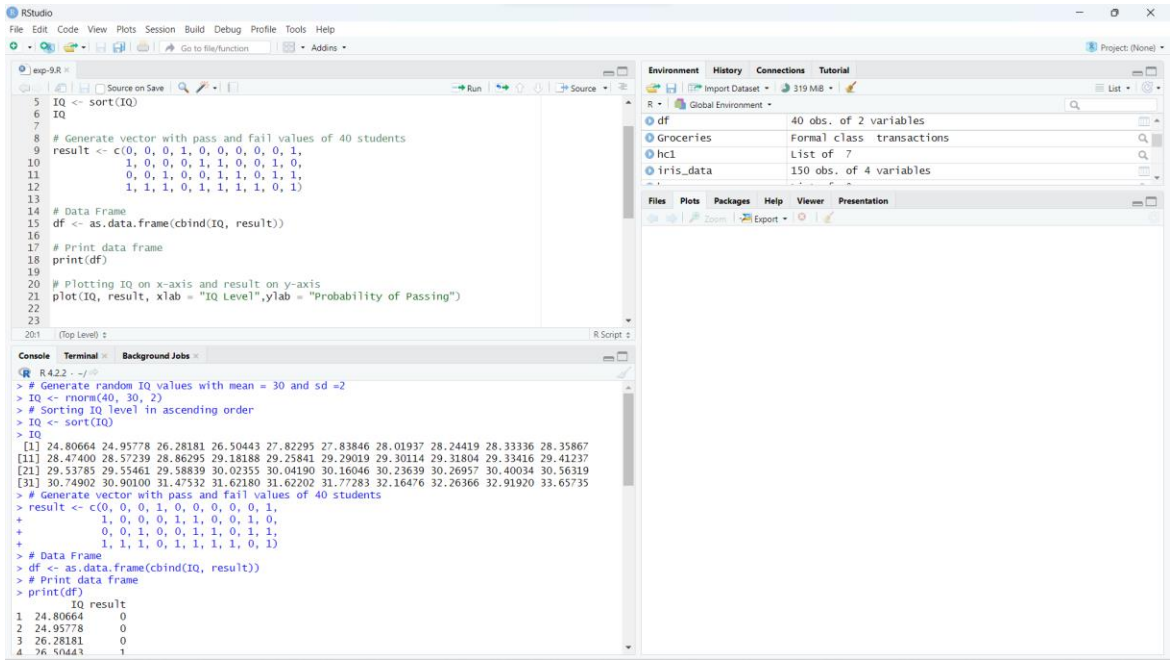
```
# Plotting IQ on x-axis and result on y-axis
plot(IQ, result, xlab = "IQ Level",ylab = "Probability of Passing")

#Linear regression
lrm <- lm(result ~ IQ)

summary(lrm)
#find the result of a person with IQ 35
a<-data.frame(IQ=35)
predRes<-predict(lrm,a)
print(predRes)

# Create a logistic model
lgm = glm(result~IQ, family=binomial, df)
# Summary of the regression model
summary(lgm)

# Create a curve based on prediction using the regression model
curve(predict(lgm, data.frame(IQ=x), type="resp"), add=TRUE)
```

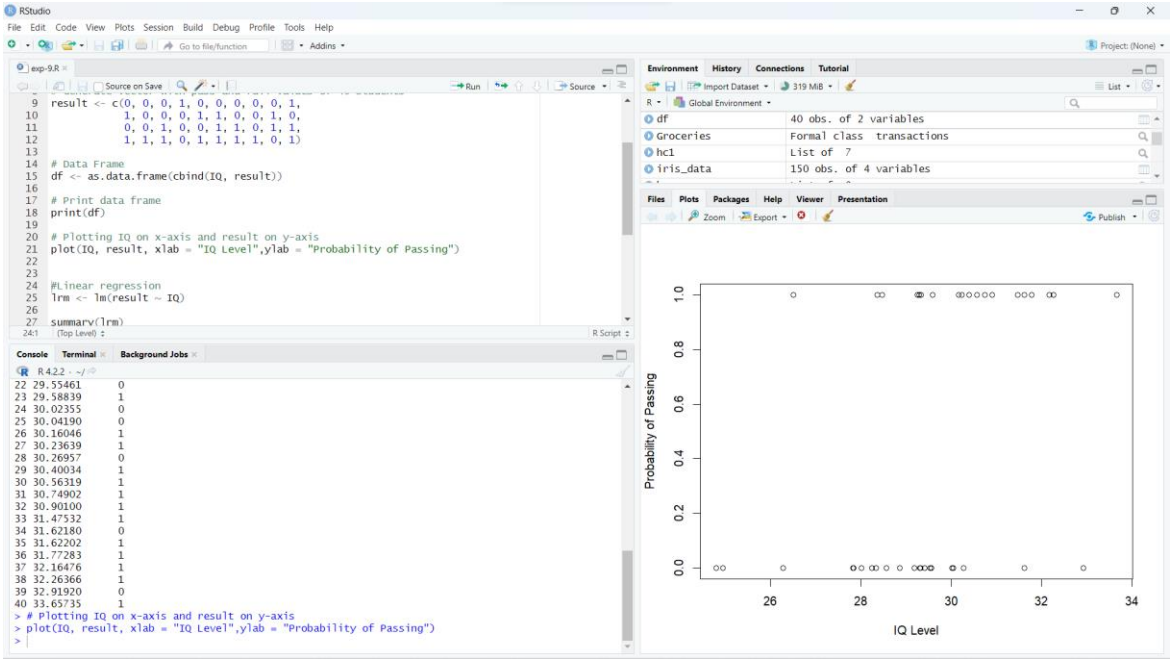
Outlier detection-

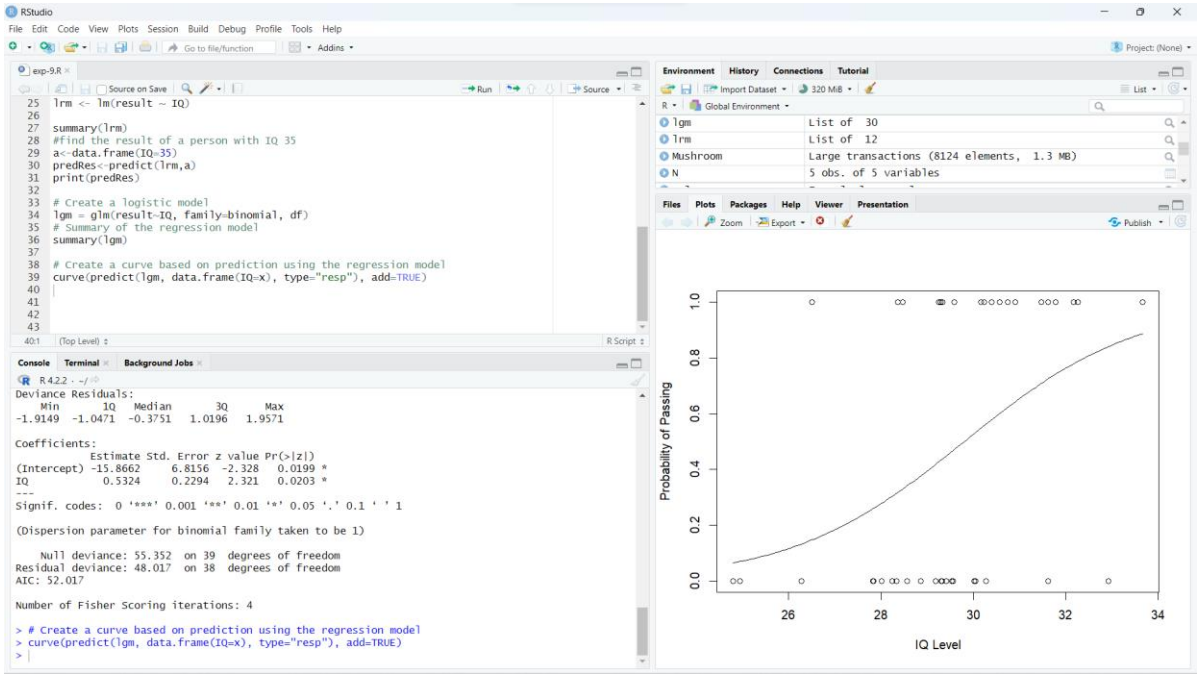
```
#creating the data containing 500 random values

data <- rnorm(500)
print(data)
#adding 10 random outliers to this data.
data[1:10] <- c(46,9,15,-90,42,50,-82,74,61,-32)

#draw boxpolot and an outlier is defined as a data point that is located outside the whiskers of the
box plot.
boxplot(data)

#remove the outlier of the provided data boxplot.stats() function in R
data <- data[!data %in% boxplot.stats(data)$out]

#draw boxplot to verify whether ouliers removed or not
boxplot(data)
```

Output-

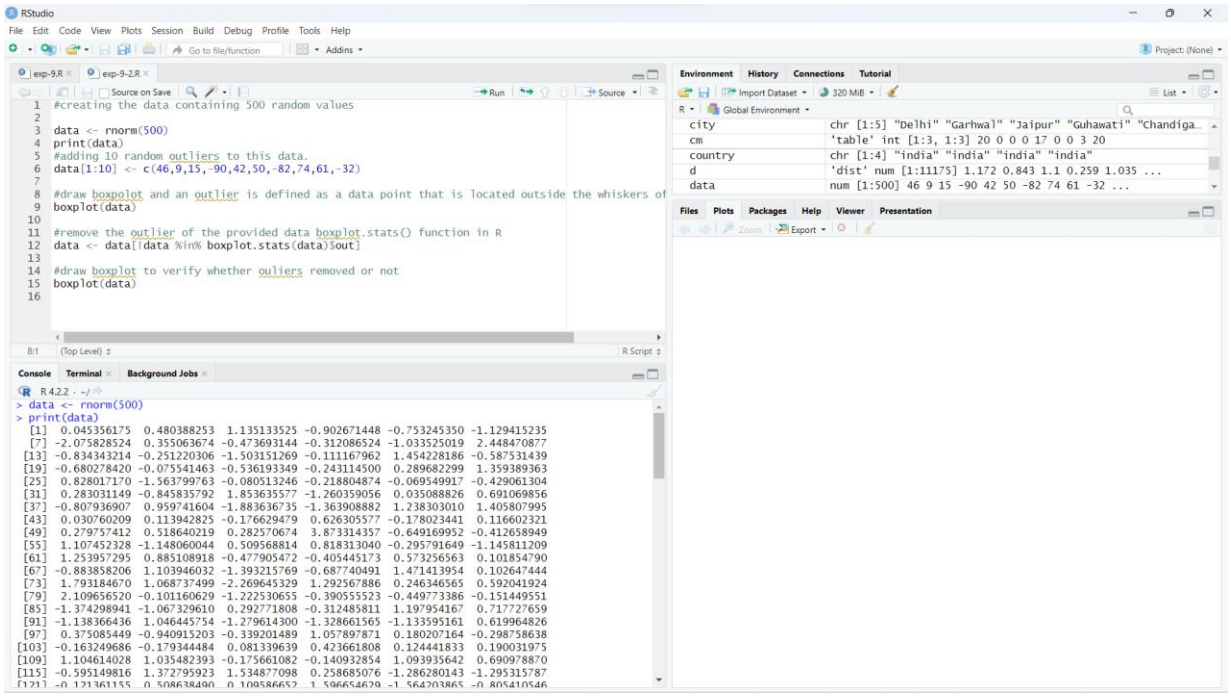
Regression Analysis-







Outlier detection-



DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING