Laxmi Charitable Trust's

Sheth L.U.J College of Arts & Sir M.V. College of Science and Commerce

Department of Information Technology (B.Sc.I.T Semester IV)

Data Analysis

Practical-XIII

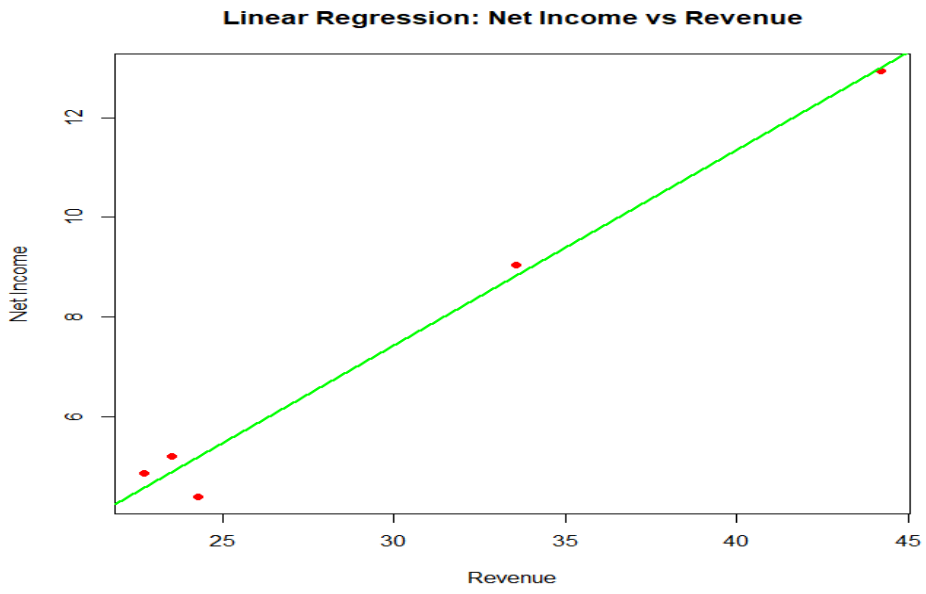
Roll No.:S006	Name:Nandini Chaudhari
Class:SYIT	Batch:01
Date of Assignment:31/01/2026	Date/Time of Submission:31/01/2026

Aim:- Performing linear regression analysis using lm() (R).

Code:-

```
# Load full dataset
qualcomm_df <- read.csv("qualcomm_data.csv")
# Check column names
print(names(qualcomm_df))
# Ensure numeric columns
qualcomm_df$Revenue <- as.numeric(gsub(",", "", qualcomm_df$Revenue))
qualcomm_df$Net_Income <- as.numeric(gsub(",", "", qualcomm_df$Net_Income))
# Fit model
model <- lm(Net_Income ~ Revenue, data = qualcomm_df)
summary(model)
# Plot
dev.new()
plot(qualcomm_df$Revenue, qualcomm_df$Net_Income,
    main = "Linear Regression: Net Income vs Revenue",
    xlab = "Revenue",
    ylab = "Net Income",
    pch = 19,
    col = "red")
abline(model, col = "green", lwd = 2)
```

Output:-



Laxmi Charitable Trust's

Sheth L.U.J College of Arts & Sir M.V. College of Science and Commerce

Department of Information Technology (B.Sc.I.T Semester IV)

Data Analysis

Practical-XIV

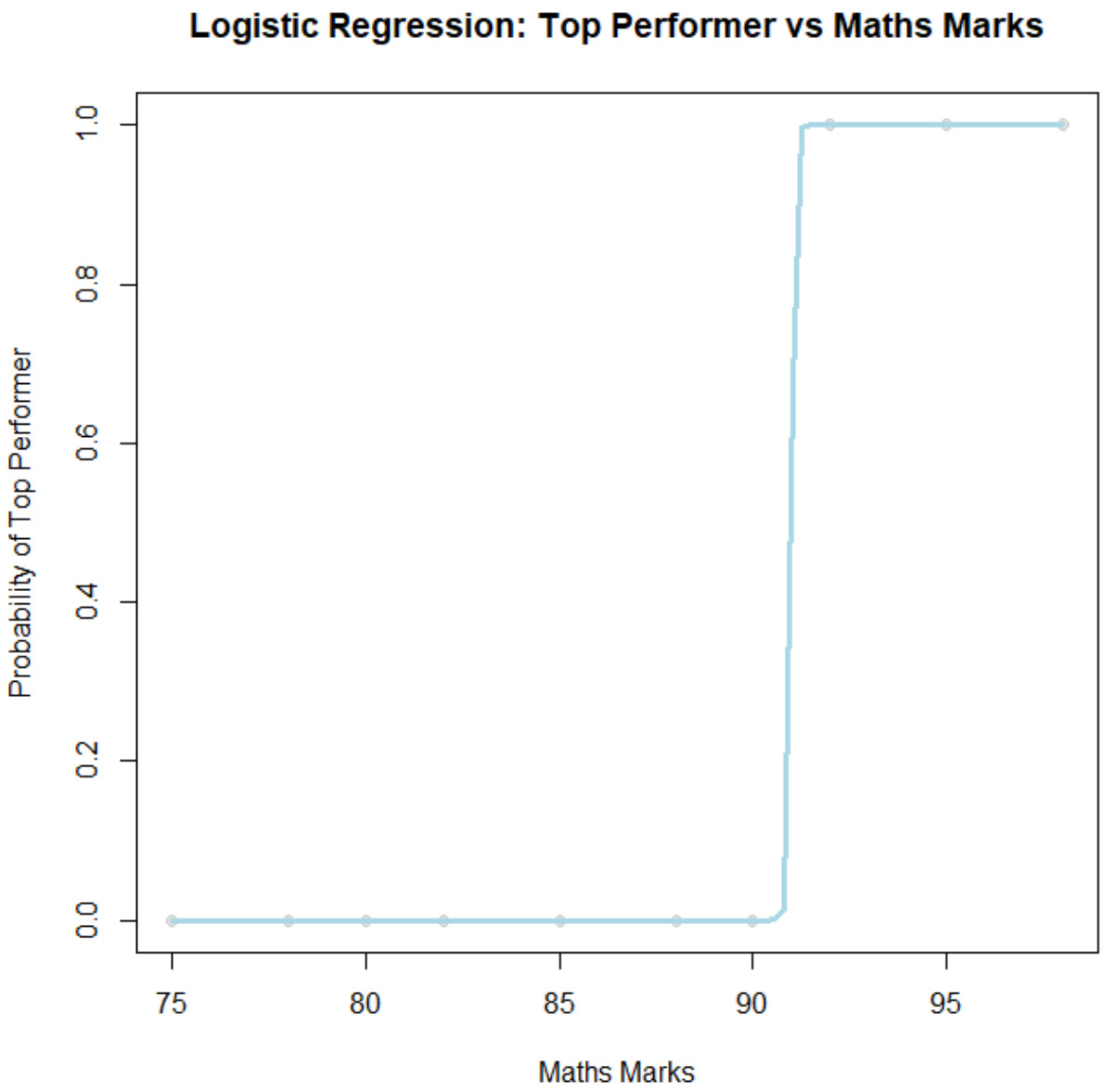
Roll No.:S006	Name:Nandini Chaudhari
Class:SYIT	Batch:01
Date of Assignment:31/01/2026	Date/Time of Submission:31/01/2026

Aim:-Performing logistic regression using glm() (R).

Code:-

```
# Creating dataset inside the code
jee_df <- data.frame(
  Maths_Marks = c(98, 95, 92, 90, 88, 85, 82, 80, 78, 75),
  Percentile  = c(99.8, 99.2, 98.5, 97.8, 96.9, 95.5, 94.2, 93.0, 91.5, 90.0)
)
# Creating binary target variable
jee_df$Top_Performer <- ifelse(jee_df$Percentile > 98, 1, 0)
# Logistic Regression Model
model <- glm(Top_Performer ~ Maths_Marks, family = binomial, data = jee_df)
# Model summary
summary(model)
# Scatter plot
plot(jee_df$Maths_Marks, jee_df$Top_Performer,
     main = "Logistic Regression: Top Performer vs Maths Marks",
     xlab = "Maths Marks",
     ylab = "Probability of Top Performer",
     col = rgb(0, 0, 0, 0.1),
     pch = 19)
# Prediction curve
x_values <- seq(min(jee_df$Maths_Marks), max(jee_df$Maths_Marks), length.out =
100)
predicted_probs <- predict(model, list(Maths_Marks = x_values), type = "response")
# Regression line
lines(x_values, predicted_probs, col = "lightblue", lwd = 3)
```

Output:-


Logistic Regression: Top Performer vs Maths Marks
Probability of Top Performer
0.0
0.2
0.4
0.6
0.8
1.0
75
80
85
90
95
Maths Marks

Laxmi Charitable Trust's

Sheth L.U.J College of Arts & Sir M.V. College of Science and Commerce

Department of Information Technology (B.Sc.I.T Semester IV)

Data Analysis

Practical-XV

Roll No.:S006	Name:Nandini Chaudhari
Class:SYIT	Batch:01
Date of Assignment:31/01/2026	Date/Time of Submission:31/01/2026

Aim:- Exporting results into external files (Excel, CSV, PDF) using write.csv() and writexl (R).

Code:-

```
# Load required package

library(writexl)

library(ggplot2)

# Step 1: Create sample data (agar df pehle se nahi hai)

df <- data.frame(

  Depression = c("Yes","No","Yes","No","Yes","No","Yes","No","Yes","No"),

  Gender =
c("Male","Female","Female","Male","Male","Female","Female","Male","Male","Female")

)

# Step 2: Create frequency table

counts <- table(df$Depression, df$Gender)

# Step 3: Show barplot on screen

barplot(counts,

        main = "Distribution of Depression by Gender",

        xlab = "Gender",

        ylab = "Count",

        col = c("#2ca02c", "#d62728"),

        legend = rownames(counts),

        beside = TRUE)


# Step 4: Save same plot to PDF

pdf("graphical_report.pdf", width = 8, height = 6)

barplot(counts,
```

```
    main = "Distribution of Depression by Gender",
    xlab = "Gender",
    ylab = "Count",
    col = c("#2ca02c", "#d62728"),
    legend = rownames(counts),
    beside = TRUE)
dev.off()
print("Graph saved to graphical_report.pdf")
```

Output:-