**Project Title**

**Students Performance Evaluation**



|  |  |  |  |
| --- | --- | --- | --- |
| 1 | 21F-9341 | Zain Babar | BCS-5E |
| 2 | 21F-9278 | M. Abubakar | BCS-5E |
| 3 | 22F-3175 | Hafiz Arsal | BAI-3B |
| 4 | 22f-3178 | Mehar Abdullah | BAI-3B |

**Course Instructor : Sir Harris Khurram**

**Project Documentation**

**1. Objective**

The primary objective of this project is to analyze the impact of various factors, including parental education, gender, race/ethnicity, lunch, and test preparation, on students' academic performance. The focus is on math scores, reading scores, and writing scores.

**2. Data Description**

**Data Source:**

The dataset used in this project contains information on students' performance,

**Variables:**

1. **Gender:** Indicates the gender of the student (male/female).
2. **Group:** Represents the ethnic or socioeconomic group to which the student belongs (e.g., Group A, Group B, etc.).
3. **Parental Education:** Describes the highest level of education attained by the student's parents (e.g., high school, associate's degree, bachelor's degree, master's degree).
4. **Lunch Type:** Specifies the type of lunch the student receives (free/reduced or standard).
5. **Test Preparation Course:** Indicates whether the student completed a test preparation course (completed or none).
6. **Math Score:** The score the student received in the math subject.
7. **Reading Score:** The score the student received in the reading subject.
8. **Writing Score:** The score the student received in the writing subject.

**Key Observations:**

1. The dataset includes a diverse set of students with varying demographic characteristics.
2. Performance is measured through scores in three subjects: math, reading, and writing.
3. Demographic factors such as gender, group, parental education, lunch type, and completion of test preparation courses are considered.
4. There is a mix of categorical variables (e.g., gender, group, parental education, lunch type, test preparation course) and numerical variables (math score, reading score, writing score).
5. Some students completed test preparation courses, and there is variability in lunch types, suggesting potential socioeconomic factors.

**Possible Analysis:**

1. Descriptive statistics can be used to summarize the central tendency and variability in scores.
2. Comparative analyses can explore differences in performance based on gender, group, parental education, and lunch type.
3. Correlation analysis can investigate relationships between different variables.
4. Regression analysis might help predict scores based on various factors.

**3.Results:**

Here are some findings based on the sample of 100 students from whole population of 1000.

1. **Gender Differences:**
   * On average, females tend to outperform males in all subjects (math, reading, writing) and overall scores.
   * This trend is consistent with findings that suggest gender-based differences in academic performance.
2. **Parental Education Impact:**
   * Students with parents holding a master's degree generally scored higher compared to those with lower parental education levels.
   * There is a positive correlation between parental education and student performance.
3. **Test Preparation Impact:**
   * Students who completed test preparation courses generally scored higher than those who did not.
   * Test preparation seems to have a positive influence on academic performance.
4. **Lunch Type Influence:**
   * Students who receive free/reduced lunch tend to have slightly lower scores compared to those with standard lunch.
   * Economic factors may play a role in academic achievement.
5. **Group Differences:**
   * There are variations in performance across different groups.
   * Group C students, on average, perform well in all subjects.
6. **Completion Status Impact:**
   * Students who completed test preparation courses generally performed better, indicating the potential effectiveness of such programs.
7. **Subject Correlation:**
   * There is a strong positive correlation between scores in math, reading, and writing, indicating consistent performance across subjects.
8. **High Achievers:**
   * Some students achieved perfect scores in all subjects, suggesting a subset of high-achieving students in the sample.
9. **Variability in Scores:**
   * There is variability in scores, with some students performing exceptionally well and others scoring lower, indicating diversity in academic performance.

**4. Codes**

In this project we have used two files UI.R and SERVER.R:

**UI.R** : this file includes all the interface code for the project including representation of histograms, boxplots, regression models and scatter plots.

**Server.R :** this file includes all functions including the implementation of all those functions.

**Codes**

**UI.R:**

library(shiny)

library(shinydashboard)

# Load the server-side logic

source("server.R")

# Define UI for application

ui <- navbarPage(

title = "Student Performance Analysis",

tabPanel("Math Scores",

plotOutput("math\_histogram"),

br(),

plotOutput("math\_boxplot"),

verbatimTextOutput("mean\_mode\_text\_math")

), # Display mean and mode

tabPanel("Reading Scores",

plotOutput("reading\_histogram"),

br(),

plotOutput("reading\_boxplot"),

verbatimTextOutput("mean\_mode\_text\_reading")

),

tabPanel("Writing Scores",

plotOutput("writing\_histogram"),

br(),

plotOutput("writing\_boxplot"),

verbatimTextOutput("mean\_mode\_text\_writing")

), # Display mean and mode

tabPanel("probability distribution",

plotOutput("math\_distribution"),

br(),

plotOutput("reading\_distribution"),

br(),

plotOutput("writing\_distribution")

),

tabPanel("regression model",

plotOutput("math\_regression"),

br(),

plotOutput("reading\_regression"),

br(),

plotOutput("writing\_regression"),

),

tabPanel("scatter PLot",

plotOutput("scatter\_plot"),

),

tags$style(HTML(".navbar { background-color: #000000; }")),

tags$style(HTML(".navbar-default .navbar-nav > li > a { color: #ffffff; }")),

tags$style(HTML(".navbar-default .navbar-nav > li > a { background-color: #000000; }")),

tags$style(HTML(".nav-tabs>li>a {color: #ffffff;}"))

)

**Server.R:**

library(shiny)

library(dplyr)

library(ggplot2)

# Read data from CSV file

students\_data <- read.csv("C:\\Users\\abdul\\OneDrive\\Desktop\\prob project\\StudentsPerformance.csv")

server <- function(input, output, session) {

# Render a histogram of math scores

output$math\_histogram <- renderPlot({

ggplot(students\_data, aes(x = math.score)) +

geom\_histogram(binwidth = 5, fill = "red", color = "black", alpha = 1) +

labs(title = "Math\_hist", x = "Math Score", y = "Frequency") +

theme\_minimal()

})

# Render a histogram of reading scores

output$reading\_histogram <- renderPlot({

ggplot(students\_data, aes(x = reading.score)) +

geom\_histogram(binwidth = 5, fill = "green", color = "black", alpha = 1) +

labs(title = "reading\_hist", x = "reading Score", y = "Frequency") +

theme\_minimal()

})

# Histogram for writing scores

output$writing\_histogram <- renderPlot({

ggplot(students\_data, aes(x = writing.score)) +

geom\_histogram(binwidth = 5, fill = "blue", color = "black", alpha = 1) +

labs(title = "writing\_hist", x = "writing score", y = "frequency") +

theme\_minimal()

})

# Probability Distribution for math scores

output$math\_distribution <- renderPlot({

ggplot(students\_data, aes(x = math.score)) +

geom\_histogram(aes(y = ..density..), binwidth = 5, fill = "red", color = "black", alpha = 0.7) +

stat\_function(fun = dnorm, args = list(mean = mean(students\_data$math.score), sd = sd(students\_data$math.score)), color = "blue") +

labs(title = "Math Probability Distribution", x = "Math Score", y = "Density") +

theme\_minimal()

})

# Probability Distribution for reading scores

output$reading\_distribution <- renderPlot({

ggplot(students\_data, aes(x = reading.score)) +

geom\_histogram(aes(y = ..density..), binwidth = 5, fill = "red", color = "black", alpha = 0.7) +

stat\_function(fun = dnorm, args = list(mean = mean(students\_data$reading.score), sd = sd(students\_data$reading.score)), color = "blue") +

labs(title = "reading Probability Distribution", x = "reading Score", y = "Density") +

theme\_minimal()})

# Probability Distribution for writing scores

output$writing\_distribution <- renderPlot({

ggplot(students\_data, aes(x = writing.score)) +

geom\_histogram(aes(y = ..density..), binwidth = 5, fill = "red", color = "black", alpha = 0.7) +

stat\_function(fun = dnorm, args = list(mean = mean(students\_data$writing.score), sd = sd(students\_data$writing.score)), color = "blue") +

labs(title = "writing Probability Distribution", x = "writing Score", y = "Density") +

theme\_minimal()})

# Regression model for math scores

output$math\_regression <- renderPlot({

ggplot(students\_data, aes(x = reading.score, y = math.score)) +

geom\_point() +

geom\_smooth(method = "lm", se = FALSE, color = "red") +

labs(title = "Regression Model for Math Scores", x = "Reading Score", y = "Math Score") +

theme\_minimal()

})

# Regression model for reading scores

output$reading\_regression <- renderPlot({

ggplot(students\_data, aes(x = writing.score, y = reading.score)) +

geom\_point() +

geom\_smooth(method = "lm", se = FALSE, color = "green") +

labs(title = "Regression Model for Reading Scores", x = "Writing Score", y = "Reading Score") +

theme\_minimal()

})

# Regression model for writing scores

output$writing\_regression <- renderPlot({

ggplot(students\_data, aes(x = math.score, y = writing.score)) +

geom\_point() +

geom\_smooth(method = "lm", se = FALSE, color = "blue") +

labs(title = "Regression Model for Writing Scores", x = "Math Score", y = "Writing Score") +

theme\_minimal()

})

# Calculate and display mean and mode for math scores

output$mean\_mode\_text\_math <- renderText({

mean\_value <- mean(students\_data$math.score)

mode\_value <- as.numeric(names(table(students\_data$math.score)[table(students\_data$math.score) == max(table(students\_data$math.score))]))

paste("Mean Math Score: ", round(mean\_value, 2),

" Mode Math Score: ", mode\_value)

})

# Calculate and display mean and mode for reading scores

output$mean\_mode\_text\_reading <- renderText({

mean\_value <- mean(students\_data$reading.score)

mode\_value <- as.numeric(names(table(students\_data$reading.score)[table(students\_data$reading.score) == max(table(students\_data$reading.score))]))

paste("Mean Reading Score: ", round(mean\_value, 2),

" Mode Reading Score: ", mode\_value)

})

# Calculate and display mean and mode for writing scores

output$mean\_mode\_text\_writing <- renderText({

mean\_value <- mean(students\_data$writing.score)

mode\_value <- as.numeric(names(table(students\_data$writing.score)[table(students\_data$writing.score) == max(table(students\_data$writing.score))]))

paste("Mean Writing Score: ", round(mean\_value, 2),

" Mode Writing Score: ", mode\_value)

})

# Render a boxplot for math scores

output$math\_boxplot <- renderPlot({

ggplot(students\_data, aes(x = 1, y = math.score)) +

geom\_boxplot(fill = "red", color = "black") +

labs(title = "Math Scores Boxplot", x = "", y = "Math Score") +

theme\_minimal()

})

# Render a boxplot for reading scores

output$reading\_boxplot <- renderPlot({

ggplot(students\_data, aes(x = 1, y = reading.score)) +

geom\_boxplot(fill = "green", color = "black") +

labs(title = "Reading Scores Boxplot", x = "", y = "Reading Score") +

theme\_minimal()

})

# Render a boxplot for writing scores

output$writing\_boxplot <- renderPlot({

ggplot(students\_data, aes(x = 1, y = writing.score)) +

geom\_boxplot(fill = "blue", color = "black") +

labs(title = "Writing Scores Boxplot", x = "", y = "Writing Score") +

theme\_minimal()

})

#scatter plot for male and female

output$scatter\_plot <- renderPlot({

ggplot(students\_data, aes(x = reading.score, y = writing.score, color = gender)) +

geom\_point() +

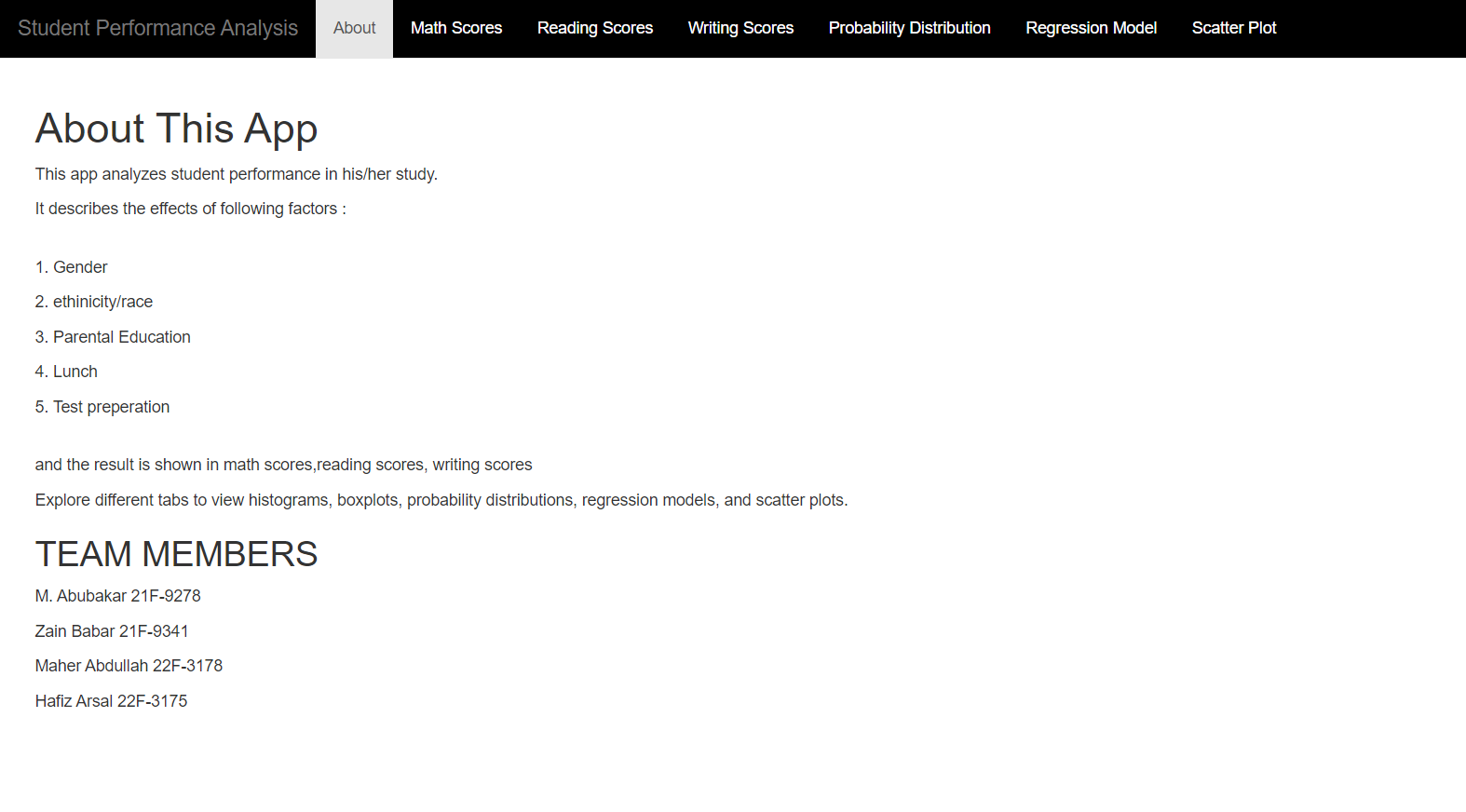
labs(title = "Reading Scores vs. Writing Scores", x = "Reading Score", y = "Writing Score", color = "Gender") +

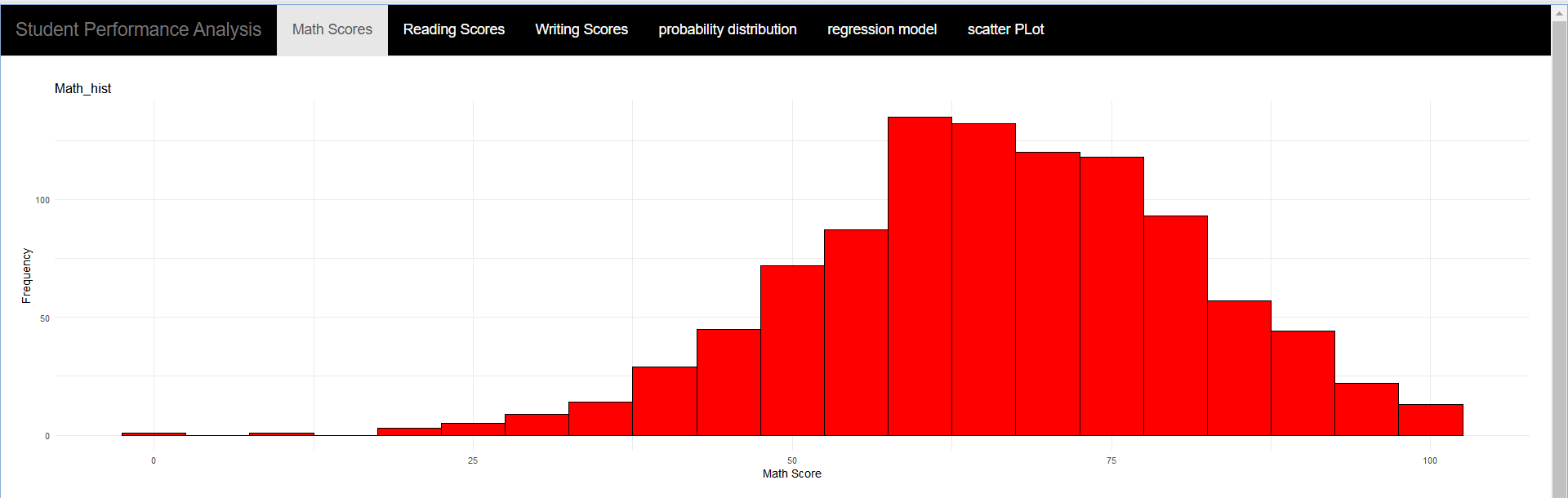
theme\_minimal()

})

}

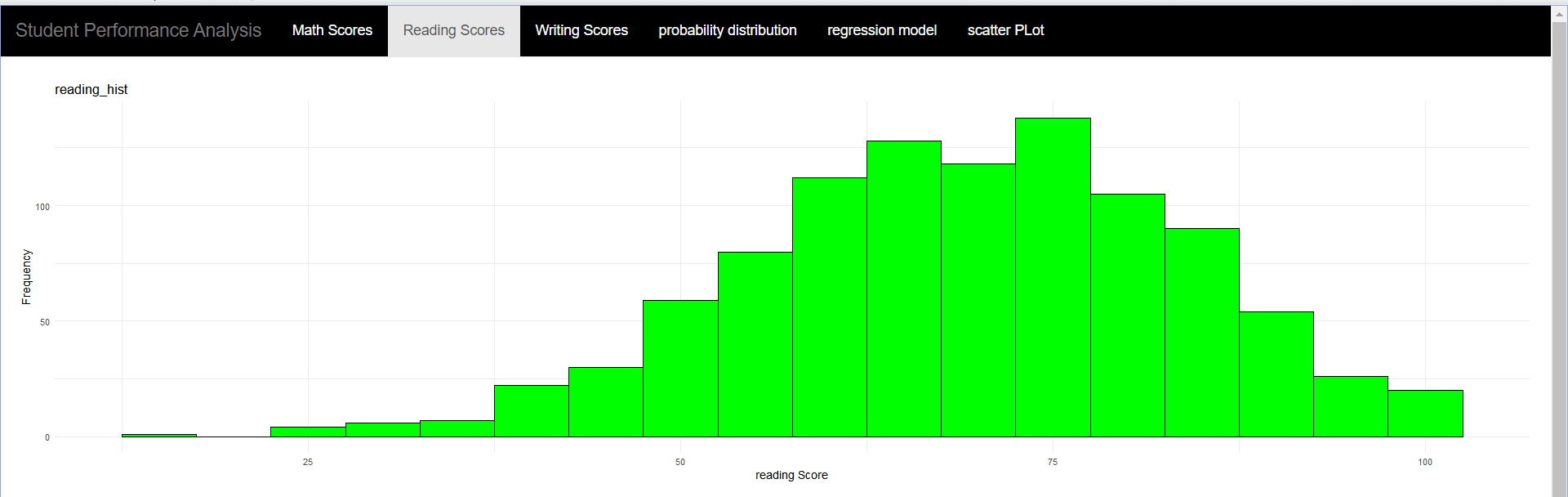
**Screenshots :**

****

****

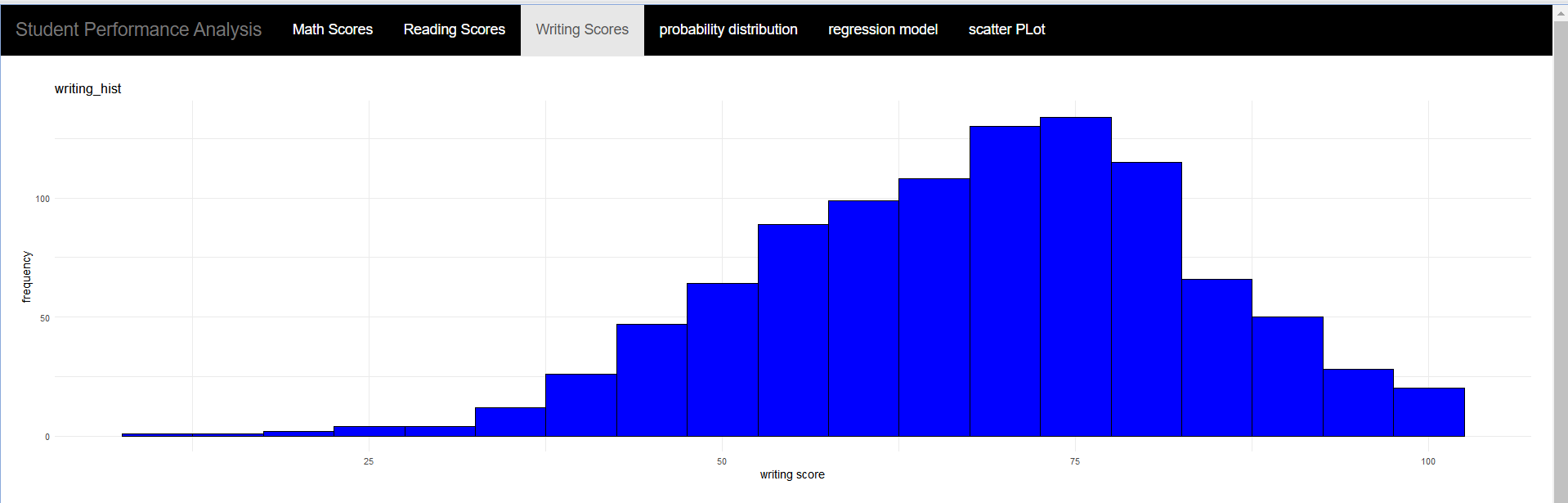
**A red line on a white background

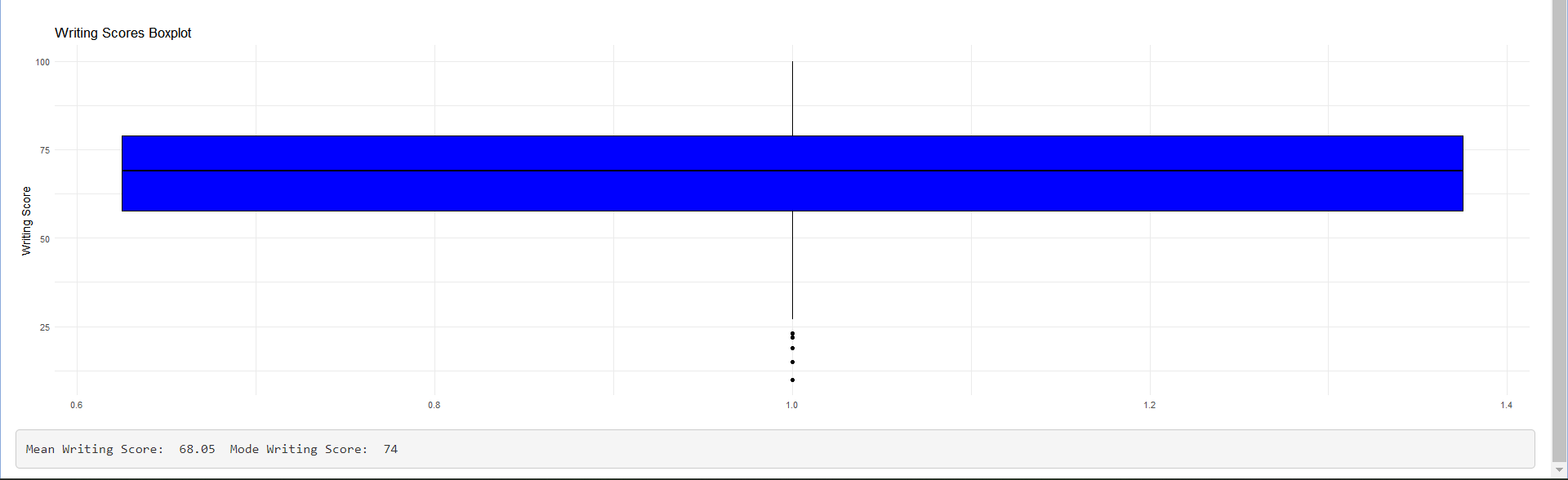
Description automatically generated**

****

**A green line on a white background

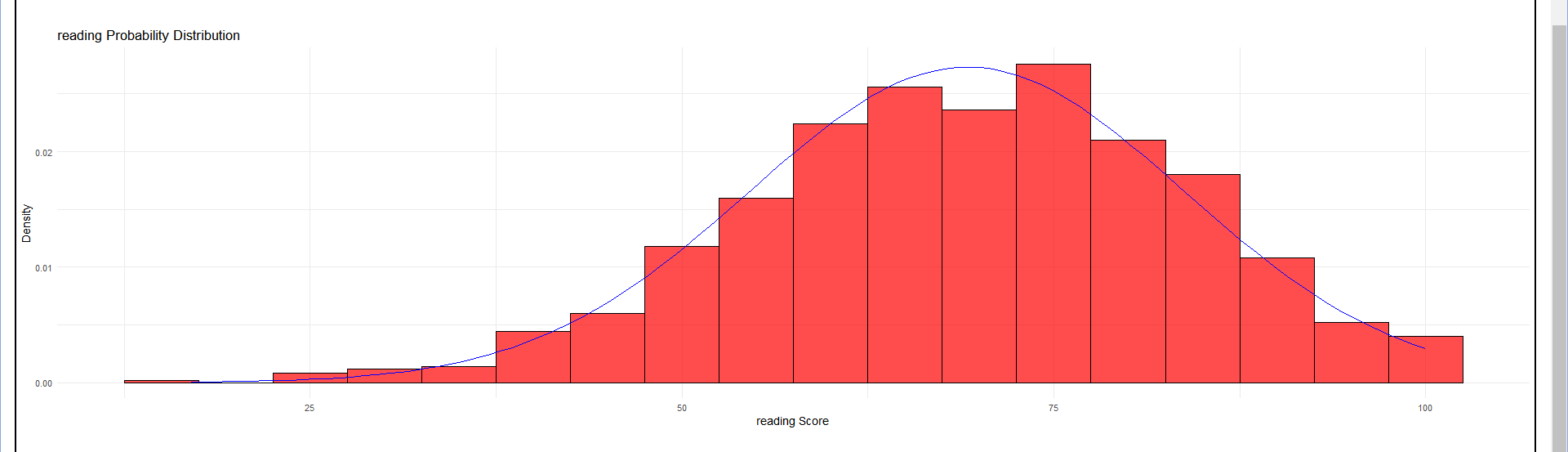
Description automatically generated**

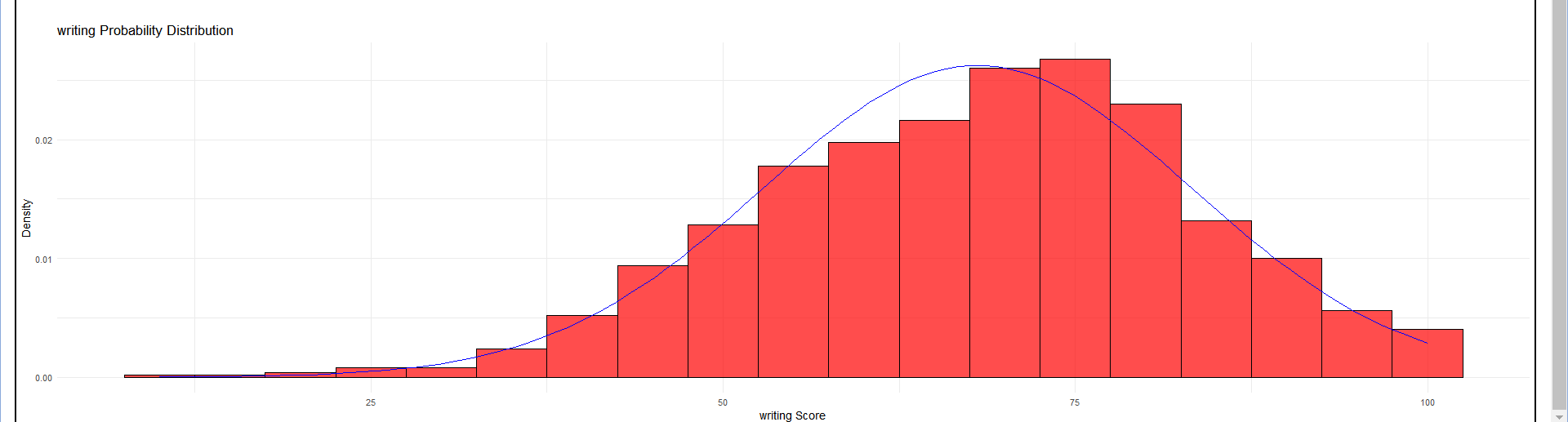
****

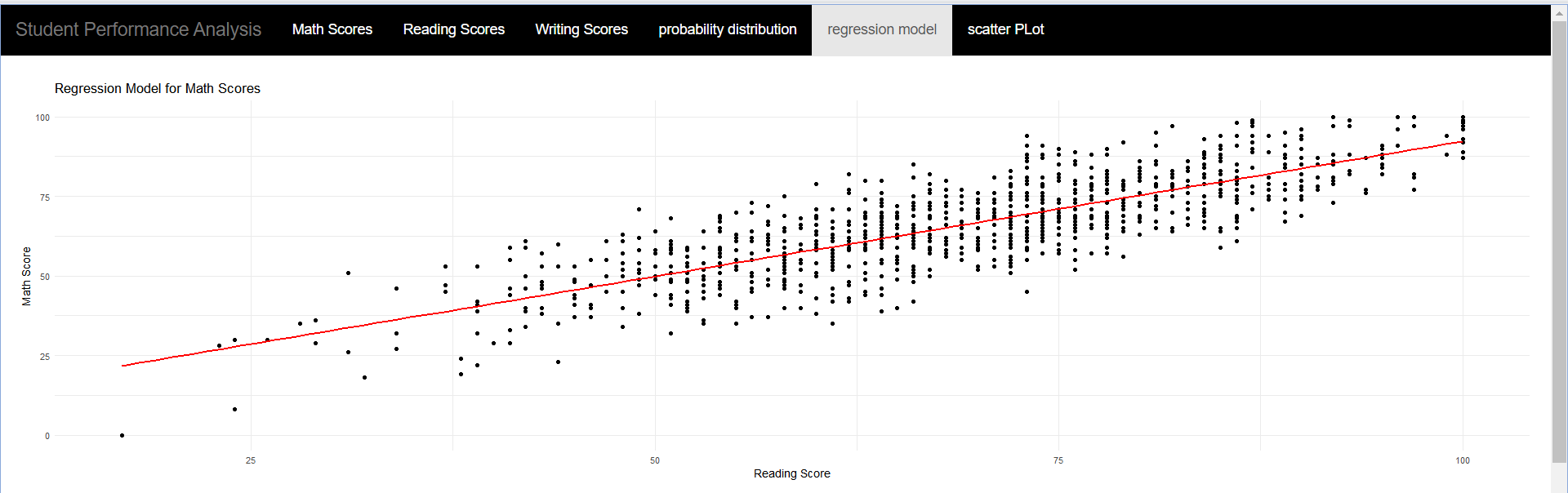
****

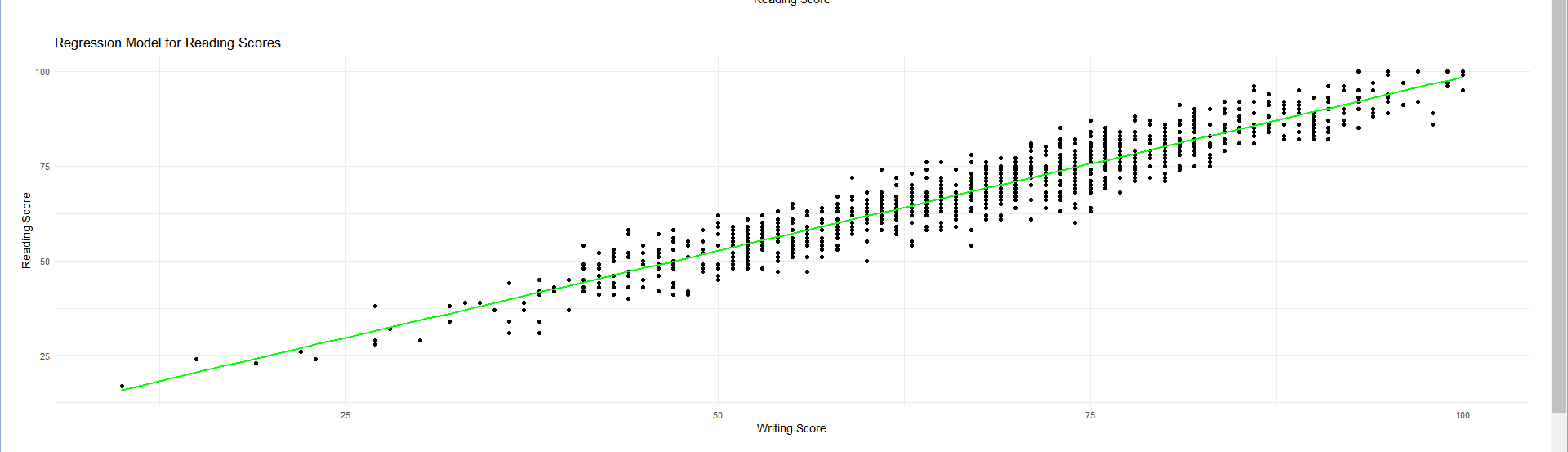
**A graph of a graph

Description automatically generated with medium confidence**

****

****

****

****

**A graph showing a line of dots

Description automatically generated**

**A graph with red and blue dots

Description automatically generated**