# Student Performance Explorer

A Comprehensive Data Analysis Application



|  |  |  |
| --- | --- | --- |
| **ROLL NO** | **NAME** | **SECTION** |
| 22F-8803 | ALI HAIDER | CS-5C |

# 1. Problem Statement

The primary issue addressed in this project is the understanding and exploration of student performance data. By analyzing scores in Mathematics, Reading, and Writing, this application aims to uncover trends and correlations in academic performance.

# 2. Objective

The objective of this work is to create an interactive Shiny application that provides visual and statistical insights into student performance across different subjects. This includes identifying trends, visualizing distributions, and analyzing relationships between scores using regression models.

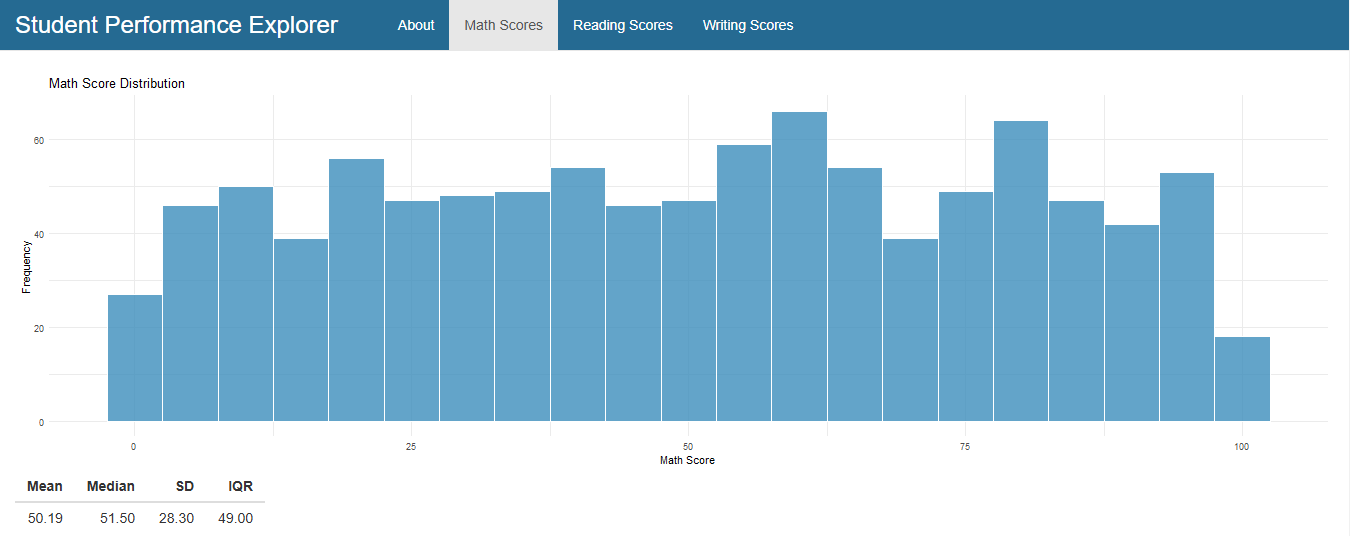
# 3. Data Description

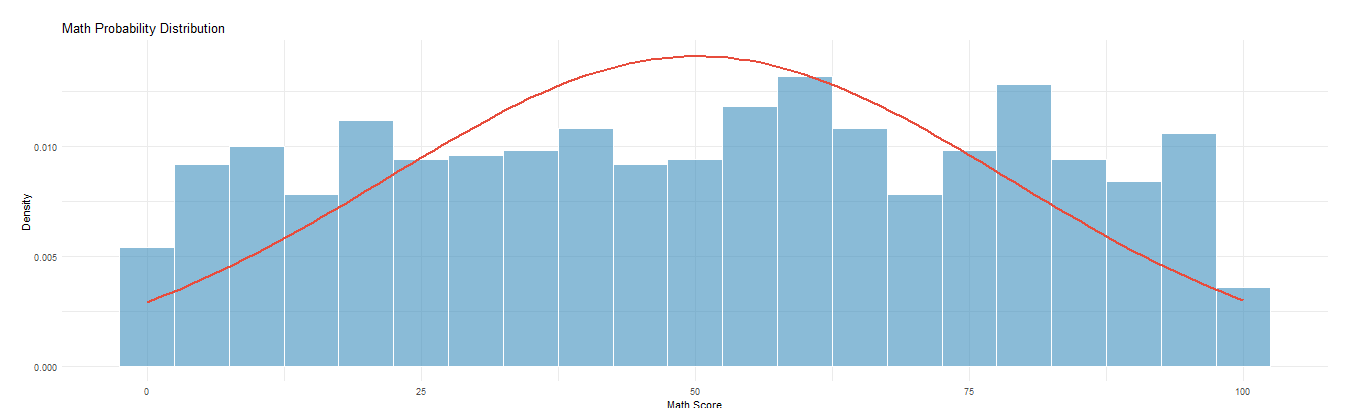
The data used in this project is sourced from a dataset titled "Modified\_StudentsPerformance.csv". It contains the following fields:  
- Gender  
- Race/Ethnicity  
- Parental Level of Education  
- Lunch Type  
- Test Preparation Course  
- Scores in Math, Reading, and Writing

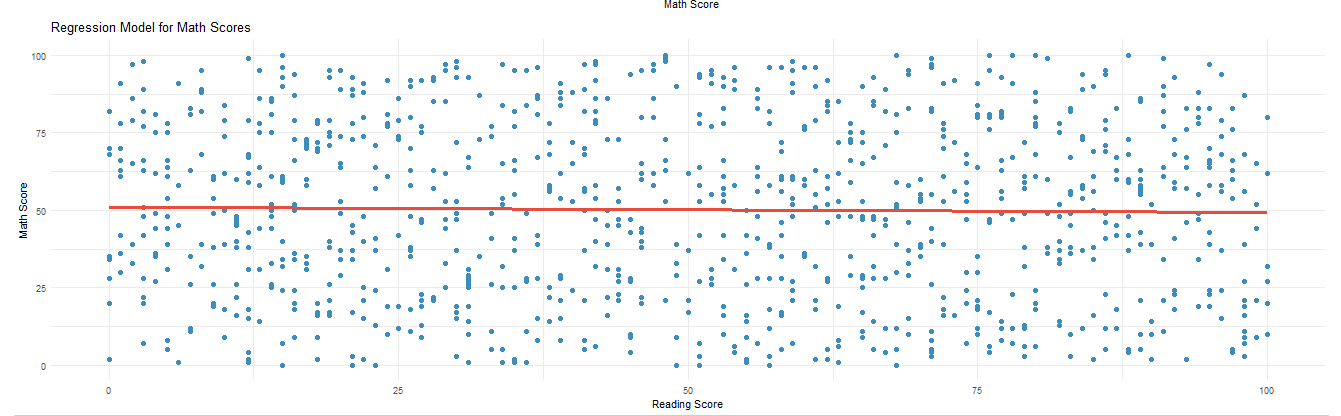
# 4. Results

## Math Scores

Below are the results for Math Scores:

1. Histogram: Visualizes the frequency distribution of Math Scores.  


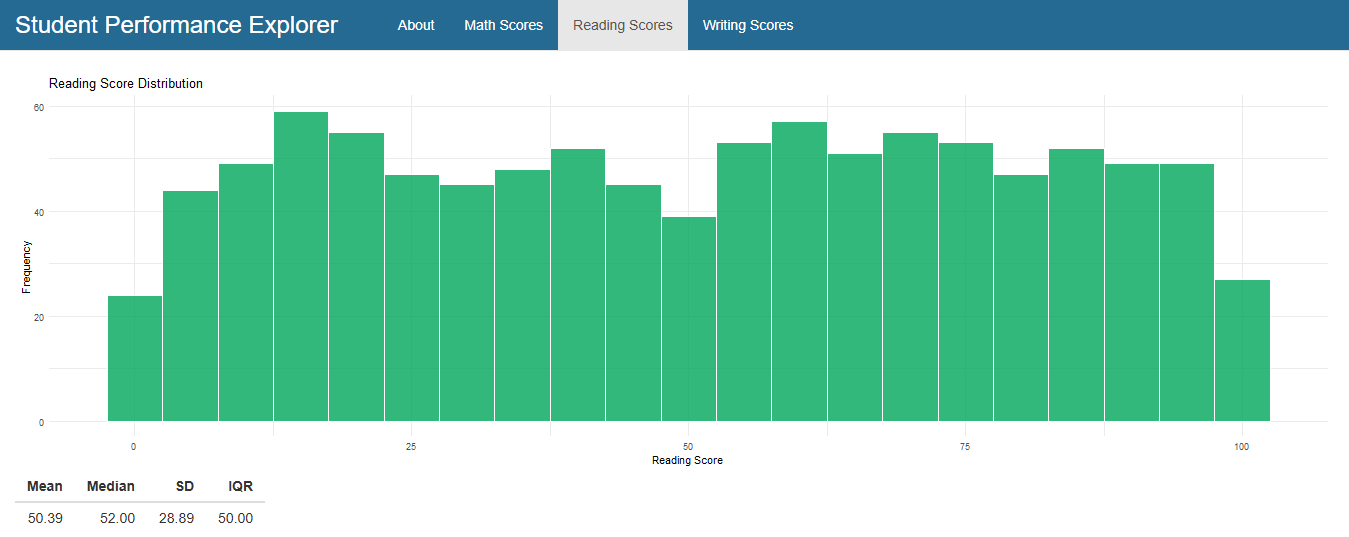
2. Probability Distribution: Shows the density of scores with a normal distribution overlay.  


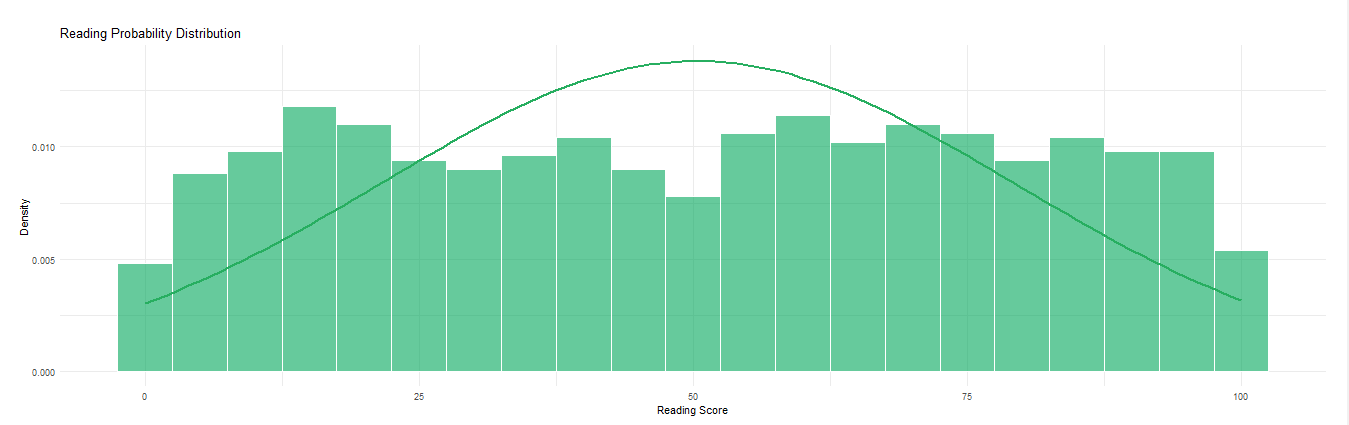
3. Regression Model: Examines the relationship between Math and Reading Scores.  


4. Confidence Intervals: Provides the confidence intervals for the regression model.  

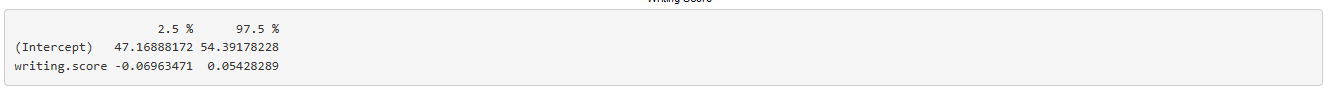

## Reading Scores

Below are the results for Reading Scores:

1. Histogram: Visualizes the frequency distribution of Reading Scores.  


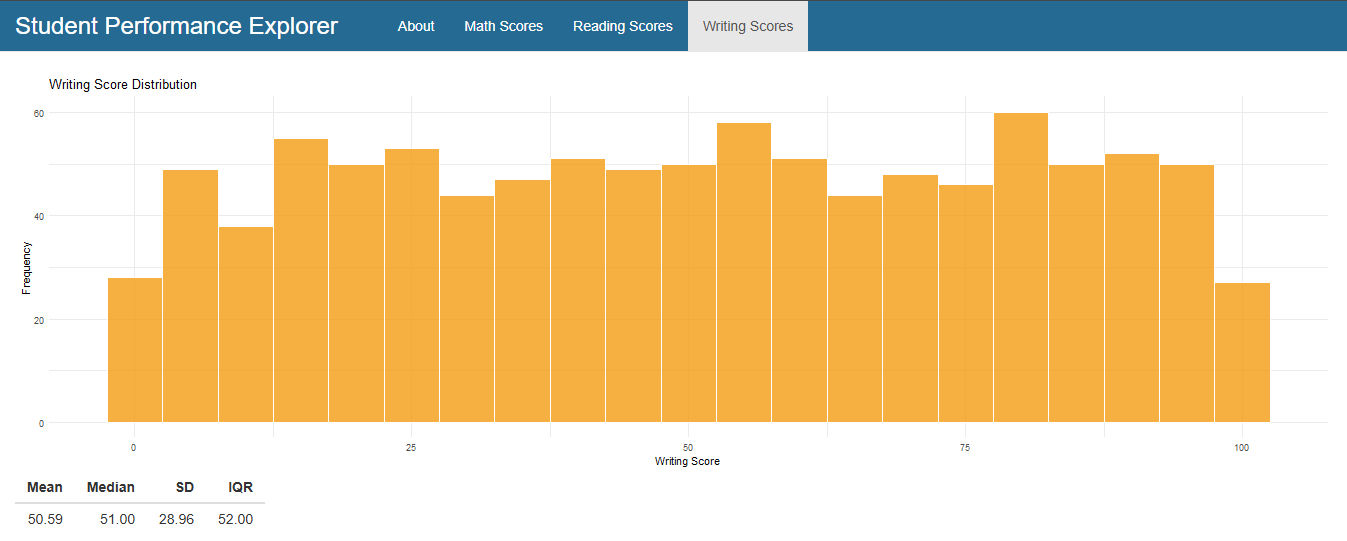
2. Probability Distribution: Shows the density of scores with a normal distribution overlay.  


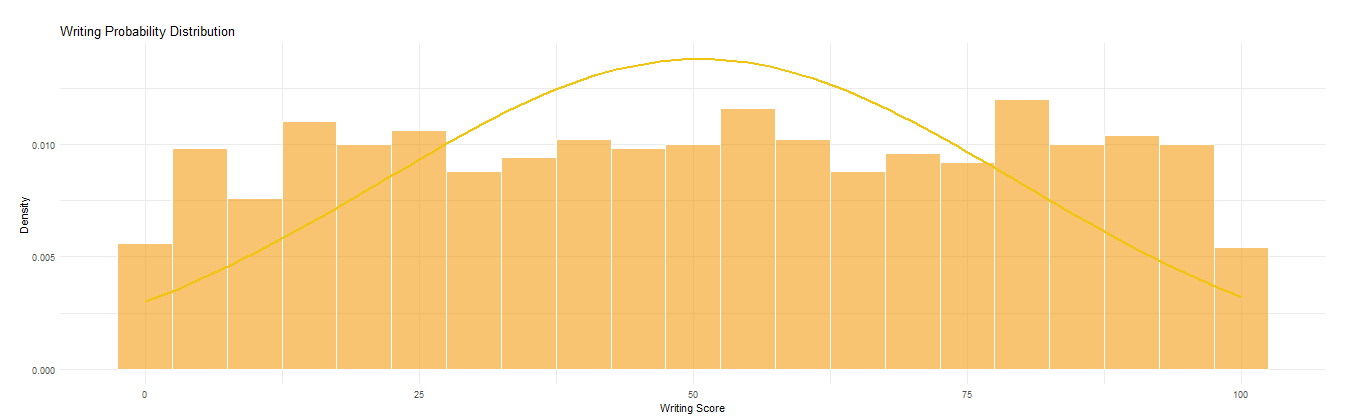
3. Regression Model: Examines the relationship between Reading and Writing Scores.  

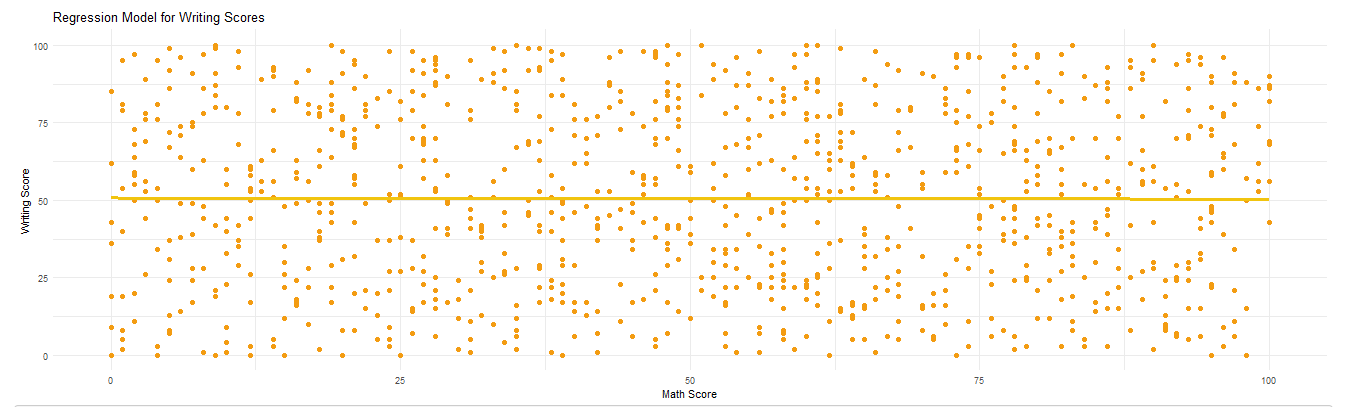

4. Confidence Intervals: Provides the confidence intervals for the regression model.  


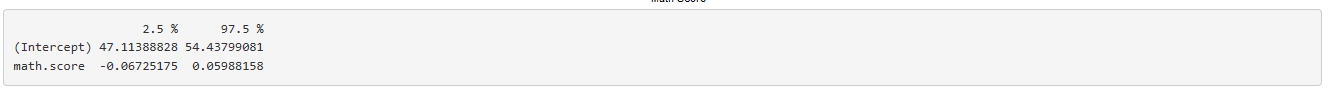
## Writing Scores

Below are the results for Writing Scores:

1. Histogram: Visualizes the frequency distribution of Writing Scores.  


2. Probability Distribution: Shows the density of scores with a normal distribution overlay.  


3. Regression Model: Examines the relationship between Writing and Math Scores.  


4. Confidence Intervals: Provides the confidence intervals for the regression model.  


# 5. Codes

The code for the application is provided below. It includes both server-side and user-interface definitions. Proper comments are included to make the logic clear.

Server Code:

library(shiny)  
library(dplyr)  
library(ggplot2)  
library(MASS) # For confidence interval calculation  
  
# Read data from CSV file  
students\_data <- read.csv("Modified\_StudentsPerformance.csv")  
  
# Define server logic required to analyze the data  
server <- function(input, output, session) {  
   
 # Summary Statistics: mean, median, sd, IQR, etc.  
 descriptive\_stats <- function(data) {  
 data\_summary <- data.frame(  
 Mean = mean(data),  
 Median = median(data),  
 SD = sd(data),  
 IQR = IQR(data)  
 )  
 return(data\_summary)  
 }  
   
 # Render a histogram of math scores  
 output$math\_histogram <- renderPlot({  
 ggplot(students\_data, aes(x = math.score)) +  
 geom\_histogram(binwidth = 5, fill = "#3c8dbc", color = "white", alpha = 0.8) +  
 labs(title = "Math Score Distribution", x = "Math Score", y = "Frequency") +  
 theme\_minimal()  
 })  
   
 # Tabular representation of Math Data  
 output$math\_table <- renderTable({  
 descriptive\_stats(students\_data$math.score)  
 })  
   
 # Probability Distribution for math scores  
 output$math\_distribution <- renderPlot({  
 ggplot(students\_data, aes(x = math.score)) +  
 geom\_histogram(aes(y = ..density..), binwidth = 5, fill = "#3c8dbc", color = "white", alpha = 0.6) +  
 stat\_function(fun = dnorm, args = list(mean = mean(students\_data$math.score), sd = sd(students\_data$math.score)), color = "#e74c3c", size = 1) +  
 labs(title = "Math Probability Distribution", x = "Math 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(color = "#3c8dbc", size = 2) +  
 geom\_smooth(method = "lm", se = FALSE, color = "#e74c3c", size = 1.2) +  
 labs(title = "Regression Model for Math Scores", x = "Reading Score", y = "Math Score") +  
 theme\_minimal()  
 })  
   
 # Confidence interval for math score regression model  
 output$math\_confidence <- renderPrint({  
 model <- lm(math.score ~ reading.score, data = students\_data)  
 confint(model)  
 })  
   
 # Repeat the above process for Reading and Writing Scores  
   
 # Reading Histogram  
 output$reading\_histogram <- renderPlot({  
 ggplot(students\_data, aes(x = reading.score)) +  
 geom\_histogram(binwidth = 5, fill = "#00a65a", color = "white", alpha = 0.8) +  
 labs(title = "Reading Score Distribution", x = "Reading Score", y = "Frequency") +  
 theme\_minimal()  
 })  
   
 output$reading\_table <- renderTable({  
 descriptive\_stats(students\_data$reading.score)  
 })  
   
 output$reading\_distribution <- renderPlot({  
 ggplot(students\_data, aes(x = reading.score)) +  
 geom\_histogram(aes(y = ..density..), binwidth = 5, fill = "#00a65a", color = "white", alpha = 0.6) +  
 stat\_function(fun = dnorm, args = list(mean = mean(students\_data$reading.score), sd = sd(students\_data$reading.score)), color = "#27ae60", size = 1) +  
 labs(title = "Reading Probability Distribution", x = "Reading Score", y = "Density") +  
 theme\_minimal()  
 })  
   
 output$reading\_regression <- renderPlot({  
 ggplot(students\_data, aes(x = writing.score, y = reading.score)) +  
 geom\_point(color = "#00a65a", size = 2) +  
 geom\_smooth(method = "lm", se = FALSE, color = "#27ae60", size = 1.2) +  
 labs(title = "Regression Model for Reading Scores", x = "Writing Score", y = "Reading Score") +  
 theme\_minimal()  
 })  
   
 output$reading\_confidence <- renderPrint({  
 model <- lm(reading.score ~ writing.score, data = students\_data)  
 confint(model)  
 })  
   
 # Writing Histogram  
 output$writing\_histogram <- renderPlot({  
 ggplot(students\_data, aes(x = writing.score)) +  
 geom\_histogram(binwidth = 5, fill = "#f39c12", color = "white", alpha = 0.8) +  
 labs(title = "Writing Score Distribution", x = "Writing Score", y = "Frequency") +  
 theme\_minimal()  
 })  
   
 output$writing\_table <- renderTable({  
 descriptive\_stats(students\_data$writing.score)  
 })  
   
 output$writing\_distribution <- renderPlot({  
 ggplot(students\_data, aes(x = writing.score)) +  
 geom\_histogram(aes(y = ..density..), binwidth = 5, fill = "#f39c12", color = "white", alpha = 0.6) +  
 stat\_function(fun = dnorm, args = list(mean = mean(students\_data$writing.score), sd = sd(students\_data$writing.score)), color = "#f1c40f", size = 1) +  
 labs(title = "Writing Probability Distribution", x = "Writing Score", y = "Density") +  
 theme\_minimal()  
 })  
   
 output$writing\_regression <- renderPlot({  
 ggplot(students\_data, aes(x = math.score, y = writing.score)) +  
 geom\_point(color = "#f39c12", size = 2) +  
 geom\_smooth(method = "lm", se = FALSE, color = "#f1c40f", size = 1.2) +  
 labs(title = "Regression Model for Writing Scores", x = "Math Score", y = "Writing Score") +  
 theme\_minimal()  
 })  
   
 output$writing\_confidence <- renderPrint({  
 model <- lm(writing.score ~ math.score, data = students\_data)  
 confint(model)  
 })  
}

UI Code:

library(shiny)  
library(shinydashboard)  
library(shinyjs)  
  
# Define UI for application  
ui <- navbarPage(  
 title = "Student Performance Explorer",  
   
 tags$style(HTML("  
 .navbar { background-color: #256a92; }  
 .navbar-default .navbar-nav > li > a { color: #ffffff; }  
 .navbar-header .navbar-brand { color: #ffffff !important; font-size: 24px; }  
   
 .about-page {  
 background-color: #f4f7f8;  
 color: #333333;  
 display: flex;  
 flex-direction: column;  
 justify-content: center;  
 align-items: center;  
 text-align: center;  
 padding: 40px 20px;  
 border-radius: 10px;  
 box-shadow: 0 4px 8px rgba(0, 0, 0, 0.1);  
 }  
 .about-page h1 {  
 color: #256a92;  
 font-size: 36px;  
 font-weight: bold;  
 margin-bottom: 20px;  
 }  
 .about-page p {  
 font-size: 18px;  
 line-height: 1.8;  
 margin: 15px 0;  
 color: #555555;  
 max-width: 800px;  
 text-align: left;  
 }  
 .about-page h2 {  
 font-size: 22px;  
 margin-top: 30px;  
 color: #256a92;  
 }  
 .about-page a {  
 color: #256a92;  
 text-decoration: none;  
 font-weight: bold;  
 }  
 .about-page a:hover {  
 text-decoration: underline;  
 }  
 ")),  
   
 tabPanel("About",  
 fluidPage(  
 div(class = "about-page",  
 h1("About This Application"),  
 p("This application provides a comprehensive analysis of student performance across various academic subjects."),  
 p("The analysis considers the following factors: Gender, Ethnicity/Race, Parental Education Level, Lunch Type, and Test Preparation Course."),  
 h2("Developed By"),  
 p("Ali Haider (Student ID: 22F-8803)")  
 )  
 )  
 ),  
   
 tabPanel("Math Scores",   
 plotOutput("math\_histogram"),  
 tableOutput("math\_table"),  
 plotOutput("math\_distribution"),  
 plotOutput("math\_regression"),  
 verbatimTextOutput("math\_confidence")  
 ),  
   
 tabPanel("Reading Scores",   
 plotOutput("reading\_histogram"),  
 tableOutput("reading\_table"),  
 plotOutput("reading\_distribution"),  
 plotOutput("reading\_regression"),  
 verbatimTextOutput("reading\_confidence") # <-- Closed the parentheses here  
 ),  
   
 tabPanel("Writing Scores",   
 plotOutput("writing\_histogram"),  
 tableOutput("writing\_table"),  
 plotOutput("writing\_distribution"),  
 plotOutput("writing\_regression"),  
 verbatimTextOutput("writing\_confidence")  
 )  
)

# 6. Conclusion

This project successfully demonstrates the use of Shiny for interactive data exploration. It provides meaningful insights into student performance, leveraging statistical analysis and visualizations to highlight key patterns and relationships between subjects.