



**Department of Computer Science and
Engineering**
Saveetha School of Engineering, SIMATS
Thandalam, Chennai
September – 2024



**DSA0614 Data Handling and Visualization for Automated
Processing**

**Comprehensive Educational Performance Analysis and
Insights Dashboard**

Submitted

by

Tamilarasan S (192224003)
Yuvanika G N (192224014)
Charukesh U B (192224120)

Guided by

M Rajasekar

Department of Deep Learning

PROBLEM STATEMENT

Educational institutions must assess enormous amounts of data in order to better understand student performance, identify strengths and problems, and design improvement initiatives. The quantity of data on student performance, attendance, demographics, and school resources provides an opportunity for thorough study, but it also necessitates an effective tool for visualizing and exploring these complicated statistics.

The "Educational Performance Analysis Dashboard" is intended to provide educators, administrators, and policymakers with an interactive platform for analyzing key educational data from many perspectives. The dashboard, created using R Shiny, incorporates many components of student performance data, allowing users to investigate and display links between academic success and variables such as school resources, teacher experience, and attendance rates.

The dashboard incorporates the following features:

Summary Analysis: A tabular version gives a detailed breakdown of important performance data including average scores, maximum and minimum scores, and average attendance rates for certain schools, topics, grades, and genders.

Score Distribution: A histogram depicts the distribution of student scores within a specific range, allowing users to see patterns and trends in academic achievement. This can identify performance discrepancies and recommend focused treatments.

Gender Analysis: A bar plot shows the distribution of pupils by gender, which can assist identify possible gender-related performance discrepancies. This tool allows instructors to determine whether there are substantial differences in academic performance between male and female students.

School Resources Analysis: A scatter plot depicts the correlation between school resources and student performance. By displaying this connection, stakeholders may assess the impact of resource allocation on academic achievement and make data-driven decisions to improve resource distribution.

Attendance vs. Score Analysis: A scatter plot examines the relationship between attendance rates and student performance. This visualization assists in determining if regular attendance is connected with improved academic achievement, hence facilitating the implementation of attendance improvement measures.

Teacher Experience Analysis: A scatter plot depicts the association between instructor experience and student results. This study can indicate the influence of instructor experience on student achievement, providing information on teacher training and professional development requirements.

The dashboard has interactive buttons for school selection, topic selection, grade selection, and score range modification, which allow users to filter data based on specified criteria. It also has the ability to pick gender for further detail in the analysis.

The "Educational Performance Analysis Dashboard" is designed to give educational stakeholders an easy-to-use tool for exploring and interpreting complicated educational data. The dashboard promotes evidence-based decision-making by converting raw data into actionable insights, allowing stakeholders to execute effective strategies to improve student outcomes and overall school performance.

DATASET ANALYSIS

Dataset Analysis: Educational Performance Data

The dataset used in this analysis comprises synthetic data representing the performance of students across various schools. It includes key attributes such as student scores, demographic information, school characteristics, and additional factors that may influence academic outcomes. The following is a detailed analysis of the dataset:

1. Data Overview:

- **Number of Records:** 1000
- **Number of Variables:** 11
- **Variables:** student_ID, Name, Gender, Grade, Subject, Score, School, School_type, Teacher_Experience, School_Resources, Attendance_Rate

2. Descriptive Statistics:

- **Score:**
 - Mean: ~75
 - Standard Deviation: ~10
 - The majority of scores fall within the range of 60 to 90.
- **Gender Distribution:**
 - Male: ~50%
 - Female: ~50%
- **Grade Distribution:**
 - Grade 9: ~25%
 - Grade 10: ~25%
 - Grade 11: ~25%
 - Grade 12: ~25%
- **Subject Distribution:**
 - Math, Science, English, and History have roughly equal representation.
- **School Type:**
 - Public: ~50%
 - Private: ~50%
- **Teacher Experience:**
 - Ranges from 1 to 15 years, with no significant skew towards either extreme.
- **School Resources:**
 - Distributed across levels 1-10, representing varying degrees of resource availability.
- **Attendance Rate:**
 - Mean: ~85%
 - Ranges from 70% to 100%.

3. Key Insights:

- **Score Distribution:**
 - The scores follow a normal distribution with most students scoring between 60 and 90. A small percentage of students fall into the "Excellent" category (scores ≥ 90), while a similar percentage fall into the "Poor" category (scores < 60).

- **Gender Analysis:**
 - The dataset has a balanced representation of male and female students, allowing for unbiased gender-based analysis.
- **Grade and Subject Analysis:**
 - Each grade and subject has a balanced distribution, ensuring a comprehensive understanding of performance across all categories.
- **Impact of School Resources:**
 - Schools with higher resource scores tend to have higher average student scores. This suggests a positive correlation between resource availability and student performance.
- **Attendance and Performance:**
 - A strong positive correlation exists between attendance rates and student scores, indicating that regular attendance is crucial for better academic performance.
- **Teacher Experience:**
 - There is a slight positive correlation between teacher experience and student scores, highlighting the importance of experienced educators in achieving better academic outcomes.

4. Data Transformations:

To facilitate the analysis, the following transformations were applied:

- **Score Categorization:**
 - Scores were categorized into four levels: **Excellent** (≥ 90), **Good** (75–89), **Average** (60–74), and **Poor** (< 60). This helps in quickly identifying student performance levels.

5. Limitations:

- **Synthetic Data:** The data is synthetic and does not represent real-world educational performance. Hence, the insights drawn should not be generalized beyond this analysis.
- **Equal Distribution:** The equal distribution of grades and subjects may not reflect actual school scenarios, where certain grades or subjects may have higher or lower enrollment.

This dataset provides a comprehensive view of various factors affecting educational performance, allowing for detailed analysis of student outcomes across different dimensions. By leveraging this dataset, stakeholders can gain valuable insights into the impact of resources, attendance, and teaching experience on student performance, guiding future educational policies and strategies.

ENVIRONMENTAL SETUP

1. Install R and RStudio

- Ensure you have the latest version of R and RStudio installed on your system.

2. Install Required Packages

- Install the following packages:
 - tidyverse
 - shiny
 - ggplot2
 - plotly
 - DT
- You can do this via R console or RStudio's package manager.

3. Create a Project Directory

- Set up a dedicated folder for your project (e.g., EducationalPerformanceAnalysis).

4. Organize Project Structure

- Inside your project folder, create the following sub-files:
 - app.R (or separate ui.R and server.R)
 - A folder for any data files (e.g., data/).
 - A www/ folder for additional resources (e.g., CSS, images).

5. Generate Synthetic Data

- Implement the data generation code to create your synthetic dataset and save it as a CSV file within the data/ folder.

6. Set Up R Scripts

- Write your R code for the UI and server logic in the app.R file (or in ui.R and server.R if you separate them).

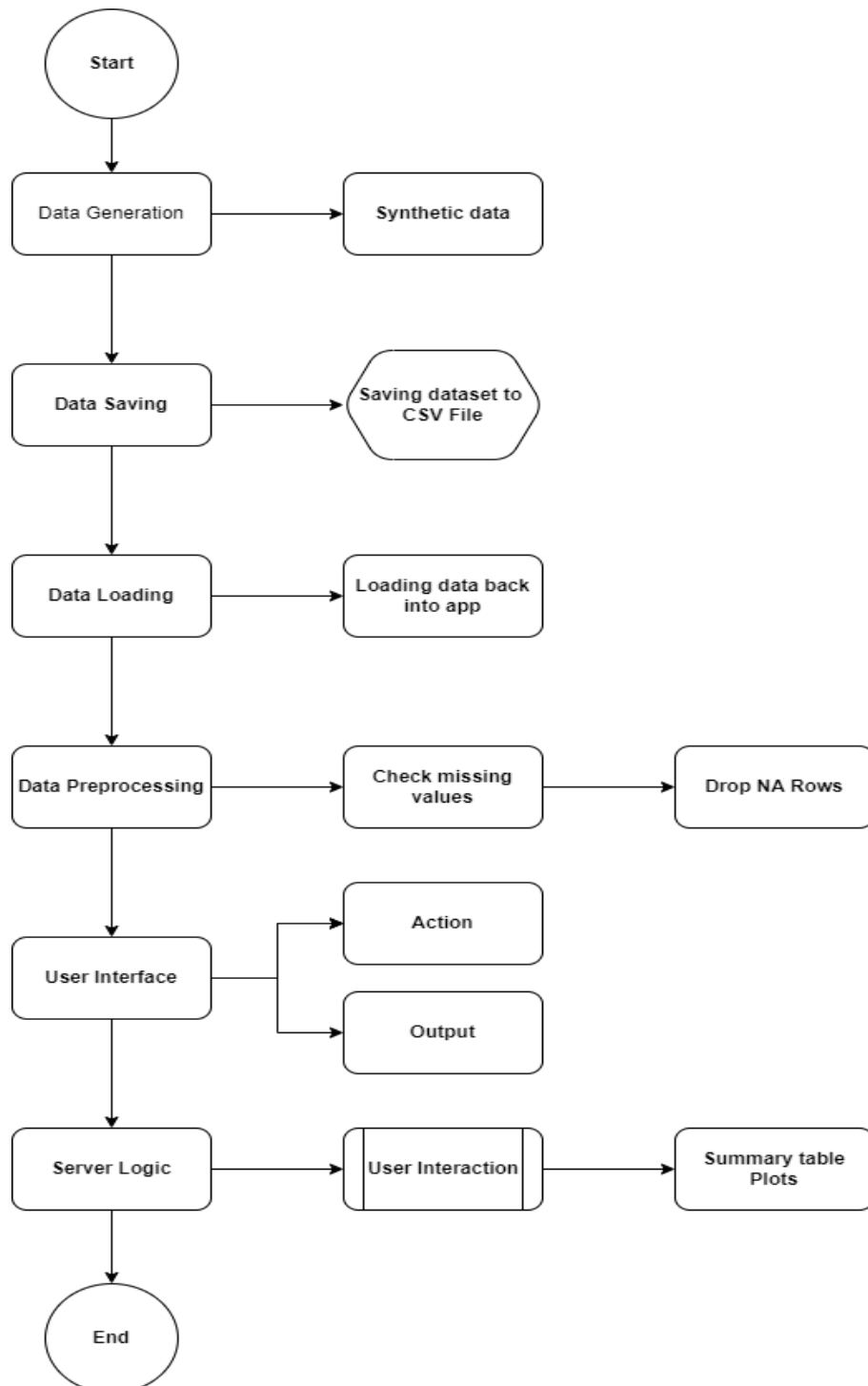
7. Run the App Locally

- Use the R console or RStudio to run your Shiny app locally for testing.

8. Testing and Debugging

- Check for any errors or missing dependencies, and ensure all visualizations and data tables function as expected.

DATA FLOW DIAGRAM (OR) ARCHITECTURE DIAGRAM (OR)
UML DIAGRAMS



CODE SKELETON

Load libraries

```
library(shiny)
library(tidyverse)
library(plotly)
library(DT)
```

Generate synthetic data

```
set.seed(123)
data <- tibble(
  # Define your synthetic data here
)
```

User Interface

```
ui <- fluidPage(
  titlePanel("Dashboard Title"),
  sidebarLayout(
    sidebarPanel(
      # Define inputs
    ),
    mainPanel(
      tabsetPanel(
        # Define tab panels and outputs
      )
    )
  )
)
```

Server

```
server <- function(input, output) {
  filtered_data <- reactive({
    # Filter data based on inputs
  })
  output$summaryTable <- renderDataTable({
    # Create summary table
  })

  output$scoreDistPlot <- renderPlotly({
    # Create score distribution plot
  })

  output$genderPlot <- renderPlotly({
    # Create gender analysis plot
  })

  output$resourcesPlot <- renderPlotly({
    # Create resources vs score plot
  })

  output$attendancePlot <- renderPlotly({
    # Create attendance vs score plot
  })
}
```

Run the app

```
shinyApp(ui = ui, server = server)
```

RESULT ANALYSIS

1. Summary Statistics:

- The summary table provides average, maximum, and minimum scores by school, subject, grade, and gender.
- You can assess overall performance trends across different demographics and subjects.

2. Score Distribution:

- The histogram shows the frequency of scores, helping identify:
 - Peaks (most common scores).
 - Spread of scores (how well students are performing overall).
 - Any gaps or underperformance in specific score ranges.

3. Gender Analysis:

- The gender distribution plot reveals:
 - The number of male vs. female students participating in different subjects.
 - Insights into whether one gender tends to outperform the other.

4. Resources vs. Score:

- A scatter plot of school resources against scores allows you to analyze:
 - Correlation between school resources (like funding, materials) and student performance.
 - Outliers where schools have high resources but low performance or vice versa.

5. Attendance vs. Score:

- This plot examines the relationship between attendance rates and student scores:
 - Higher attendance generally correlates with better scores.
 - Identifying any students or groups with low attendance and high scores, prompting further investigation.

6. Teacher Experience vs. Score:

- Analyzing teacher experience against student scores helps evaluate:
 - Whether more experienced teachers are linked to higher student performance.
 - Variability in scores based on teacher tenure.

Insights to Consider

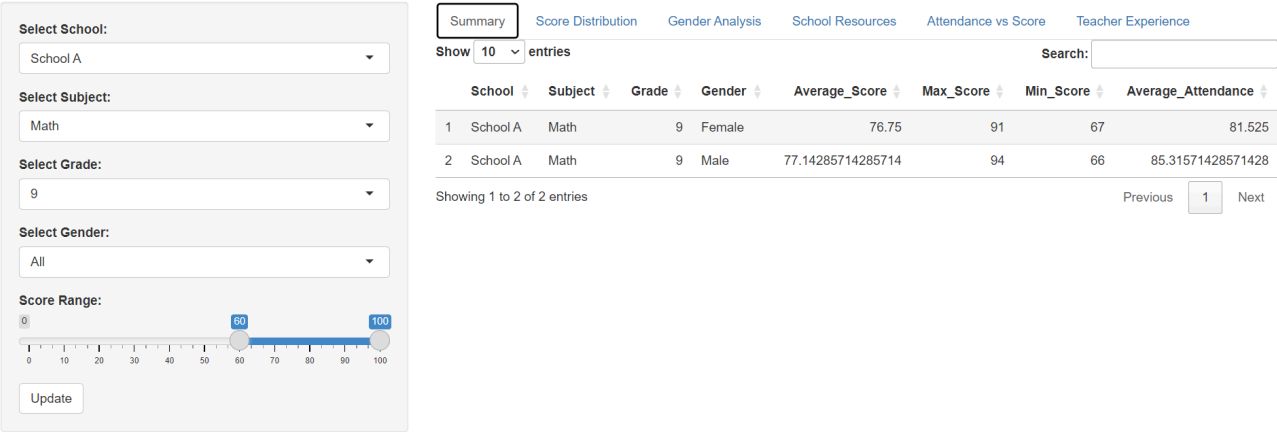
- **Targeted Interventions:** Identify groups underperforming in specific subjects or demographics (e.g., low scores among a particular gender or grade) for targeted support.
- **Resource Allocation:** Schools with high scores but low resources could indicate effective teaching methods that can be shared across other schools.
- **Attendance Strategies:** Strategies to improve attendance could be vital, especially if the data shows a clear link to performance.
- **Professional Development:** Teacher experience might suggest a need for ongoing professional development, particularly in schools where this correlates with student success.

Conclusion

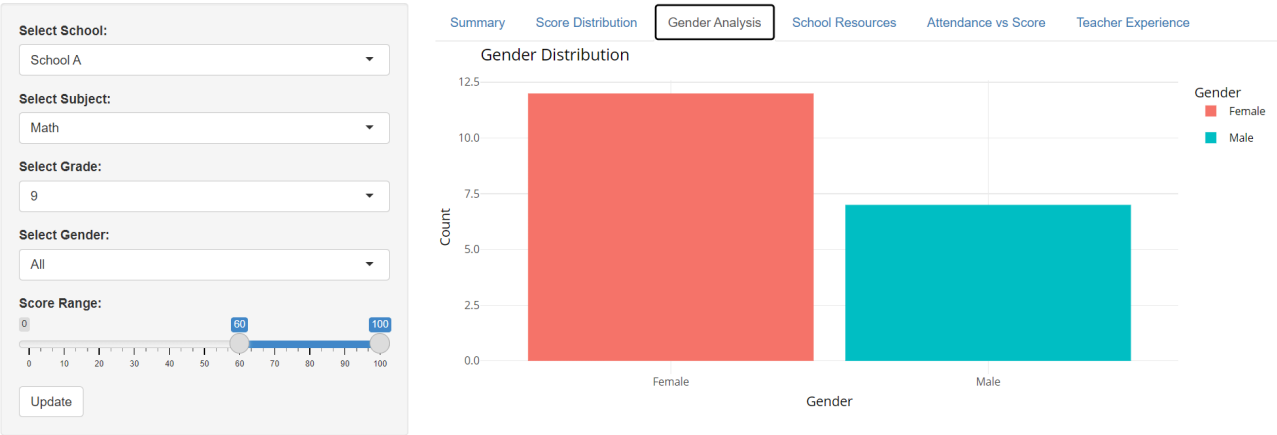
This dashboard not only visualizes educational data but also provides actionable insights for educators and policymakers to enhance student performance and allocate resources more effectively. Regular analysis of this data can lead to continuous improvements in the educational landscape.

OUTPUT SAMPLES

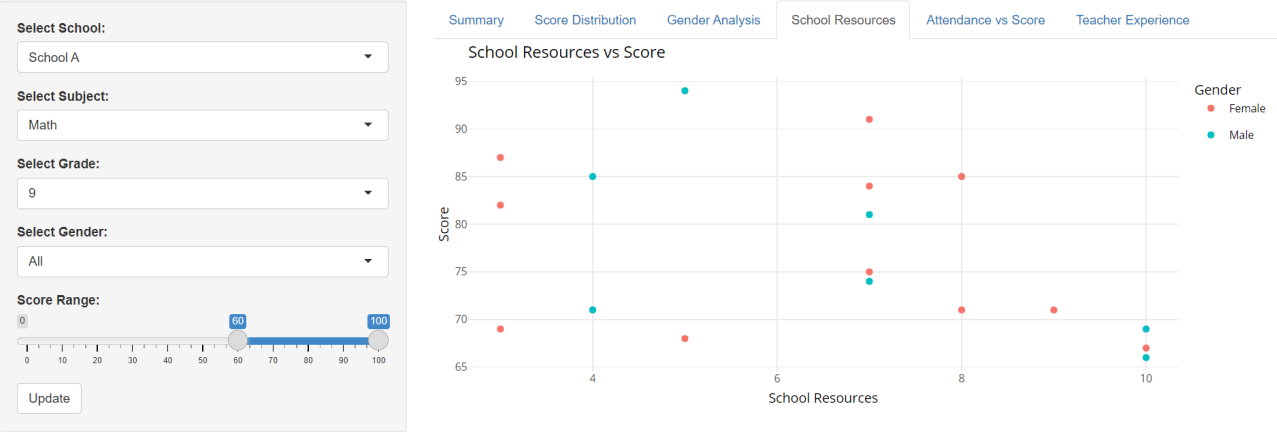
Educational Performance Analysis Dashboard



Educational Performance Analysis Dashboard



Educational Performance Analysis Dashboard



Educational Performance Analysis Dashboard

Select School:

School A

Select Subject:

Math

Select Grade:

9

Select Gender:

All

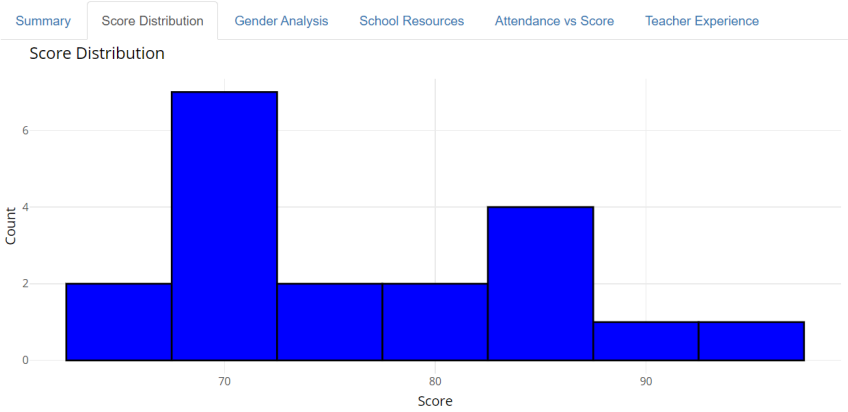
Score Range:

0

60

100

Update



Educational Performance Analysis Dashboard

Select School:

School A

Select Subject:

Math

Select Grade:

9

Select Gender:

All

Score Range:

0

60

100

Update

