# INTERACTIVE TOOL FOR VISUALIZING MARKS DISTRIBUTION IN A CLASS

**A PROJECT REPORT**

***Submitted by***

**B R VIMAL AANANTH (2303811714821058)**

***in partial fulfillment of requirements for the award of the course***

**AGI1252-FUNDAMENTALS OF DATA SCIENCE USING R**

***in***

**ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING**

**K. RAMAKRISHNAN COLLEGE OF TECHNOLOGY**

(An Autonomous Institution, affiliated to Anna University Chennai and Approved by AICTE, New Delhi)

**SAMAYAPURAM – 621 112**

**JUNE 2025**

# K. RAMAKRISHNAN COLLEGE OF TECHNOLOGY (AUTONOMOUS)

**SAMAYAPURAM – 621 112**

## BONAFIDE CERTIFICATE

Certified that this project report on **"INTERACTIVE TOOL FOR VISUALIZING MARKS DISTRIBUTION IN A CLASS"** is the bonafide work of **B R VIMAL AANANTH (2303811714821058)** who carried out the project work during the academic year 2024 - 2025 under my supervision.

**SIGNATURE**

Dr.T.AVUDAIAPPAN, M.E.,Ph.D.,

**HEAD OF THE DEPARTMENT**

ASSOCIATE PROFESSOR

Department of Artificial Intelligence

K.Ramakrishnan College of Technology (Autonomous)

Samayapuram–621112.

**SIGNATURE**

Mrs.E.SRI SANTHOSHINI, M.E.,

**SUPERVISOR**

ASSISTANT PROFESSOR

Department of Artificial Intelligence

K.Ramakrishnan College of Technology (Autonomous)

Samayapuram–621112.

Submitted for the viva-voce examination held on 02.06.2025

**INTERNAL EXAMINER**

**EXTERNAL EXAMINER**

# DECLARATION

I declare that the project report on **"INTERACTIVE TOOL FOR VISUALIZING MARKS DISTRIBUTION IN A CLASS"** is the result of original work done by me and best of my knowledge, similar work has not been submitted to **"ANNA UNIVERSITY CHENNAI"** for the requirement of Degree of **BACHELOR OF ENGINEERING**. This project report is submitted on the partial fulfilment of the requirement of the completion of the course **AGI1252- FUNDAMENTALS OF DATA SCIENCE USING R.**

.

**Signature**

**B R VIMAL AANANTH**

Place: Samayapuram

Date: 02-06-2025

## ACKNOWLEDGEMENT

It is with great pride that I express our gratitude and in-debt to our institution "**K.Ramakrishnan College of Technology (Autonomous)**", for providing us with the opportunity to do this project.

I glad to credit honourable chairman **Dr. K. RAMAKRISHNAN**, **B.E.,** for having provided for the facilities during the course of our study in college.

I would like to express our sincere thanks to our beloved Executive Director **Dr. S. KUPPUSAMY, MBA, Ph.D.,** for forwarding to our project and offering adequate duration in completing our project.

I would like to thank **Dr. N. VASUDEVAN, M.Tech., Ph.D.,** Principal, who gave opportunity to frame the project the full satisfaction.

I whole heartily thanks to **Dr. T. AVUDAIAPPAN, M.E.,Ph.D.,** Head of the department, **ARTIFICIAL INTELLIGENCE** for providing his encourage pursuing this project.

I express our deep expression and sincere gratitude to our project supervisor **Mrs.E.SRI SANTHOSHINI, M.E.,** Department of **ARTIFICIAL INTELLIGENCE,** for her incalculable suggestions, creativity, assistance and patience which motivated us to carry out this project.

I render our sincere thanks to Course Coordinator and other staff members for providing valuable information during the course.

I wish to express our special thanks to the officials and Lab Technicians of our departments who rendered their help during the period of the work progress.

**INSTITUTE**

**Vision:**

- To serve the society by offering top-notch technical education on par with global standards.

**Mission:**

- Be a center of excellence for technical education in emerging technologies by exceeding the needs of industry and society.
- Be an institute with world class research facilities.
- Be an institute nurturing talent and enhancing competency of students to transform them as all – round personalities respecting moral and ethical values.

**DEPARTMENT**

**Vision:**

- To become a renowned hub for AIML technologies to producing highly talented globally recognizable technocrats to meet industrial needs and societal expectation.

**Mission**

- To impart advanced education in AI and Machine Learning, built upon a foundation in Computer Science and Engineering.
- To foster Experiential learning equips students with engineering skills to tackle real-world problems.
- To promote collaborative innovation in AI, machine learning, and related research and development with industries.
- To provide an enjoyable environment for pursuing excellence while upholding strong personal and professional values and ethics.

## PROGRAM EDUCATIONAL OBJECTIVES (PEO)

- **PEO1:** Excel in technical abilities to build intelligent systems in the fields of AI & ML in order to find new opportunities.
- **PEO2:** Embrace new technology to solve real-world problems, whether alone or as a team, while prioritizing ethics and societal benefits.
- **PEO3:** Accept lifelong learning to expand future opportunities in research and product development.

## PROGRAM SPECIFIC OUTCOMES (PSO)

- **PSO1:** Expertise in tailoring ML algorithms and models to excel in designated applications and fields.
- **PSO2:** Ability to conduct research, contributing to machine learning advancements and innovations that tackle emerging societal challenges.

## PROGRAM OUTCOMES (POs)

Engineering students will be able to:

1. **Engineering knowledge:** Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
2. **Problem analysis:** Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences
3. **Design/development of solutions:** Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations

4. **Conduct investigations of complex problems:** Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions

5. **Modern tool usage:** Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations

6. **The engineer and society:** Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice

7. **Environment and sustainability:** Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development

8. **Ethics:** Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

9. **Individual and team work:** Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

10. **Communication:** Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

11. **Project management and finance:** Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

12. **Life-long learning:** Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

# ABSTRACT

In the modern educational environment, data-driven decision-making plays a pivotal role in assessing and enhancing student performance. Traditional methods of analyzing student marks such as using spreadsheets or manual records are often time-consuming, lack interactivity, and fail to provide real-time insights. To address these challenges, we propose the development of an interactive tool using R programming for visualizing the distribution of marks within a class. This tool is designed to simplify academic performance analysis through the integration of statistical computation and dynamic visualization. Built using R technologies like Shiny, ggplot2, plotly, and dplyr, the proposed system enables educators to load student marks from various formats (CSV, Excel), perform automatic calculations of statistical measures (mean, median, standard deviation, percentiles), and generate interactive visual representations such as histograms and box plots. The tool supports real-time filtering and data exploration, allowing users to analyze performance trends by subject, student, or score range. Unlike static desktop tools, this web-based platform ensures accessibility, user-friendliness, and efficient report generation, with export options for both visual content and summary reports. The implementation successfully bridges the gap between raw student data and actionable insights. By transforming routine data analysis into an engaging and efficient process, it empowers teachers to identify learning patterns, track academic progress, and make informed decisions. This innovative approach enhances the overall quality of education by supporting timely interventions and personalized feedback, making it a valuable asset in the evolving landscape of academic analytics.

## ABSTRACT WITH POs AND PSOs MAPPING

## CO 5 : BUILD DATABASES FOR SOLVING REAL-TIME PROBLEMS.

| ABSTRACT | POs MAPPED | PSOs MAPPED |
|---|---|---|
| In academic institutions, the effective analysis of student marks is essential for evaluating performance and supporting data-driven teaching strategies. Traditional methods using paper-based or static spreadsheet tools are inefficient and offer limited scope for interactive exploration. To overcome these limitations, this project introduces an interactive, web-based tool developed using R programming, specifically leveraging Shiny, ggplot2, plotly, and dplyr libraries. The proposed tool allows users to upload student marks data, compute statistical measures like mean, median, percentiles, and standard deviation, and visualize the results through interactive plots such as histograms and box plots. The application enables real-time filtering, dynamic updates, and report generation in PDF/HTML formats, offering both functionality and user convenience. This empowers educators to gain immediate insights, identify academic trends, and take timely actions for performance improvement. The tool requires no programming knowledge to operate, making it widely accessible and scalable across institutions. By integrating interactive visualization with statistical analysis, this solution transforms traditional performance tracking into an intuitive, efficient, and intelligent educational tool. | **PO1 -3**<br>**PO2 -3**<br>**PO3 -3**<br>**PO4 -3**<br>**PO5 -3**<br>**PO6 -3**<br>**PO7 -3**<br>**PO8 -3**<br>**PO9 -3**<br>**PO10 -3**<br>**PO11-3**<br>**PO12 -3** | **PSO1 -3**<br>**PSO2 -3**<br>**PSO3 -3** |

Note: 1- Low, 2-Medium, 3- High

# TABLE OF CONTENTS

# LIST OF FIGURES

# CHAPTER 1
# INTRODUCTION

## 1 INTRODUCTION

In the current academic environment, analyzing student performance is essential for informed teaching and continuous improvement. Traditional methods like paper records and spreadsheets are time-consuming and lack interactivity. To overcome these limitations, this project introduces an interactive visualization tool developed using R programming. The tool enables educators to upload student marks, compute statistical measures such as mean, median, and standard deviation, and visualize the data using dynamic charts like histograms and box plots. Built with Shiny, ggplot2, plotly, and dplyr, the system allows real-time filtering and easy export of reports, making academic data analysis efficient, accurate, and user-friendly.

## 1.1 OBJECTIVE

The primary objective of this project is to develop an interactive tool that simplifies and enhances the analysis of student academic performance. It aims to enable educators to easily upload student marks in various formats, perform automated statistical calculations, and generate insightful visualizations. The system strives to support real-time data filtering and customization, allowing users to focus on specific subjects, score ranges, or individual students. Another key goal is to provide the ability to export comprehensive reports in accessible formats like PDF and HTML, facilitating effective communication and record-keeping. Overall, the project seeks to improve the accuracy, efficiency, and accessibility of student performance analysis to aid better educational decision-making.

## 1.2 OVERVIEW

This project addresses the limitations of traditional performance evaluation methods by introducing a web-based application built using R programming. It allows users to upload datasets in formats such as CSV or Excel and instantly processes the data to display statistical measures like mean, median, standard deviation, and percentiles. The tool generates visualizations in the form of histograms, box plots, and bar graphs, which update dynamically based on user-selected filters such as subject, score range, or student group. Additionally, the system supports report generation in PDF and HTML formats, enabling educators to store and share performance summaries effectively. With its real-time interactivity and automation features, the tool serves as a comprehensive solution for academic performance monitoring.

## 1.3 TECHNOLOGY USED

This project is developed using R programming, chosen for its strong capabilities in statistical analysis and data visualization. The interactive dashboard is built with Shiny, allowing real-time responsiveness to user input and seamless interaction. Visualizations are created with ggplot2 for elegant static charts and enhanced with plotly to add interactive features like zooming, filtering, and dynamic tooltips that improve data exploration. Efficient data manipulation is achieved through dplyr, which supports handling large datasets smoothly and enables complex data transformations with simple commands. For importing student marks, readr and readxl libraries are used to process CSV and Excel files respectively, ensuring compatibility with common educational data formats. Together, these technologies work cohesively to provide a robust, efficient, and user-friendly platform that facilitates comprehensive and insightful student performance analysis.

# CHAPTER 2

# PROJECT METHODOLOGY

## 2.1 PROPOSED WORK

The proposed system is a web-based interactive tool designed to automate the process of analyzing and visualizing student marks. Built using R programming and its supporting libraries, the tool provides a simple yet powerful interface for educators to upload, analyze, and interpret academic data. Upon uploading a marks dataset in CSV or Excel format, the system performs automatic statistical analysis to compute values such as mean, median, percentiles, and standard deviation. The core functionality of the system revolves around its ability to generate dynamic, real-time graphs—such as histograms, box plots, and bar charts—that change based on user inputs or filters. Educators can filter data by subject, score range, or specific students to obtain tailored visual feedback. The interface is built using the Shiny framework, which allows for real-time responsiveness, and is integrated with ggplot2 and plotly for both static and interactive plotting. The system also supports exporting reports in PDF or HTML formats, making it useful for documentation and communication purposes. The tool is particularly useful in classroom and institutional settings where quick, visual interpretation of student performance is essential. It eliminates the need for manual chart creation and offers an intuitive, user-friendly platform for both teaching and administrative staff.

## 2.2 BLOCK DIAGRAM

# BLOCK DIAGRAM

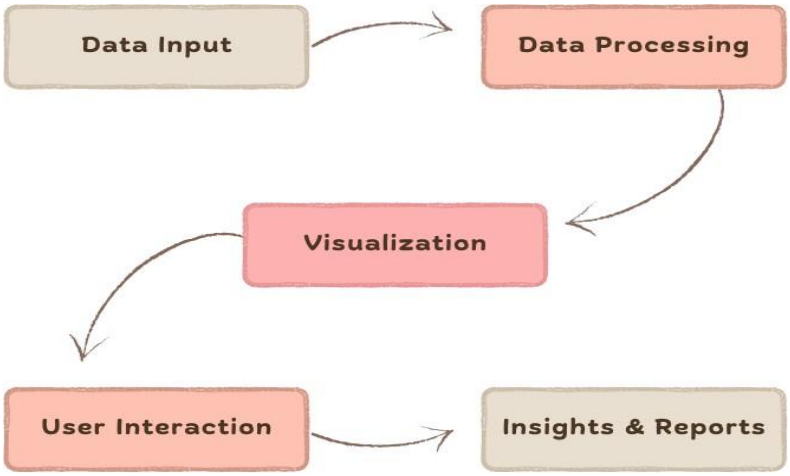


**Figure no. : 2.2.1 – Block Diagram For Interactive Tool For Visualizing Marks Distribution In A Class**

# CHAPTER 3

# MODULE DESCRIPTION

The Interactive Tool for Visualizing Marks Distribution is developed using modular design principles to ensure scalability, clarity, and maintainability. Each module is responsible for a specific function such as data handling, statistical computation, visualization, interactivity, or report generation. Below is a detailed description of each module:

## 3.1. DATA LOADING & PREPROCESSING MODULE

This module handles the import of student marks from external sources such as CSV or Excel files. It uses R libraries like readr and readxl to load the data efficiently into memory. Once imported, the data is preprocessed using the dplyr package to ensure it is structured, cleaned, and ready for analysis. This includes handling missing values, converting data types, and ensuring consistency in score entries across subjects and students.

## 3.2. STATISTICAL ANALYSIS MODULE

This module performs the core computations that form the basis of performance evaluation. It calculates essential statistical measures such as mean, median, mode, standard deviation, and percentiles for each subject or for the overall student group. These metrics help educators understand class-level performance trends, identify outliers, and make informed decisions regarding academic intervention.

## 3.3. VISUALIZATION MODULE

This module is responsible for transforming numerical data into insightful graphical formats. It utilizes ggplot2 for generating high-quality static

plots and plotly for interactive charts. The types of visualizations include histograms (for frequency distribution), box plots (for spread and outlier detection), and bar charts (for comparative analysis across subjects or students). These visuals provide a quick and intuitive understanding of academic performance at various levels.

### 3.4. INTERACTIVE DASHBOARD MODULE

Built using the Shiny framework, this module forms the user interface of the system. It allows users to interact with the dataset through dropdowns, sliders, and search bars. Based on user input (such as selecting a subject or score range), the corresponding statistical outputs and visualizations update in real time. This enhances user engagement and makes the analysis more personalized and targeted.

### 3.5. REPORT GENERATION MODULE

This module allows users to generate and export analysis results in the form of PDF or HTML reports. These reports include both graphical and tabular summaries, which can be used for academic documentation, parent meetings, or administrative reviews. The export functionality ensures that insights are preserved and shared easily, even outside the digital platform.

# CHAPTER 4

# CONCLUSION & FUTURE SCOPE

The project “Interactive Tool for Visualizing Marks Distribution in a Class” successfully demonstrates how data analysis and visualization can be integrated into an educational setting using R programming. The system enables educators to efficiently upload, process, and interpret student marks data through real-time statistical analysis and dynamic visualizations. With the use of R libraries such as Shiny, ggplot2, plotly, and dplyr, the tool provides an intuitive and accessible interface that simplifies performance tracking and supports data-driven academic decision-making. The implementation of interactive dashboards and exportable reports bridges the gap between raw numerical data and meaningful educational insights. This eliminates the limitations of traditional spreadsheet-based analysis, reduces manual effort, and enhances the accuracy of evaluations. Furthermore, the modular architecture of the tool ensures flexibility and scalability, making it suitable for institutions of various sizes and teaching approaches. Overall, the project meets its objectives by providing a reliable, efficient, and user-friendly solution for academic performance analysis, reinforcing the importance of technology in modern education.

## FUTURE SCOPE

While the current version of the system addresses core functionalities effectively, there are several areas where the tool can be further enhanced to increase its impact and usability:

- **Student Performance Prediction**: Integrate machine learning algorithms to forecast student outcomes based on historical data.

- **Mobile App Support**: Develop a mobile-compatible version of the dashboard to allow access on smartphones and tablets.
- **Cloud Integration**: Enable cloud-based storage and processing to support centralized access and multi-user collaboration.
- **Multilingual Interface**: Add support for regional languages to increase accessibility for teachers and institutions across diverse locations.
- **Integration with Learning Management Systems (LMS)**: Enable direct import/export with platforms like Moodle or Google Classroom.
- **Notification System**: Include automated alerts for low-performing students or sudden score drops.
- **Advanced Visualizations**: Add heatmaps, pie charts, and percentile bands to support deeper data interpretation.
- **Access Roles and Permissions**: Implement user roles such as Admin, Teacher, and Principal for secure, role-based access.

By implementing these future enhancements, the tool can evolve into a comprehensive academic analytics platform that not only visualizes marks but also drives student success through predictive and adaptive learning insights.

# CHAPTER 5

# APPENDIX A - SOURCE CODE

```
# Load necessary libraries

library(shiny)

library(dplyr)

library(ggplot2)

library(plotly)

library(readr)

# Define UI

ui <- fluidPage(

  titlePanel("Interactive Tool for Visualizing Marks Distribution"),

  sidebarLayout(

    sidebarPanel(

      fileInput("file1", "Choose CSV File", accept = ".csv"),

      uiOutput("subject_ui"),

      checkboxInput("show_stats", "Show Statistics", TRUE)

    ),

    mainPanel(

      tabsetPanel(

        tabPanel("Summary Table", tableOutput("summary_table")),

        tabPanel("Histogram", plotlyOutput("hist_plot")),

        tabPanel("Box Plot", plotlyOutput("box_plot")),

        tabPanel("Bar Chart", plotlyOutput("bar_plot")),


tabPanel("Density Plot", plotlyOutput("density_plot")),

        tabPanel("Statistics", verbatimTextOutput("stats_text"))
```

```
)
    )
  )
)
# Define server logic
server <- function(input, output, session) {
 # Reactive expression to read the CSV file
 data <- reactive({
    req(input$file1)
    read_csv(input$file1$datapath)
  })
  # Dynamic UI for subject selection
  output$subject_ui <- renderUI({
  req(data())
    subjects <- names(data())[-1]  # Assuming the first column is Student Name/ID
    selectInput("subject", "Select Subject", choices = subjects)
  })
  # Reactive expression for selected data
  selected_data <- reactive({
  req(input$subject)
    df <- data()

df <- df %>% select(Student = 1, Marks = input$subject)
    df <- na.omit(df)
    df
  })
  # Render summary table
```

```
output$summary_table <- renderTable({
  head(selected_data())
 })
 # Render histogram plot
 output$hist_plot <- renderPlotly({
  p <- ggplot(selected_data(), aes(x = Marks)) +
   geom_histogram(binwidth = 5, fill = "skyblue", color = "black") +
   theme_minimal() +
   labs(title = paste("Histogram -", input$subject), x = "Marks", y = "Count")
  ggplotly(p)
 })
 # Render box plot
 output$box_plot <- renderPlotly({
  p <- ggplot(selected_data(), aes(y = Marks)) +
   geom_boxplot(fill = "orange") +
   theme_minimal() +
   labs(title = paste("Boxplot -", input$subject), y = "Marks")
  ggplotly(p)

 })
 # Render bar chart (average marks per student)
 output$bar_plot <- renderPlotly({
  p <- ggplot(selected_data(), aes(x = Student, y = Marks)) +
   geom_bar(stat = "identity", fill = "lightgreen") +
   theme_minimal() +
   labs(title = paste("Average Marks per Student -", input$subject), x = "Student", y = "Marks") +
   theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

```
ggplotly(p)
  })
  # Render density plot
  output$density_plot <- renderPlotly({
    p <- ggplot(selected_data(), aes(x = Marks)) +
      geom_density(fill = "lightblue", alpha = 0.5) +
      theme_minimal() +
      labs(title = paste("Density Plot -", input$subject), x = "Marks", y = "Density")
    ggplotly(p)
  })
  # Render statistics text
  output$stats_text <- renderPrint({
  req(input$show_stats)
    marks <- selected_data()$Marks
    cat("Mean:", mean(marks), "\n")

  cat("Median:", median(marks), "\n")
    cat("Standard Deviation:", sd(marks), "\n")
    cat("Minimum:", min(marks), "\n")
    cat("Maximum:", max(marks), "\n")
    cat("25th Percentile:", quantile(marks, 0.25), "\n")
    cat("75th Percentile:", quantile(marks, 0.75), "\n")
  })
}
# Run the application
shinyApp(ui = ui, server = server)
```

# APPENDIX B – SCREENSHOTS

## DATA LOADING & PREPROCESSING

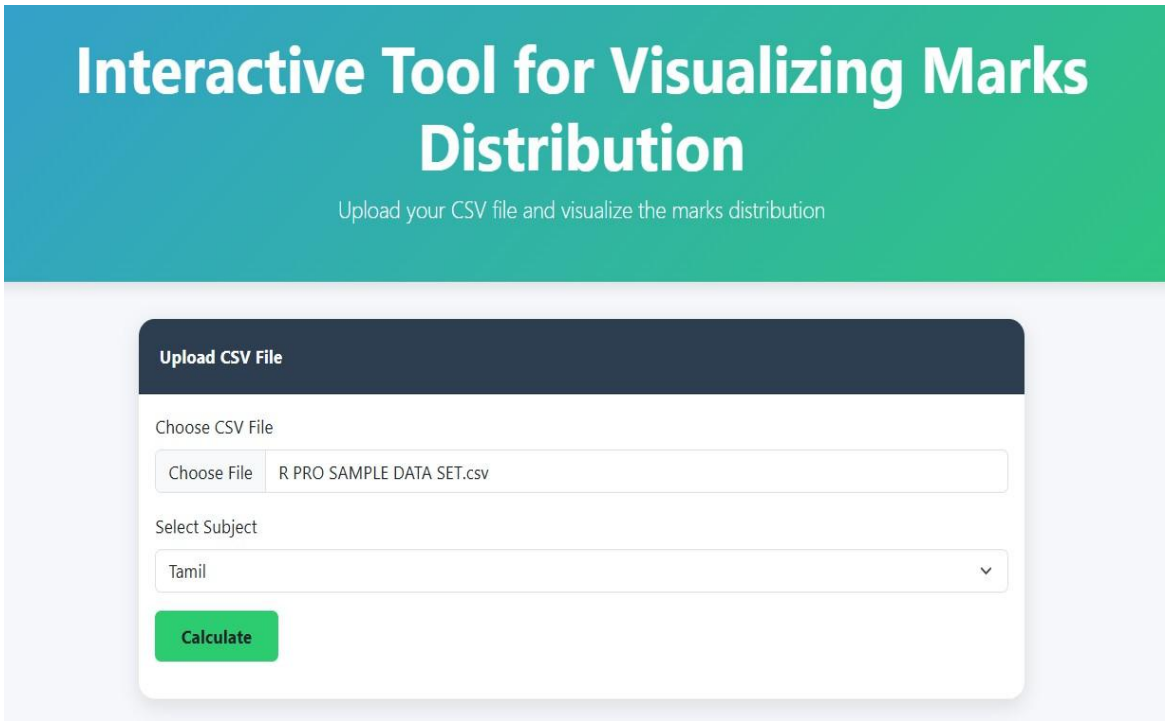


**Figure no. : B.1 – Data Loading & Preprocessing**

## STATISTICAL ANALYSIS

Statistics

| Statistic | Value |
|---|---|
| Mean | 65.88 |
| Median | 68.50 |
| Standard Deviation | 24.31 |
| Minimum | 0 |
| Maximum | 96 |
| 25th Percentile (Q1) | 53.25 |
| 75th Percentile (Q3) | 88.00 |

**Figure no. : B.2 – Statistical Analysis**

**INTERACTIVE DASHBOARD**

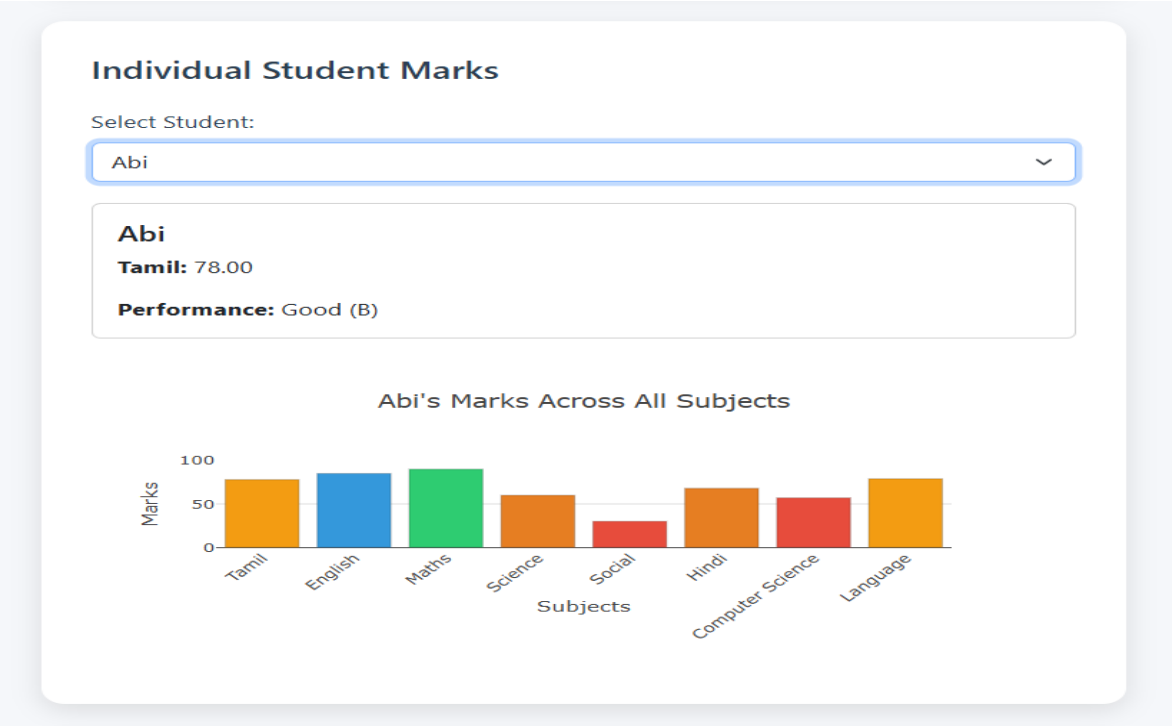


**Figure no. : B.3 – Interactive Dashboard**

**REPORT GENERATION**

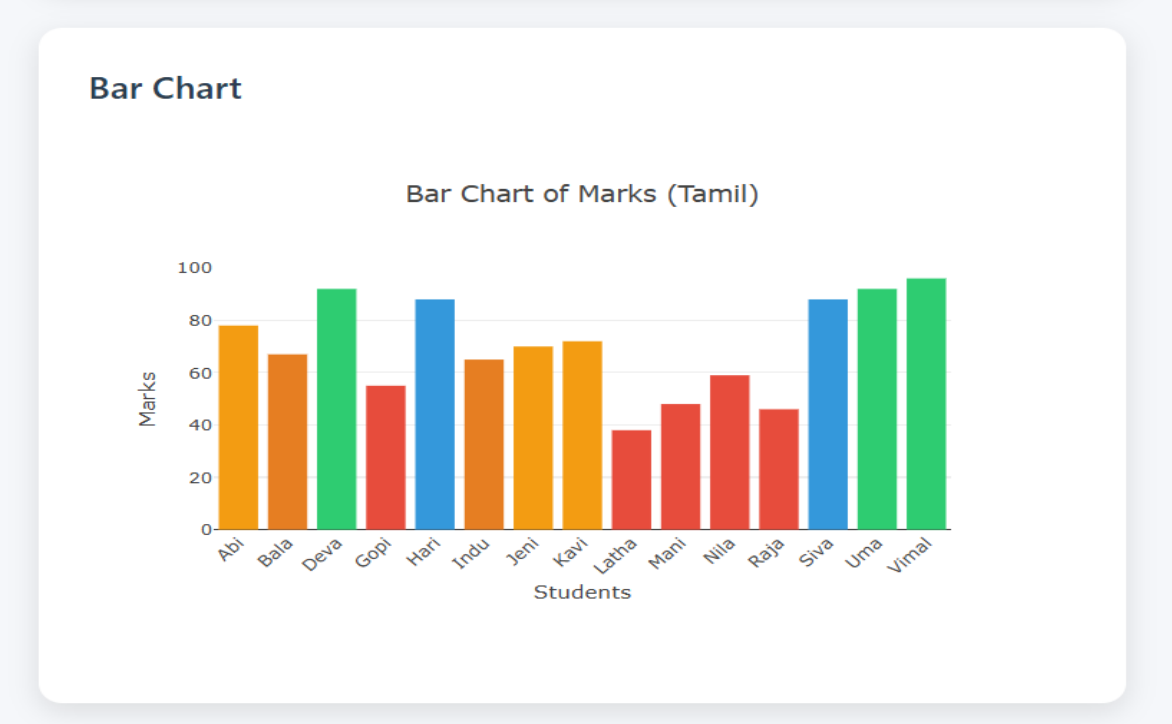


**Figure no. : B.4 – Report Generation**

## REFERENCES

### 1. Books


- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- Chang, W. (2021). *R Graphics Cookbook: Practical Recipes for Visualizing Data* (2nd ed.). O'Reilly Media.
- Kabacoff, R. I. (2015). *R in Action: Data Analysis and Graphics with R*. Manning Publications.
- Grolemund, G., & Wickham, H. (2016). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. O'Reilly Media.


### 2. Websites


- R Project Official Site: https://www.r-project.org
- Shiny (RStudio): https://shiny.posit.co/
- ggplot2 Documentation: https://ggplot2.tidyverse.org/
- plotly for R: https://plotly.com/r/
- dplyr Reference: https://dplyr.tidyverse.org/
- readxl and readr Documentation: https://readxl.tidyverse.org/
- W3Schools – HTML Reference: https://www.w3schools.com/html/


### 3. YouTube Videos


- Introduction to Shiny: https://www.youtube.com/watch?v=2v9gZsS9B7s

- Data Visualization in R:

  https://www.youtube.com/watch?v=RJAEpVG8owY

- R for Beginners – Crash Course:

  https://www.youtube.com/watch?v=_V8eKsto3Ug

**4. Reports**

- UNESCO (2022). *Data Use in Education: From Assessment to Action*.
- Ministry of Education (India). (2023). *National Report on Education Technology and Digital Learning*.
- OECD. (2020). *Innovating Education and Educating for Innovation: The Power of Data Analytics*.