

Sheth L.U.J. & Sir M.V. College

Practical 11 Report: Analysis of Student Exam Scores

Aim:- To perform graphical analysis of student exam scores and related factors such as hours studied, sleep hours, attendance percentage, and previous scores, using R and ggplot2.

Dataset Description

The dataset `student_exam_scores.csv` contains the following variables:

Column Name	Description
<code>student_id</code>	Unique identifier for each student
<code>hours_studied</code>	Number of hours studied by the student
<code>sleep_hours</code>	Number of hours of sleep per day
<code>attendance_percent</code>	Attendance percentage of the student
<code>previous_scores</code>	Marks obtained in previous exams
<code>exam_score</code>	Marks obtained in the current exam

Name :- Priya Gupta

Roll no. :- So81

Sheth L.U.J. & Sir M.V. College

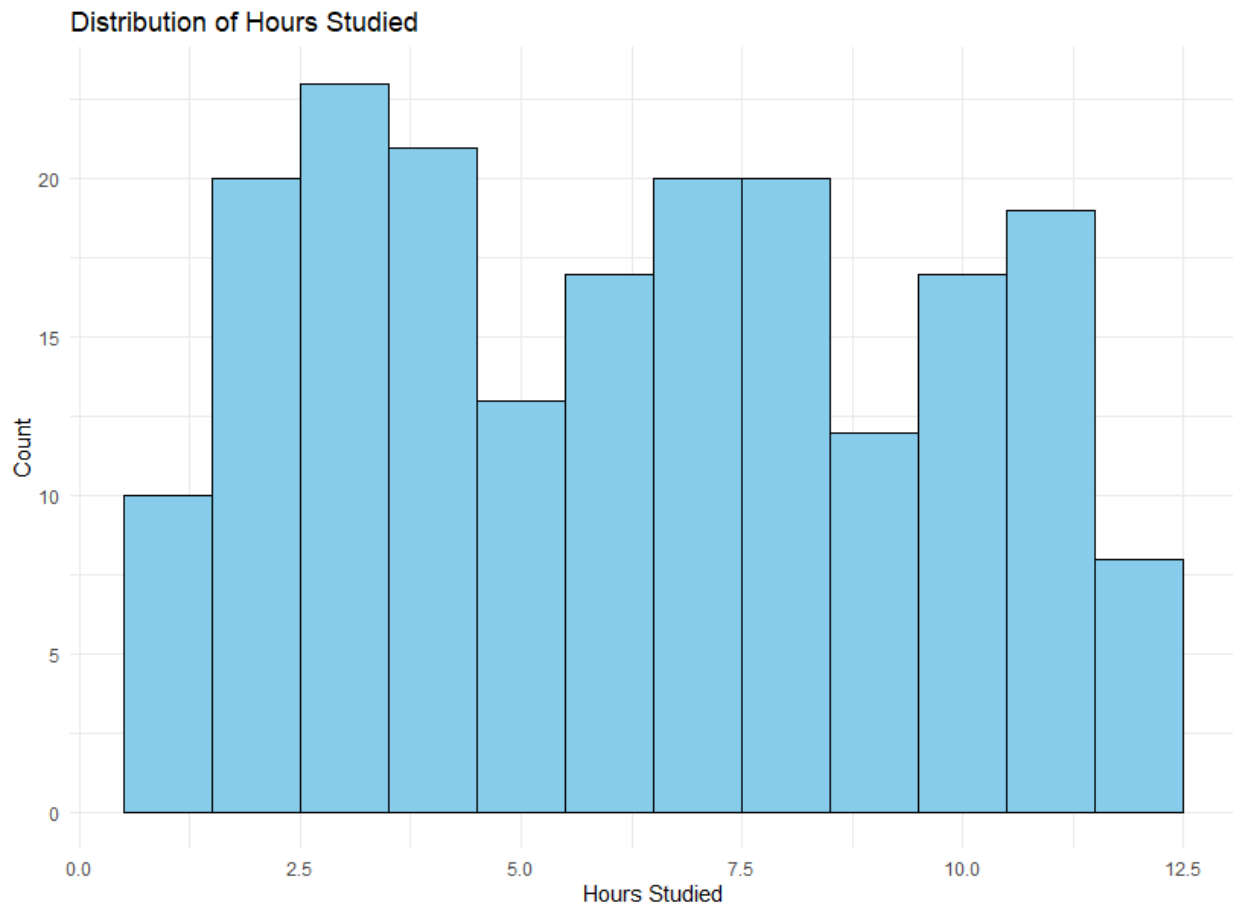
Stepwise Graphical Analysis

Step 1: Histogram of Hours Studied

Observation:

Most students studied between 3 to 9 hours. This helps us understand study patterns before exams.

Graph:



Name :- Priya Gupta

Roll no. :- So81

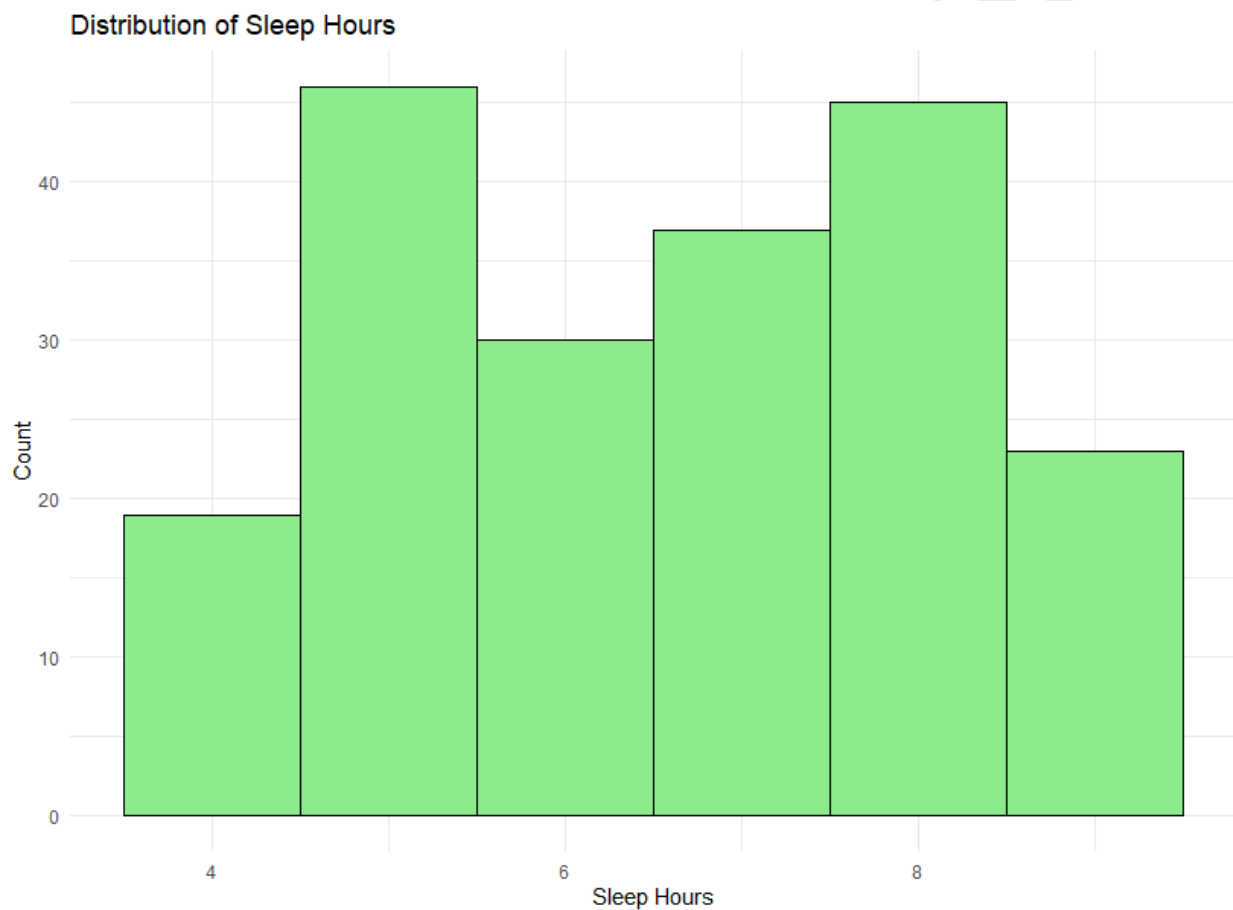
Sheth L.U.J. & Sir M.V. College

Step 2: Histogram of Sleep Hours

Observation:

Sleep hours ranged from 4 to 9 hours. Most students slept around 5–8 hours. Adequate sleep may influence exam performance.

Graph:



Name :- Priya Gupta

Roll no. :- So81

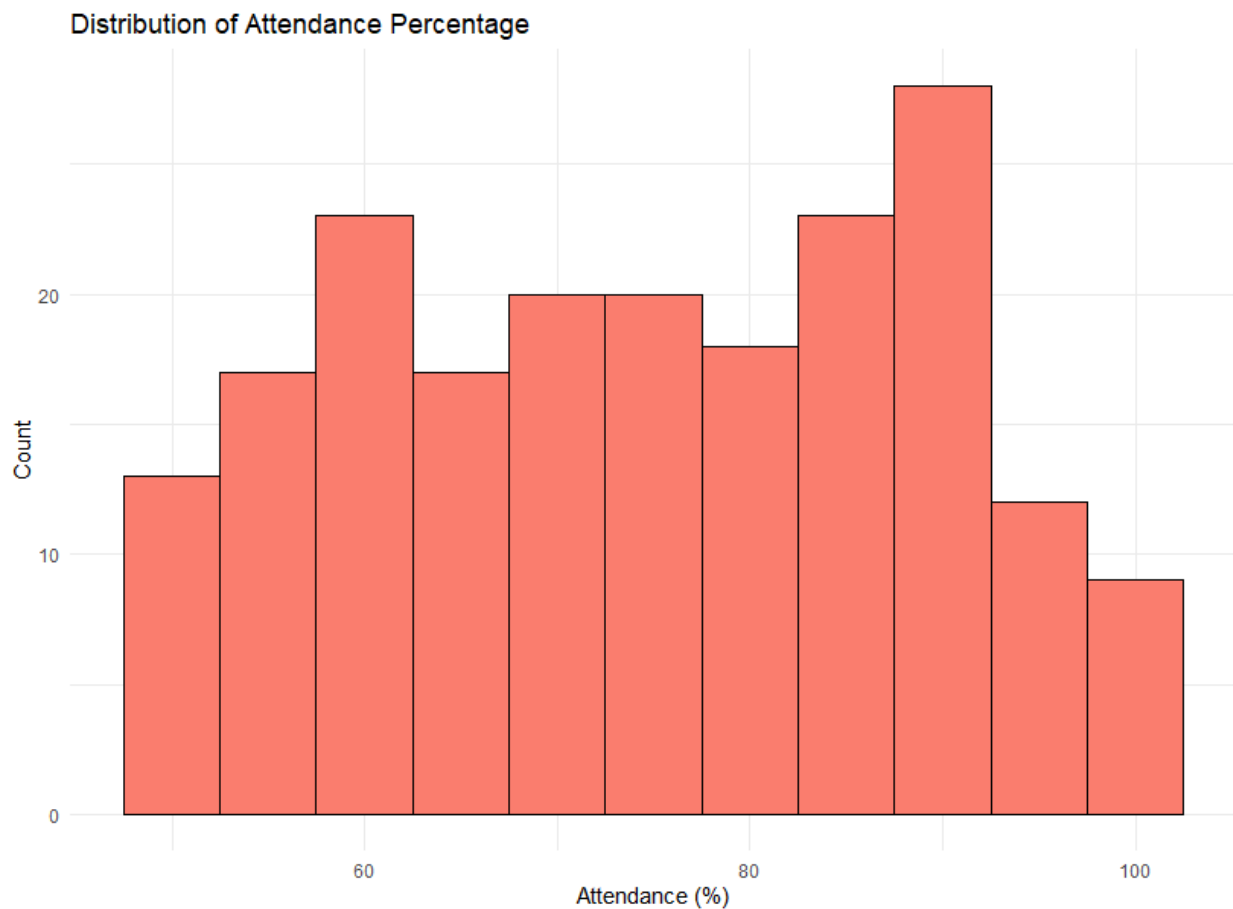
Sheth L.U.J. & Sir M.V. College

Step 3: Histogram of Attendance Percentage

Observation:

Attendance varied from 50% to 95%. Most students attended between 60–90% classes.

Graph:



Name :- Priya Gupta

Roll no. :- So81

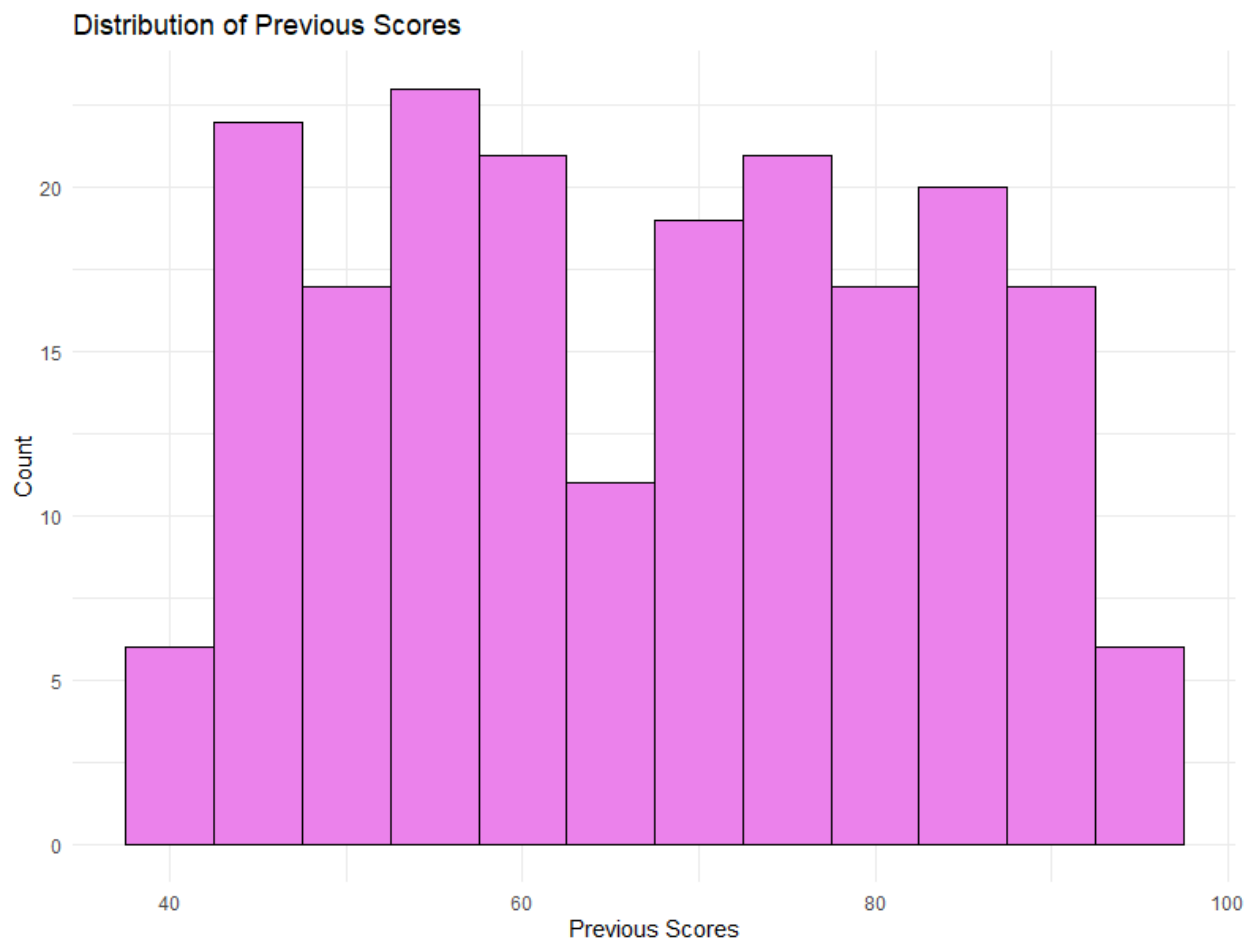
Sheth L.U.J. & Sir M.V. College

Step 4: Histogram of Previous Scores

Observation:

Previous scores ranged from 45 to 86 marks. This gives a baseline for student performance trends.

Graph:



Name :- Priya Gupta

Roll no. :- So81

Sheth L.U.J. & Sir M.V. College

Step 5: Histogram of Exam Scores

Observation:

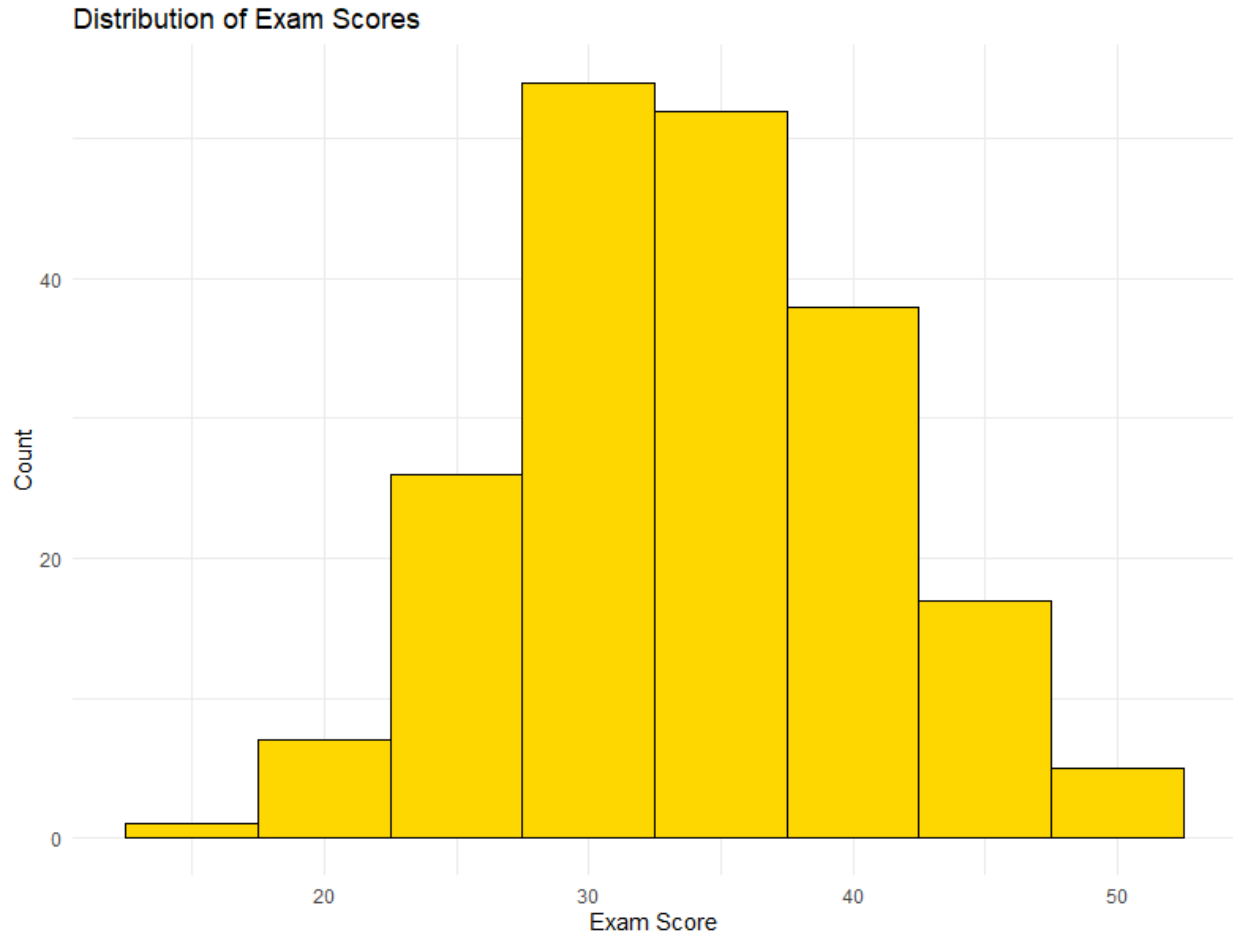
Exam scores ranged from 25 to 40 marks. Understanding distribution helps in analyzing overall student performance.

Graph:

Name :- Priya Gupta

Roll no. :- So81

Sheth L.U.J. & Sir M.V. College



Step 6: Boxplot of Exam Scores by Attendance (Grouped)

Observation:

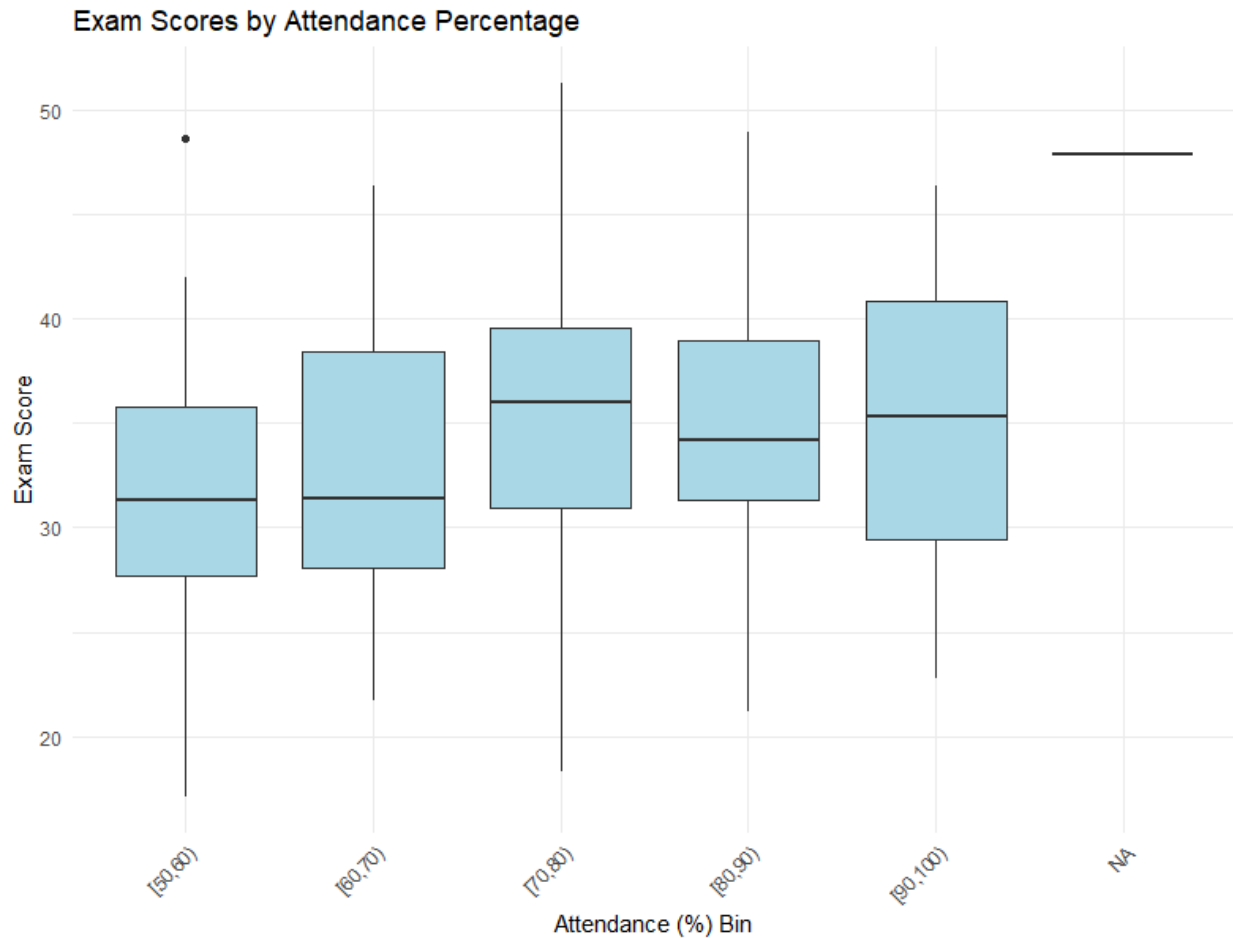
Students with higher attendance tend to score better. Binning attendance in 10% intervals shows trends in exam scores.

Graph:

Name :- Priya Gupta

Roll no. :- So81

Sheth L.U.J. & Sir M.V. College



Step 7: Boxplot of Exam Scores by Hours Studied (Grouped)

Observation:

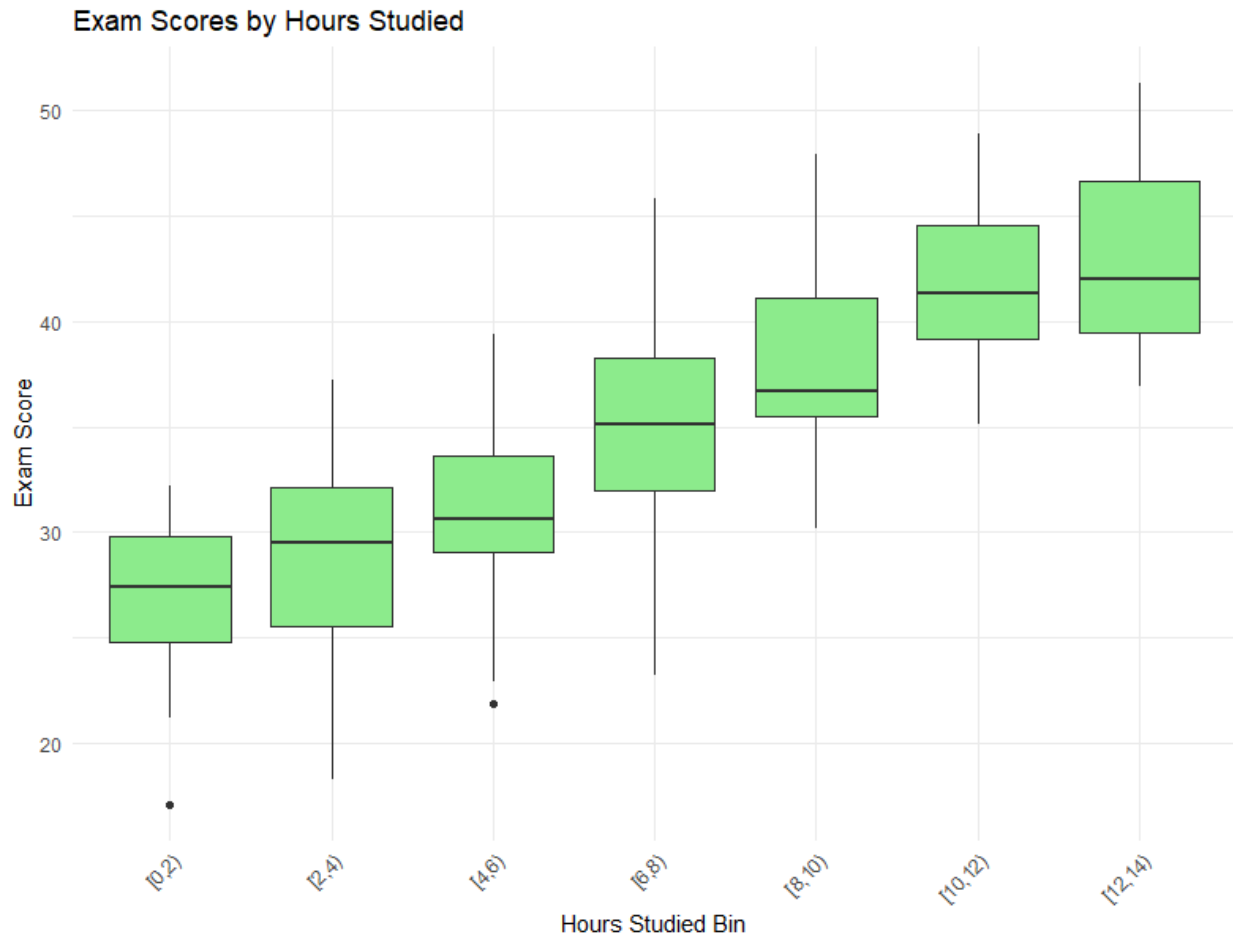
Students who studied more hours generally obtained higher scores. Binning hours studied helps visualize this trend.

Graph:

Name :- Priya Gupta

Roll no. :- So81

Sheth L.U.J. & Sir M.V. College



Conclusion

- Hours Studied: Higher study hours positively influence exam scores.
- Sleep Hours: Students with adequate sleep (6–8 hours) tend to perform better.
- Attendance: High attendance is associated with higher exam scores.

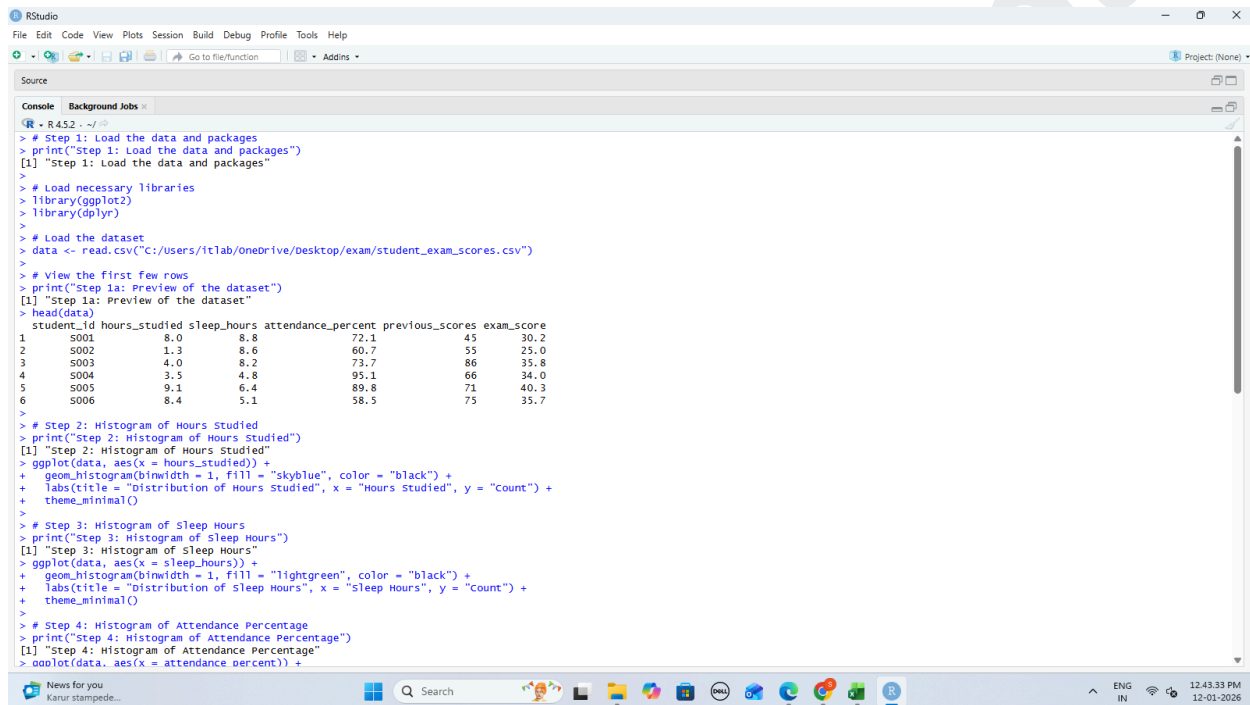
Name :- Priya Gupta

Roll no. :- So81

Sheth L.U.J. & Sir M.V. College

- Previous Scores: Baseline performance can predict current exam performance.
- Graphical analysis using histograms and boxplots provides clear insights into student behavior and academic outcomes.

Screenshots



```
R - R452 - ~/ -  
File Edit Code View Plots Session Build Debug Profile Tools Help  
Go to file/function Addins  
Project (None)  
Source  
Console Background Jobs  
R - R452 - ~/ -  
> # Step 1: Load the data and packages  
> print("Step 1: Load the data and packages")  
[1] "Step 1: Load the data and packages"  
>  
> # Load necessary libraries  
> library(ggplot2)  
> library(dplyr)  
>  
> # Load the dataset  
> data <- read.csv("C:/Users/itlab/OneDrive/Desktop/exam/student_exam_scores.csv")  
>  
> # View the first few rows  
> print("Step 1a: Preview of the dataset")  
[1] "Step 1a: Preview of the dataset"  
> head(data)  
  student_id hours_studied sleep_hours attendance_percent previous_scores exam_score  
1      S001           8.0           8.8           72.1           45          30.2  
2      S002           1.3           8.6           60.7           55          25.0  
3      S003           4.0           8.2           73.7           86          35.8  
4      S004           3.5           4.8           95.1           66          34.0  
5      S005           9.1           6.4           89.8           71          40.3  
6      S006           8.4           5.1           58.5           75          35.7  
>  
> # Step 2: Histogram of Hours Studied  
> print("Step 2: Histogram of Hours Studied")  
[1] "Step 2: Histogram of Hours Studied"  
> ggplot(data, aes(x = hours_studied)) +  
+   geom_histogram(binwidth = 1, fill = "skyblue", color = "black") +  
+   labs(title = "Distribution of Hours Studied", x = "Hours Studied", y = "Count") +  
+   theme_minimal()  
>  
> # Step 3: Histogram of Sleep Hours  
> print("Step 3: Histogram of Sleep Hours")  
[1] "Step 3: Histogram of Sleep Hours"  
> ggplot(data, aes(x = sleep_hours)) +  
+   geom_histogram(binwidth = 1, fill = "lightgreen", color = "black") +  
+   labs(title = "Distribution of Sleep Hours", x = "Sleep Hours", y = "Count") +  
+   theme_minimal()  
>  
> # Step 4: Histogram of Attendance Percentage  
> print("Step 4: Histogram of Attendance Percentage")  
[1] "Step 4: Histogram of Attendance Percentage"  
> ggplot(data, aes(x = attendance_percent)) +
```

Name :- Priya Gupta

Roll no. :- So81

Sheth L.U.J. & Sir M.V. College

The screenshot displays the RStudio interface with the following components:

- Top Bar:** Includes the RStudio logo and menu options: File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Toolbar:** Contains icons for file operations (Open, Save, Print, etc.) and a search bar.
- Source Panel:** Shows the R script being edited. The code includes comments for each step and uses the `ggplot2` package for data visualization.
- Console Panel:** Displays the output of the R script, showing the execution of each step and the resulting plots.
- Plots Panel:** Currently empty, but it would display the histograms and boxplots generated by the code.
- Taskbar:** At the bottom of the screen, showing the Windows taskbar with various application icons and the system clock.

The R code in the Source panel is as follows:

```
# Step 4: Histogram of Attendance Percentage
> print("Step 4: Histogram of Attendance Percentage")
[1] "Step 4: Histogram of Attendance Percentage"
> ggplot(data, aes(x = attendance_percent)) +
+   geom_histogram(binwidth = 5, fill = "salmon", color = "black") +
+   labs(title = "Distribution of Attendance Percentage", x = "Attendance (%)", y = "Count") +
+   theme_minimal()
>
> # Step 5: Histogram of Previous Scores
> print("Step 5: Histogram of Previous Scores")
[1] "Step 5: Histogram of Previous Scores"
> ggplot(data, aes(x = previous_scores)) +
+   geom_histogram(binwidth = 5, fill = "violet", color = "black") +
+   labs(title = "Distribution of Previous Scores", x = "Previous Scores", y = "Count") +
+   theme_minimal()
>
> # Step 6: Histogram of Exam Scores
> print("Step 6: Histogram of Exam Scores")
[1] "Step 6: Histogram of Exam Scores"
> ggplot(data, aes(x = exam_score)) +
+   geom_histogram(binwidth = 5, fill = "gold", color = "black") +
+   labs(title = "Distribution of Exam Scores", x = "Exam Score", y = "Count") +
+   theme_minimal()
>
> # Step 7: Boxplot of Exam Scores by Attendance Percentage (grouped in 10% bins)
> print("Step 7: Boxplot of Exam Scores by Attendance Percentage")
[1] "Step 7: Boxplot of Exam Scores by Attendance Percentage"
> data <- data %>%
+   mutate(attendance_bin = cut(attendance_percent, breaks = seq(0, 100, 10), right = FALSE))
>
> ggplot(data, aes(x = attendance_bin, y = exam_score)) +
+   geom_boxplot(fill = "lightblue") +
+   labs(title = "Exam Scores by Attendance Percentage", x = "Attendance (%) Bin", y = "Exam Score") +
+   theme_minimal() +
+   theme(axis.text.x = element_text(angle = 45, hjust = 1))
>
> # Step 8: Boxplot of Exam Scores by Hours Studied (grouped in 2-hour bins)
> print("Step 8: Boxplot of Exam Scores by Hours Studied")
[1] "Step 8: Boxplot of Exam Scores by Hours Studied"
> data <- data %>%
+   mutate(hours_bin = cut(hours_studied, breaks = seq(0, max(hours_studied)+2, 2), right = FALSE))
>
```

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function

Addins

Project: (None)

Source

Console Background Jobs

```
R - R452 ~\>  
+ theme_minimal()  
>  
> # Step 5: Histogram of Previous Scores  
> print("Step 5: Histogram of Previous Scores")  
[1] "Step 5: Histogram of Previous Scores"  
> ggplot(data, aes(x = previous_scores)) +  
+   geom_histogram(binwidth = 5, fill = "violet", color = "black") +  
+   labs(title = "Distribution of Previous Scores", x = "Previous Scores", y = "Count") +  
+   theme_minimal()  
>  
> # Step 6: Histogram of Exam Scores  
> print("Step 6: Histogram of Exam Scores")  
[1] "Step 6: Histogram of Exam Scores"  
> ggplot(data, aes(x = exam_score)) +  
+   geom_histogram(binwidth = 5, fill = "gold", color = "black") +  
+   labs(title = "Distribution of Exam Scores", x = "Exam Score", y = "Count") +  
+   theme_minimal()  
>  
> # Step 7: Boxplot of Exam Scores by Attendance Percentage (grouped in 10% bins)  
> print("Step 7: Boxplot of Exam Scores by Attendance Percentage")  
[1] "Step 7: Boxplot of Exam Scores by Attendance Percentage"  
> data <- data %>%  
+   mutate(attendance_bin = cut(attendance_percent, breaks = seq(0, 100, 10), right = FALSE))  
>  
> ggplot(data, aes(x = attendance_bin, y = exam_score)) +  
+   geom_boxplot(fill = "lightblue") +  
+   labs(title = "Exam Scores by Attendance Percentage", x = "Attendance (%) Bin", y = "Exam Score") +  
+   theme_minimal() +  
+   theme(axis.text.x = element_text(angle = 45, hjust = 1))  
>  
> # Step 8: Boxplot of Exam Scores by Hours Studied (grouped in 2-hour bins)  
> print("Step 8: Boxplot of Exam Scores by Hours Studied")  
[1] "Step 8: Boxplot of Exam Scores by Hours Studied"  
> data <- data %>%  
+   mutate(hours_bin = cut(hours_studied, breaks = seq(0, max(hours_studied)+2, 2), right = FALSE))  
>  
> ggplot(data, aes(x = hours_bin, y = exam_score)) +  
+   geom_boxplot(fill = "lightgreen") +  
+   labs(title = "Exam Scores by Hours Studied", x = "Hours Studied Bin", y = "Exam Score") +  
+   theme_minimal() +  
+   theme(axis.text.x = element_text(angle = 45, hjust = 1))  
>
```

News for you Kanur stampede...

Search

DELL O Microsoft Edge Google Chrome VS Code R

ENG IN 12:44:17 PM 12-01-2020

Name :- Priya Gupta

Roll no. :- So81

Sheth L.U.J. & Sir M.V. College

S081 Priya Gupta

Name :- Priya Gupta

Roll no. :- So81