Name: Prateek P

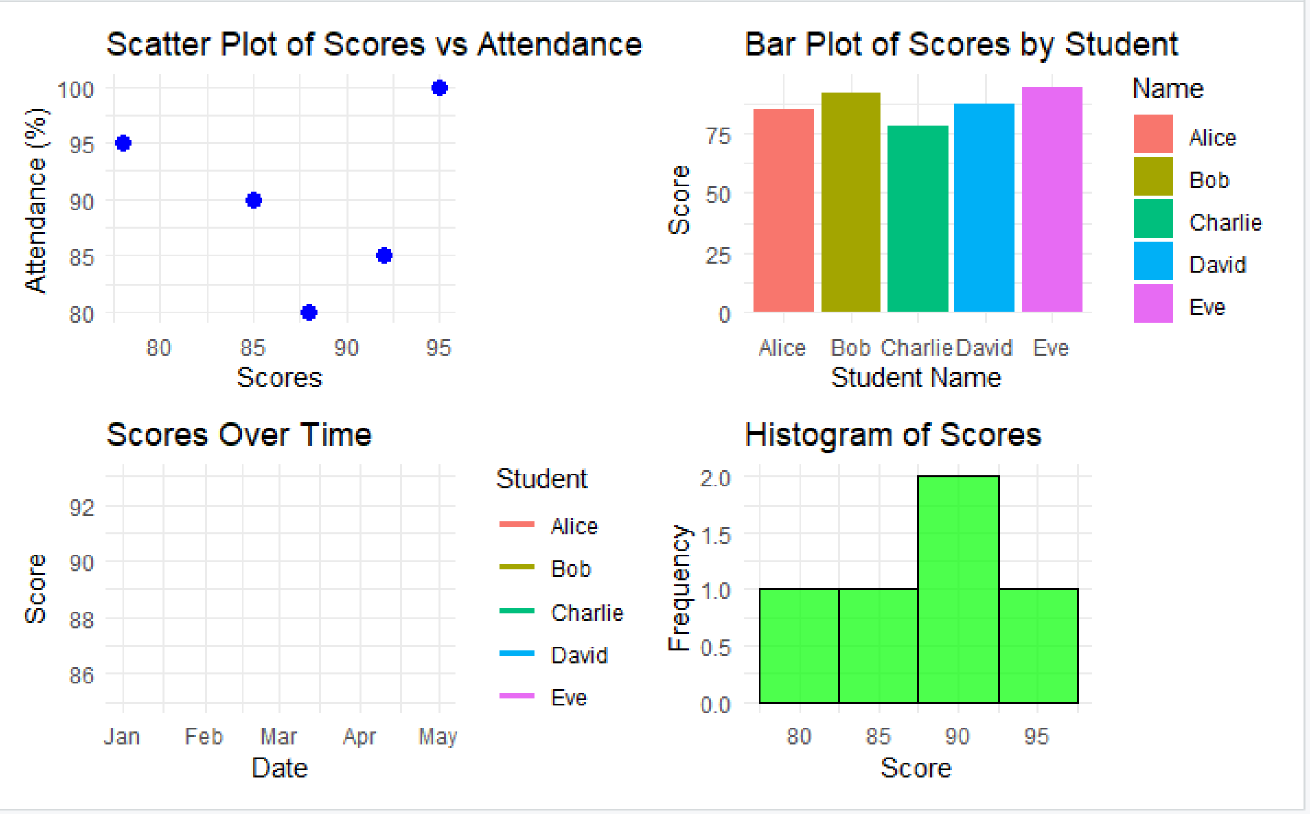
USN : 1MS22CI050

Program 9

```r
# Load required libraries
library(ggplot2)
library(patchwork)  # for combining plots
library(plotly)     # for interactive plots

# Prepare example data for student performance
student_data <- data.frame(
  Name = c("Alice", "Bob", "Charlie", "David", "Eve"),
  Score = c(85, 92, 78, 88, 95),
  Attendance = c(90, 85, 95, 80, 100)
)
student_data

# i. Create a scatter plot to visualize the relationship between scores and attendance percentages
scatter_plot <- ggplot(student_data, aes(x = Score, y = Attendance)) +
  geom_point(color = "blue", size = 3) +
  labs(title = "Scatter Plot of Scores vs Attendance", x = "Scores", y = "Attendance (%)") +
  theme_minimal()
print(scatter_plot)

# ii. Generate a bar plot to show the distribution of scores among different students
bar_plot <- ggplot(student_data, aes(x = Name, y = Score, fill = Name)) +
  geom_bar(stat = "identity") +
  labs(title = "Bar Plot of Scores by Student", x = "Student Name", y = "Score") +
  theme_minimal()
print(bar_plot)

# iii. Create a line plot to display the trend of scores over time (assuming scores collected over time for each student)
# Creating a time-based trend data (just for illustration purposes)
time_data <- data.frame(
  Date = seq(as.Date("2023-01-01"), by = "months", length.out = 5),
  Score = c(85, 90, 88, 93, 91),
  Student = c("Alice", "Bob", "Charlie", "David", "Eve")
)

line_plot <- ggplot(time_data, aes(x = Date, y = Score, color = Student)) +
  geom_line(size = 1.2) +
  labs(title = "Scores Over Time", x = "Date", y = "Score") +
  theme_minimal()
```

```
print(line_plot)

# iv. Generate a histogram to visualize the distribution of scores
histogram_plot <- ggplot(student_data, aes(Score)) +
  geom_histogram(binwidth = 5, fill = "green", color = "black", alpha = 0.7) +
  labs(title = "Histogram of Scores", x = "Score", y = "Frequency") +
  theme_minimal()
print(histogram_plot)

# You can combine these plots using patchwork if you'd like to display them
together
scatter_plot + bar_plot + line_plot + histogram_plot
```

Output



Program 10

```
# Load required library
library(dplyr)

# Sample data frame
df <- data.frame(
  ID = 1:10,
  Name = c("John", "Jane", "Sam", "Lucy", "Paul", "Anna", "Steve", "Kate",
"Tom", "Emma"),
  Age = c(28, 34, 45, 23, 36, 29, 41, 38, 30, 24),
  Salary = c(50000, 60000, 55000, 40000, 65000, 48000, 70000, 72000, 43000,
51000),
```

```r
  Department = c("HR", "IT", "Finance", "HR", "IT", "Finance", "IT", "HR", "Finance", "IT")
)

# 1. Select specific columns
df_selected <- df %>%
  select(Name, Age, Salary)

# 2. Filter rows based on a condition (Age > 30)
df_filtered <- df %>%
  filter(Age > 30)

# 3. Create new variables with mutate (e.g., calculate salary increase)
df_mutated <- df %>%
  mutate(Salary_Increment = Salary * 0.1)

# 4. Group the data by Department and summarize (e.g., average salary)
df_grouped <- df %>%
  group_by(Department) %>%
  summarize(Average_Salary = mean(Salary), Count = n())

# 5. Arrange the data by Salary in descending order
df_arranged <- df %>%
  arrange(desc(Salary))

# 6. Join two data frames (creating a second data frame for the example)
df2 <- data.frame(
  ID = 1:5,
  Bonus = c(2000, 3000, 2500, 1800, 2200)
)

# Perform a left join with df2 to add bonus information to df
df_joined <- df %>%
  left_join(df2, by = "ID")

# Display the results of each step
print("Selected Columns:")
print(df_selected)

print("Filtered Data (Age > 30):")
print(df_filtered)

print("Mutated Data (Salary Increment):")
print(df_mutated)

print("Grouped by Department (Average Salary):")
print(df_grouped)

print("Arranged by Salary (Descending):")
print(df_arranged)

print("Joined Data with Bonus:")
print(df_joined)
```

Output

```
   ID  Name Age Salary Department Bonus
1   1  John  28  50000         HR  2000
2   2  Jane  34  60000         IT  3000
3   3   Sam  45  55000    Finance  2500
4   4  Lucy  23  40000         HR  1800
5   5  Paul  36  65000         IT  2200
6   6  Anna  29  48000    Finance    NA
7   7 Steve  41  70000         IT    NA
8   8  Kate  38  72000         HR    NA
9   9   Tom  30  43000    Finance    NA
10 10  Emma  24  51000         IT    NA
```