Program 12

Input:

# Install and load necessary libraries

if (!requireNamespace("writexl", quietly = TRUE)) {

install.packages("writexl")

}

library(tidyverse)

library(readxl)

library(writexl)

# Load the dataset from Excel

ipl\_data <- read\_excel("C:/Users/ADMIN/Desktop/R lab/ipl.xlsx")

# Check the structure of the dataset

str(ipl\_data)

# Output summary statistics of numerical columns

summary(ipl\_data)

# Check for NAs in numeric columns

sapply(ipl\_data, function(x) sum(is.na(x) & is.numeric(x)))

# Calculate and display the total number of matches

total\_matches <- nrow(ipl\_data)

cat("Total Number of Matches:", total\_matches, "\n")

# Determine the number of unique teams

unique\_teams <- union(unique(ipl\_data$Team1), unique(ipl\_data$Team2))

num\_unique\_teams <- length(unique\_teams)

cat("Number of Unique Teams:", num\_unique\_teams, "\n")

# List the unique teams

cat("Unique Teams:", unique\_teams, "\n")

# Calculate the number of matches won by each team

matches\_won <- table(ipl\_data$Winner)

cat("Matches Won by Each Team:\n", matches\_won, "\n")

# Compute average total runs scored

if ("Total.Runs" %in% colnames(ipl\_data)) {

avg\_runs <- mean(ipl\_data$Total.Runs, na.rm = TRUE)

cat("Average Total Runs Scored:", avg\_runs, "\n")

} else {

cat("Total.Runs column not found in the dataset.\n")

}

# Calculate average total wickets taken

if ("Total.Wickets" %in% colnames(ipl\_data)) {

avg\_wickets <- mean(ipl\_data$Total.Wickets, na.rm = TRUE)

cat("Average Total Wickets Taken:", avg\_wickets, "\n")

} else {

cat("Total.Wickets column not found in the dataset.\n")

}

# Identify and display the most frequently used venue for matches

if ("Venue" %in% colnames(ipl\_data)) {

most\_frequent\_venue <- names(which.max(table(ipl\_data$Venue)))

cat("Most Frequently Used Venue:", most\_frequent\_venue, "\n")

} else {

cat("Venue column not found in the dataset.\n")

}

# Create a bar plot to visualize the number of matches won by each team

library(ggplot2)

ggplot(data = data.frame(Team = names(matches\_won), Wins = as.numeric(matches\_won)),

aes(x = reorder(Team, -Wins), y = Wins, fill = Team)) +

geom\_bar(stat = "identity") +

labs(title = "Number of Matches Won by Each Team",

x = "Team",

y = "Number of Wins") +

theme\_minimal()

Output:

tibble [5 × 8] (S3: tbl\_df/tbl/data.frame)

$ Match ID : num [1:5] 1 2 3 4 5

$ Date : POSIXct[1:5], format: "2022-02-15" "2022-06-24" "2023-05-06" ...

$ Team1 : chr [1:5] "Tigers" "Royals" "Majestics" "Titans" ...

$ Team2 : chr [1:5] "KKR" "RCB" "Victors" "Avengers" ...

$ Venue : chr [1:5] "Bangalore" "Mysore" "Dubai" "Hyderabad" ...

$ Winner : chr [1:5] "Tigers" "Royals" "Victors" "Titans" ...

$ Total.Runs : num [1:5] 120 200 50 49 89

$ Total.wickets: num [1:5] 3 1 5 6 2

Total Number of Matches: 5

Number of Unique Teams: 10

Unique Teams: Tigers Royals Majestics Titans Battlers KKR RCB Victors Avengers Batsmen

Matches Won by Each Team:

1 1 1 1 1

Average Total Runs Scored: 101.6

Total.Wickets column not found in the dataset.

Most Frequently Used Venue: Australia

> # Create a bar plot to visualize the number of matches won by each team

> library(ggplot2)

> ggplot(data = data.frame(Team = names(matches\_won), Wins = as.numeric(matches\_won)),

+ aes(x = reorder(Team, -Wins), y = Wins, fill = Team)) +

+ geom\_bar(stat = "identity") +

+ labs(title = "Number of Matches Won by Each Team",

+ x = "Team",

+ y = "Number of Wins") +

+ theme\_minimal()

Histograph:

