R Notebook

#FIFA World Cup(1930-2014) Project

# Loading all necessary libraries  
library(ggplot2)  
#ggplot2 for creating advanced visualizations (e.g., bar plots, line charts, etc.)  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

#dplyr for data manipulation (e.g., filtering, grouping, summarizing data)  
library(plotly)

##   
## Attaching package: 'plotly'

## The following object is masked from 'package:ggplot2':  
##   
## last\_plot

## The following object is masked from 'package:stats':  
##   
## filter

## The following object is masked from 'package:graphics':  
##   
## layout

#plotly for creating interactive visualizations  
library(viridis)

## Loading required package: viridisLite

#Load viridis for color palettes that are visually appealing.  
library(maps)

##   
## Attaching package: 'maps'

## The following object is masked from 'package:viridis':  
##   
## unemp

# maps for geographic data and map-based visualizations  
library(ggthemes)  
# ggthemes for additional themes and styles for ggplot2 plots  
library(tidyverse)

## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
## ✔ forcats 1.0.0 ✔ stringr 1.5.1  
## ✔ lubridate 1.9.4 ✔ tibble 3.2.1  
## ✔ purrr 1.0.2 ✔ tidyr 1.3.1  
## ✔ readr 2.1.5

## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ plotly::filter() masks dplyr::filter(), stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ✖ purrr::map() masks maps::map()  
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

# For data manipulation and visualization  
# tidyverse is a collection of R packages designed for data science workflows

# Took the data from kaggle (1930-2014)

#FIFA World Cup(1930-2014)

#This sets the working directory in R to the specified path.  
setwd("/Users/ajaykommineni/Downloads/archive")  
#This is useful when we want to load or save files without specifying the full path every time.  
  
wc\_matches <- read.csv("wcmatches.csv")  
#reads the CSV file named "wcmatches.csv" from the working directory.  
# View the first few rows of the dataset  
head(wc\_matches)

## year country city stage home\_team away\_team home\_score away\_score  
## 1 1930 Uruguay Montevideo Group 1 France Mexico 4 1  
## 2 1930 Uruguay Montevideo Group 4 Belgium United States 0 3  
## 3 1930 Uruguay Montevideo Group 2 Brazil Yugoslavia 1 2  
## 4 1930 Uruguay Montevideo Group 3 Peru Romania 1 3  
## 5 1930 Uruguay Montevideo Group 1 Argentina France 1 0  
## 6 1930 Uruguay Montevideo Group 1 Chile Mexico 3 0  
## outcome win\_conditions winning\_team losing\_team date month dayofweek  
## 1 H France Mexico 1930-07-13 Jul Sunday  
## 2 A United States Belgium 1930-07-13 Jul Sunday  
## 3 A Yugoslavia Brazil 1930-07-14 Jul Monday  
## 4 A Romania Peru 1930-07-14 Jul Monday  
## 5 H Argentina France 1930-07-15 Jul Tuesday  
## 6 H Chile Mexico 1930-07-16 Jul Wednesday

# Summary statistics  
summary(wc\_matches)

## year country city stage   
## Min. :1930 Length:900 Length:900 Length:900   
## 1st Qu.:1970 Class :character Class :character Class :character   
## Median :1990 Mode :character Mode :character Mode :character   
## Mean :1987   
## 3rd Qu.:2006   
## Max. :2018   
## home\_team away\_team home\_score away\_score   
## Length:900 Length:900 Min. : 0.000 Min. :0.000   
## Class :character Class :character 1st Qu.: 0.000 1st Qu.:0.000   
## Mode :character Mode :character Median : 1.000 Median :1.000   
## Mean : 1.569 Mean :1.262   
## 3rd Qu.: 2.000 3rd Qu.:2.000   
## Max. :10.000 Max. :8.000   
## outcome win\_conditions winning\_team losing\_team   
## Length:900 Length:900 Length:900 Length:900   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
## date month dayofweek   
## Length:900 Length:900 Length:900   
## Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character   
##   
##   
##

#it shows statistics like minimum, maximum, mean, median, and quartiles.  
str(wc\_matches)

## 'data.frame': 900 obs. of 15 variables:  
## $ year : int 1930 1930 1930 1930 1930 1930 1930 1930 1930 1930 ...  
## $ country : chr "Uruguay" "Uruguay" "Uruguay" "Uruguay" ...  
## $ city : chr "Montevideo" "Montevideo" "Montevideo" "Montevideo" ...  
## $ stage : chr "Group 1" "Group 4" "Group 2" "Group 3" ...  
## $ home\_team : chr "France" "Belgium" "Brazil" "Peru" ...  
## $ away\_team : chr "Mexico" "United States" "Yugoslavia" "Romania" ...  
## $ home\_score : int 4 0 1 1 1 3 0 0 1 6 ...  
## $ away\_score : int 1 3 2 3 0 0 4 3 0 3 ...  
## $ outcome : chr "H" "A" "A" "A" ...  
## $ win\_conditions: chr "" "" "" "" ...  
## $ winning\_team : chr "France" "United States" "Yugoslavia" "Romania" ...  
## $ losing\_team : chr "Mexico" "Belgium" "Brazil" "Peru" ...  
## $ date : chr "1930-07-13" "1930-07-13" "1930-07-14" "1930-07-14" ...  
## $ month : chr "Jul" "Jul" "Jul" "Jul" ...  
## $ dayofweek : chr "Sunday" "Sunday" "Monday" "Monday" ...

#It shows the number of rows and columns, the data type of each column, and a preview of the data.  
# Check for missing values  
colSums(is.na(wc\_matches))

## year country city stage home\_team   
## 0 0 0 0 0   
## away\_team home\_score away\_score outcome win\_conditions   
## 0 0 0 0 0   
## winning\_team losing\_team date month dayofweek   
## 169 169 0 0 0

# - The `is.na()` function checks for missing values (`NA`) in the data frame.  
# - `colSums()` sums up the number of missing values for each column.  
# - This helps identify which columns have missing data and how many missing values are present.

#This sets the working directory in R to the specified path.  
setwd("/Users/ajaykommineni/Downloads/archive")  
#This is useful when we want to load or save files without specifying the full path every time  
# Load the dataset  
worldcups <- read.csv("worldcups.csv")  
# View the first few rows of the dataset  
head(worldcups)

## year host winner second third fourth  
## 1 1930 Uruguay Uruguay Argentina USA Yugoslavia  
## 2 1934 Italy Italy Czechoslovakia Germany Austria  
## 3 1938 France Italy Hungary Brazil Sweden  
## 4 1950 Brazil Uruguay Brazil Sweden Spain  
## 5 1954 Switzerland West Germany Hungary Austria Uruguay  
## 6 1958 Sweden Brazil Sweden France West Germany  
## goals\_scored teams games attendance  
## 1 70 13 18 434000  
## 2 70 16 17 395000  
## 3 84 15 18 483000  
## 4 88 13 22 1337000  
## 5 140 16 26 943000  
## 6 126 16 35 868000

# Summary statistics  
summary(worldcups)

## year host winner second   
## Min. :1930 Length:21 Length:21 Length:21   
## 1st Qu.:1958 Class :character Class :character Class :character   
## Median :1978 Mode :character Mode :character Mode :character   
## Mean :1977   
## 3rd Qu.:1998   
## Max. :2018   
## third fourth goals\_scored teams   
## Length:21 Length:21 Min. : 70.0 Min. :13.00   
## Class :character Class :character 1st Qu.: 89.0 1st Qu.:16.00   
## Mode :character Mode :character Median :126.0 Median :16.00   
## Mean :121.3 Mean :21.76   
## 3rd Qu.:146.0 3rd Qu.:32.00   
## Max. :171.0 Max. :32.00   
## games attendance   
## Min. :17.00 Min. : 395000   
## 1st Qu.:32.00 1st Qu.: 943000   
## Median :38.00 Median :1774022   
## Mean :42.86 Mean :1898122   
## 3rd Qu.:64.00 3rd Qu.:2724604   
## Max. :64.00 Max. :3568567

str(worldcups)

## 'data.frame': 21 obs. of 10 variables:  
## $ year : int 1930 1934 1938 1950 1954 1958 1962 1966 1970 1974 ...  
## $ host : chr "Uruguay" "Italy" "France" "Brazil" ...  
## $ winner : chr "Uruguay" "Italy" "Italy" "Uruguay" ...  
## $ second : chr "Argentina" "Czechoslovakia" "Hungary" "Brazil" ...  
## $ third : chr "USA" "Germany" "Brazil" "Sweden" ...  
## $ fourth : chr "Yugoslavia" "Austria" "Sweden" "Spain" ...  
## $ goals\_scored: int 70 70 84 88 140 126 89 89 95 97 ...  
## $ teams : int 13 16 15 13 16 16 16 16 16 16 ...  
## $ games : int 18 17 18 22 26 35 32 32 32 38 ...  
## $ attendance : int 434000 395000 483000 1337000 943000 868000 776000 1614677 1673975 1774022 ...

# Check for missing values  
colSums(is.na(worldcups))

## year host winner second third fourth   
## 0 0 0 0 0 0   
## goals\_scored teams games attendance   
## 0 0 0 0

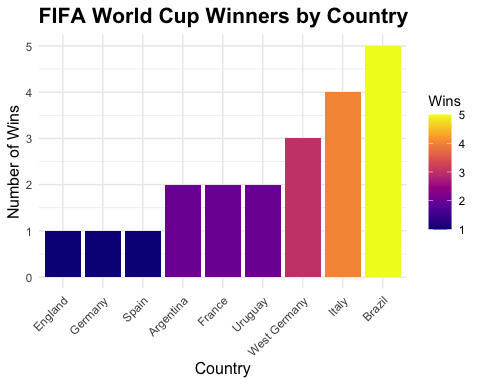
## After analysing the data of FIFA world cup. These are some visualization

# ————————–

# Visualization 1: World Cup Winners (Bar Chart)

#bar plot showing the number of FIFA World Cup wins by country # ————————–

winners <- worldcups %>%  
#Grouping the data by the `winner` column (countries that have won the World Cup)  
 group\_by(winner) %>%  
 summarize(Wins = n()) %>%  
#this will count the number of wins for each country using `n()` and store it in a new column `Wins`  
 arrange(desc(Wins))  
# Sorts the data in descending order of wins (most wins at the top)  
  
ggplot(winners, aes(x = reorder(winner, Wins), y = Wins, fill = Wins)) +  
#Creating bar plot to visualize the number of World Cup wins by country  
# Using the Wins column for the y-axis (number of wins)  
# Also Using the `Wins` column to fill the bars with color (gradient based on number of wins)  
 geom\_bar(stat = "identity") +  
# Create a bar plot where the height of the bars corresponds to the `Wins` values  
 scale\_fill\_viridis\_c(option = "plasma") +  
 labs(title = "FIFA World Cup Winners by Country",  
 # Adding a title to the plot and labels  
 x = "Country",  
 y = "Number of Wins") +  
 theme\_minimal() +  
 theme(axis.text.x = element\_text(angle = 45, hjust = 1),  
 # Rotating x-axis labels by 45 degrees and align them properly  
 plot.title = element\_text(size = 16, face = "bold"),  
 axis.title = element\_text(size = 12))



# Customize the axis title font size  
  
#Findings  
# The plot effectively highlights the dominance of Brazil, Germany, and Italy in the FIFA World Cup, while also showing the achievements of other countries. .It provides a clear visual representation of the distribution of World Cup wins and underscores the concentration of success among a small group of nations.  
#The distribution of wins is skewed, with a few countries (Brazil, Germany, Italy) dominating the tournament, while others have won only once or twice.

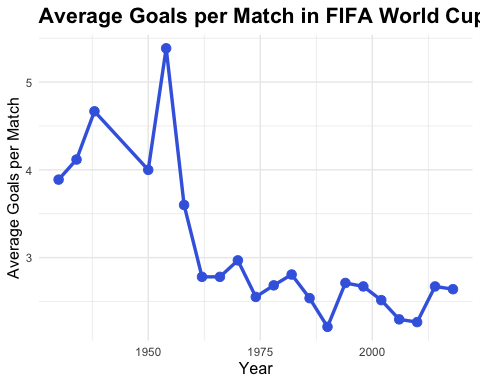
# ————————–

# Visualization 2: Average Goals per Match Over Time (Line Chart)

# ————————–

# Calculate the average goals per match for each World Cup edition  
worldcups <- worldcups %>%  
 mutate(AvgGoals = goals\_scored / games)  
#adds a new column to the `worldcups` data frame.  
#Avg goals is calculated by dividing the total goals scored (`goals\_scored`) by the number of games played (`games`).  
  
#Creating a line plot to visualize the average goals per match over time  
ggplot(worldcups, aes(x = year, y = AvgGoals)) +  
 geom\_line(color = "royalblue", size = 1.2) +  
 #Adds a line to the plot  
 geom\_point(color = "royalblue", size = 3) +  
 #Add points to the plot  
 labs(title = "Average Goals per Match in FIFA World Cups Over Time",  
# Adding a title to the plot and labels  
 x = "Year",  
 y = "Average Goals per Match") +  
 theme\_minimal() +  
 theme(plot.title = element\_text(size = 16, face = "bold"),  
 axis.title = element\_text(size = 12))

## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.  
## ℹ Please use `linewidth` instead.  
## This warning is displayed once every 8 hours.  
## Call `lifecycle::last\_lifecycle\_warnings()` to see where this warning was  
## generated.



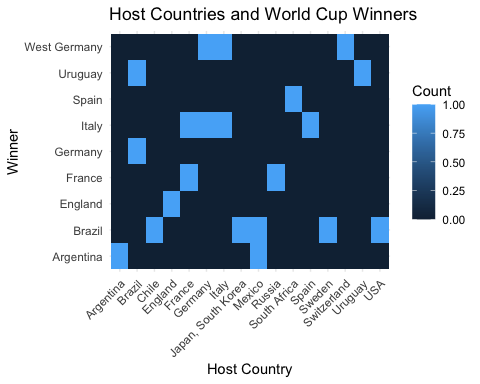
#Findings  
#There is no clear upward or downward trend over time, indicating that the average goals per match have remained relatively stable.  
#In recent timesthe tactical changes (e.g., more defensive strategies, better goalkeeping, and improved team organization) have influenced scoring patterns.  
#The stability in average goals per match in the modern era suggests that football has become more balanced, with teams focusing on both attacking and defensive strategies.

# ————————–

# Visualization 3: heatmap between host countries and World Cup winners

# ————————–

# Create a table of host and winner  
host\_winner\_table <- table(worldcups$host, worldcups$winner)  
# table function creates a contingency table that counts the occurrences of each combination of host and winner  
# Convert to a data frame for ggplot  
host\_winner\_df <- as.data.frame(host\_winner\_table)  
colnames(host\_winner\_df) <- c("Host", "Winner", "Count")  
#The `as.data.frame()` function converts the table into a data frame, which is easier to work with in ggplot2.  
#The resulting data frame has three columns:host, winner, and Freq is count.  
#Creating a heatmap  
ggplot(host\_winner\_df, aes(x = Host, y = Winner, fill = Count)) +  
#x = Host: Use the Host column for the x-axis (host countries).  
#y = Winner: Use the Winner column for the y-axis (winning countries).  
#fill = Count: Use the Count column to determine the color intensity of the tiles.  
 geom\_tile() +  
 labs(title = "Host Countries and World Cup Winners",  
# Adding a title to the plot and labels  
 x = "Host Country",  
 y = "Winner") +  
 theme\_minimal() +  
 theme(axis.text.x = element\_text(angle = 45, hjust = 1))



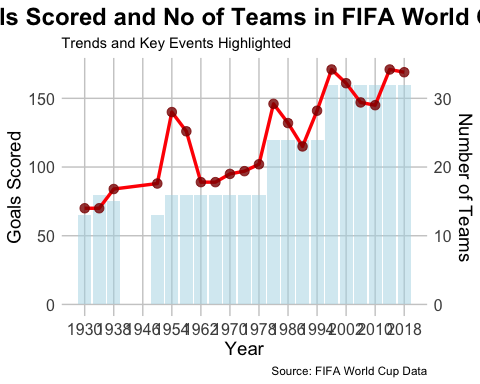
#Findings  
#The heatmap visualizes the relationship between host countries and World Cup winners  
#shows that host countries have won the World Cup multiple times  
#suggests that home advantage plays a significant role in a team's success, as hosts often perform better due to familiarity with conditions, fan support.  
#somecountries won the World Cup without hosting

# ————————–

# Visualization 4: goals scored and the number of teams played in FIFA World Cups over time

# ————————–

#Creating a combined plot to visualize goals scored and number of teams over time  
ggplot(worldcups, aes(x = year)) +  
# Add a bar plot for the number of teams  
 geom\_bar(aes(y = teams \* 5), stat = "identity", fill = "lightblue", alpha = 0.5) +   
# Scale teams for visibility  
# Add the main line for goals scored  
#Uses the actual values of teams for the bar heights.  
#Sets the bar color to light blue.  
#Makes the bars semi-transparent for better visibility of overlapping elements.  
 geom\_line(aes(y = goals\_scored), color = "red", size = 1.2) +  
# Add points for each World Cup  
 geom\_point(aes(y = goals\_scored), color = "darkred", size = 3, alpha = 0.8) +  
# Highlight specific years (e.g., highest and lowest goals scored)  
 theme\_minimal() +  
 theme(  
 plot.title = element\_text(size = 18, face = "bold", hjust = 0.5),  
 axis.title = element\_text(size = 14),  
 axis.text = element\_text(size = 12),  
 panel.grid.major = element\_line(color = "gray80"),  
 panel.grid.minor = element\_blank()  
 ) +  
# Adding titles and labels  
 labs(  
 title = "Goals Scored and No of Teams in FIFA World Cups",  
 subtitle = "Trends and Key Events Highlighted",  
 x = "Year",  
 y = "Goals Scored",  
 caption = "Source: FIFA World Cup Data"  
 ) +  
# Adjust the x-axis to show every 4 years (World Cup frequency)  
 scale\_x\_continuous(breaks = seq(1930, 2018, by = 8)) +  
# Add a secondary y-axis for the number of teams  
 scale\_y\_continuous(  
 name = "Goals Scored",  
 sec.axis = sec\_axis(~ . / 5, name = "Number of Teams") # Scale back to original team numbers  
 )



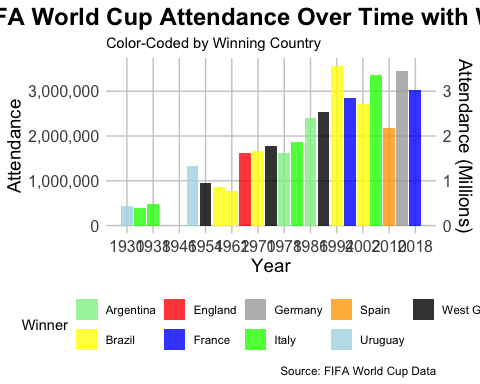
#Findings  
#The red line shows the total goals scored in each World Cup edition.  
#There is a general upward trend in goals scored over time, with some fluctuation due to increase in number of teams   
#The increase in teams correlates with the growth of the tournament and the inclusion of more countries.  
#The plot provides a comprehensive view of how goals scored and the number of teams have evolved in FIFA World Cups over time. It highlights key trends, such as the expansion of the tournament and the impact on goal totals, while also showing the fluctuations in playing styles. This analysis can help football analysts and fans understand the historical dynamics of the World Cup

# ————————–

# Visualization 5: FIFA World Cup attendance over time, with bars by the winning country.

# ————————–

# Creating a color palette for the winning countries  
winner\_colors <- c(  
 "Uruguay" = "lightblue",  
 "Italy" = "green",  
 "West Germany" = "black",  
 "Brazil" = "yellow",  
 "England" = "red",  
 "Argentina" = "lightgreen",  
 "Germany" = "darkgray",  
 "France" = "blue",  
 "Spain" = "orange"  
)  
# Create the plot  
ggplot(worldcups, aes(x = year, y = attendance, fill = winner)) +  
# Add bars for attendance, colored by the winning country  
#year column for the x-axis (World Cup editions).  
#attendancecolumn for the y-axis (total attendance).  
#winner column to fill the bars with colors based on the winning country.  
 geom\_bar(stat = "identity", alpha = 0.8) +  
#Uses the actual values of `attendance` for the bar heights  
 # Customize the theme  
 theme\_minimal() +  
# Applies a minimal theme for a clean and modern look.  
  
 theme(  
 plot.title = element\_text(size = 18, face = "bold", hjust = 0.5),  
 axis.title = element\_text(size = 14),  
 axis.text = element\_text(size = 12),  
 panel.grid.major = element\_line(color = "gray80"),  
 panel.grid.minor = element\_blank(),  
 legend.position = "bottom"  
 ) +  
# Adding titles and labels  
 labs(  
 title = "FIFA World Cup Attendance Over Time with Winners",  
 subtitle = "Color-Coded by Winning Country",  
 x = "Year",  
 y = "Attendance",  
 fill = "Winner",  
 caption = "Source: FIFA World Cup Data"  
 ) +  
# Adjust the x-axis to show every 4 years (World Cup frequency)  
 scale\_x\_continuous(breaks = seq(1930, 2018, by = 8)) +  
# Format the y-axis to show attendance in millions  
 scale\_y\_continuous(labels = scales::comma, sec.axis = sec\_axis(~ . / 1e6, name = "Attendance (Millions)")) +  
# Use the custom color palette for the bars  
 scale\_fill\_manual(values = winner\_colors)



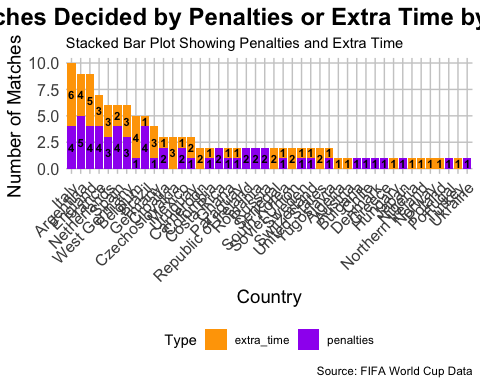
#Findings  
#The attendance has generally increased over time, reflecting the growing popularity of the World Cup.  
#It highlights key trends, such as the increase in attendance and the impact of hosting, while also showing the historical dominance of certain countries. fans understand the dynamics of World Cup attendance and its relationship with winning countries.  
#Future hosts can use this data to plan stadium capacities and marketing strategies to maximize attendance.

# ————————–

# Visualization 6:Matches Decided by Penalties or Extra Time by Country”

# ————————–

# Count matches with extra time or penalties by team  
extra\_time\_penalties <- wc\_matches %>%  
 filter(grepl("AET|penalties", win\_conditions)) %>%  
#Filters the dataset to include only matches that went to extra time (AET) or penalties.  
# Combine home and away teams into one column  
 pivot\_longer(cols = c(home\_team, away\_team), names\_to = "team\_type", values\_to = "team") %>%  
#reshapes the data by combining the home\_team and away\_team columns into a single column called team.  
 group\_by(team) %>%  
 #Stores the team names in the new column team.  
 summarise(  
 penalties = sum(grepl("penalties", win\_conditions)),  
 #counts the number of matches   
 extra\_time = sum(grepl("AET", win\_conditions))  
# grep is used again to count occurrences of "penalties" and "AET" in the `win\_conditions` column.  
 ) %>%  
 # Reshape the data for stacked bar plot  
 pivot\_longer(cols = c(penalties, extra\_time), names\_to = "type", values\_to = "count")  
#Creates a new column type to indicate whether the count is for penalties or extra time.  
# Create the plot  
ggplot(extra\_time\_penalties, aes(x = reorder(team, -count), y = count, fill = type)) +  
  
 geom\_bar(stat = "identity", position = "stack") +  
#Uses the actual values of count for the bar heights.  
#Stacks the bars for penalties and extra time on top of each other.  
# Add text labels for the total count  
 geom\_text(  
 aes(label = ifelse(count > 0, count, "")),  
 position = position\_stack(vjust = 0.5), color = "black", size = 3, fontface = "bold"  
 ) +  
 theme\_minimal() +  
 theme(  
 plot.title = element\_text(size = 18, face = "bold", hjust = 0.5),  
 axis.title = element\_text(size = 14),  
 axis.text = element\_text(size = 12),  
 axis.text.x = element\_text(angle = 45, hjust = 1),  
 panel.grid.major = element\_line(color = "gray80"),  
 panel.grid.minor = element\_blank(),  
 legend.position = "bottom"  
 ) +  
 # Add titles and labels  
 labs(  
 title = "Matches Decided by Penalties or Extra Time by Country",  
 subtitle = "Stacked Bar Plot Showing Penalties and Extra Time",  
 x = "Country",  
 y = "Number of Matches",  
 fill = "Type",  
 caption = "Source: FIFA World Cup Data"  
 ) +  
 # Use a custom color palette  
 scale\_fill\_manual(values = c("penalties" = "purple", "extra\_time" = "orange"))



#Findings  
#plot provides a clear visualization of the number of matches decided by penalties or extra time for each country. plot suggests that teams in high-pressure situations and underscores the importance of preparation for penalty shootouts and extra time  
#Countries with a high number of matches decided by penalties or extra time are often those that consistently reach the knockout stages of the tournament.Teams can use this data to analyze their performance in high-pressure situations and improve their strategies for penalty shootouts and extra time

# ————————–

# Visualization 7:unique countries that have played Fifa World cup (1930-2014)

# ————————–

# Extract unique countries that have played (home\_team and away\_team)  
participating\_countries <- unique(c(wc\_matches$home\_team, wc\_matches$away\_team))  
# Get world map data  
world\_map <- map\_data("world")  
#retrieves geographic data for the world map, including coordinates for each country.  
  
# Standardize country names to match the world map data  
participating\_countries <- case\_when(  
 participating\_countries == "USA" ~ "United States",  
 participating\_countries == "England" ~ "UK",  
 participating\_countries == "Czechoslovakia" ~ "Czech Republic",  
 participating\_countries == "West Germany" ~ "Germany",  
 participating\_countries == "Soviet Union" ~ "Russia",  
 participating\_countries == "Yugoslavia" ~ "Serbia",  
 participating\_countries == "Dutch West Indies" ~ "Netherlands",  
 participating\_countries == "FR Yugoslavia" ~ "Serbia",  
 TRUE ~ participating\_countries  
)  
#is used to standardize country names to match the naming conventions in the world map data.  
# Create a data frame for participating countries  
participating\_data <- data.frame(region = participating\_countries, played = TRUE)  
#A logical column set to TRUE for all countries that have participated in the World Cup.  
# Merge with world map data  
world\_map <- world\_map %>%  
 left\_join(participating\_data, by = "region")

## Warning in left\_join(., participating\_data, by = "region"): Detected an unexpected many-to-many relationship between `x` and `y`.  
## ℹ Row 33859 of `x` matches multiple rows in `y`.  
## ℹ Row 70 of `y` matches multiple rows in `x`.  
## ℹ If a many-to-many relationship is expected, set `relationship =  
## "many-to-many"` to silence this warning.

# Plot the map  
ggplot() +  
 geom\_polygon(data = world\_map, aes(x = long, y = lat, group = group, fill = played), color = "white", size = 0.1) +  
#draws the world map using the world\_map data.  
#Maps the longitude long, latitude lat, and group to the plot, and fills the polygons based on the played column.  
#Sets the border color of the countries to white.  
#Sets the border thickness to 0.1.  
 scale\_fill\_manual(values = c("TRUE" = "Blue", "FALSE" = "lightgray"), na.value = "lightgray", name = "Played in World Cup") +  
 labs(title = "Countries That Have Played in the FIFA World Cup",  
 subtitle = "All participating countries since 1930") +  
 theme\_void() +  
 theme(plot.title = element\_text(size = 16, face = "bold"),  
 legend.position = "bottom")



#Places the legend at the bottom of the plot.  
  
  
#Findings  
#The blue-colored countries represent those that have participated in the World Cup at least once.  
#The light gray-colored countries represent those that have never participated.  
#The map shows that most countries in Europe, South America, North America, and Africa have participated in the World Cup.