Untitled

2024-07-10

# Load necessary libraries  
library(readxl)

## Warning: package 'readxl' was built under R version 4.4.1

library(dplyr)

## Warning: package 'dplyr' was built under R version 4.4.1

##   
## 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

library(ggplot2)

## Warning: package 'ggplot2' was built under R version 4.4.1

# Load the data  
file\_path\_swathi <- "C:\\Users\\ASUS\\Downloads\\swathi.xlsx"  
data\_swathi <- read\_excel(file\_path\_swathi)

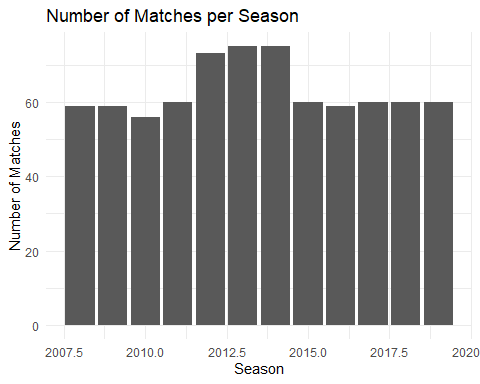
# Summary of the dataset  
summary(data\_swathi)

## id ipl season city   
## Min. : 1.0 Min. :2008 Length:756   
## 1st Qu.: 189.8 1st Qu.:2011 Class :character   
## Median : 378.5 Median :2013 Mode :character   
## Mean : 1792.2 Mean :2014   
## 3rd Qu.: 567.2 3rd Qu.:2016   
## Max. :11415.0 Max. :2019   
## date team1 team2   
## Min. :2008-04-18 00:00:00.00 Length:756 Length:756   
## 1st Qu.:2011-04-16 00:00:00.00 Class :character Class :character   
## Median :2013-05-09 12:00:00.00 Mode :character Mode :character   
## Mean :2013-10-07 06:22:51.43   
## 3rd Qu.:2016-05-18 06:00:00.00   
## Max. :2019-05-12 00:00:00.00   
## toss\_winner toss\_decision result dl\_applied   
## Length:756 Length:756 Length:756 Min. :0.00000   
## Class :character Class :character Class :character 1st Qu.:0.00000   
## Mode :character Mode :character Mode :character Median :0.00000   
## Mean :0.02513   
## 3rd Qu.:0.00000   
## Max. :1.00000   
## winner win\_by\_runs win\_by\_wickets player\_of\_match   
## Length:756 Min. : 0.00 Min. : 0.000 Length:756   
## Class :character 1st Qu.: 0.00 1st Qu.: 0.000 Class :character   
## Mode :character Median : 0.00 Median : 4.000 Mode :character   
## Mean : 13.28 Mean : 3.351   
## 3rd Qu.: 19.00 3rd Qu.: 6.000   
## Max. :146.00 Max. :10.000   
## venue umpire1 umpire2 Column1   
## Length:756 Length:756 Length:756 Length:756   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##

# Count of matches by season  
matches\_by\_season <- data\_swathi %>%  
 group\_by(`ipl season`) %>%  
 summarise(count = n())  
print(matches\_by\_season)

## # A tibble: 12 × 2  
## `ipl season` count  
## <dbl> <int>  
## 1 2008 59  
## 2 2009 59  
## 3 2010 56  
## 4 2011 60  
## 5 2012 73  
## 6 2013 75  
## 7 2014 75  
## 8 2015 60  
## 9 2016 59  
## 10 2017 60  
## 11 2018 60  
## 12 2019 60

# Plotting  
# Number of matches per season  
ggplot(matches\_by\_season, aes(x = `ipl season`, y = count)) +  
 geom\_bar(stat = "identity") +  
 theme\_minimal() +  
 labs(title = "Number of Matches per Season", x = "Season", y = "Number of Matches")



# Average runs scored in winning matches  
average\_runs <- data\_swathi %>%  
 group\_by(`ipl season`) %>%  
 summarise(avg\_runs = mean(`win\_by\_runs`, na.rm = TRUE))

ggplot(average\_runs, aes(x = `ipl season`, y = avg\_runs)) +  
 geom\_line() +  
 geom\_point() +  
 theme\_minimal() +  
 labs(title = "Average Runs Scored in Winning Matches by Season", x = "Season", y = "Average Runs")

