Liga MX - Study Case

Jair Herrera

2025-05-17

Load Libraries

**library**(dplyr)

## Warning: package 'dplyr' was built under R version 4.4.3

##

## 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 4.4.3

**library**(lubridate)

## Warning: package 'lubridate' was built under R version 4.4.3

##

## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':

##

## date, intersect, setdiff, union

**library**(tidyr)

## Warning: package 'tidyr' was built under R version 4.4.3

**library**(ggimage) *# For adding logos*

## Warning: package 'ggimage' was built under R version 4.4.3

Dataset Info

This dataset was obtained in Kaggle created by Gerardo Jaime Escareño.  
It can be found here [link](https://www.kaggle.com/datasets/gerardojaimeescareo/ligamx-matches-2016-2022?resource=download).

Load and Clean Data

df <- read.csv("C:/Users/Jair/Documents/2016\_2024\_Liga\_MX.csv")

df$date <- as.Date(df$date)

df$year <- year(df$date)

df$home\_team <- gsub("Le[oó]n", "Leon", df$home\_team)

df$away\_team <- gsub("Le[oó]n", "Leon", df$away\_team)

df$home\_team <- gsub("Atl[eé]tico San Luis", "Atletico San Luis", df$home\_team)

df$away\_team <- gsub("Atl[eé]tico San Luis", "Atletico San Luis", df$away\_team)

df <- df %>%

mutate(result = case\_when(

home\_goals > away\_goals ~ "HomeWin",

home\_goals < away\_goals ~ "AwayWin",

home\_goals == away\_goals ~ "Draw"

))

Team Statistics

home\_wins <- df %>%

filter(result == "HomeWin") %>%

count(team = home\_team, name = "home\_wins")

away\_wins <- df %>%

filter(result == "AwayWin") %>%

count(team = away\_team, name = "away\_wins")

home\_losses <- df %>%

filter(result == "AwayWin") %>%

count(team = home\_team, name = "home\_losses")

away\_losses <- df %>%

filter(result == "HomeWin") %>%

count(team = away\_team, name = "away\_losses")

home\_draws <- df %>%

filter(result == "Draw") %>%

count(team = home\_team, name = "home\_draws")

away\_draws <- df %>%

filter(result == "Draw") %>%

count(team = away\_team, name = "away\_draws")

total\_stats <- full\_join(home\_wins, away\_wins, by = "team") %>%

full\_join(home\_losses, by = "team") %>%

full\_join(away\_losses, by = "team") %>%

full\_join(home\_draws, by = "team") %>%

full\_join(away\_draws, by = "team") %>%

mutate(

total\_wins = coalesce(home\_wins, 0) + coalesce(away\_wins, 0),

total\_losses = coalesce(home\_losses, 0) + coalesce(away\_losses, 0),

total\_draws = coalesce(home\_draws, 0) + coalesce(away\_draws, 0)

) %>%

select(team, total\_wins, total\_losses, total\_draws) %>%

arrange(desc(total\_wins))

print(head(total\_stats, 22))

## team total\_wins total\_losses total\_draws

## 1 Club America 178 73 99

## 2 Tigres UANL 164 80 101

## 3 Monterrey 157 85 92

## 4 Cruz Azul 141 95 85

## 5 Leon 134 103 79

## 6 Pachuca 127 100 85

## 7 Toluca 126 104 85

## 8 Guadalajara Chivas 116 103 96

## 9 Santos Laguna 113 112 88

## 10 U.N.A.M. - Pumas 105 115 92

## 11 Club Tijuana 98 124 77

## 12 Atlas 94 132 81

## 13 Necaxa 91 118 89

## 14 Puebla 90 130 82

## 15 Club Queretaro 74 130 85

## 16 Atletico San Luis 63 93 41

## 17 Monarcas 51 53 38

## 18 FC Juarez 45 94 43

## 19 Mazatlán 38 76 41

## 20 Veracruz 22 77 21

## 21 Lobos Buap 20 37 11

## 22 Jaguares de Chiapas 7 20 7

Case Study Plot with Logos

A graph of a graph showing the number of different outcomes

AI-generated content may be incorrect.