

PremierLeague

Adithya Venghatesan 23206237

Final Project Part 1: Analysis

Database Used: English Premier League 2021-22 Match Data

from: <https://data.world/evangower/english-premier-league-2021-22-match-data>

```
# Import necessary libraries  
library(dplyr)
```

Warning: package 'dplyr' was built under R version 4.3.2

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(lubridate)
```

Attaching package: 'lubridate'

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

date, intersect, setdiff, union

```
# Read data from "soccer21-22.csv" file.
premData <- read.csv("soccer21-22.csv")
head(premData)
```

	Date	HomeTeam	AwayTeam	FTHG	FTAG	FTR	HTHG	HTAG	HTR	Referee	HS
1	13/08/2021	Brentford	Arsenal	2	0	H	1	0	H	M Oliver	8
2	14/08/2021	Man United	Leeds	5	1	H	1	0	H	P Tierney	16
3	14/08/2021	Burnley	Brighton	1	2	A	1	0	H	D Coote	14
4	14/08/2021	Chelsea	Crystal Palace	3	0	H	2	0	H	J Moss	13
5	14/08/2021	Everton	Southampton	3	1	H	0	1	A	A Madley	14
6	14/08/2021	Leicester	Wolves	1	0	H	1	0	H	C Pawson	9

	AS	HST	AST	HF	AF	HC	AC	HY	AY	HR	AR
1	22	3	4	12	8	2	5	0	0	0	0
2	10	8	3	11	9	5	4	1	2	0	0
3	14	3	8	10	7	7	6	2	1	0	0
4	4	6	1	15	11	5	2	0	0	0	0
5	6	6	3	13	15	6	8	2	0	0	0
6	17	5	3	6	10	5	4	1	2	0	0

Dataset cleaning

```
premData$Date <- as.Date(premData$Date, format = "%d/%m/%Y")
premData$month <- month(premData$Date)
# premData %>% relocate(month)
head(premData)
```

	Date	HomeTeam	AwayTeam	FTHG	FTAG	FTR	HTHG	HTAG	HTR	Referee	HS
1	2021-08-13	Brentford	Arsenal	2	0	H	1	0	H	M Oliver	8
2	2021-08-14	Man United	Leeds	5	1	H	1	0	H	P Tierney	16
3	2021-08-14	Burnley	Brighton	1	2	A	1	0	H	D Coote	14
4	2021-08-14	Chelsea	Crystal Palace	3	0	H	2	0	H	J Moss	13
5	2021-08-14	Everton	Southampton	3	1	H	0	1	A	A Madley	14
6	2021-08-14	Leicester	Wolves	1	0	H	1	0	H	C Pawson	9

	AS	HST	AST	HF	AF	HC	AC	HY	AY	HR	AR	month
1	22	3	4	12	8	2	5	0	0	0	0	8

2	10	8	3	11	9	5	4	1	2	0	0	8
3	14	3	8	10	7	7	6	2	1	0	0	8
4	4	6	1	15	11	5	2	0	0	0	0	8
5	6	6	3	13	15	6	8	2	0	0	0	8
6	17	5	3	6	10	5	4	1	2	0	0	8

General Analysis of the League

```
matchesPlayed <- nrow(premData)
cat("Number of matches played:",matchesPlayed,"\n")
```

Number of matches played: 380

```
matchdays <- length(unique(premData$Date))
cat("Number of unique Matchdays:",matchdays,"\n")
```

Number of unique Matchdays: 123

```
numberReferees <- length(unique(premData$Referee))
cat("Number of referees:",numberReferees,"\n")
```

Number of referees: 22

```
head(premData)
```

	Date	HomeTeam	AwayTeam	FTHG	FTAG	FTR	HTHG	HTAG	HTR	Referee	HS
1	2021-08-13	Brentford	Arsenal	2	0	H	1	0	H	M Oliver	8
2	2021-08-14	Man United	Leeds	5	1	H	1	0	H	P Tierney	16
3	2021-08-14	Burnley	Brighton	1	2	A	1	0	H	D Coote	14
4	2021-08-14	Chelsea	Crystal Palace	3	0	H	2	0	H	J Moss	13
5	2021-08-14	Everton	Southampton	3	1	H	0	1	A	A Madley	14
6	2021-08-14	Leicester	Wolves	1	0	H	1	0	H	C Pawson	9
AS HST AST HF AF HC AC HY AY HR AR month											
1	22	3	4	12	8	2	5	0	0	0	8
2	10	8	3	11	9	5	4	1	2	0	8
3	14	3	8	10	7	7	6	2	1	0	8
4	4	6	1	15	11	5	2	0	0	0	8
5	6	6	3	13	15	6	8	2	0	0	8
6	17	5	3	6	10	5	4	1	2	0	8

Final Project Part 2: Package

Final Project Part 3: Functions