

teams-cpmr

Charles Adedotun

2022-11-15

```
## Installing package into '/Users/runner/work/_temp/Library'
## (as 'lib' is unspecified)

##
## The downloaded binary packages are in
## /var/folders/24/8k48jl6d249_n_qfxwsl6xvm0000gn/T//RtmppGAy1E/downloaded_packages

## -- Attaching packages ----- tidyverse 1.3.2 --
## v ggplot2 3.4.0      v purrr 0.3.5
## v tibble 3.1.8       v dplyr 1.0.10
## v tidyr 1.2.1        v stringr 1.4.1
## v readr 2.1.3        v forcats 0.5.2
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

```
raw_data <- read.csv("../scraped-data/discipline.csv")
raw_data <- as_tibble(raw_data)
```

```
raw_columns <- raw_data %>% select(Team, Matches, Yellow.Cards)
head(raw_columns)
```

```
## # A tibble: 6 x 3
##   Team                Matches Yellow.Cards
##   <chr>                <int>      <int>
## 1 " Wolverhampton Wanderers "      15         29
## 2 " Nottingham Forest "          15         36
## 3 " Everton "                  15         33
## 4 " West Ham United "           15         19
## 5 " Southampton "              15         24
## 6 " Aston Villa "              15         37
```

```
new_table <- raw_columns %>% mutate(yellow_card_ratio = Yellow.Cards / Matches)
new_table <- new_table[order(-new_table$yellow_card_ratio),]
head(new_table)
```

```
## # A tibble: 6 x 4
##   Team                Matches Yellow.Cards yellow_card_ratio
##   <chr>                <int>      <int>          <dbl>
## 1 " Manchester United "      14         37          2.64
```

```
## 2 " Aston Villa "          15          37          2.47
## 3 " Nottingham Forest "    15          36          2.4
## 4 " Fulham "              15          36          2.4
## 5 " Crystal Palace "      14          33          2.36
## 6 " Everton "             15          33          2.2
```

```
data <- new_table %>% select(Team, yellow_card_ratio)
head(data)
```

```
## # A tibble: 6 x 2
##   Team                yellow_card_ratio
##   <chr>                <dbl>
## 1 " Manchester United "          2.64
## 2 " Aston Villa "              2.47
## 3 " Nottingham Forest "          2.4
## 4 " Fulham "                  2.4
## 5 " Crystal Palace "           2.36
## 6 " Everton "                  2.2
```

```
write.csv(data, "../cleaned-data/teams-cpmr.csv", row.names = FALSE)
```