

teams-cpmr

Charles Adedotun

2022-06-20

```
## 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//RtmpvNCFbx/downloaded_packages

## -- Attaching packages ----- tidyverse 1.3.1 --

## v ggplot2 3.3.6      v purrr 0.3.4
## v tibble 3.1.7       v dplyr 1.0.9
## v tidyr 1.2.0        v stringr 1.4.0
## v readr 2.1.2        v forcats 0.5.1

## -- 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 " Norwich City "                     38             55
## 2 " Watford "                           38             57
## 3 " Burnley "                           38             68
## 4 " Wolverhampton Wanderers "          38             59
## 5 " Leeds United "                     38            101
## 6 " Brighton and Hove Albion "          38             73

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 " Leeds United "      38        101           2.66
## 2 " Newcastle United "  38         79           2.08
## 3 " Aston Villa "      38         79           2.08
## 4 " Everton "          38         78           2.05
## 5 " Manchester United " 38         75           1.97
## 6 " Brighton and Hove Albion " 38         73           1.92
```

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

```
## # A tibble: 6 x 2
##   Team                yellow_card_ratio
##   <chr>                <dbl>
## 1 " Leeds United "      2.66
## 2 " Newcastle United "  2.08
## 3 " Aston Villa "      2.08
## 4 " Everton "          2.05
## 5 " Manchester United "  1.97
## 6 " Brighton and Hove Albion " 1.92
```

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