teams-cpmr

Charles Adedotun

2022-11-10

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

```
##
##   There is a binary version available but the source version is later:
##           binary source needs_compilation
## rmarkdown   2.17   2.18             FALSE
```

```
## installing the source package 'rmarkdown'
```

```
## -- 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 "      14           26
## 2 " Nottingham Forest "             14           35
## 3 "  Everton "                      14           33
## 4 "  West Ham United "              14           16
## 5 "  Southampton "                  14           22
## 6 "  Aston Villa "                  14           30
```

```
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 "      13           36               2.77
## 2 " Nottingham Forest "      14           35               2.5
## 3 "  Fulham "                14           35               2.5
## 4 "  Everton "               14           33               2.36
## 5 "  Crystal Palace "        13           29               2.23
## 6 "  Chelsea "               13           28               2.15
```

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

```
## # A tibble: 6 x 2
##   Team                    yellow_card_ratio
##   <chr>                               <dbl>
## 1 "  Manchester United "               2.77
## 2 " Nottingham Forest "               2.5
## 3 "  Fulham "                         2.5
## 4 "  Everton "                        2.36
## 5 "  Crystal Palace "                 2.23
## 6 "  Chelsea "                        2.15
```

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