

referees-cpmr

Charles Adedotun

2022-08-16

```
## 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//RtmpMCdaod/downloaded_packages

## -- Attaching packages ----- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr 0.3.4
## v tibble 3.1.8       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/referees.csv")
raw_data <- as_tibble(raw_data)
```

```
head(raw_data)
```

```
## # A tibble: 6 x 4
##   Referee      Matches Red.cards Yellow.cards
##   <chr>         <int>     <int>       <int>
## 1 Robert Jones      2         0         7
## 2 Michael Oliver    2         0         6
## 3 Anthony Taylor    2         2         8
## 4 Stuart Attwell    1         0         4
## 5 Peter Bankes      1         0         6
## 6 John Brooks       1         0         6
```

```
raw_columns <- raw_data %>% select(Referee, Matches, Yellow.cards)
head(raw_columns)
```

```
## # A tibble: 6 x 3
##   Referee      Matches Yellow.cards
##   <chr>         <int>       <int>
## 1 Robert Jones      2         7
## 2 Michael Oliver    2         6
## 3 Anthony Taylor    2         8
## 4 Stuart Attwell    1         4
## 5 Peter Bankes      1         6
## 6 John Brooks       1         6
```

```
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
##   Referee      Matches Yellow.cards yellow_card_ratio
##   <chr>         <int>         <int>         <dbl>
## 1 Peter Banks     1             6             6
## 2 John Brooks     1             6             6
## 3 Craig Pawson    1             5             5
## 4 Graham Scott    1             5             5
## 5 Paul Tierney    1             5             5
## 6 Anthony Taylor  2             8             4
```

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

```
## # A tibble: 6 x 2
##   Referee      yellow_card_ratio
##   <chr>         <dbl>
## 1 Peter Banks     6
## 2 John Brooks     6
## 3 Craig Pawson    5
## 4 Graham Scott    5
## 5 Paul Tierney    5
## 6 Anthony Taylor  4
```

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