

referees-cpmr

Charles Adedotun

2022-10-07

```
## 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//Rtmp4RJ0jr/downloaded_packages

## -- Attaching packages ----- tidyverse 1.3.2 --
## v ggplot2 3.3.6      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/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 " Michael Oliver "      7         0         23
## 2 " Anthony Taylor "      7         4         23
## 3 " Paul Tierney "        6         2         25
## 4 " Stuart Attwell "      5         2         18
## 5 " Simon Hooper "        5         0         15
## 6 " Robert Jones "        5         1         22
```

```
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 " Michael Oliver "      7         23
## 2 " Anthony Taylor "      7         23
## 3 " Paul Tierney "        6         25
## 4 " Stuart Attwell "      5         18
## 5 " Simon Hooper "        5         15
## 6 " Robert Jones "        5         22
```

```
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 " John Brooks "      3             16             5.33
## 2 " Peter Bankes "     4             21             5.25
## 3 " Michael Salisbury " 3             15              5
## 4 " Chris Kavanagh "   1              5              5
## 5 " Graham Scott "     1              5              5
## 6 " Craig Pawson "     5             23             4.6
```

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

```
## # A tibble: 6 x 2
##   Referee      yellow_card_ratio
##   <chr>         <dbl>
## 1 " John Brooks "      5.33
## 2 " Peter Bankes "     5.25
## 3 " Michael Salisbury "  5
## 4 " Chris Kavanagh "    5
## 5 " Graham Scott "      5
## 6 " Craig Pawson "     4.6
```

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