

referees-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//Rtmp6DNSKq/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/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 "      12         1         39
## 2 " Anthony Taylor "      12         5         44
## 3 " Paul Tierney "       12         2         45
## 4 " Simon Hooper "       10         0         28
## 5 " Stuart Attwell "      9         2         29
## 6 " Robert Jones "        9         1         34
```

```
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 "      12         39
## 2 " Anthony Taylor "      12         44
## 3 " Paul Tierney "       12         45
## 4 " Simon Hooper "       10         28
## 5 " Stuart Attwell "      9         29
## 6 " Robert Jones "        9         34
```

```
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 " Robert Madley "           1           6             6
## 2 " Chris Kavanagh "          5          26            5.2
## 3 " Michael Salisbury "       7          33            4.71
## 4 " Peter Bankes "           7          30            4.29
## 5 " John Brooks "            7          29            4.14
## 6 " Craig Pawson "           8          33            4.12
```

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

```
## # A tibble: 6 x 2
##   Referee      yellow_card_ratio
##   <chr>          <dbl>
## 1 " Robert Madley "           6
## 2 " Chris Kavanagh "         5.2
## 3 " Michael Salisbury "      4.71
## 4 " Peter Bankes "           4.29
## 5 " John Brooks "            4.14
## 6 " Craig Pawson "           4.12
```

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