

referees-cpmr

Charles Adedotun

2022-04-22

```
## 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//Rtmp7WSNG3/downloaded_packages

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

## v ggplot2 3.3.5      v purrr 0.3.4
## v tibble 3.1.6       v dplyr 1.0.8
## 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 Anthony Taylor      23        4        89
## 2 Paul Tierney        23        2        94
## 3 Martin Atkinson     22        2        53
## 4 Jonathan Moss       22        4        52
## 5 Michael Oliver      22        4        66
## 6 Craig Pawson        22        2        95

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 Anthony Taylor      23      89
## 2 Paul Tierney        23      94
## 3 Martin Atkinson     22      53
## 4 Jonathan Moss       22      52
## 5 Michael Oliver      22      66
## 6 Craig Pawson        22      95
```

```
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 Michael Salisbury     2         11           5.5
## 2 John Brooks           3         16           5.33
## 3 David Coote          19         90           4.74
## 4 Peter Bankes          11         51           4.64
## 5 Kevin Friend          14         62           4.43
## 6 Craig Pawson          22         95           4.32
```

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

```
## # A tibble: 6 x 2
##   Referee      yellow_card_ratio
##   <chr>          <dbl>
## 1 Michael Salisbury     5.5
## 2 John Brooks          5.33
## 3 David Coote           4.74
## 4 Peter Bankes          4.64
## 5 Kevin Friend          4.43
## 6 Craig Pawson          4.32
```

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