referees-cpmr

Charles Adedotun

2022-05-12

```
## 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//Rtmp93SweK/downloaded_packages
```

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

```
## v ggplot2 3.3.6      v purrr   0.3.4
## v tibble  3.1.7      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()
```

```r
raw_data <- read.csv("../scraped-data/referees.csv")
raw_data <- as_tibble(raw_data)
```

```r
head(raw_data)
```

```
## # A tibble: 6 x 4
##   Referee         Matches Red.cards Yellow.cards
##   <chr>             <int>     <int>        <int>
## 1 Craig Pawson         25         2          108
## 2 Jonathan Moss        24         4           57
## 3 Michael Oliver       24         5           72
## 4 Anthony Taylor       24         4           93
## 5 Paul Tierney         24         2           98
## 6 Martin Atkinson      23         2           56
```

```r
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 Craig Pawson        25         108
## 2 Jonathan Moss       24          57
## 3 Michael Oliver      24          72
## 4 Anthony Taylor      24          93
## 5 Paul Tierney        24          98
## 6 Martin Atkinson     23          56
```

```r
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         4           21              5.25
## 2 David Coote        20           93              4.65
## 3 Peter Bankes       12           55              4.58
## 4 Kevin Friend       16           71              4.44
## 5 Craig Pawson       25          108              4.32
## 6 Paul Tierney       24           98              4.08
```

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

```
## # A tibble: 6 x 2
##   Referee      yellow_card_ratio
##   <chr>                    <dbl>
## 1 John Brooks               5.25
## 2 David Coote               4.65
## 3 Peter Bankes              4.58
## 4 Kevin Friend              4.44
## 5 Craig Pawson              4.32
## 6 Paul Tierney              4.08
```

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