referees-cpmr

Charles Adedotun

2022-08-30

```
## 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//RtmpXBf0lT/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 1.4.1
## v readr   2.1.2      v forcats 0.5.2
## -- 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 Stuart Attwell      3         1            6
## 2 Peter Bankes        3         0           14
## 3 Jarred Gillett      3         0           12
## 4 Andy Madley         3         0            8
## 5 Michael Oliver      3         0           10
## 6 Craig Pawson        3         0           15
```

```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 Stuart Attwell      3            6
## 2 Peter Bankes        3           14
## 3 Jarred Gillett      3           12
## 4 Andy Madley         3            8
## 5 Michael Oliver      3           10
## 6 Craig Pawson        3           15
```

```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 Michael Salisbury       2           11               5.5
## 2 Craig Pawson            3           15               5
## 3 John Brooks             2           10               5
## 4 Graham Scott            1            5               5
## 5 Peter Bankes            3           14               4.67
## 6 Paul Tierney            3           14               4.67
```

```r
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 Craig Pawson                    5
## 3 John Brooks                     5
## 4 Graham Scott                    5
## 5 Peter Bankes                    4.67
## 6 Paul Tierney                    4.67
```

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