referees-cpmr

Charles Adedotun

2022-09-13

```
## 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//RtmpxU2ZvL/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   1.0.10
## v tidyr   1.2.1      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 Michael Oliver       5         0           16
## 2 Anthony Taylor       5         2           17
## 3 Paul Tierney         5         2           21
## 4 Stuart Attwell       4         1           13
## 5 Peter Bankes         4         0           21
## 6 Simon Hooper         4         0           12
```

```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 Michael Oliver       5           16
## 2 Anthony Taylor       5           17
## 3 Paul Tierney         5           21
## 4 Stuart Attwell       4           13
## 5 Peter Bankes         4           21
## 6 Simon Hooper         4           12
```

```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              3           16              5.33
## 2 Peter Bankes             4           21              5.25
## 3 Graham Scott             1            5              5
## 4 Craig Pawson             4           19              4.75
## 5 Michael Salisbury        3           14              4.67
## 6 Paul Tierney             5           21              4.2
```

```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                    5.33
## 2 Peter Bankes                   5.25
## 3 Graham Scott                   5
## 4 Craig Pawson                   4.75
## 5 Michael Salisbury              4.67
## 6 Paul Tierney                   4.2
```

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