referees-cpmr

Charles Adedotun

2022-09-04

```
## 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//RtmplOdNto/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   1.2.0      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()
```

```
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 Peter Bankes         4         0           21
## 2 Michael Oliver       4         0           13
## 3 Craig Pawson         4         0           19
## 4 Anthony Taylor       4         2           13
## 5 Paul Tierney         4         2           15
## 6 Stuart Attwell       3         1            6
```

```
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 Peter Bankes         4           21
## 2 Michael Oliver       4           13
## 3 Craig Pawson         4           19
## 4 Anthony Taylor       4           13
## 5 Paul Tierney         4           15
## 6 Stuart Attwell       3            6
```

```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 Peter Bankes           4           21             5.25
## 3 John Brooks            2           10                5
## 4 Graham Scott           1            5                5
## 5 Craig Pawson           4           19             4.75
## 6 Jarred Gillett         3           12                4
```

```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 Peter Bankes                  5.25
## 3 John Brooks                      5
## 4 Graham Scott                     5
## 5 Craig Pawson                  4.75
## 6 Jarred Gillett                   4
```

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