referees-cpmr

Charles Adedotun

2022-08-03

```
## 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//RtmpFzpEtT/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 1.4.0
## 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       28         5          104
## 2 Paul Tierney         27         5          107
## 3 Martin Atkinson      26         2           65
## 4 Mike Dean            26         2           82
## 5 Michael Oliver       26         7           77
## 6 Craig Pawson         26         2          109
```

```
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       28          104
## 2 Paul Tierney         27          107
## 3 Martin Atkinson      26           65
## 4 Mike Dean            26           82
## 5 Michael Oliver       26           77
## 6 Craig Pawson         26          109
```

```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        18           78              4.33
## 5 Craig Pawson        26          109              4.19
## 6 Darren England      15           60              4
```

```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                4.33
## 5 Craig Pawson                4.19
## 6 Darren England              4
```

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