referees-cpmr

Charles Adedotun

2022-10-25

```
## 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//RtmpF2IUBh/downloaded_packages

## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr   0.3.5
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## v readr   2.1.3      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 " Michael Oliver "    10         1           31
## 2 " Anthony Taylor "    10         5           35
## 3 " Paul Tierney "       9         2           39
## 4 " Simon Hooper "       8         0           25
## 5 " Stuart Attwell "     7         2           23
## 6 " Robert Jones "       7         1           28
```

```
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 "    10           31
## 2 " Anthony Taylor "    10           35
## 3 " Paul Tierney "       9           39
## 4 " Simon Hooper "       8           25
## 5 " Stuart Attwell "     7           23
## 6 " Robert Jones "       7           28
```

```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 "    5           24               4.8
## 2 " Craig Pawson "         7           31               4.43
## 3 " Paul Tierney "         9           39               4.33
## 4 " Peter Bankes "         6           26               4.33
## 5 " John Brooks "          5           21               4.2
## 6 " Robert Jones "         7           28               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 "            4.8
## 2 " Craig Pawson "                 4.43
## 3 " Paul Tierney "                 4.33
## 4 " Peter Bankes "                 4.33
## 5 " John Brooks "                  4.2
## 6 " Robert Jones "                 4
```

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