referees-cpmr

Charles Adedotun

2022-11-04

```
## 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//RtmpgjvS3t/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 "    11         1           32
## 2 " Anthony Taylor "    11         5           38
## 3 " Paul Tierney "      10         2           41
## 4 " Simon Hooper "       9         0           26
## 5 " Robert Jones "       8         1           28
## 6 " Andy Madley "        8         0           24
```

```
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 "    11           32
## 2 " Anthony Taylor "    11           38
## 3 " Paul Tierney "      10           41
## 4 " Simon Hooper "       9           26
## 5 " Robert Jones "       8           28
## 6 " Andy Madley "        8           24
```

```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 " Robert Madley "         1            6                 6
## 2 " Michael Salisbury "     6           31              5.17
## 3 " Craig Pawson "          7           31              4.43
## 4 " Peter Bankes "          6           26              4.33
## 5 " Paul Tierney "         10           41               4.1
## 6 " John Brooks "           6           24                 4
```

```r
data <- new_table %>% select(Referee, yellow_card_ratio)
head(data)
```

```
## # A tibble: 6 x 2
##   Referee             yellow_card_ratio
##   <chr>                           <dbl>
## 1 " Robert Madley "                   6
## 2 " Michael Salisbury "            5.17
## 3 " Craig Pawson "                 4.43
## 4 " Peter Bankes "                 4.33
## 5 " Paul Tierney "                  4.1
## 6 " John Brooks "                     4
```

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