Documentation

Nimit Sharma

5/8/2021

This is a baseball hall of fame players app

This is a simple app that allows users (baseball fans) to query Hall of Fame data and view top hitters and pitchers. The two simple metrics were used (Batting Average and Earned Run Average (ERA).

The use can query this data by selecting the country where the player was born, then select a range of height (inches) and weight (lbs). Finally the user can select the sort order based on their interest - Hitter vs. Pitcher.

The database was built using Lahman package. Pease see the libraries used below

```
# Import libraries
library(shiny)
library(shinythemes)
library(shinyWidgets)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.1.1     v dplyr   1.0.5
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(Lahman)
```

The data frame was constructed using various joins and filters

See code below:

```
a1<-Batting %>%
    group_by(playerID)%>%
    summarise(avg_HR = round(mean(HR, na.rm = TRUE)),
              career_HR = sum(HR, na.rm = TRUE),
              sum_H = sum(H, na.rm = TRUE),
              sum_AB = sum(AB, na.rm = TRUE))%>%
```

```
    mutate(bat_avg = round(sum_H/sum_AB,3))%>%
    select(playerID, bat_avg, avg_HR, career_HR) %>%
    arrange(desc(bat_avg))

a2<-a1%>%
    inner_join(HallOfFame, a1, by=c("playerID")) %>%
    filter((inducted == "Y") & (category == "Player"))%>%
    select(playerID, bat_avg, avg_HR, career_HR, yearID, category)

a3<- inner_join(a2, People, by=c("playerID")) %>%
    select(playerID, yearID, nameFirst, nameLast,   weight, height, birthCountry,
           bat_avg, avg_HR, career_HR, yearID, category)

p1<-Pitching %>%
    group_by(playerID)%>%
    summarise(avg_ERA = mean(ERA))

a4<-a3%>%
    left_join(p1, a3, by=c("playerID")) %>%
    select(playerID, yearID, nameFirst, nameLast,   weight, height, birthCountry,
           bat_avg, avg_HR, career_HR, avg_ERA, yearID, category)

df<-rename(a4, PlayerFirstName=nameFirst,
           PlayerLastName=nameLast,
           Year=yearID,
           Weight=weight,
           Height=height,
           CountryOfBirth=birthCountry,
           BattingAverage=bat_avg,
           AverageHomeRuns=avg_HR,
           CareerHomeRuns=career_HR,
           AverageERA=avg_ERA,
           Category=category)
```