

Lesson1

Kevin Cummiskey

January 8, 2020

Welcome to MA388 Sabermetrics!

The purpose of today is to get to know each other and get started looking at some baseball data. If you're not proficient in R yet, no worries, please feel free to ask me or get help from someone around you.

Who is your favorite baseball player?

Find your favorite player's **playerID** in the Master table and write it below. This ID will be used to access this player's data in other databases. See my code below for some assistance.

```
#my favorite player was Wade Boggs, UNTIL he became a Yankee
```

```
library(Lahman)
```

```
library(tidyverse)
```

```
#get Boggs' player ID
```

```
Master %>%
```

```
  filter(nameLast == "Boggs", nameFirst == "Wade")
```

```
##   playerID birthYear birthMonth birthDay birthCountry birthState
## 1 boggs001      1958          6       15          USA          NE
##   birthCity deathYear deathMonth deathDay deathCountry deathState
## 1   Omaha      NA          NA       NA          <NA>          <NA>
##   deathCity nameFirst nameLast   nameGiven weight height bats throws
## 1   <NA>      Wade   Boggs Wade Anthony   190    74    L     R
##      debut  finalGame  retroID  bbrefID deathDate  birthDate
## 1 1982-04-10 1999-08-27 boggw001 boggs001    <NA> 1958-06-15
```

Find out something interesting about your player using the Batting, Pitching, or Fielding tables. Feel free to use my code below... if you don't understand the code at this point, it's okay!

#I want to see if Wade Boggs had a higher batting average with the Red Sox or Yankees.

#Extract Wade Boggs seasonal batting data

```
boggs = Batting %>%  
  filter(playerID == "boggs01")
```

#calculate batting average by team

```
boggs %>%  
  group_by(teamID) %>%  
  summarize(BA = sum(H)/sum(AB))
```

```
## # A tibble: 3 x 2
```

```
##   teamID    BA
```

```
##   <fct>   <dbl>
```

```
## 1 BOS    0.338
```

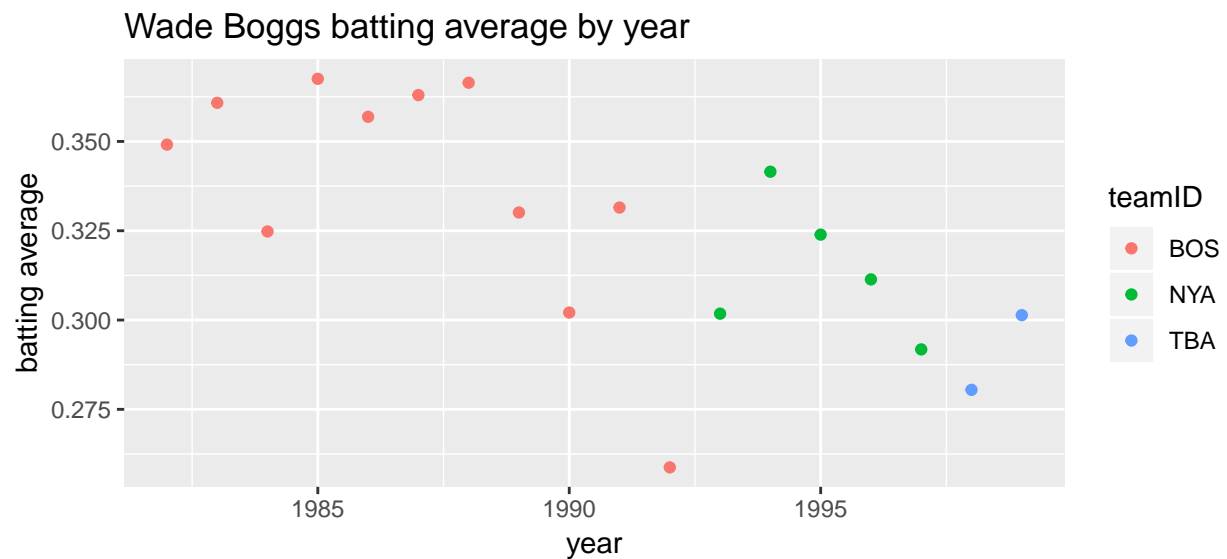
```
## 2 NYA    0.313
```

```
## 3 TBA    0.289
```

As I suspected, Wade Boggs was a better player when he was with the Red Sox!

#plot his batting average by season

```
boggs %>%  
  ggplot(aes(x = yearID, y = H/AB, color = teamID)) +  
  geom_point() +  
  labs(x = "year",  
       y = "batting average",  
       title = "Wade Boggs batting average by year")
```



Please fill out the following information and return it to me before the end of class.

- First Name:
 - Last Name:
 - Name you want me to call you:
 - Phone number I can reach you:
 - Favorite baseball team:
 - Academic Major:
 - Are there specific statistical methods you are interested in learning more about?
-
- Are there specific areas of sabermetrics you are interested in learning more about?