

# Lab 4

Install and load the package *Lahman*, which will give you access to the dataset *Teams*

```
install.packages("Lahman")  
library(Lahman)
```

- Produce a subset of the data that has the following characteristics:
  - Only one team (your choice)
  - data from 1980 to present (or as present as the dataset gets)
  - Includes 5 columns: name, yearID, W, L, R, RA

(The variables above correspond to the team name, the year, wins, losses, runs scored, and runs allowed)

- Make sure you select a team that is currently still around, or it probably won't be interesting (see a list of current teams [here](#)).
- Create a new variable corresponding to the winning percentage for the team you chose over time

$$w_{pct} = \frac{wins}{wins + losses}$$

- Order by winning percentage: Least to greatest
- Order by winning percentage: greatest to least
- Compute the mean and standard deviation of winning percentage
- With the full dataset
  - compute the average and standard deviation of winning percentage for each team.
  - Order by highest winning percentage