

Getting Multiple Years of Retrosheet Game Logs

Kevin Cummiskey

January 27, 2020

Obtaining Retrosheet Game Logs

In the Chapter 4 exercises, you will have to obtain multiple years of retrosheet game logs. Here are some instructions to help you.

1. Get the game log headers from the author's github.

```
library(tidyverse)

urlRemote = "https://raw.githubusercontent.com/"
pathGithub = "maxtoki/baseball_R/master/data/"
fileName = "game_log_header.csv"
game_log_headers <- read_csv(file = paste0(urlRemote,pathGithub,fileName))
```

2. Go to <https://www.retrosheet.org/gamelogs/index.html> and download the years you want. You can also get an entire decades as well.
3. Find the .zip file you downloaded and unzip it. Delete the .zip file.
4. Get the file names to upload.

```
# this is the path of the folder where you downloaded the game logs.
my_path = "C:/Users/kevin.cummiskey/Documents/MA388_GameLogs/"
files = list.files(path = my_path)
files = paste(my_path,files, sep = "")
files
```

```
## [1] "C:/Users/kevin.cummiskey/Documents/MA388_GameLogs/GL2000.TXT"
## [2] "C:/Users/kevin.cummiskey/Documents/MA388_GameLogs/GL2001.TXT"
## [3] "C:/Users/kevin.cummiskey/Documents/MA388_GameLogs/GL2002.TXT"
## [4] "C:/Users/kevin.cummiskey/Documents/MA388_GameLogs/GL2003.TXT"
## [5] "C:/Users/kevin.cummiskey/Documents/MA388_GameLogs/GL2004.TXT"
## [6] "C:/Users/kevin.cummiskey/Documents/MA388_GameLogs/GL2005.TXT"
## [7] "C:/Users/kevin.cummiskey/Documents/MA388_GameLogs/GL2006.TXT"
## [8] "C:/Users/kevin.cummiskey/Documents/MA388_GameLogs/GL2007.TXT"
## [9] "C:/Users/kevin.cummiskey/Documents/MA388_GameLogs/GL2008.TXT"
## [10] "C:/Users/kevin.cummiskey/Documents/MA388_GameLogs/GL2009.TXT"
```

5. Read in the game logs and combine into one data frame.

```
#read in data
game_logs = files %>%
  map(read_csv,
       col_names = names(game_log_headers),
       na = character()) %>%
  reduce(rbind)
```