For this project, I looked at the batting statistics for Baseball players and teams over the years. The file was Batting.csv

I wanted to know:

What were the total hits, runs, homeruns, RBIs, etc… over the course of the entire MLB existence? (I created a function that took keys to find the totals for the stats that mattered to the “user”, in this case me)

What were the averages of those statistics per player over the course of time?

Total Runs per year?

Total Run per team?

The average, standard deviation, min and max total runs per year

Description of Steps Taken:

* I uploaded the csv file using DictReader, and created a list object of dictionaries, called batting.
* I munged the data by using parse\_maybe\_int and parse\_date functions, to make sure the types for ints could be summed.
* I created functions sum\_columns and divide\_dictionary to sum any of the columns in the dictionary list that I wanted (the statistics of interest) and then divide those by the total length of the file to find the averages per player
* I used defaultdict to group the runs by team and by year
* I used numpy to find the average, standard deviation, min and max of those totals

Summary:

Runs by Year:

Mean: 13139.9166667

Std: 5295.37158556

Min:1904

Max: 24971