

Overview: For this scenario, we have been asked to write an application that will be used to provide information about baseball player statistics. Approach this problem as if it is an application going to production. We don't expect it to be perfect (no production code is), but we also don't want you to hack together a throw-away script. This should be representative of something that you would be comfortable releasing to a production environment. Also, spend whatever amount of time you think is reasonable. If you don't get all the requirements completed, that's ok. Just do the best you can with the time that you have available. You may use whatever gems, frameworks and tools that you think are appropriate, just provide any special setup instructions when you submit your solution.

Assumptions: All requests currently are based on data in the hitting file. Future requests of the system will require data from a pitching file as well. Consider this in the design.

Requirements: When the application is run, use the provided data and calculate the following results and write them to STDOUT.

- 1) Most improved batting average(hits / at-bats) from 2009 to 2010. Only include players with at least 200 at-bats.
- 2) Slugging percentage for all players on the Oakland A's (teamID = OAK) in 2007.
- 3) Who was the AL and NL triple crown winner for 2011 and 2012. If no one won the crown, output "(No winner)"

Formulas:

Batting average = hits / at-bats

Slugging percentage = $((\text{Hits} - \text{doubles} - \text{triples} - \text{home runs}) + (2 * \text{doubles}) + (3 * \text{triples}) + (4 * \text{home runs})) / \text{at-bats}$

Triple crown winner – The player that had the highest batting average AND the most home runs AND the most RBI in their league. It's unusual for someone to win this, but there was a winner in 2012. “Officially” the batting title (highest league batting average) is based on a minimum of 502 plate appearances. The provided dataset does not include plate appearances. It also does not include walks so plate appearances cannot be calculated. Instead, use a constraint of a minimum of 400 at-bats to determine those eligible for the league batting title.

Data: All the necessary data is available in the two csv files attached:

Batting-07-12.csv – Contains all the batting statistics from 2007-2012.

Column header key:

AB – at-bats

H – hits

2B – doubles

3B – triples

HR – home runs

RBI – runs batted in

Master-small.csv – Contains the demographic data for all baseball players in history through 2012.

Please note: We are looking for you to demonstrate your knowledge related to common software practices to include reusability, portability, and encapsulation – to name a few. Work submitted should be in project form and implemented as you were implementing any production solution.