MAT 3672 Project Fall 2019 Professor Natov: WAR
Due: On or before Wednesday 16 October 2019
No late projects accepted

WAR! No, not the kind with battles, but rather Wins Above Replacement. There is lots of data generated by baseball players; books such as *Money Ball* discuss how the management of teams, went from inuition to analytics. Every professional baseball team makes extensive use of statistics to meaure player performance.

A relatively new statistic is now considered the most important, and that is WAR. Wins Above Replacement indicates additional wins a team should expect. For example, Mike Trout has a WAR score of 10.2. If an average player on a baseball team were placed by Mike Trout, that team should average about 10.2 additional wins (during 162 games). If that seems a little vague, that's because it is.

In this project we are going to reverse engineer WAR using multivariate linear regression. More specifically, we will look at 2018 ESPN WAR values for a number of players. The goal is to determine a formula for computing it. There are many variables, and our regression analysis will help pick out those that are significant in determining WAR. At the time of this writing, ESPN WAR does not have a formula, so we have a chance to be among the first to derive one.

After the multivartiate regression analysis, we can use our formula on 2019 data and see how well it matches.

Write a conclusion which includes: your interpretation of the results, and what assumptions we made. What applications do we see for WAR? What projects might follow from this work?