Final Project Proposal: Baseball Operations

Seth Friman, Sophie Lefebvre, Grace Michael, Tara Sawhney, Jason Stitt

Motivation

Baseball is America's pastime. Collectively, in the MLB alone, the teams are worth around $66 billion. In our group, we have baseball fans and future baseball-professional hopefuls. We believe that the outcome of our project could theoretically be used by baseball front offices to assist in organizing rosters and setting lineups during the season. This could be a helpful project to have to show future recruiters as it shows a combination of our personal interests with our technical skills and knowledge.

Goals and objectives

Our code will create a player management system to project rosters and lineups for a randomly generated baseball game. We plan to complete this project in Jupyter Notebook, due namely to the visualizations we hope to produce. Baseball produces a lot of statistics, so the system can always be fine-tuned to take more and more into account and further improve the accuracy of the developer. We hypothesize that we will be able to accurately produce a sufficient, competitive roster of baseball players for any generated matchup.

Data sources

https://www.mlb.com/stats/

This will be our primary data source as it hosts all current and historic game and player stats in the MLB.

https://www.milb.com/stats/

This will be our secondary source of data. It shows the same information as the previous source, however, for Minor League (MiLB) teams and players.

https://baseballsavant.mlb.com/statcast_search

This is a site that accesses the previously mentioned MLB data via specific filtering measures so we can more easily and accurately check our data.

Platform architecture

We currently believe that the platform architecture for this project will be a re-usable software library. Our expected deliverables include our finalized game rosters, our starting lineups, and a series of visuals to support these selections. We are also

considering running a ‘baseball game’ with the rosters to see who the potential winner would be.

Repo Link:

https://github.ccs.neu.edu/jstitt/DS3500project