A major input into our models is prospect status. Prospect lists are published by Baseball America (others have begun to publish list’s but Baseball America has the most historical data) on a yearly basis. There are two main prospect lists that Pando is concerned with, the top 100 overall prospects and the top 10 prospects within a given organization. This project is focused on the latter. These lists are usually published throughout the subsequent offseason and can be viewed as a pseudo-grade based on a player’s most recent performance. This causes a lag effect – the most recent lists are based on the previous year’s performance. For example, prospect lists published in 2018 are evaluations of a player’s 2017 performance. This lag requires Pando to predict the individuals that will reside on these prospect lists. Therefore, the goal of this project is to predict the prospect status of **pitchers** that played in **2018**. In the attached csv, this is called “org\_top10\_post”. For the sake of this project a simplifying assumption is made:

1. Whether a prospect is a top 10 organizational prospect is a yes / no question (represented as 1 for yes and 0 for no). You are not asked to project where a player actually falls on the list.

This project has the attached data:

1. Spreadsheet in csv format of pitcher seasons who debuted between 1999 and 2018. This document includes production metrics for both a pitcher’s minor league and major league statistics (if they made the major leagues).
2. Spreadsheet in csv format detailing the features / columns included with the project

**Useful learning materials (for those unfamiliar with baseball)**

Baseball overview 1: <https://www.youtube.com/watch?v=skOsApsF0jQ>

Baseball overview 2: <https://www.youtube.com/watch?v=I8VGW0C_GO4>

Minor League Baseball (overview): <http://probaseballinsider.com/what-is-minor-league-baseball/>

Minor League Baseball (video): <https://www.youtube.com/watch?v=4ct5z5dcVWc>

MLB draft: <https://en.wikipedia.org/wiki/Major_League_Baseball_draft>

Baseball data: <https://en.wikipedia.org/wiki/Sabermetrics>