**Baseball Project Summary**

**Where do most players come from?**

We found that most players in baseball history were born in the US with 86.9% of the population being American born. This makes sense, because Major League Baseball has all their teams in North America and the closer a player is to a team, the easier and cheaper it is to scout them and sign them to a contract.

We then looked at the trend of Internationally born players by their debut year to see if the game was becoming more global. We found that the total percentage of Internationally born players increased for every decade with nearly 30% of players that debuted in the 2010s as being foreign born.

**What non-US countries produce the most players?**

We found that the Dominican Republic, Venezuela, Canada and Puerto Rico produce the most players in the Major Leagues. This is not surprising, because many teams have baseball academies in the D.R. and Venezuela that are used to sign and develop teenagers into Major Leaguers for a fraction of the cost compared to complexes in the United States. It also makes sense that Puerto Rico and Canada have high numbers, because they are the rare foreign countries that are subject to the amateur draft every year. The amateur draft ensures that a player’s signing bonus is restricted so it is usually cheaper to draft these players than to sign a player from Japan or Korea where they have prospering leagues that can offer more lucrative contracts to their native players. This makes it harder to lure these players away from their homeland and ensures that they would need bigger contract offers to leave the comforts of their home.

**What US regions produce the most players?**

Next, we divided the US into nine different regions to see where most US players come from. I was surprised to see that almost no Major League players were from the Mountain region. This is probably due to the smaller population and colder weather in the North, but I would have thought that the states of Nevada and Arizona would be able to boost the total number of players. Las Vegas and Phoenix are two well known areas for producing High School talent, but our data does not indicate this to be the case historically.

**Does size matter for Homers and Strikeouts?**

The last thing we wanted to examine was if player size significantly impacted career performance. The first relationship we analyzed was career homeruns vs. BMI. In theory, bigger players should have higher homerun totals than smaller players. However, we did not find a strong correlation between BMI and career homeruns. We calculated a .14 R-squared value. The correlation was probably weak, because there are many reasons why a player may not be able to accumulate home runs other than body size. Player durability and overall skill level could keep a player’s homerun total down while a player with less power overall could play twice as long and still accumulate the same amount of career homeruns. We may have gotten a stronger correlation if we looked at average homeruns per season instead of career totals.

The next relationship we examined was career strikeouts for pitchers vs. height. There is a belief that taller pitchers accumulate more strikeouts than shorter pitchers, so we wanted to see the correlation between the two variables. Again, we found a very weak correlation between career strikeouts and height with an R-squared value of .08. Our correlation was probably very weak for the same reasons of our previous experiment. Many things can go wrong during a player’s career, so it may have made more sense to look for a relationship between strikeouts per season and height instead.