Jake Mitchell

4-16-2018

Sabermetrics-Hoenigman

**Sabermetrics Final Project: Description**

**Rough plan for statistic:**

My goal for my statistic is to quantify a player’s baserunning intelligence. There are a lot of statistics that quantify a players base running ability, but until statcast data was made public we had no way to determine how much or a player’s performance on the bases was due to physical ability vs mental ability. A player’s sprint speed is highly indicative of their ability to run as shown by Lindsey Alder of Deadspin <https://deadspin.com/sprint-speed-helps-tell-us-whos-good-at-baserunning-and-1796424570>. This means that if we can find a league average measure of base running ability relative to their sprint speed, a players baserunning distance from this average will be indicative of their base running intelligence.

I will calculate a player’s BsR (from the fangraphs formula), and divide it by that players plate appearances for that range to get a rate statistic. I will then do some regression to find an equation that provides an expected BsR/PA based on a give sprint speed. Then player by player I will find the difference between their personal BsR/PA and their expected based on their sprint speed (their physical ability). I would say that this difference will be a decent representation of the players base running mental ability as it gives us their benefit after accounting for their physical abilities.

**How to get the data:**

I will be using the pybaseball package for python. This will hopefully allow me to get the data from the lahman database as well as all statcast data. In order to get the sprint speeds for individual players I will need to grab them from the sprint speed leaderboard at baseball savant as there is no api to download the information. Hopefully these will allow me to get all of the data I need. I will also likely need to hard code the values for things like the runs per stolen base and runs per caught stealing as they will be difficult to calculate at run time. I may try and make these available as inputs at run time.

**How to present the statistic:**

This is going to be the difficult part of this project for me as I have never built a web application. I do however have a lot of experience with python so I am hoping to make a Django Python website. I plan on hosting it using Heroku. This will allow me to focus on the backend and statistic portion of the assignment rather than fighting with the web application. I don’t know how pretty it will be, but it will hopefully be very easy to use. I don’t know how well it displays matplotlib plots, but that would be ideal to have them visualize the information and highlight the player you selected.