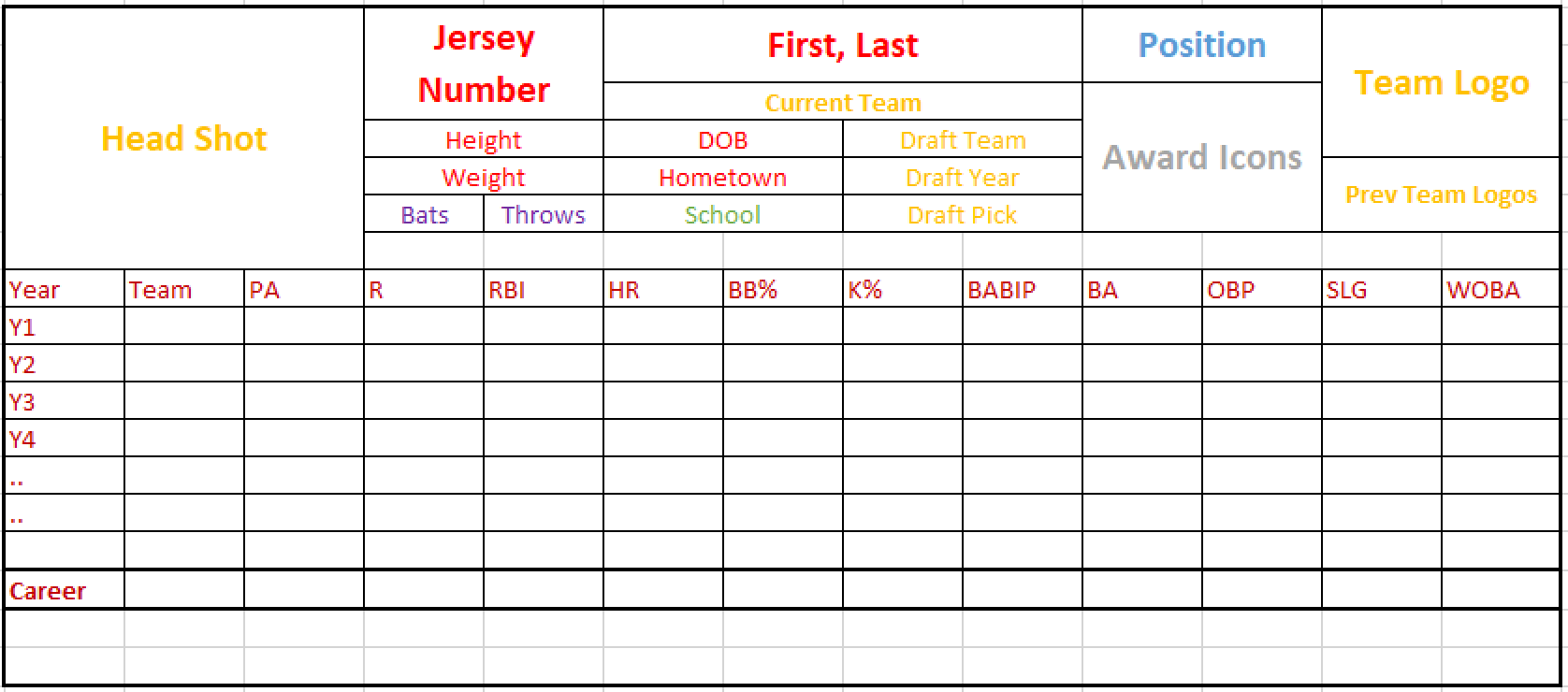
Example Data Project

*<your name>*

# Design



# Explanation

I chose to design a card for batters, and based it off of a typical baseball card. Baseball cards value is derived from tradition and for that reason, I chose to make minimal changes to the traditional baseball stats. Part of the allure is comparing the statistics from one card to another, across all different eras and drastically changing the stats on the card would make that impossible. Modern, more advanced statistics are certainly better at assessing the true talent of a player, but I think of baseball cards as more of a way to remember what happened, and for that reason, I think the customary descriptive statistics are appropriate. The changes I made were to change walks and strikeouts to rate stats to provide better context and adding wOBA, as a relatively understandable “all-compassing” offensive stat.

# Methodology

## Web Services

I started by looking at different baseball cards to figure out what I wanted to put on the card. After sketching it out, I explored the STATS API docs to find what endpoints had the required information. I was able to get all of the required information via two API calls.

### People

The people endpoint has almost everything related to a player, so the difficult part was exploring the hydrations and stat groups and types to determine what the options were and which were relevant. The documentation doesn’t include much information about what the options are or what is included with each, so it took some time to experiment with the different options. After figuring out the hydrations required (awards, education, stats), the stat group (hitting), and stat type (yearByYear), the most difficult part was unnesting the responses to get the data into a flat structure.

### Draft

The draft endpoint required the draft year and the personId. The person endpoint did not have detailed draft information available, so it required a second API call. For some reason, the player headshot is included in the draft endpoint and not the person endpoint, so that is being used as well. Because we are using such a small portion of the response from the draft endpoint, I took advantage of the fields parameter, which allows the web service call to specify which fields it wants in the response.

## Data Model

The data model is relatively simple. Some of the data, like the biographical information, is one observation per player. All of that data is combined into the person table, which is treated as the master record. The stats, awards, and teams are all variable-length per player, so those are each their own table, each of which has a many-to-one relationship with person.