**Metadata Guide**

**leagues\_MLB\_2015-value-batting\_players\_value\_batting.xlsx**

Baseball-reference.com. (2015). *2015 Major League Value Batting* [Exportable data set]. Retrieved from http://www.baseball-reference.com/leagues/MLB/2015-value-batting.shtml.

Date downloaded: 02 February 2017

To obtain:

Visit http://www.baseball-reference.com/leagues/MLB/2015-value-batting.shtml and scroll down to sortable Value Batting statistics for individual players. Just above the data, on the right side, is an “Export” option. Hovering over this option will display a message that says “Get a downloadable file suitable for Excel.” Click this option, and the file will download and open in Excel automatically. When saving the file, save it as a .xlsx workbook. Use the “find and replace” function in Excel to remove trailing pound signs and asterisks from player’s names. In the “find what” query, enter “~#” and replace all of these pound signs with nothing (i.e., click “Replace all” without entering anything into the “replace with” query). Repeat this process with “~\*” in the “find what” query.

Variables of interest:

Name

This variable comprises all players’ names.

Age

This variable indicates the age of a player at the conclusion of the season. Its units are years.

WAR

This variable indicates Wins Above Replacement, which measures the amount of additional wins a player contributes to his team over a replacement level player at his position.

**leagues\_MLB\_2014-value-batting\_players\_value\_batting.xlsx**

Baseball-reference.com. (2014). *2014 Major League Value Batting* [Exportable data set]. Retrieved from http://www.baseball-reference.com/leagues/MLB/2014-value-batting.shtml.

Date downloaded: 02 February 2017

To obtain:

Instructions for obtaining this dataset are identical to the instructions provided for the first hitter data entry.

Variables of interest:

The variables in this dataset used in the construction of or included in my final dataset are the exact same as those listed in the first hitter data entry.

**leagues\_MLB\_2013-value-batting\_players\_value\_batting.xlsx**

Baseball-reference.com. (2013). *2013 Major League Value Batting* [Exportable data set]. Retrieved from http://www.baseball-reference.com/leagues/MLB/2013-value-batting.shtml.

Date downloaded: 02 February 2017

To obtain:

Instructions for obtaining this dataset are identical to the instructions provided for the first hitter data entry.

Variables of interest:

The variables in this dataset used in the construction of or included in my final dataset are the exact same as those listed in the first hitter data entry.

**leagues\_MLB\_2012-value-batting\_players\_value\_batting.xlsx**

Baseball-reference.com. (2012). *2012 Major League Value Batting* [Exportable data set]. Retrieved from http://www.baseball-reference.com/leagues/MLB/2012-value-batting.shtml.

Date downloaded: 02 February 2017

To obtain:

Instructions for obtaining this dataset are identical to the instructions provided for the first hitter data entry.

Variables of interest:

The variables in this dataset used in the construction of or included in my final dataset are the exact same as those listed in the first hitter data entry.

**leagues\_MLB\_2011-value-batting\_players\_value\_batting.xlsx**

Baseball-reference.com. (2011). *2011 Major League Value Batting* [Exportable data set]. Retrieved from http://www.baseball-reference.com/leagues/MLB/2011-value-batting.shtml.

Date downloaded: 02 February 2017

To obtain:

Instructions for obtaining this dataset are identical to the instructions provided for the first hitter data entry.

Variables of interest:

The variables in this dataset used in the construction of or included in my final dataset are the exact same as those listed in the first hitter data entry.

**leagues\_MLB\_2010-value-batting\_players\_value\_batting.xlsx**

Baseball-reference.com. (2010). *2010 Major League Value Batting* [Exportable data set]. Retrieved from http://www.baseball-reference.com/leagues/MLB/2010-value-batting.shtml.

Date downloaded: 02 February 2017

To obtain:

Instructions for obtaining this dataset are identical to the instructions provided for the first hitter data entry.

Variables of interest:

The variables in this dataset used in the construction of or included in my final dataset are the exact same as those listed in the first hitter data entry.

**2015\_Contract\_Data.xlsx**

ESPN.com. (2015). *2015 MLB Free agents* [Webpage pasted into excel]. Retrieved from http://www.espn.com/mlb/freeagents/\_/year/2015.

Date downloaded: 02 February 2017

To obtain:

Visit http://www.espn.com/mlb/freeagents/\_/year/2015. The data are not exportable via a direct download, but they can be easily copied and pasted into excel. To do this, highlight all the data on the page, including variable titles. Open a blank Excel workbook, click cell A1, and select the “paste special” option at the top left of the screen. Choose “text.” The dataset will then be pasted as text into the workbook.

Variables of interest:

PLAYER

This variable comprises all players’ names.

POS

This variable indicates each player’s position as an abbreviation. For example, “SS” stands for shortstop. This variable is important because all pitchers were dropped from the data (i.e., all players whose POS variables were SP, P, or RP).

STATUS

This variable indicates a player’s contractual status and can be either Signed, FA (free agent), or Retired. All players whose status was not Signed were dropped from the data.

YRS

This variable indicates the length of a player’s contract. It is measured in years and is equal to the “length” variable in the final dataset.

DOLLARS

This variable indicates the total amount of money the contract is worth and is measured in dollars. This variable is equal to the “totalvalue” variable in the final dataset. Further, “avgannual” in the final dataset is equal to totalvalue/length.

**2014\_Contract\_Data.xlsx**

ESPN.com. (2014). *2014 MLB Free agents* [Webpage pasted into excel]. Retrieved from http://www.espn.com/mlb/freeagents/\_/year/2014.

Date downloaded: 02 February 2017

To obtain:

Instructions for obtaining this dataset are identical to the instructions provided for the first contract data entry.

Variables of interest:

The variables in this dataset used in the construction of or included in my final dataset are the exact same as those listed in the first contract data entry.

**2013\_Contract\_Data.xlsx**

ESPN.com. (2013). *2013 MLB Free agents* [Webpage pasted into excel]. Retrieved from http://www.espn.com/mlb/freeagents/\_/year/2013.

Date downloaded: 02 February 2017

To obtain:

Instructions for obtaining this dataset are identical to the instructions provided for the first contract data entry.

Variables of interest:

The variables in this dataset used in the construction of or included in my final dataset are the exact same as those listed in the first contract data entry.

**2012\_Contract\_Data.xlsx**

ESPN.com. (2012). *2012 MLB Free agents* [Webpage pasted into excel]. Retrieved from http://www.espn.com/mlb/freeagents/\_/year/2012.

Date downloaded: 02 February 2017

To obtain:

Instructions for obtaining this dataset are identical to the instructions provided for the first contract data entry.

Variables of interest:

The variables in this dataset used in the construction of or included in my final dataset are the exact same as those listed in the first contract data entry.

**Player\_Race.xlsx**

Masella, John. (2016). *Player Race.* Created by author. Filename: Player\_Race.xlsx.

(File created by author.)

Date created: 18 November 2016.

To obtain:

Contact the author at jmasella@haverford.edu or at johnmasella95@yahoo.com.

Variables of interest:

Name

This variable comprises the names of all players included in the 150 player-year observations in the final dataset.

Race

This variable indicates the race of a player and can assume “A” for Asian, “B” for black, “H” for Hispanic, and “W” for white. In the final dataset, a dummy exists that is equal to one if a player’s Race variable is A, B, or H, and equal to zero if a player’s Race variable is W. In this way, players are broken down into non-white and white categories.