

The Baseball Database

1 Introduction

This dataset has 2454 baseball players with 12 sheets to describe their statistics. You don't have to analyze all attributes of the data. Select a point of view you are interested and use the visualization skills taught in class such as brush, link and highlight, to explore the dataset!

2 Data Tables

The data follow these general principles. Each player is assigned a unique number (playerID). All of the information relating to that player is tagged with his playerID. The playerIDs are in PlayerList, and are linked to names and birthdates in the MASTER table.

<http://bbs.3dmgame.com/thread-3172776-1-1.html>, it's reference for baseball terms in Chinese.

The database is comprised of the following main tables:

- MASTER - Player names, DOB, and biographical info
- Batting - batting statistics
- Pitching - pitching statistics
- Fielding - fielding statistics

It is supplemented by these tables:

- AllstarFull - All-Star appearances
- BattingPost - post-season batting statistics
- PitchingPost - post-season pitching statistics
- FieldingPost - post-season fielding data
- Salaries - player salary data
- Appearances - details on the positions a player appeared at
- CollegePlaying - list of players and the colleges they attended
- PlayeList - All players

3 Detail Description

Sections 3.1 through 3.12 of this document describe each of the tables in detail and the fields that each contains.

3.1 MASTER table

playerID	A unique code assigned to each player. The playerID links the data in this file with records in the other files.
----------	--

birthYear	Year player was born
birthMonth	Month player was born
birthday	Day player was born
birthCountry	Country where player was born
birthState	State where player was born
birthCity	City where player was born
deathYear	Year player died
deathMonth	Month player died
deathDay	Day player died
deathCountry	Country where player died
deathState	State where player died
deathCity	City where player died
nameFirst	Player's first name
nameLast	Player's last name
nameGiven	Player's given name (typically first and middle)
weight	Player's weight in pounds
height	Player's height in inches
bats	Player's batting hand (left, right, or both)
throws	Player's throwing hand (left or right)
debut	Date that player made first major league appearance
finalGame	Date that player made first major league appearance (blank if still active)
retroID	ID used by retrosheet
bbrefID	ID used by Baseball Reference website

3.2 Batting Table

playerID	Player ID code
yearID	Year
stint	player's stint (order of appearances within a season)
teamID	Team
lgID	League
G	Games
AB	At Bats
R	Runs
H	Hits
2B	Doubles
3B	Triples
HR	Homeruns
RBI	Runs Batted In
SB	Stolen Bases
CS	Caught Stealing
BB	Base on Balls
SO	Strikeouts
IBB	Intentional walks

HBP	Hit by pitch
SH	Sacrifice hits
SF	Sacrifice flies
GIDP	Grounded into double plays

3.3 Pitching table

playerID	Player ID code
yearID	Year
stint	player's stint (order of appearances within a season)
teamID	Team
lgID	League
W	Wins
L	Losses
G	Games
GS	Games Started
CG	Complete Games
SHO	Shutouts
SV	Saves
IPOuts	Outs Pitched (innings pitched x 3)
H	Hits
ER	Earned Runs
HR	Homeruns
BB	Walks
SO	Strikeouts
BAOpp	Opponent's Batting Average
ERA	Earned Run Average
IBB	Intentional Walks
WP	Wild Pitches
HBP	Batters Hit By Pitch
BK	Balks
BFP	Batters faced by Pitcher
GF	Games Finished
R	Runs Allowed
SH	Sacrifices by opposing batters
SF	Sacrifice flies by opposing batters
GIDP	Grounded into double plays by opposing batter

3.4 Fielding Table

playerID	Player ID code
yearID	Year
stint	player's stint (order of appearances within a season)

teamID	Team
lgID	League
Pos	Position
G	Games
GS	Games Started
InnOuts	Time played in the field expressed as outs
PO	Putouts
A	Assists
E	Errors
DP	Double Plays
PB	Passed Balls (by catchers)
WP	Wild Pitches (by catchers)
SB	Opponent Stolen Bases (by catchers)
CS	Opponents Caught Stealing (by catchers)
ZR	Zone Rating

3.5 AllstarFull table

playerID	Player ID code
YearID	Year
gameNum	Game number (zero if only one All-Star game played that season)
gameID	Retrosheet ID for the game idea
teamID	Team
lgID	League
GP	1 if Played in the game
startingPos	If player was game starter, the position played

3.6 BattingPost table

yearID	Year
round	Level of playoffs
playerID	Player ID code
teamID	Team
lgID	League
G	Games
AB	At Bats
R	Runs
H	Hits
2B	Doubles
3B	Triples
HR	Homeruns
RBI	Runs Batted In

SB	Stolen Bases
CS	Caught stealing
BB	Base on Balls
SO	Strikeouts
IBB	Intentional walks
HBP	Hit by pitch
SH	Sacrifices
SF	Sacrifice flies
GIDP	Grounded into double plays

3.7 PitchingPost table

playerID	Player ID code
yearID	Year
round	Level of playoffs
teamID	Team
lgID	League
W	Wins
L	Losses
G	Games
GS	Games Started
CG	Complete Games
SHO	Shutouts
SV	Saves
IPOuts	Outs Pitched (innings pitched x 3)
H	Hits
ER	Earned Runs
HR	Homeruns
BB	Walks
SO	Strikeouts
BAOpp	Opponents' batting average
ERA	Earned Run Average
IBB	Intentional Walks
WP	Wild Pitches
HBP	Batters Hit By Pitch
BK	Balks
BFP	Batters faced by Pitcher
GF	Games Finished
R	Runs Allowed
SH	Sacrifice Hits allowed
SF	Sacrifice Flies allowed
GIDP	Grounded into Double Plays

3.8 Salaries table

yearID	Year
teamID	Team
lgID	League
playerID	Player ID code
salary	Salary

3.9 FieldingPost table

playerID	Player ID code
yearID	Year
teamID	Team
lgID	League
round	Level of playoffs
Pos	Position
G	Games
GS	Games Started
InnOuts	Time played in the field expressed as outs
PO	Putouts
A	Assists
E	Errors
DP	Double Plays
TP	Triple Plays
PB	Passed Balls
SB	Stolen Bases allowed (by catcher)
CS	Caught Stealing (by catcher)

3.10 Appearances table

yearID	Year
teamID	Team
lgID	League
playerID	Player ID code
G_all	Total games played
GS	Games started
G_batting	Games in which player batted
G_defense	Games in which player appeared on defense
G_p	Games as pitcher
G_c	Games as catcher

G_1b	Games as firstbaseman
G_2b	Games as secondbaseman
G_3b	Games as thirdbaseman
G_ss	Games as shortstop
G_lf	Games as leftfielder
G_cf	Games as centerfielder
G_rf	Games as right fielder
G_of	Games as outfielder
G_dh	Games as designated hitter
G_ph	Games as pinch hitter
G_pr	Games as pinch runner

3.11 CollegePlaying table

playerid	Player ID code
schoolID	school ID code
year	year

3.12 PalyerList table

playerID	Player ID code
----------	----------------