

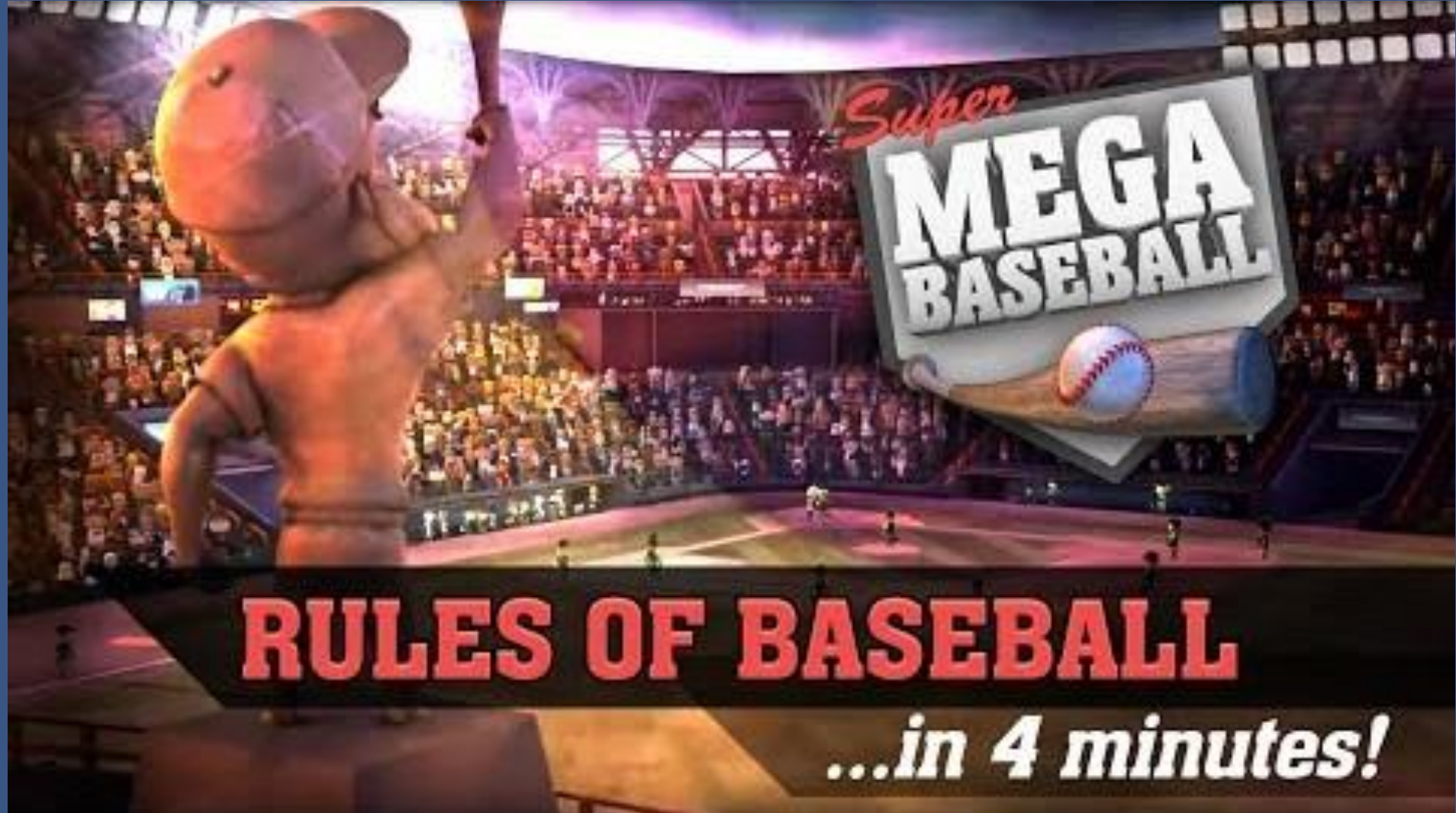
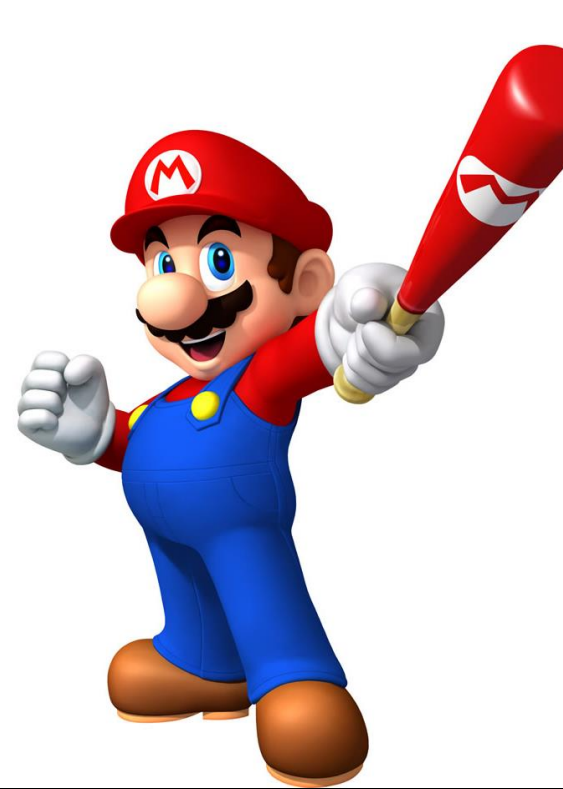


Baseball I



Produced by Dr. Mario | UNC STOR 390

Overview of Baseball

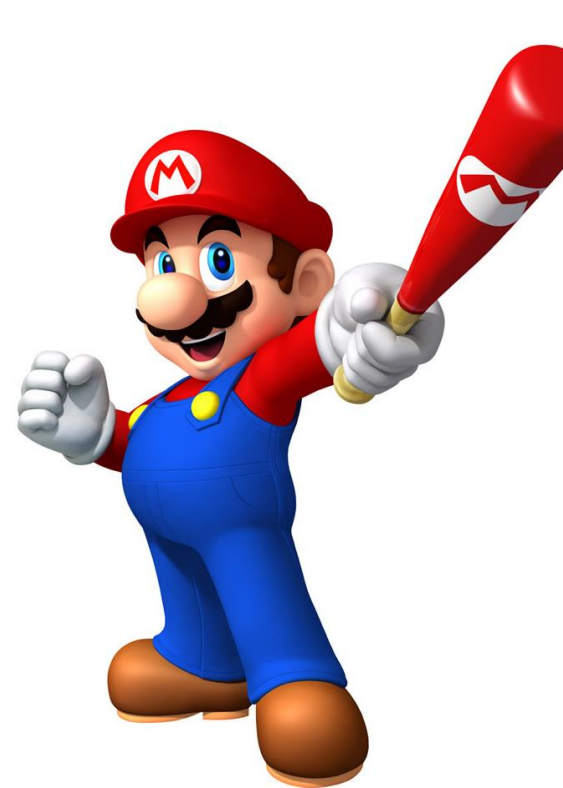




History of Sabermetrics

- **Society of American Baseball Research (SABR)**
 - August 10, 1971 in Cooperstown, NY
 - Founded by Bob Davids
 - 16 Original "Statihistorians"
 - Publication: *Baseball Research Journal*
 - Yearly Dues: \$65 (Standard) or \$25 (Student)
- **Objectives of SABR**
 - To foster the study of baseball as a significant American social and athletic institution.
 - To establish an accurate historical account of baseball through the years.
 - To facilitate the dissemination of baseball research information.
 - To stimulate the best interests of baseball as our national pastime.
 - To cooperate in safeguarding the proprietary interests of individual research efforts of members of the Society.



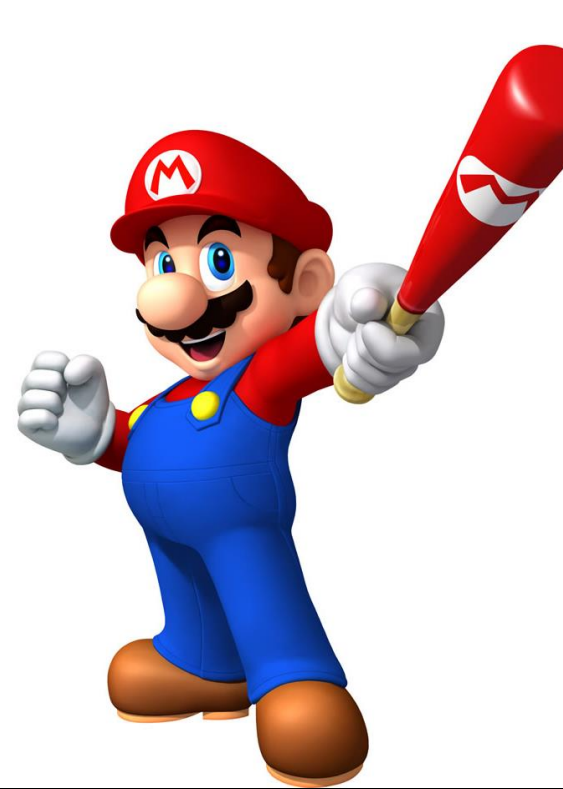


History of Sabermetrics

- **Bill James**
 - Author and Researcher Focused on Baseball
 - Publication: *Baseball Abstracts* (Annual Since 1977)
 - Term Coined in 1980: "Saber" = "SABR"
 - Sabermetrics Is the Search for Objective Knowledge About Baseball
 - One of Time Magazines 100 Most Influential (2006)
- **Sabermetrics**
 - Using Mathematics and Statistics to Understand Baseball
 - Development of Advanced Metrics to Measure Performance
 - Textbooks by Gabriel Costa, Michael Huber, and John Saccoman
 - *Understanding Sabermetrics* (2008)
 - *Practicing Sabermetrics* (2009)



History of Sabermetrics



- *Moneyball: The Art of Winning an Unfair Game*
 - Authored by Michael Lewis (2003)
 - About Oakland Athletics MLB Season in 2002
 - General Manager Billy Beane
 - Finished with the Same Number of Wins as New York Yankees
 - Payroll Difference: \$44M to \$125M
 - Statistical Analyses Proved Flaws in Classic Statistics
 - Identified Undervalued Players Using Different Metrics such as On-Base Percentage
 - Used Statistics to Modify Playing Style (Devaluation of Steals)



History of Sabermetrics





History of Sabermetrics

- Jim Albert
 - Statistics Professor at Bowling Green State University
 - *Sabermetrics: The Past, The Present, and The Future* (2010)
 - *Analyzing Baseball Data with R*: (2014)
- Useful Websites
 - www.baseballprospectus.com
 - www.hardballtimes.com
 - www.retrosheet.org
- Free Data Sources
 - Lahman Database
 - Retrosheet
 - PITCH F/X





Classic and Modern Metrics

- **Measuring Batting**

- Batting Average (AVG)

$$AVG = \frac{H}{AB}$$

- On-base Percentage (OBP)

$$OBP = \frac{H + BB + HBP}{AB + BB + HBP + SF}$$

- Slugging Percentage (SLG)

$$SLG = \frac{S + 2D + 3T + 4HR}{AB}$$

- OPS

$$OPS = OBP + SLG$$

H = Hit

AB = At-bat

BB = Walk

HBP = Hit by Pitch

SF = Sacrifice Fly

S = Single

D = Double

T = Triple

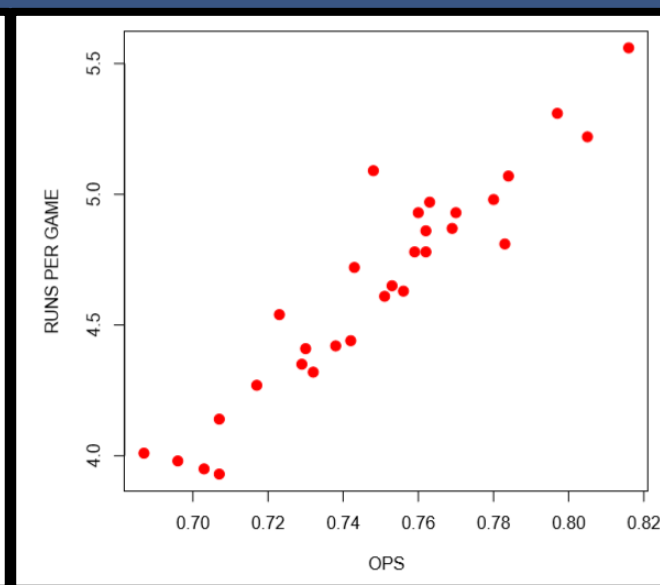
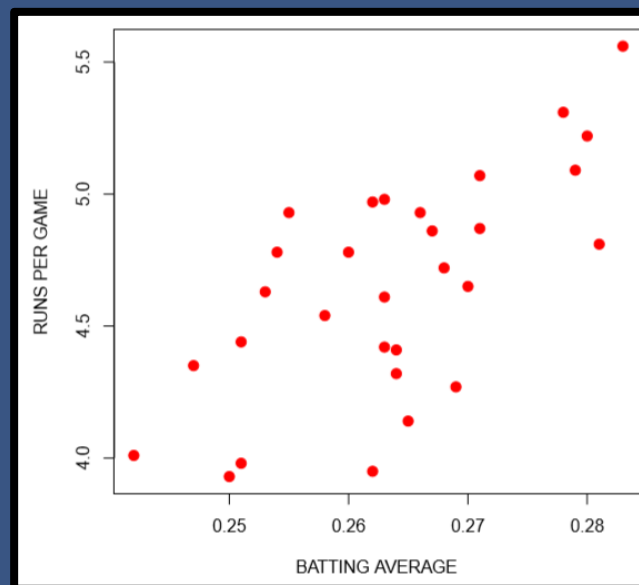
HR = Home Run





Classic and Modern Metrics

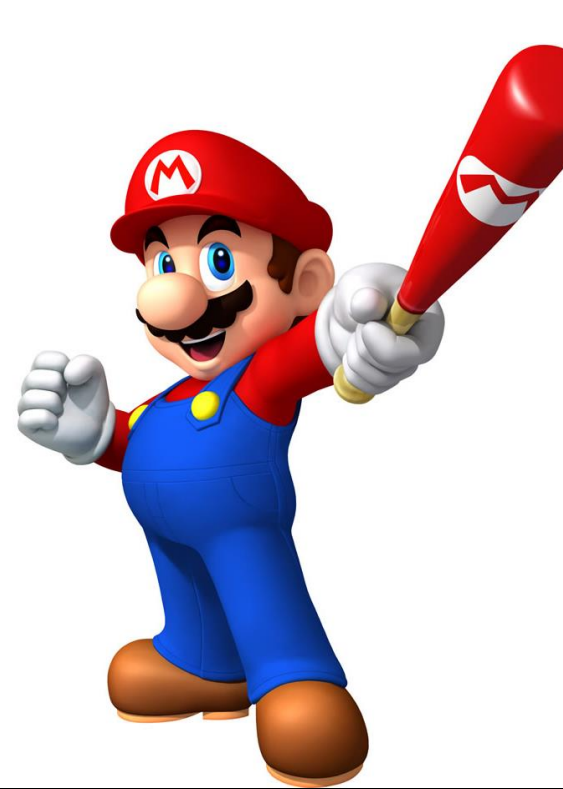
- Measuring Batting
 - Runs Per Game (R/G)



H = Hit
AB = At-bat
BB = Walk
HBP = Hit by Pitch
SF = Sacrifice Fly
S = Single
D = Double
T = Triple
R = Run
G = Game



Classic and Modern Metrics



- Measuring Pitching
 - Win Percentage (WIN%)

$$WIN\% = \frac{W}{W + L}$$

- Earned Run Average (ERA)

$$ERA = 9 \times \frac{ER}{IP}$$

- Strikeout Rate (K/9)

$$K/9 = 9 \times \frac{K}{IP}$$

- Defense-Independent Pitching Statistic (DIPS)
- DICE

$$DICE = 3 + \frac{13 \times HR + 3(BB \times HBP) - 2 \times K}{IP}$$

W = Win

L = Loss

ER = Earned Run

IP = Innings

K = Strikeout

HR = Home Run

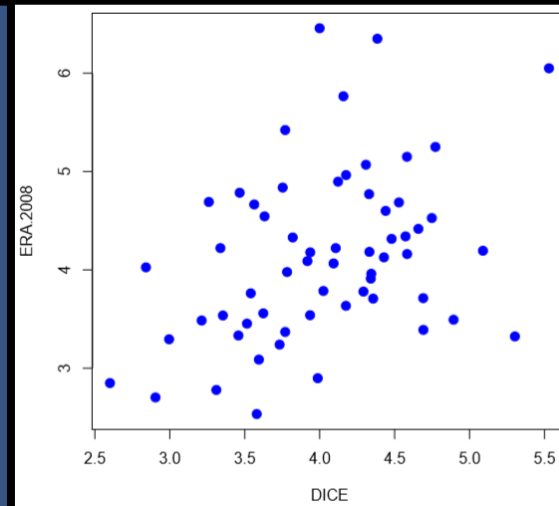
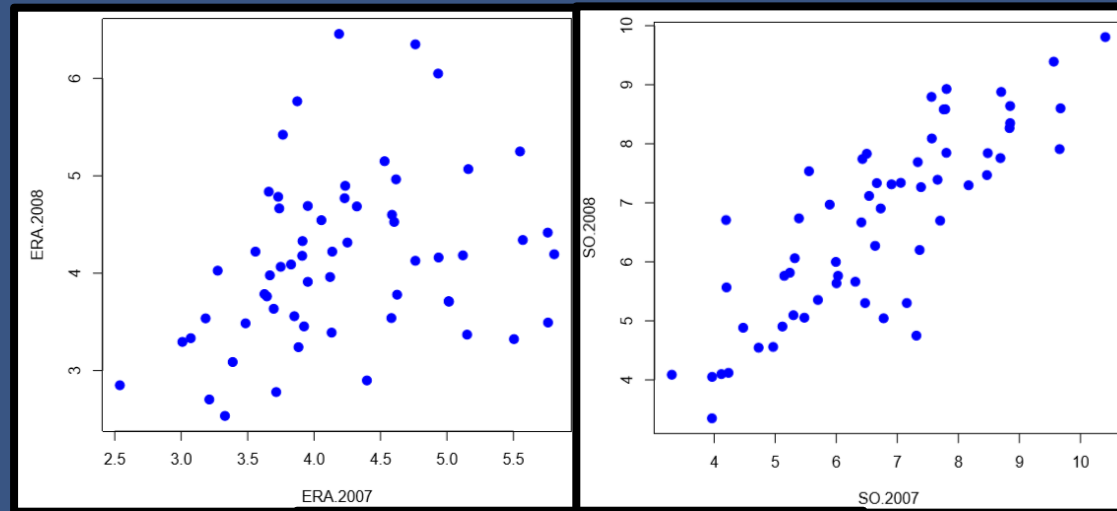
BB = Walk

HBP = Hit by Pitch



Classic and Modern Metrics

- Measuring Pitching



W = Win

L = Loss

ER = Earned Run

IP = Innings

K = Strikeout

HR = Home Run

BB = Walk

HBP = Hit by Pitch



Classic and Modern Metrics



- Measuring Fielding
 - Fielding Percentage (FLD%)

$$FLD\% = \frac{PO + A}{PO + A + E}$$

- Range Factor (RF)

$$RF = 9 \times \frac{PO + A}{IP}$$

- Omar Vizquel and Derek Jeter

PO = Put-out

A = Assists

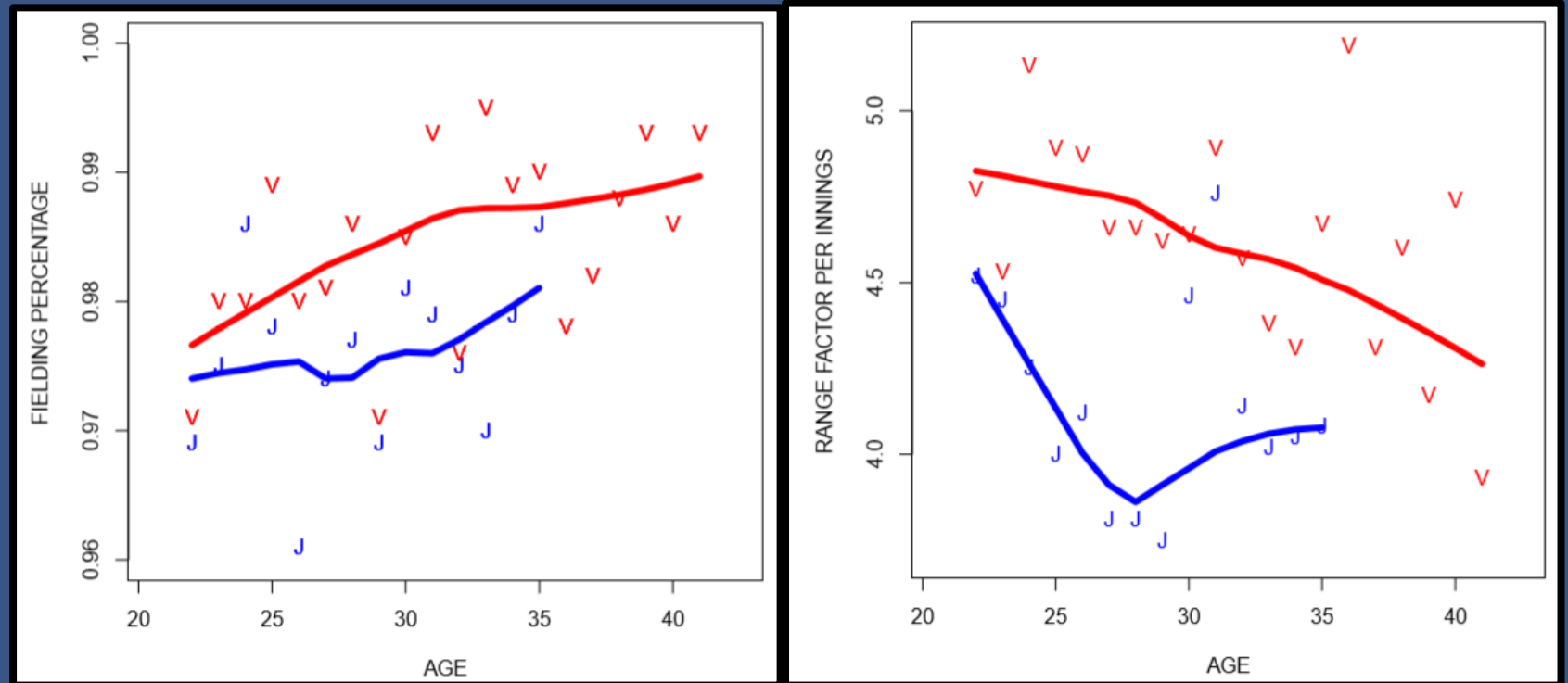
E = Errors



Classic and Modern Metrics

PO = Put-out
A = Assists
E = Errors

- Measuring Fielding
 - Omar Vizquel and Derek Jeter
 - Omar (9 Golden Gloves)
 - Derek (3 Golden Gloves)





Pythagorean Theorem

WP = Win %

RS = Runs Scored

RA = Runs Allowed

- Classic Pythagorean Theorem
 - Relationship Between the Sides of a Right Triangle
 - $a^2 = b^2 + c^2$
- What is Known: More Runs = More Wins
- Relationship Between Runs and Wins?
 - Bill James' Pythagorean Method

$$WP \approx \frac{RS^2}{RS^2 + RA^2}$$

- Example: Kansas City in 2014 World Series

- 651 Runs Scored
- 624 Runs Allowed

$$89 \text{ Wins} \approx 162 \times \frac{651^2}{651^2 + 624^2} = 84.43$$



Pythagorean Theorem

WP = Win %
RS = Runs Scored
RA = Runs Allowed

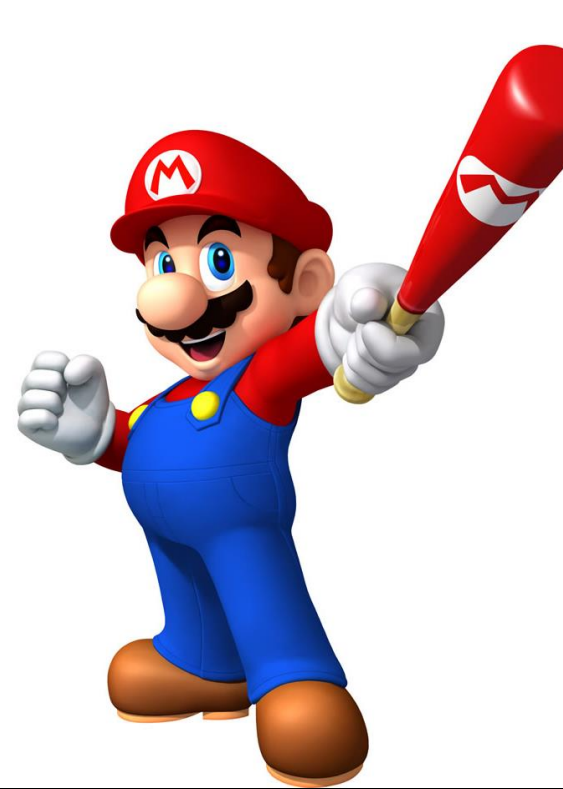
- Optimization of Relationship
 - What is the Best Choice of α ?

$$WP = \frac{RS^\alpha}{RS^\alpha + RA^\alpha} + \epsilon \longrightarrow \text{\#Blessed}$$

- Minimization of Sum of Squared Errors
- Optimal: $\alpha = 1.82$
- Alternative Expression

$$WP = \frac{(RS/RA)^\alpha}{(RS/RA)^\alpha + 1} + \epsilon$$

- Useful for Forecasting Playoff Series Winners
 - Pythagorean Method: 53.8% Accurate
 - Games Won Approach: 50% Accurate



Pythagorean Theorem

WP = Win %
RS = Runs Scored
RA = Runs Allowed

- Useful for Valuing Players in Trades

- Example: Cleveland Indians
- Currently: RS=870 and RA=800
- Trade Bing Crosby (100 Runs)
- For Frank Sinatra (120 Runs)
- Difference: +20 Runs
- Before Trade:

$$WP \approx \frac{\left(\frac{870}{800}\right)^{1.82}}{\left(\frac{870}{800}\right)^{1.82} + 1} = 0.538$$

- After Trade:

$$WP \approx \frac{\left(\frac{890}{800}\right)^{1.82}}{\left(\frac{890}{800}\right)^{1.82} + 1} = 0.548$$





Final Inspiration

There's No Crying in Statistics.

- Mahatma Mario