

# Questioning the Truth: Sabermetrics and Baseball

---

by John Aleshunas

Math & Computer Science Department

Walker School of Business and  
Technology

Who are you going to believe, me or  
your lying eyes?

- Groucho Marx

# Full Disclosure Disclaimer

While this presentation appears to be about baseball, it is actually about mathematics, computer science, and critical thinking.

# An Example...

**The Situation:** One out into the 8<sup>th</sup> inning of the 7<sup>th</sup> game of the 2003 ALCS. Boston leads New York 5-2. The Yankees begin a rally versus the Red Sox ace, Pedro Martinez.

**The Data:** Throughout the 2003 season Martinez was a ticking bomb after passing the 100-pitch mark.

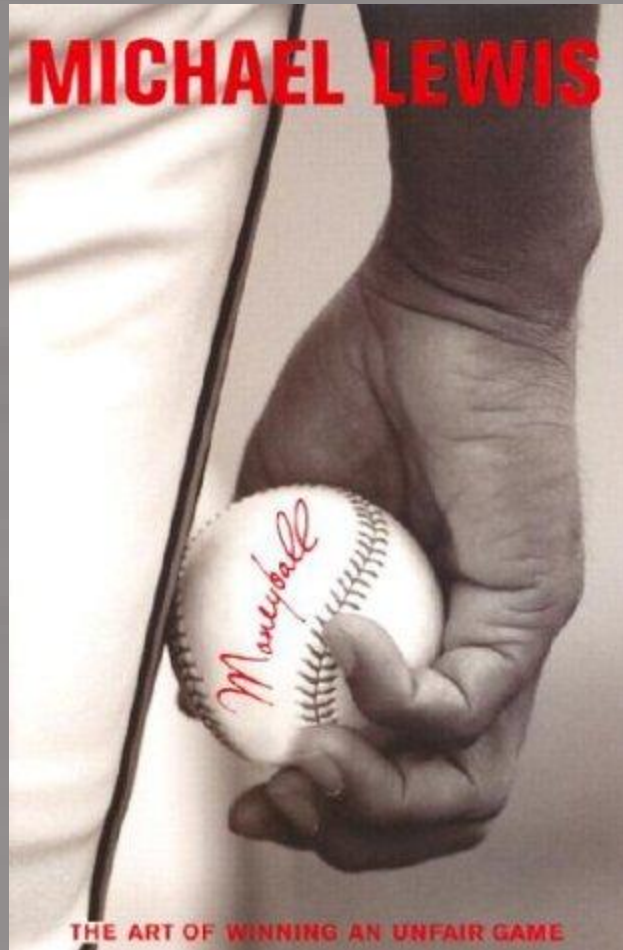
**The Call:** The Boston manager, Grady Little, decides to keep Martinez in the game. Little held fast to his instinct to stick with the best pitcher on the team.

**The Result:** The Yankees pound Martinez for three runs to tie the score and went on to win the game and the series.

Grady Little called it “his gut”. It cost him his job.

In God we trust; all others must have data.

# Moneyball and Sabermetrics



# Is it worth it?

## Moneyball's Billy Beane To Sign 5-Year Extension, Still Makes Less Than Brad Pitt

Cork Gaines | February 08, 2012 | 🔥 1,599 | 💬 1

In an interview with Jon Erlichman of *Bloomberg West*, Oakland A's owner Lew Wolff stated that general manager Billy Beane has agreed to a five-year extension that will keep him with the A's through the 2019 season. Wolff announced that he was also extending team president Mike Crowley.

You can see video of the interview at [Bloomberg.com](http://Bloomberg.com)

While details of the contract were not released, it is assumed that Beane will make much less than the reported [\\$10 million Brad Pitt received up front](#) to play Billy Beane in the movie "Moneyball." Brian Cashman, the general manager of the New York Yankees makes approximately \$3 million per season. And Theo Epstein makes \$3.5-4.0 million per season as team president of the Chicago Cubs.



(Wikimedia Commons)

Whether Beane actually stays with the A's through 2019 is another story. [According to Buster Olney of ESPN.com](#), Beane is free to pursue other jobs. That could include the Los Angeles Dodgers, who will soon have new owners with deep pockets.

# A Digression...

- **Six Sigma** is a business management strategy originally developed by Motorola, USA in 1986.
- Uses a five-phase improvement process: **DMAIC**
  - **Define** the process
  - Determine **Measures** [metrics] to quantify the process
  - Collect the data and **Analyze** it
  - Generate solutions that **Improve** the process
  - Use this methodology to **Control** the process



# The value of metrics

The present method of scoring the game and preparing scores for publication is faulty to the extreme, and it is calculated to drive players into playing for their records rather than for their sides.

Henry Chadwick 1880



# George Lindsey



# George Lindsey

□ 2 June 1920 – 6 September 2011

□ Canadian Army Officer

□ Order of Canada

□ PhD Nuclear Physics,  
Queen's University

□ President, Canadian  
Operations Research Society

## *Operations Research*

July–August 1963

### AN INVESTIGATION OF STRATEGIES IN BASEBALL

George R. Lindsey

*SACLANT ASW Research Center, La Spezia, Italy*

(Received November 20, 1962)

The advisability of a particular strategy must be judged not only in terms of the situation on the bases and the number of men out, but also with regard to the inning and the score. Two sets of data taken from a large number of major league games are used to give (1) the dependence of the probability of winning the game on the score and inning, and (2) the distribution of runs scored between the arrival of a new batter at the plate in each of twenty-four situations and the end of the half-inning. The first shows that the runs adding most to the probability of winning are the one tying the score and the one putting the batting team one ahead. All other runs are slightly less valuable early in the game, and become increasingly less valuable as the game proceeds. By combining the two sets of data, the situations are determined in which an intentional base on balls, a double play allowing a run to score, a sacrifice, and an attempted steal are advisable strategies, if average players are concerned. An index of batting effectiveness based on the contribution to run production in average situations is developed.

IN THE course of most baseball games a number of occasions arise in which the managers of the competing teams are called upon to select one out of two or more possible strategies.

For example, when the batting team has one or more runners on base, alternative strategies are

- (a) have the runner(s) try to steal a base, or
- (b) rely on a safe hit by the batter to advance the runner(s)

If there are less than two out, a third possible strategy would be

- (c) have the batter try to sacrifice, thus advancing the runner(s) at the expense of another out

For the fielding team, a choice always available, but most attractive with runners on second and third base but first base empty, and a dangerous batter at the plate, is

- (a) to issue an intentional base on balls to the batter, instead of
- (b) pitching to the batter

# How to Win a Baseball Game

□ The defensive team must:

- Create outs
- Prevent runs

□ The offensive team must:

- Prevent outs
- Create runs

# An Investigation of Strategies in Baseball

		Bases Occupied							
		0	1	2	3	1,2	1,3	2,3	1,2,3
Outs	0	0.46	0.81	1.19	1.39	1.47	1.94	1.96	2.22
	1	0.24	0.50	0.67	0.98	0.94	1.12	1.56	1.64
	2	0.10	0.22	0.30	0.36	0.40	0.53	0.69	0.82

□ Sacrifice Bunt

□ Stolen Base

□ Intentional Walk

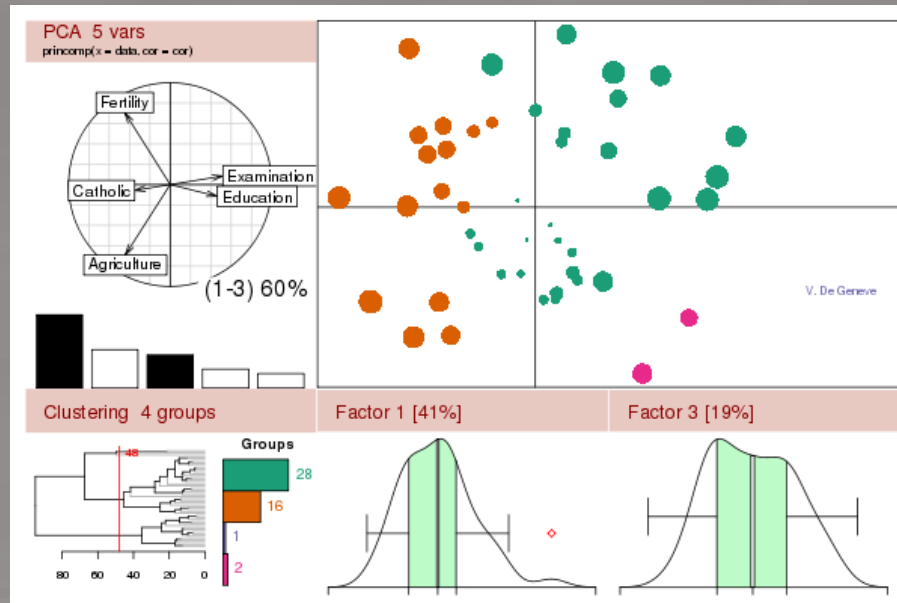
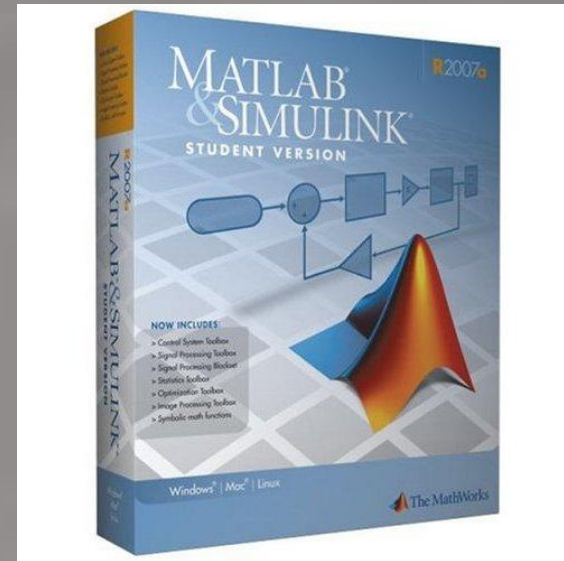
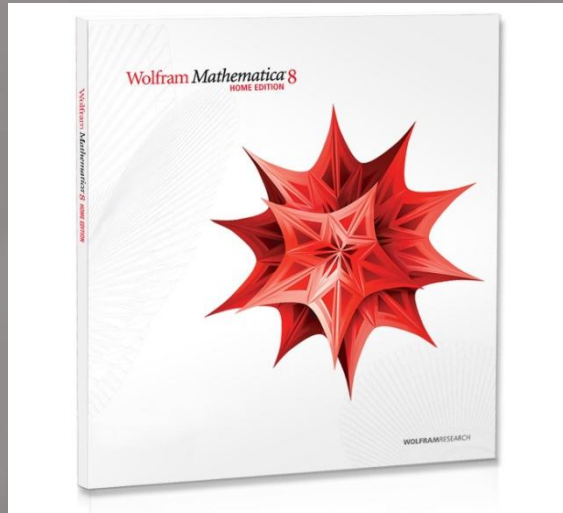
# Let's go to the data...

The logo for RETROSHEET, featuring the word "RETROSHEET" in a green, serif font, enclosed within a green outline of a baseball bat.

RETROSHEET



# Analytical Tools



# The importance of data and analysis

- They separate what we know we know from what we think we know
- Analysis can answer many questions
- It is important to ask the correct question



# Another example:

□ “..and he fouls off another pitch. Molina’s having a good at-bat here.”

□ Questions:

□ What is an average at-bat?

□ What is the typical number of pitches for an average at-bat?

□ A game has how many pitches?

□ Does any of this analysis say anything about the designated hitter rule?

# Computational Exercise

# Some Notable Individuals

□ Allan Roth

□ Earnshaw Cook

□ Bill James

□ Earl Weaver

□ Dick Cramer

□ John Thorn

□ Pete Palmer

□ Stephen Jay Gould

□ Sandy Alderson

□ Craig Wright

□ Paul DePodesta

□ Dave Smith

# Back to Moneyball

Sabermetrics and the stock market are becoming ever more intertwined. Morningstar fund analysts Have been instructed to read Bill James' old Writings to learn how to question accepted truth Through statistics, and to apply those lessons to Handicapping mutual funds.

The Numbers Game by Alan Schwartz

# Additional Applications of Analysis

[Find a Job](#) | [M&S Wine](#) | [Our Papers](#) | [Feedback](#)

Sunday, Feb 26 2012 12AM 5°C 3AM 5°C [5-Day Forecast](#)

# MailOnline

# Science & Tech

[Home](#) | [U.K. Home](#) | [News](#) | [Sport](#) | [U.S. Showbiz](#) | [Femail](#) | [Health](#) | **Science** | [Money](#) | [RightMinds](#) | [Travel](#) | [Columnists](#)

[Science Home](#) | [Pictures](#) | [Gadgets Gifts and Toys Store](#)

[Login](#)

## Hope for HIV vaccine thanks to mathematical formula that tracks the Stock Market

By [TED THORNHILL](#)

Last updated at 4:55 PM on 24th February 2012

[Comments \(0\)](#) | [Share](#)

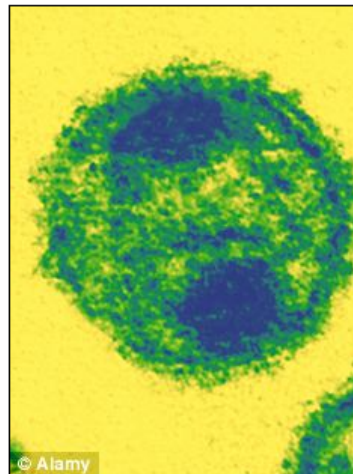
HIV has eluded vaccine-makers for 30 years in part due to the virus' extreme ability to mutate – but now scientists have identified a promising strategy for a vaccine using a mathematical formula that tracks stock market prices.

Researchers from the Massachusetts Institute of Technology (MIT) and the Ragon Institute in Cambridge, in the same state, think there's a way of trapping the virus with two choices – being destroyed by the immune system or destroying itself.

Vaccines prime the immune system to target molecular signatures associated with a particular pathogen.

But HIV's ability to mutate has made it difficult to identify reliable vaccine targets.

In their search for a new type of target, the team from the Ragon Institute targeted mutating amino acids. These are known to destroy HIV-infected cells.



**HIV: The disease has so far defeated scientists through its ability to mutate**

[Site](#) [Web](#) [bing](#)

**Headlines**

**Most Read**

- ▶ **Born to be mild! The electric motorbike that has all the creature comforts of a car**
- ▶ Mind where you're going: The skateboard controlled by brainwaves that moves wherever you think it should
- ▶ **Global warming 'could make us shorter' after horses are found to have shrunk the LAST time the world heated up**
- ▶ Girls beat boys at arithmetic... because they have better language skills
- ▶ **Hope for HIV vaccine thanks to mathematical formula that tracks the Stock Market**
- ▶ Range Rover goes topless! Iconic four-wheel drive brand gives motorists the chance to get the wind in their hair with new £50,000 model
- ▶ **Struggling to spk: Firms send new arrivals on courses to stop them using mobile shorthand in conversation**
- ▶ First contact... or the start of World War III? Amateur

# Discussion and Questions

John Aleshunas

[jalesh@webster.edu](mailto:jalesh@webster.edu)

<http://mercury.webster.edu/aleshunas>