

# 42.510 Statistical Learning for Data Science

## Tutorial 3

BASEBALL ANALYTICS

Fall 2025



SINGAPORE UNIVERSITY OF  
TECHNOLOGY AND DESIGN

# Baseball analytics

## Moneyball — *Michael Lewis*

- Oakland Athletics (The A's)
- *A Major League Baseball (MLB)* team from the American League (West).
- Won ALCS 1989 - 1991 (top the American League).
- Won the World Series in 1990.
- One of the biggest payrolls in 1991.
- 1995 - change of ownership - payroll slashed.
- The slashed budget becomes the norm for quite a few years.

# Moneyball

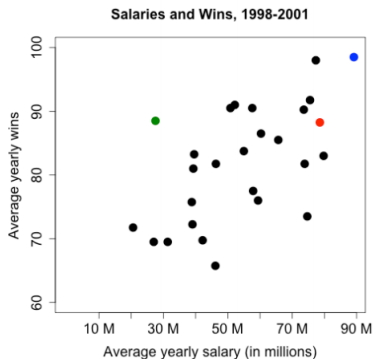
- Despite this the A's remain competitive in the MLB season games and perform very well.
- How?
- Use of analytics by General Manager Billy Beane.
- In 2002, the A's lose Jason Giambi (star player, MVP for 2000) and two other stars.
- Yet they make it to the playoffs, as well win 20 consecutive games (1st time in MLB history).

# Major League Baseball and playoffs

- Consists of two leagues, each consisting of 15 teams
  - American League: Boston Red Sox, NY Yankees , Oakland A's, ...
  - National League: SF Giants, St. L Cardinals, NY Mets, ...
- Every season each team plays 162 games within their own league.
- The top 5 teams (according to wins) in each league at the end of the season make it to the Play-offs.
- Post-season matches are knockout and leads to Leagues Division Series and World Series winners (which is not our concern now).
- The goal for the A's was to make it to the play-offs.

# Rules of Baseball

- <http://www.youtube.com/watch?v=sk0sApsF0jQ>



The average number of wins for each team in Major League Baseball (MLB) plotted as a function of their average yearly payroll (from 1998-2001). The green dot is the Oakland Athletics, the were winning a large number of games with one of the lowest payrolls. The red dot is the Boston Red Sox, and the blue dot is the New York Yankees, the richest teams in MLB.

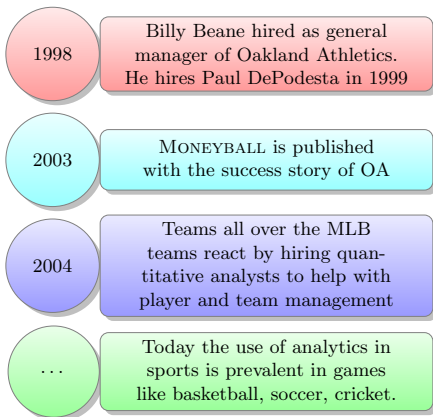
## Win percentages

Year	Won	Lost	Played	Win %
1998	74	88	162	45.6
1999	87	75	162	53.7
2000	91	71	162*	56.2
2001	102	60	162*	63.0
<b>2002</b>	<b>103</b>	<b>59</b>	<b>162*</b>	63.5
2003	96	66	162*	59.3
2004	91	71	162*	56.2

- Record for Billy Beane as a manager for the A's. \* indicates the years the Oakland Athletics made it to the playoffs
- 2002 is the year when three of their star players including Jason Giambi (MVP in 2000) left the A's.

# Paul DePodesta

Paul DePodesta, a Harvard Economics helped Billy Beane use quantitative tools to manage Oakland Athletics.



# Moneyball (2011)

- Medieval thinking: <http://www.youtube.com/watch?v=TpBcwG0v080>
- What is the problem? [http://www.youtube.com/watch?v=pWgyy\\_rlmag](http://www.youtube.com/watch?v=pWgyy_rlmag)
- Breaking biases: <http://www.youtube.com/watch?v=yGf6LNWY9AI>



# Baseball Statistics

- Batting Average (BA): how often a batter reaches base by hitting safely.

$$BA = \frac{\text{NUMBER OF HITS}}{\text{NUMBER OF AT BATS}}$$

- On Base Percentage (OBP): how often a batter reaches base. It is the fraction of plate appearances (includes at bats, walks, hit by pitch and sacrifice flies) in which the player reached base successfully through either a hit, walk or otherwise (hit by pitch).
- Slugging Percentage (SLG):

$$SLG = \frac{\text{TOTAL BASES}}{\text{NUMBER OF AT BATS}}$$

This counts doubles twice as much as singles, homeruns four times as much as singles and so on. The value is between 0 and 4 (hence not really a percentage!!)

- On base plus slugging (OPS):

$$OPS = OBP + SLG$$

- <http://www.baseball-reference.com>

# Oakland Athletics in 2002

- Lost star Jasom Giambi to New York Yankees.
- Lost two other stars stars Johnny Damon and Jason Isringhausen.
- Replaced them by:
  - an injured* Scott Hatteberg (with good OBP),
  - an unorthodox pitcher* Chad Bradford,
  - past his prime* David Justice,
  - all for much smaller prices.
- Won 103 games and made it to the playoffs.
- Won 20 consecutive games (an MLB record).