



Winning with R (and friends)

How data analysts affect the standings in sports championships



E-R

Agenzia sanitaria e
sociale regionale







E-R

Agenzia sanitaria e
sociale regionale



Baseball Prospectus

[R] ANNOUNCEMENT: Chapman & Hall/CRC: The R Series

Calver, Rob | Wed, 19 Jan 2011 01:44:03 -0800

We are pleased to announce the launch of a new series of books on R.

Chapman & Hall/CRC: The R Series

Aims and Scope

This book series reflects the recent rapid growth in the development and application of R, the programming language and software environment for statistical computing and graphics. R is now widely used in academic research, education, and industry. It is constantly growing, with new versions of the core software released regularly and more than 2,600 packages available. It is difficult for the documentation to keep pace with the expansion of the software, and this vital book series provides a forum for the publication of books covering many aspects of the development and application of R.

The scope of the series is wide, covering three main threads:

- Applications of R to specific disciplines such as biology, epidemiology, genetics, engineering, finance, and the social sciences.
- Using R for the study of topics of statistical methodology, such as linear and mixed modeling, time series, Bayesian methods, and missing data.
- The development of R, including programming, building packages, and graphics.

The books will appeal to programmers and developers of R software, as well as applied statisticians and data analysts in many fields. The books will feature detailed worked examples and R code fully integrated into the text, ensuring their usefulness to researchers, practitioners and students.

The Mail Archive



[🏠 The Mail Archive home](#)

[📄 r-help - all messages](#)

[📄 r-help - about the list](#)

[⏪ Previous message](#)

[⏩ Next message](#)





The R Series – Call for proposals



The R Series – Call for proposals

- Applications of R to specific disciplines such as biology, epidemiology, genetics, engineering, finance, and the social sciences.

The R Series – Call for proposals

- Applications of R to specific disciplines such as biology, epidemiology, genetics, engineering, finance, and the social sciences.
- Using R for the study of topics of statistical methodology, such as linear and mixed modeling, time series, Bayesian methods, and missing data.

The R Series – Call for proposals

- Applications of R to specific disciplines such as biology, epidemiology, genetics, engineering, finance, and the social sciences.
- Using R for the study of topics of statistical methodology, such as linear and mixed modeling, time series, Bayesian methods, and missing data.
- The development of R, including programming, building packages, and graphics.



The R Series – Call for proposals

- Applications of R to specific disciplines such as biology, epidemiology, genetics, engineering, finance, and the social sciences.
- Using R for the study of topics of statistical methodology, such as linear and mixed modeling, time series, Bayesian methods, and missing data.
- The development of R, including programming, building packages, and graphics.



The R Series – Upcoming books

- Applications of R to specific disciplines such as biology, epidemiology, genetics, engineering, finance, and the social sciences.



The R Series – Upcoming books

- Analyzing Baseball Data with R
(January 2014)



Is this cool?

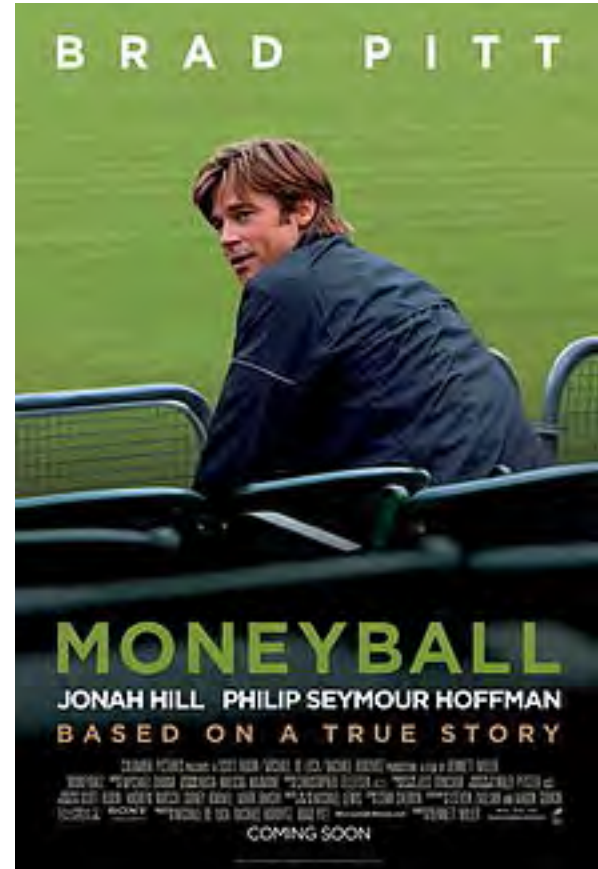
Is this cool?

- A best-seller book



Is this cool?

- A movie nominated for six Academy Awards
 - (starring Brad Pitt)



Is this cool?

- A Simpsons' Episode





Is this new?

Is this new?



Is this new?





Why prevalent in baseball?

Why prevalent in baseball?

- Stats available since the XIX century

Why prevalent in baseball?

- Stats available since the XIX century
- “Discrete” events

Why prevalent in baseball?

- Stats available since the XIX century
- “Discrete” events
- Long season = fewer sample issues



What is this about?

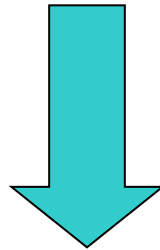


What is this about?

- Correctly evaluating players' contribution toward wins

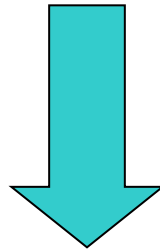
What is this about?

- Correctly evaluating players' contribution toward wins



What is this about?

- Correctly evaluating players' contribution toward wins



- Finding market inefficiencies



Success stories



Success stories

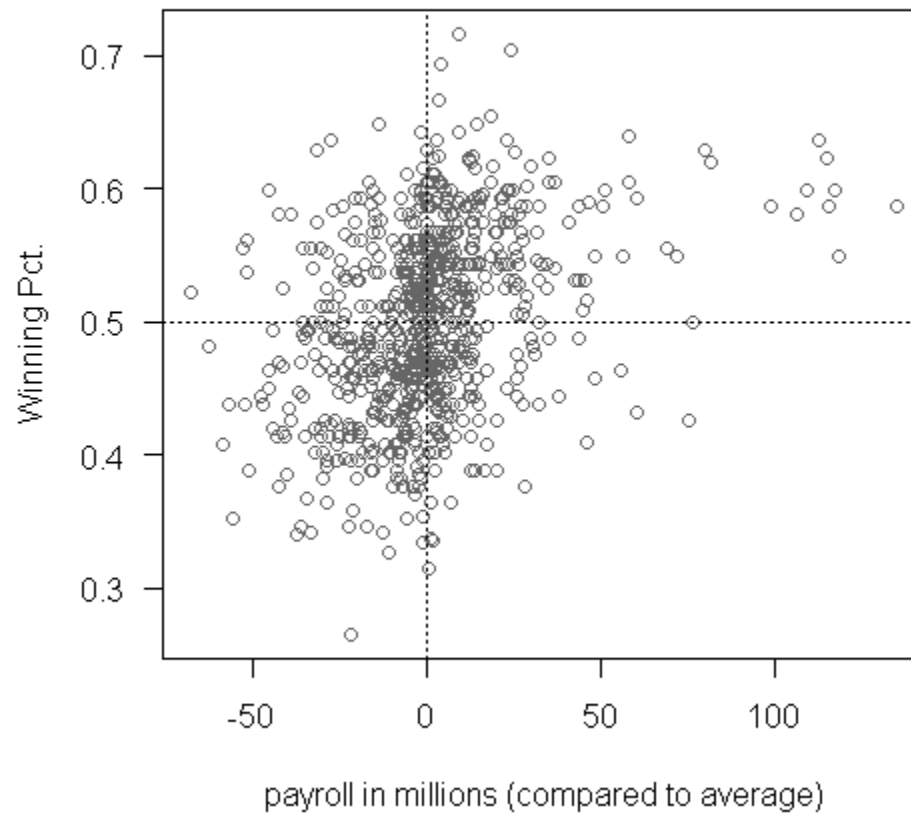
- 2000 - 2002 Oakland Athletics

Success stories

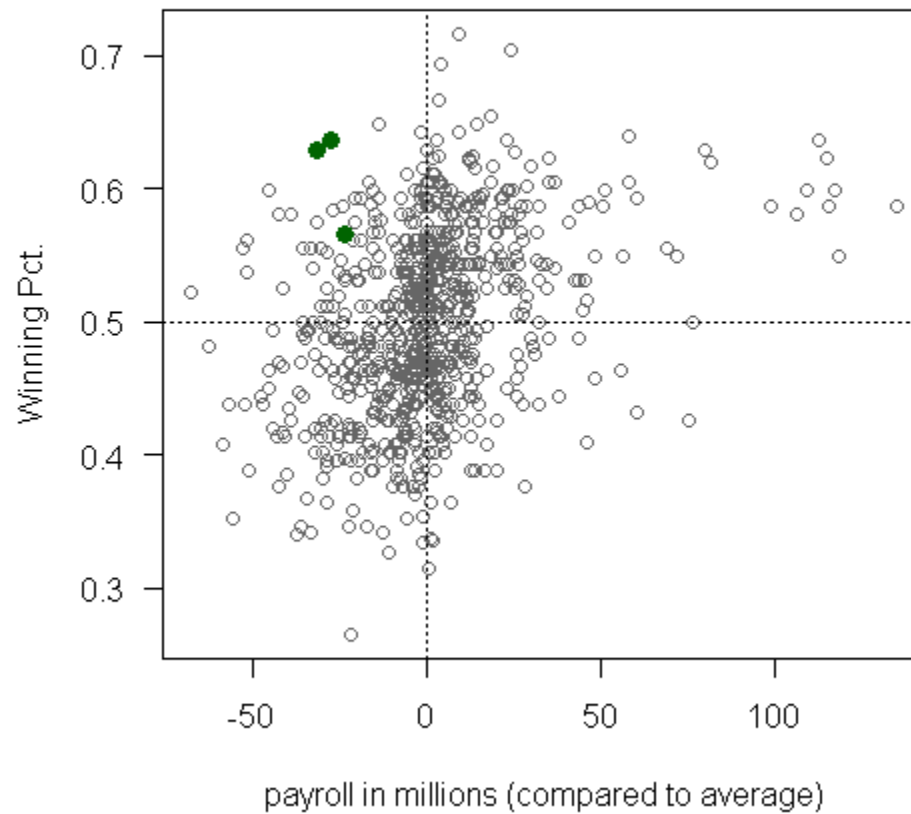
- 2000 - 2002 Oakland Athletics



Success stories



Success stories



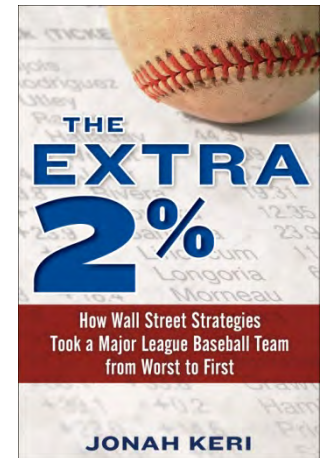


Success stories

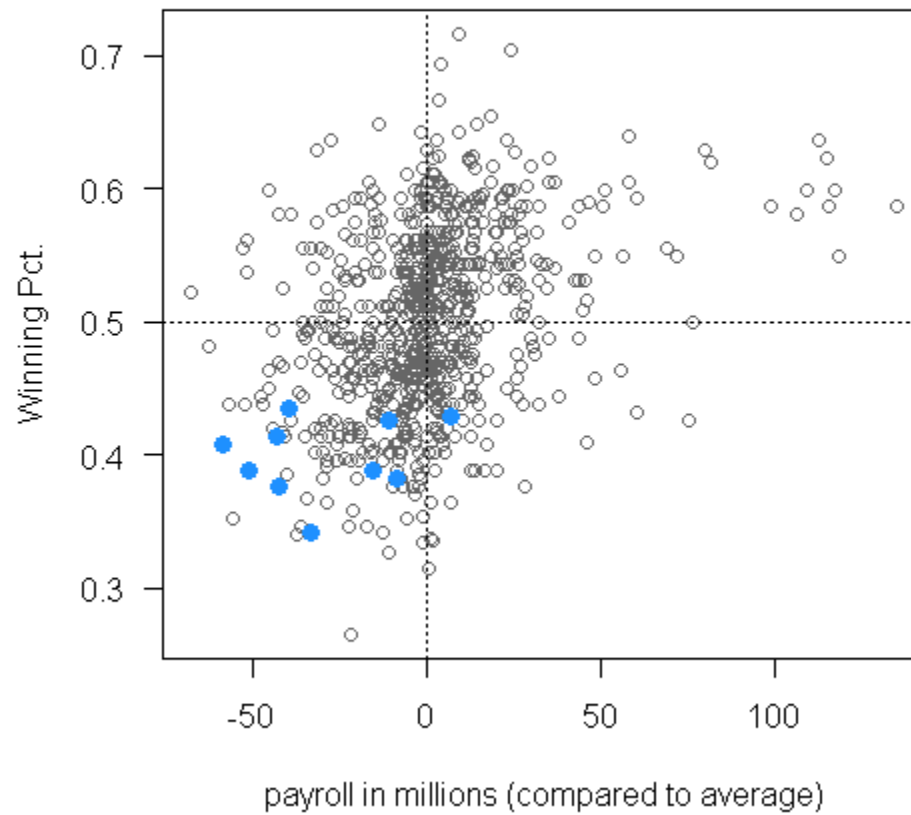
- 2008 - ... Tampa Bay Rays

Success stories

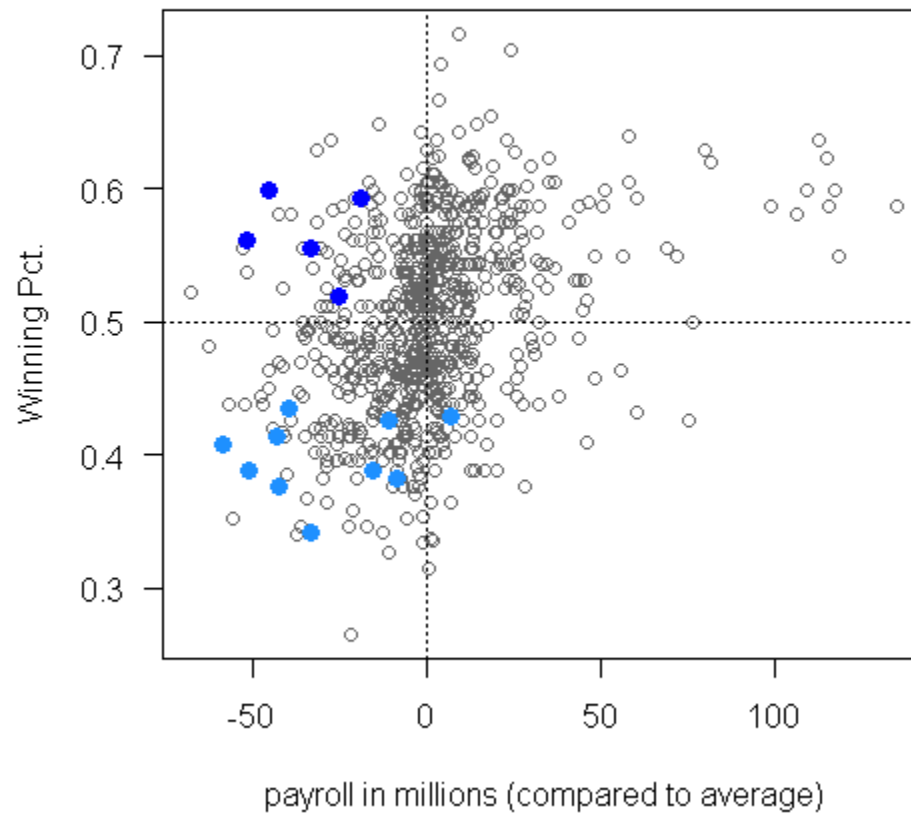
- 2008 - ... Tampa Bay Rays



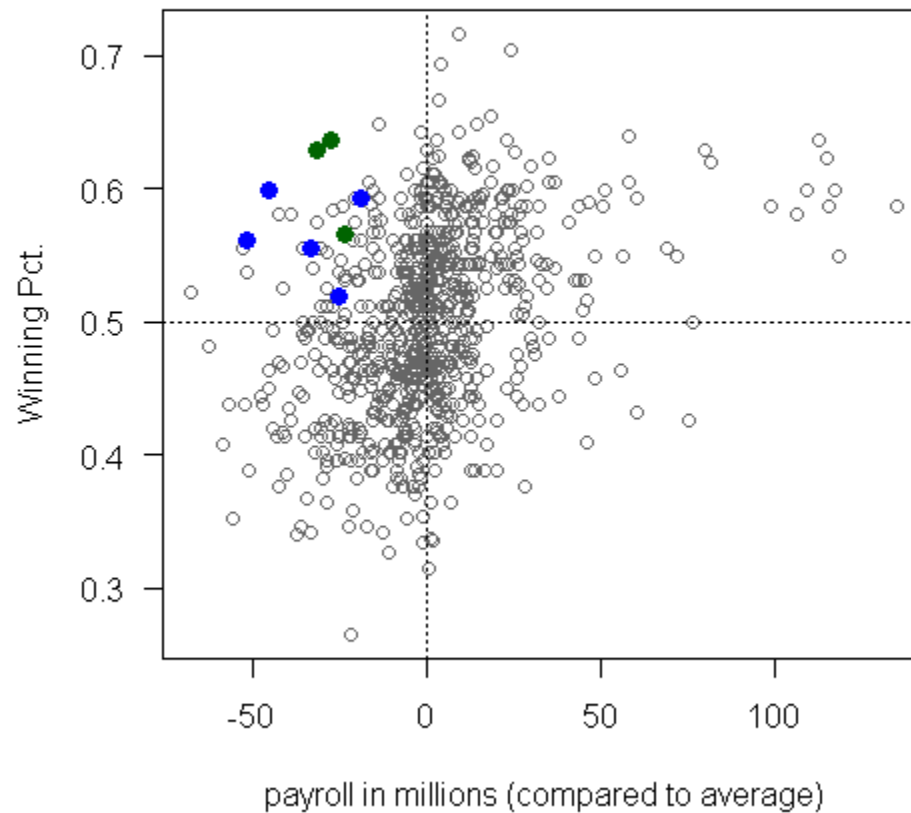
Success stories



Success stories



Success stories





Success stories

- 2000 - ... Dallas Mavericks

Success stories

- 2000 - ... Dallas Mavericks
 - Before: Winning Percentage = 40%

Success stories

- 2000 - ... Dallas Mavericks
 - Before: Winning Percentage = 40%
 - After: Winning Percentage = 69%



Stuff done with R

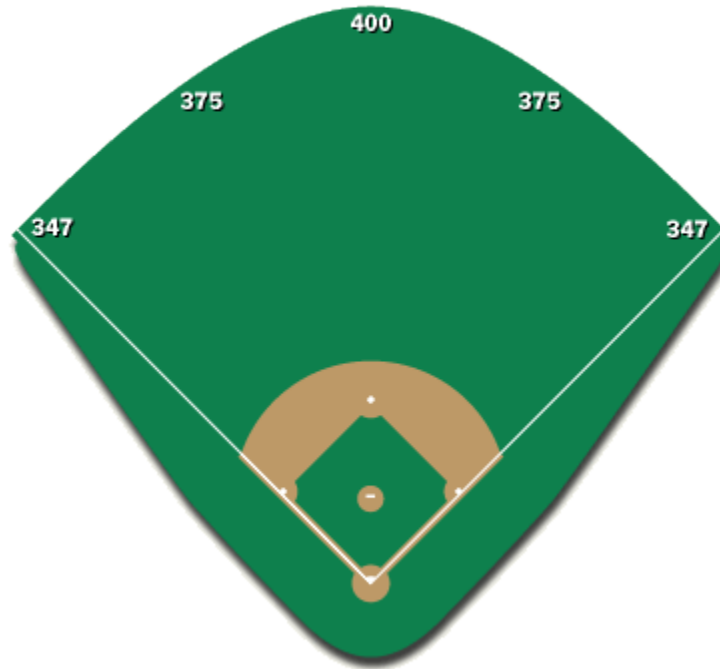


Stuff done with R

- Baseball defensive alignment

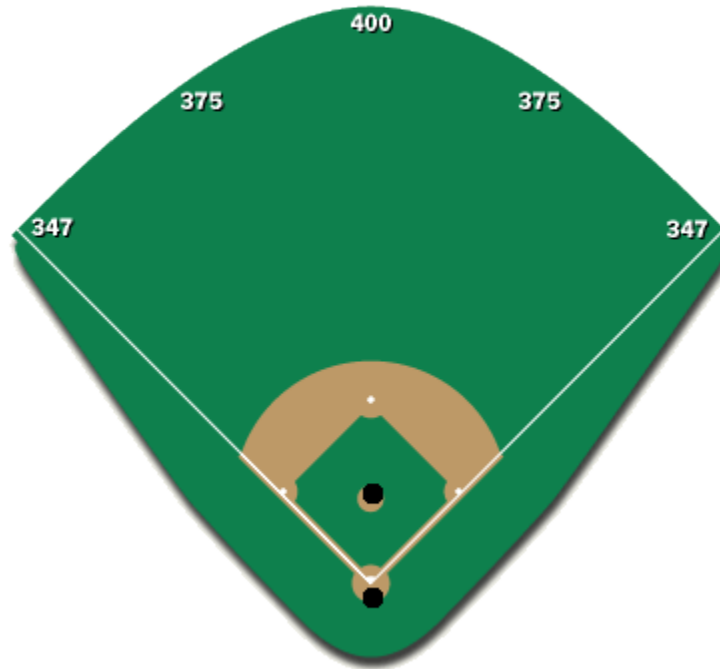
Stuff done with R

- Baseball defensive alignment



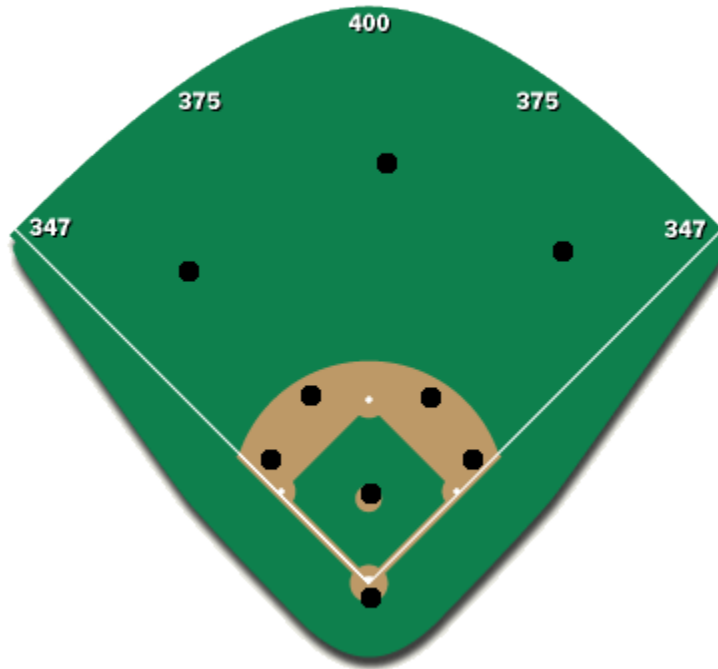
Stuff done with R

- Baseball defensive alignment



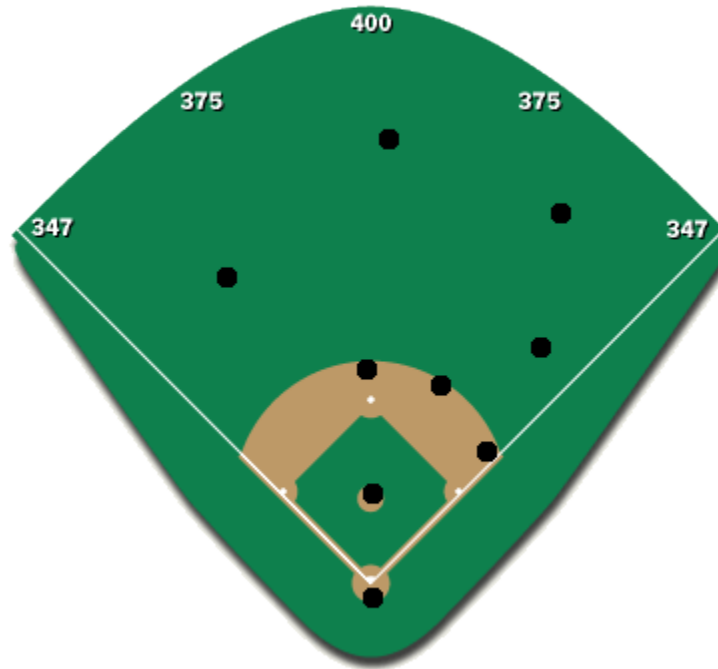
Stuff done with R

- Baseball defensive alignment



Stuff done with R

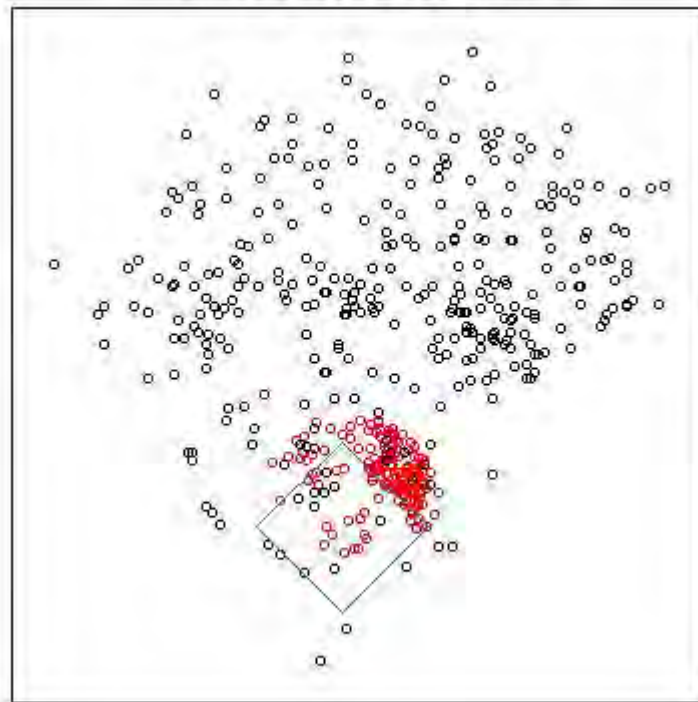
- Baseball defensive alignment



Stuff done with R

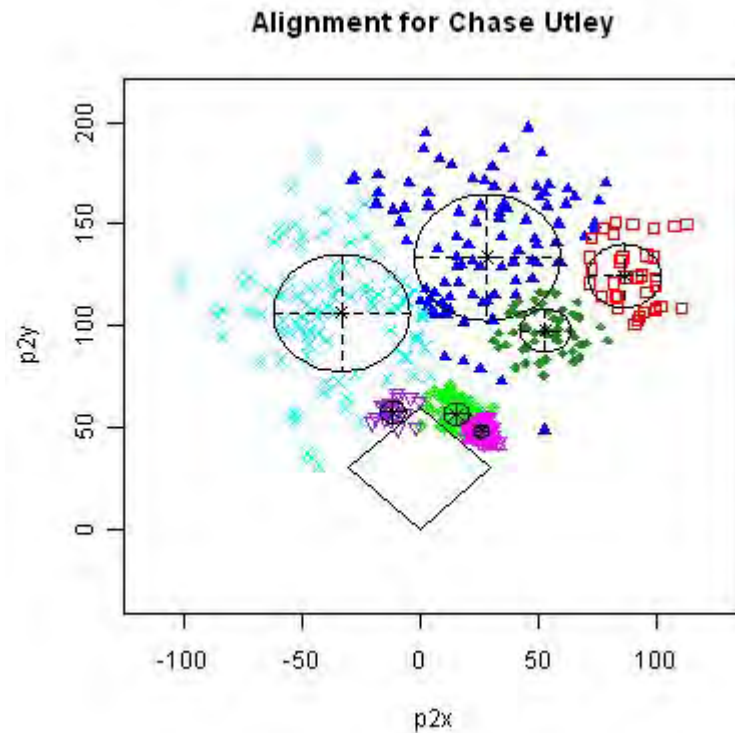
- Baseball defensive alignment

Chase Utley's spray chart



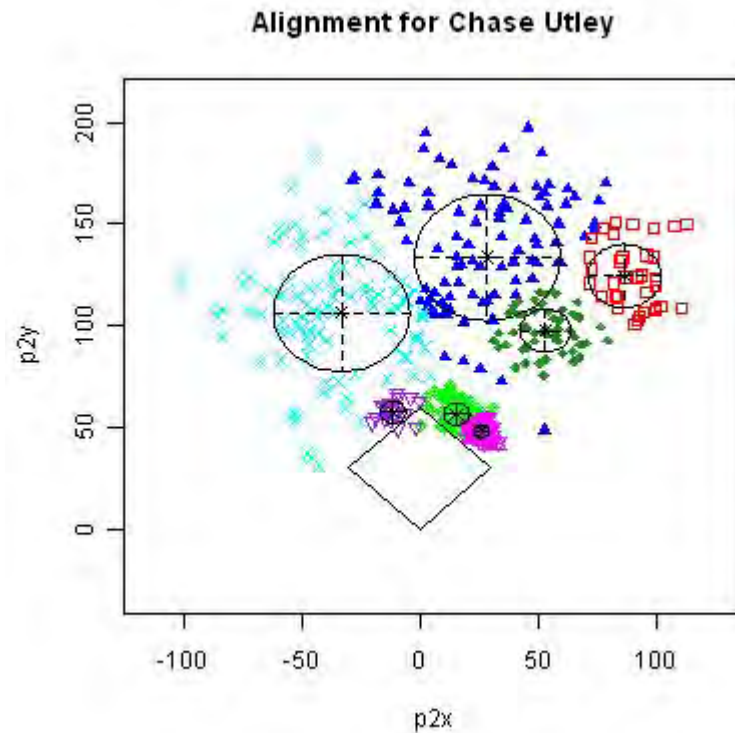
Stuff done with R

- Baseball defensive alignment



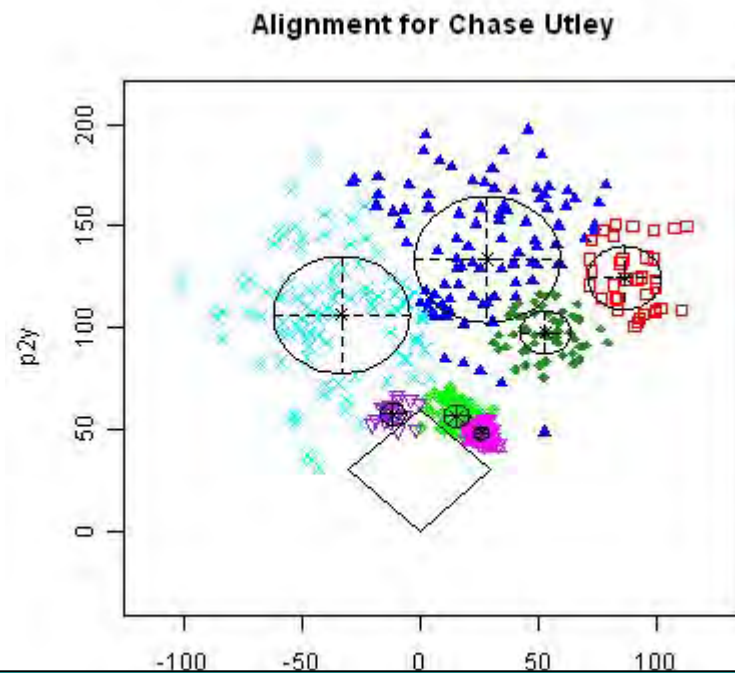
Stuff done with R

- Baseball defensive alignment



Stuff done with R

- Baseball defensive alignment

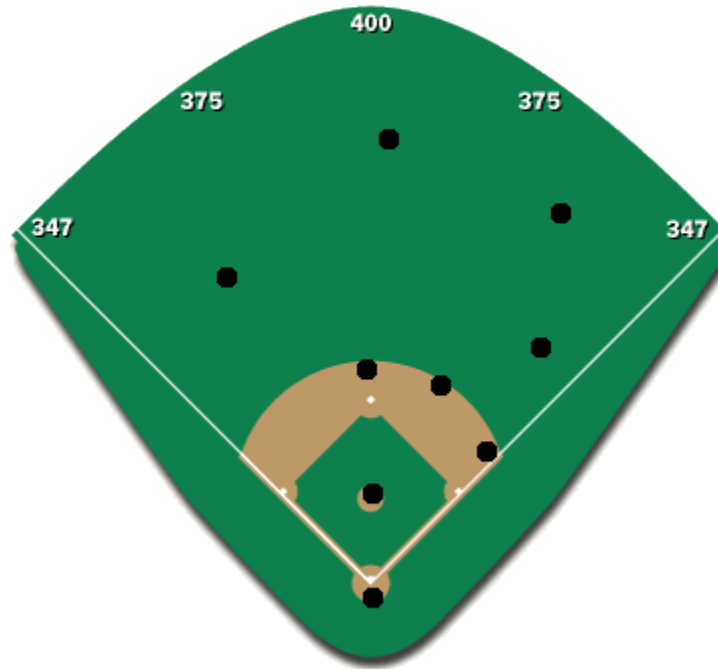


Mclust

Source: Prof Pepper's Assistant Blog

Stuff done with R

- Baseball defensive alignment



Stuff done with R

- Football field goal decisions

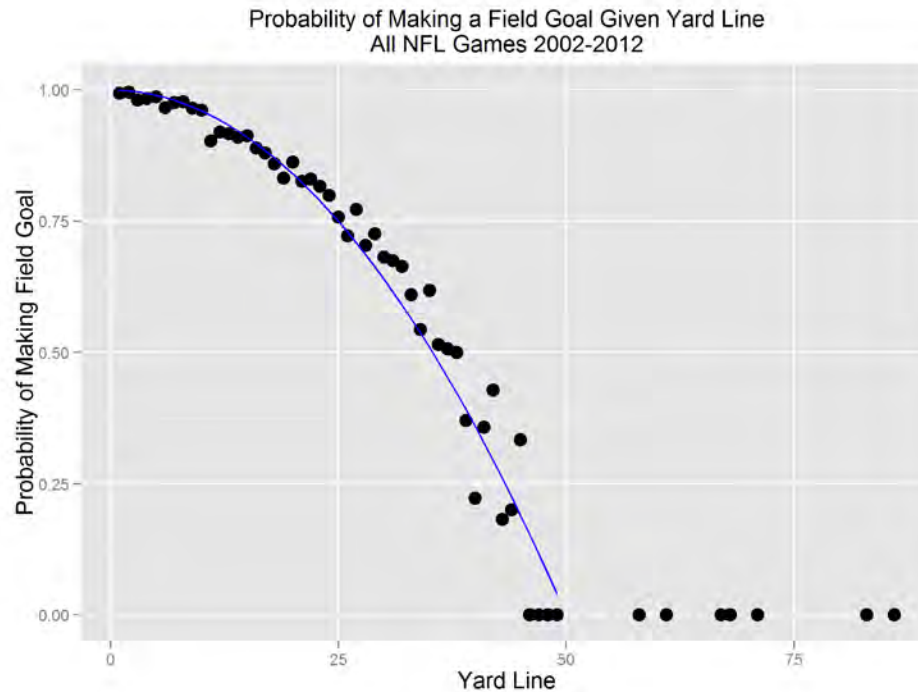
Stuff done with R

- Football field goal decisions



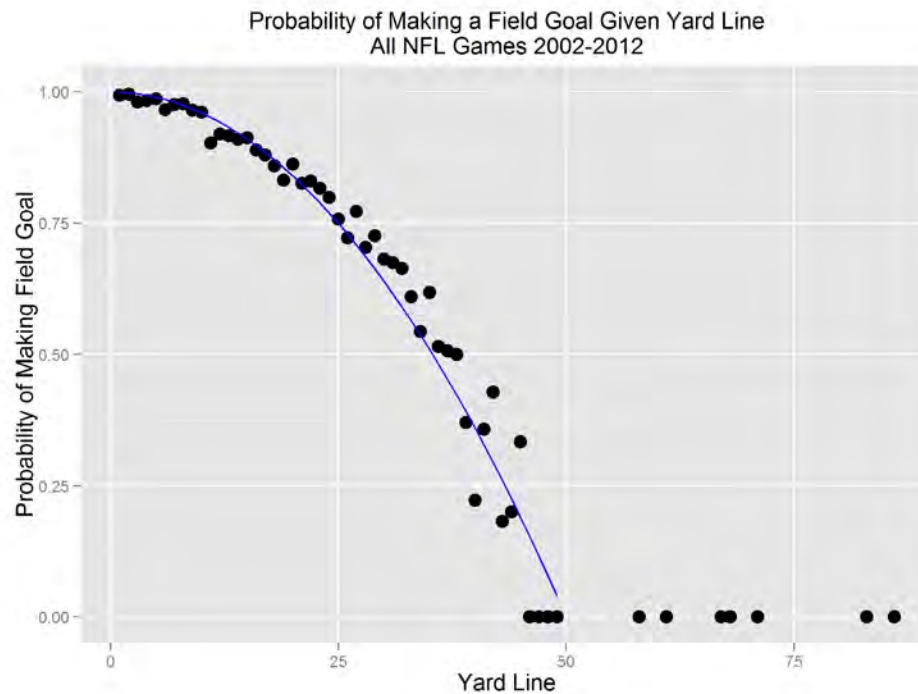
Stuff done with R

- Football field goal decisions



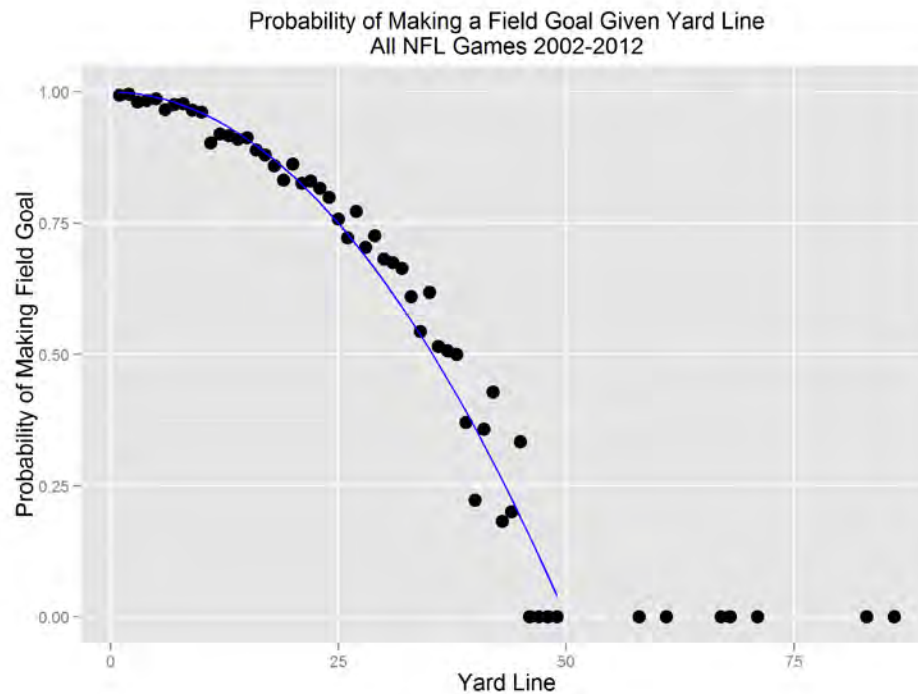
Stuff done with R

- Football field goal decisions



Stuff done with R

- Football field goal decisions



ggplot2

Source: Decision Science News



Stuff done with R

- Football field goal (icing the kicker)

Stuff done with R

- Football field goal (icing the kicker)
 - +5mph of wind → -3% success rate
 - +5°F → +1% success rate
 - Playing at home → no effect
 - (controlling for the kicker)

Stuff done with R

- Football field goal (icing the kicker)
 - +5mph of wind → -3% success rate
 - +5°F → +1% success rate
 - Playing at home → no effect
 - (controlling for the kicker)

 - “icing the kicker” → no significant effect

Stuff done with R

- Football field goal (icing the kicker)
 - +5mph of wind → -3% success rate
 - +5°F → +1% success rate
 - Playing at home → no effect
 - (controlling for the kicker)
 - “icing the kicker” → no significant effect



lme4

Stuff done with R

- Football field goal (icing the kicker)
 - +5mph of wind → -3% success rate
 - +5°F → +1% success rate
 - Playing at home → no effect
 - (controlling for the kicker)
 - “icing the kicker” → no significant effect



lme4

Source: Prof Pepper's Assistant Blog

Stuff done with R

- Basketball individual defense

Stuff done with R

HOME PAGE TODAY'S PAPER VIDEO MOST POPULAR TIMES TOPICS

Subscribe: Digital / Home Delivery Log In Register Now

The New York Times Magazine

Search All NYTimes.com

WORLD U.S. N.Y. / REGION BUSINESS TECHNOLOGY SCIENCE HEALTH SPORTS OPINION ARTS STYLE TRAVEL JOBS REAL ESTATE AUTOS

THE TIMES MAGAZINE T MAGAZINE KEY PLAY

Advertise on NYTimes.com

The No-Stats All-Star



Robert Seale for The New York Times

Statistical Anomaly His greatness is not marked in box scores or at slam-dunk contests, but on the court Shane Battier makes his team better, often much better, and his opponents worse, often much worse.

By MICHAEL LEVMS
Published: February 13, 2009

Out of Duke University. . . . A 6-foot-8-inch forward. . . .

He had more or less admitted to me that this part of his job left him cold. 'It's the same thing

Max Marchi

Related Letters: The No-Stats All-Star (March 1, 2009)

TWITTER LINKEDIN SIGN IN TO E-MAIL PRINT SHARE

More Articles in Magazine »

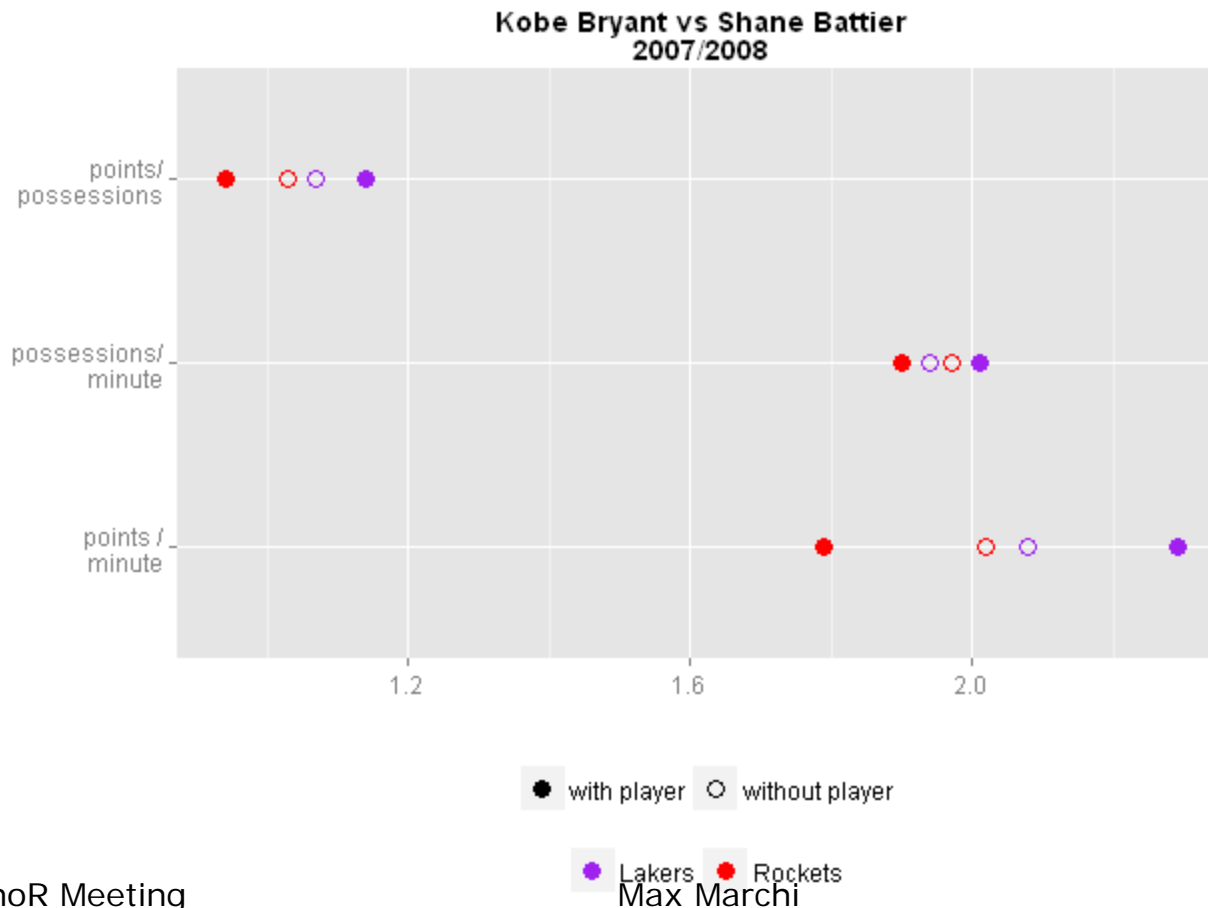
MOST POPULAR

E-MAILED BLOGGED SEARCHED VIEWED

1. A.D.H.D. Seen in 11% of U.S. Children as Diagnoses Rise
2. Opinion: State-Wrecked: The Corruption of Capitalism in America
3. The Stone: On Being Catholic
4. Op-Ed Contributor: The Tar Sands Disaster
5. PAUL KRUGMAN: Lessons From a Comeback
6. THOMAS L. FRIEDMAN: Need a Job? Invent It
7. Nature's Drone, Pretty and Deadly
8. Is Giving the Secret to Getting Ahead?

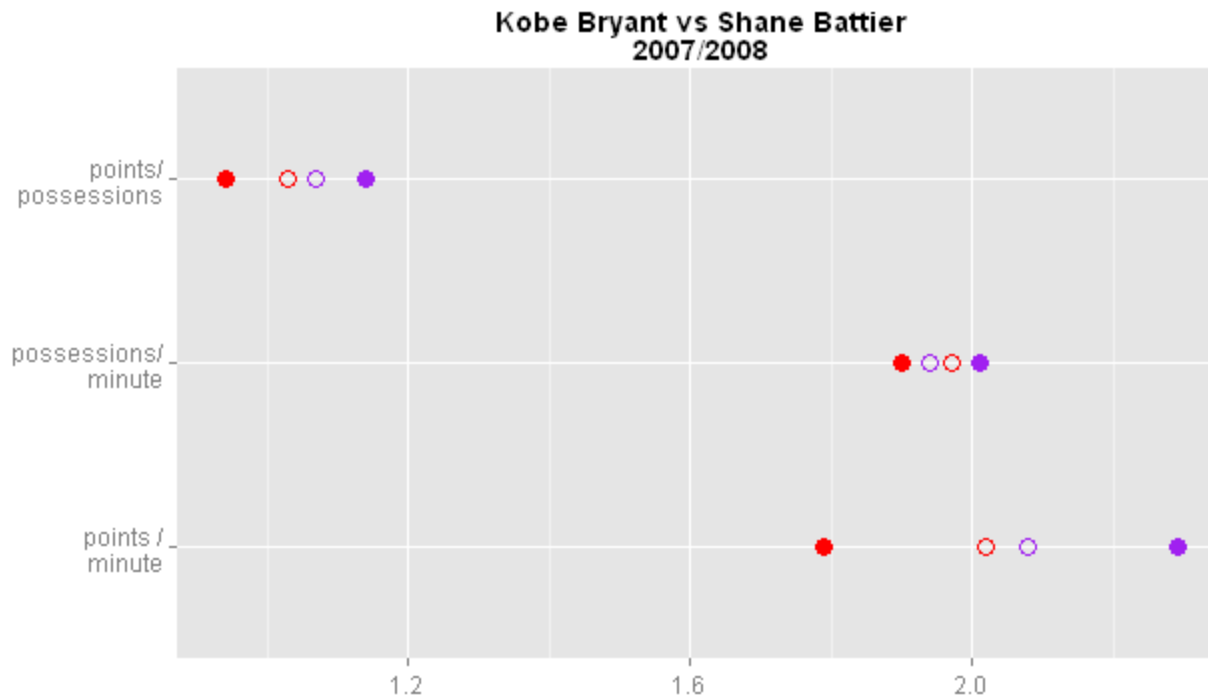
Stuff done with R

○ Basketball individual defense



Stuff done with R

○ Basketball individual defense



Adaptation from: Prof Pepper's Assistant Blog

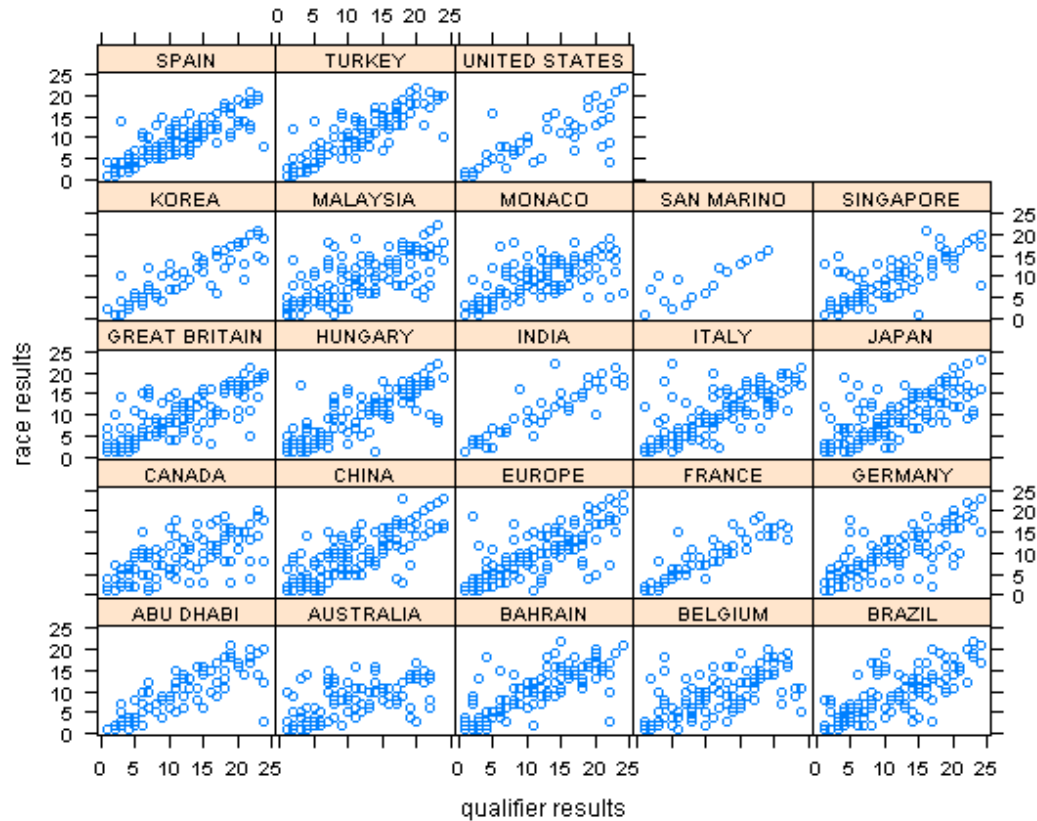


Stuff done with R

- F.1: qualifier vs race results

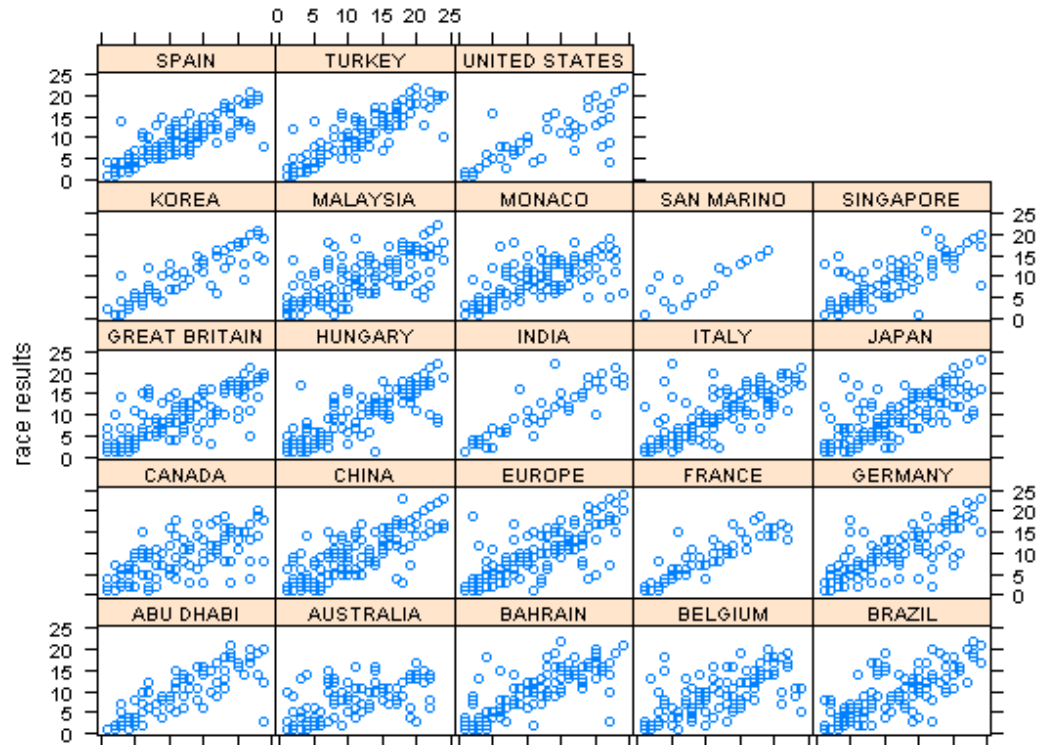
Stuff done with R

- F.1: qualifier vs race results



Stuff done with R

- F.1: qualifier vs race results



Adaptation from: OUseful.Info, the blog...

Shiny Apps

- <http://glimmer.rstudio.com/pssguy/leagueTables/>
- <http://www.rodrigozamith.com/2013/03/20/visualizing-season-performance-by-ncaa-tournament-teams/>
- <http://cpsievert.github.com/pitchRx/demo/>



Job postings requiring R knowledge

...that I have heard of...

Job postings requiring R knowledge

...that I have heard of...





R friends



R friends

- MySQL (or other DBMS)



R friends

- MySQL (or other DBMS)
- Python (or other programming languages)

R friends

- MySQL (or other DBMS)
- Python (or other programming languages)
 - parsing/scraping... regular expressions... XML/JSON



Conflict of interests disclosure

Analyzing Baseball Data with R



Conflict of interests disclosure

Analyzing Baseball Data with R
by

Jim Albert

Conflict of interests disclosure

Analyzing Baseball Data with R

by

Max Marchi

and

Jim Albert

References

- Moneyball – by Michael Lewis
- Moneyball (movie) – directed by Bennett Miller
- Moneybart – The Simpsons episode #467 (3rd of season 22)
- The Extra 2% - by Jonah Keri

- <http://profpeppersassistant.blogspot.it/2009/02/refining-shift.html>
- <http://www.decisionsciencenews.com/2013/01/28/football-geeks-your-10705-field-goals-are-ready/>
- <http://profpeppersassistant.blogspot.it/2013/01/icing-kicker.html>
- http://www.nytimes.com/2009/02/15/magazine/15Battier-t.html?pagewanted=all&_r=0
- <http://profpeppersassistant.blogspot.it/2009/03/moneybasketball.html>
- <http://blog.ouseful.info/2013/01/30/f1stats-visually-comparing-qualifying-and-grid-positions-with-race-classification/>



Grazie

max.marchi@gmail.com