



Cutting Edge Predictive Analytics

Spark Summit
June 2017

Eric Siegel, Ph.D.
Founder, Predictive Analytics World
Author, *Predictive Analytics*

For more information and citations regarding the examples in this presentation, see the book "Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die" by Eric Siegel (<http://www.thepredictionbook.com>). Most of the various examples shown are covered in the book - so, for greater detail about each case study named, see its reference/citation - search by organization name within the book's Notes PDF, available online at <http://www.PredictiveNotes.com>.



**Delivering on the
promise of data science**

San Francisco | Chicago | New York | Washington D.C. | London | Berlin

2017 Events:

PAW Chicago—June
PAW New York—Oct
PAW Gov, Berlin, London...

*“Stimulating talks and
prolific networking...
PAW met the highest
expectations.”*

- Marianna Dizik,
Google



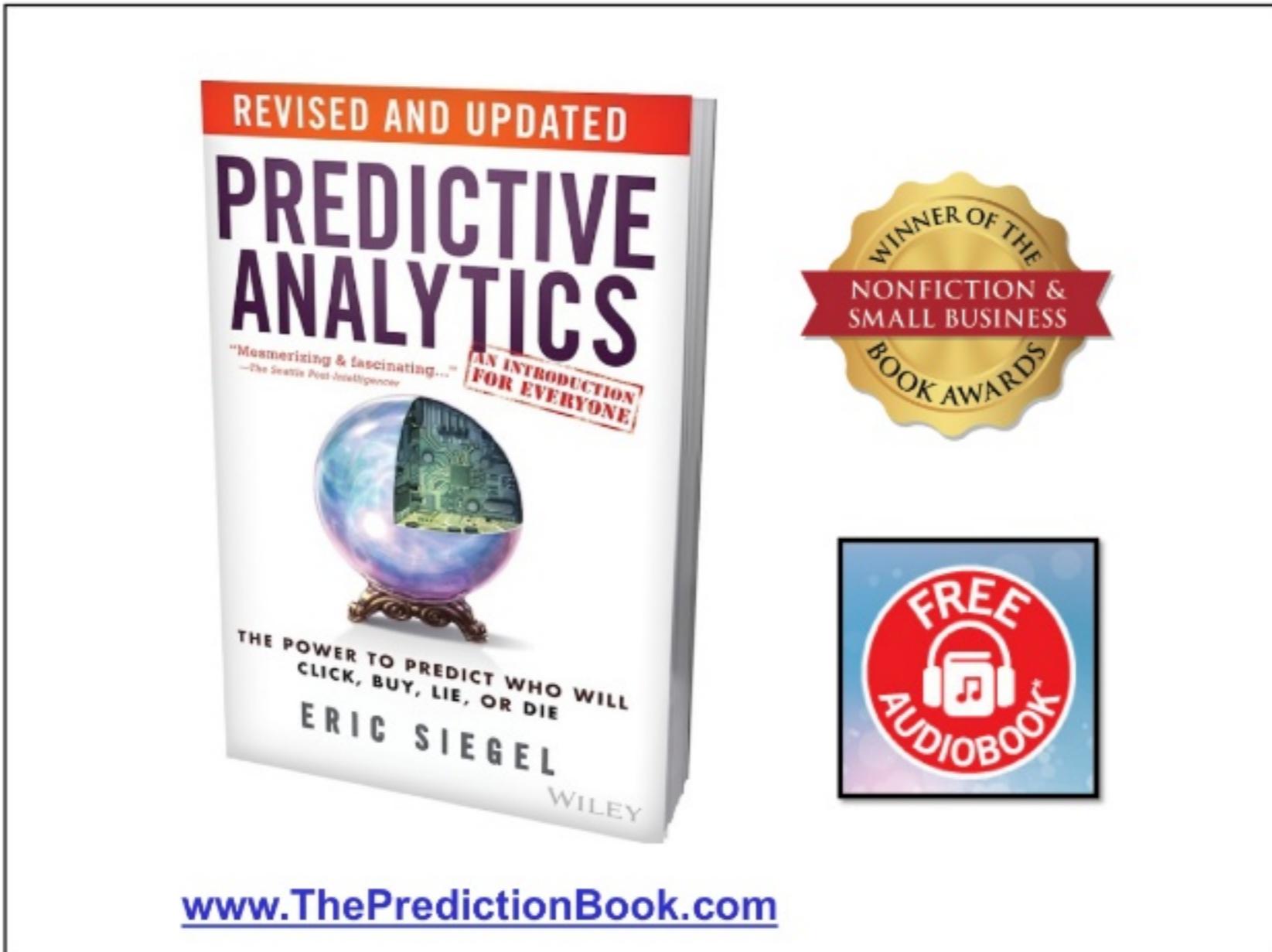
www.pawcon.com

With 10 events a year globally, Predictive Analytics World delivers vendor-neutral sessions across verticals such as banking, financial services, e-commerce, entertainment, government, healthcare, manufacturing, high technology, insurance, non-profits, publishing, and retail.

Predictive Analytics World industry events include PAW Business, PAW Financial, PAW Government, PAW Healthcare, PAW Workforce, and PAW Manufacturing.

Why bring together such a wide range of endeavors? No matter how you use predictive analytics, the story is the same: Predictively scoring customers, employees, students, voters, patients, equipment, and other organizational elements optimizes performance. Predictive analytics initiatives across industries leverage the same core predictive modeling technology, share similar project overhead and data requirements, and face common process challenges and analytical hurdles.

Info: <http://www.predictiveanalyticsworld.com>



Attendees may receive a copy of the audiobook version at no charge by simply emailing me directly to request: eric@predictionimpact.com



Predictive analytics educational/infotainment geek rap video:

<http://www.PredictThis.org>

Agenda

- 0) What is predictive analytics
- 1) Uplift modeling — *persuasion*
- 2) Sound science — *avoid BS prediction*

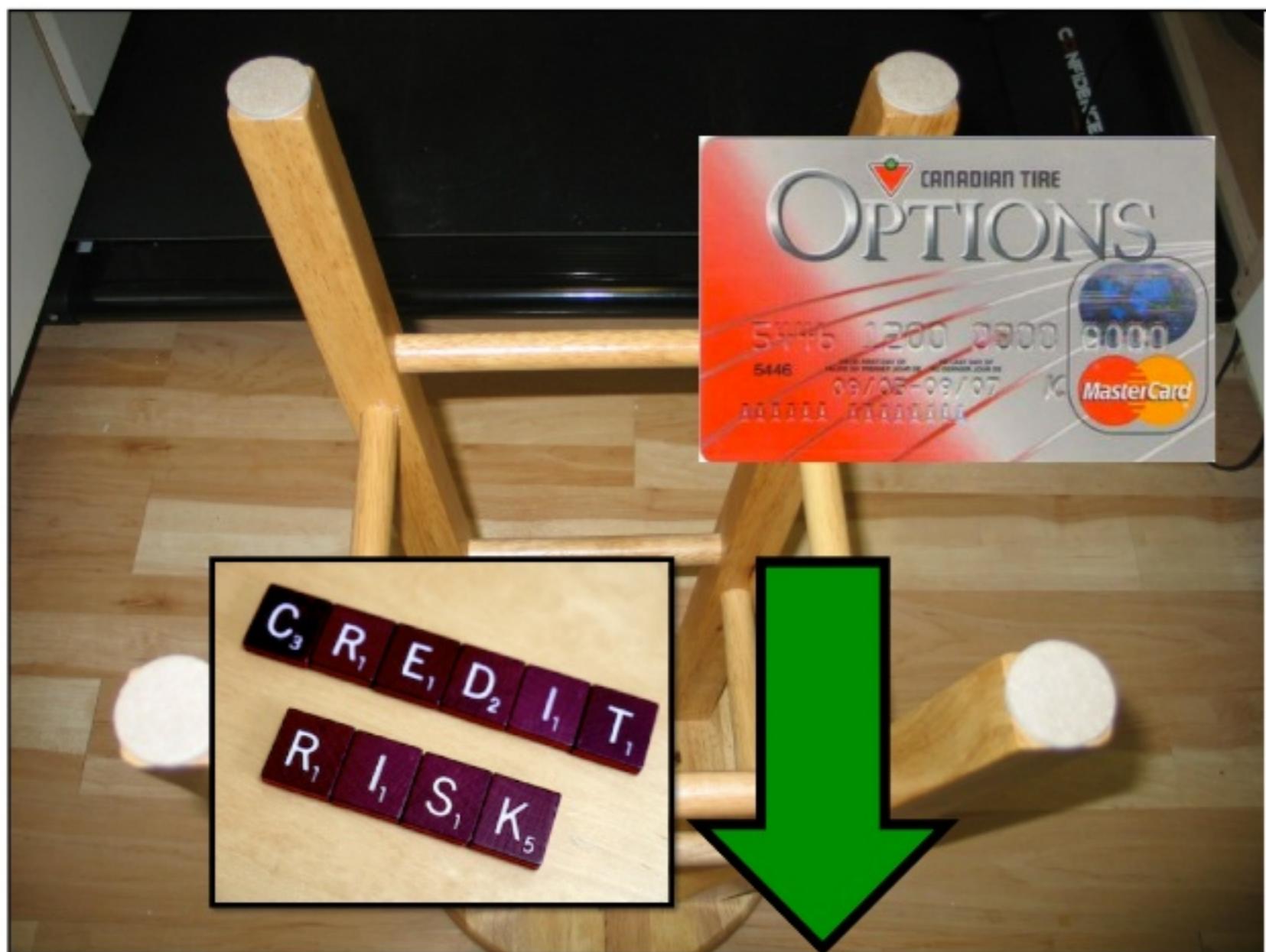
Predictive Analytics World



0) What is predictive analytics



Canadian Tire examples, from "What Does Your Credit-Card Company Know About You?" New York Times, May 12, 2009. <http://www.nytimes.com/2009/05/17/magazine/17credit-t.html>



Canadian Tire examples, from "What Does Your Credit-Card Company Know About You?" New York Times, May 12, 2009. <http://www.nytimes.com/2009/05/17/magazine/17credit-t.html>



People who “like” curly fries on Facebook are “more intelligent.”

Reference for most examples in this presentation are in the Notes PDF for Eric Siegel's book, "Predictive Analytics." For each example's reference/citation, search by organization name within the book's Notes PDF, available at www.PredictiveNotes.com



Neighborhoods of San Francisco that exhibit higher rates of certain crimes also exhibit higher demand for Uber rides.

The Data Effect:

Data is always predictive.

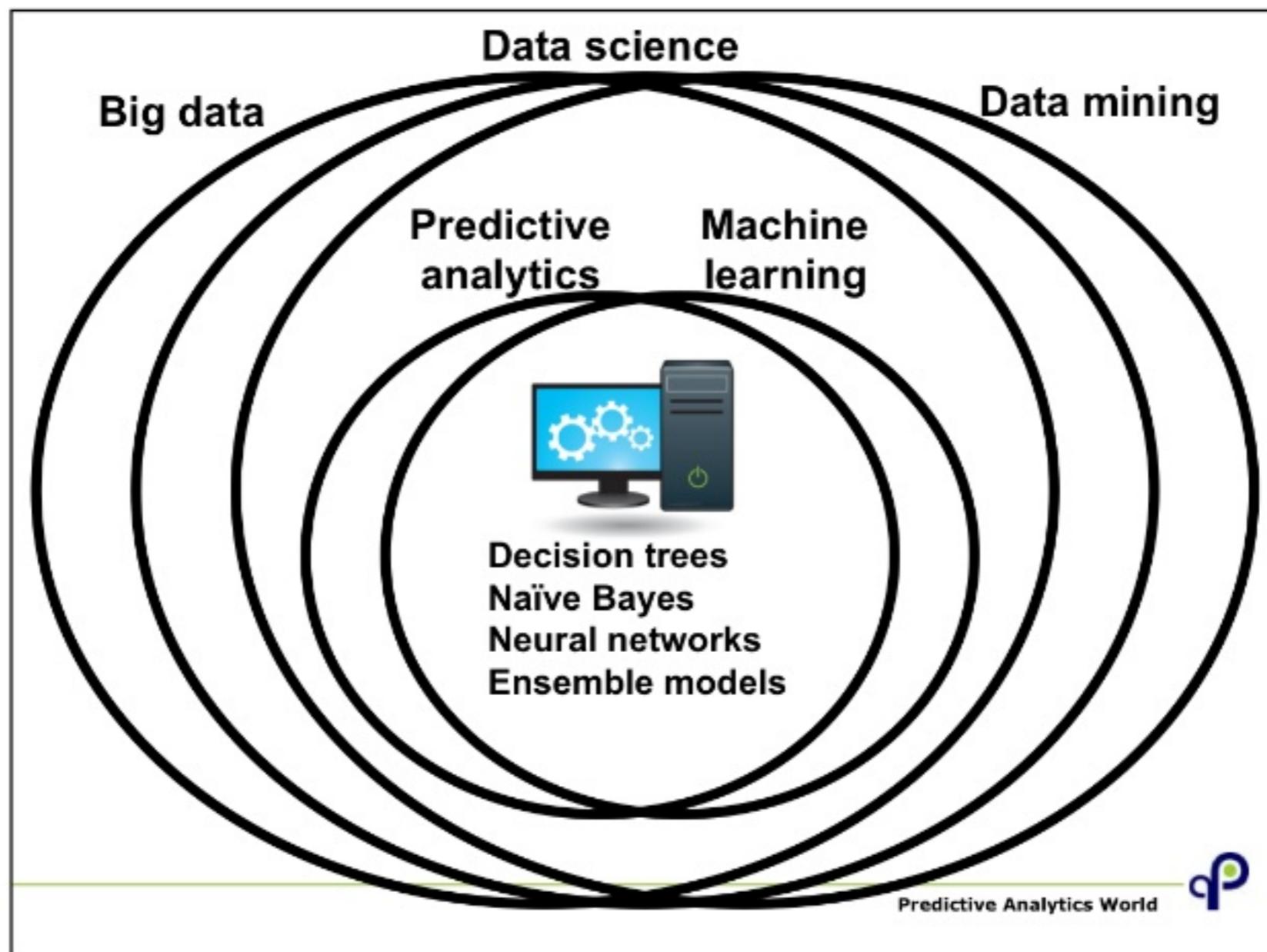
Predictive Analytics:

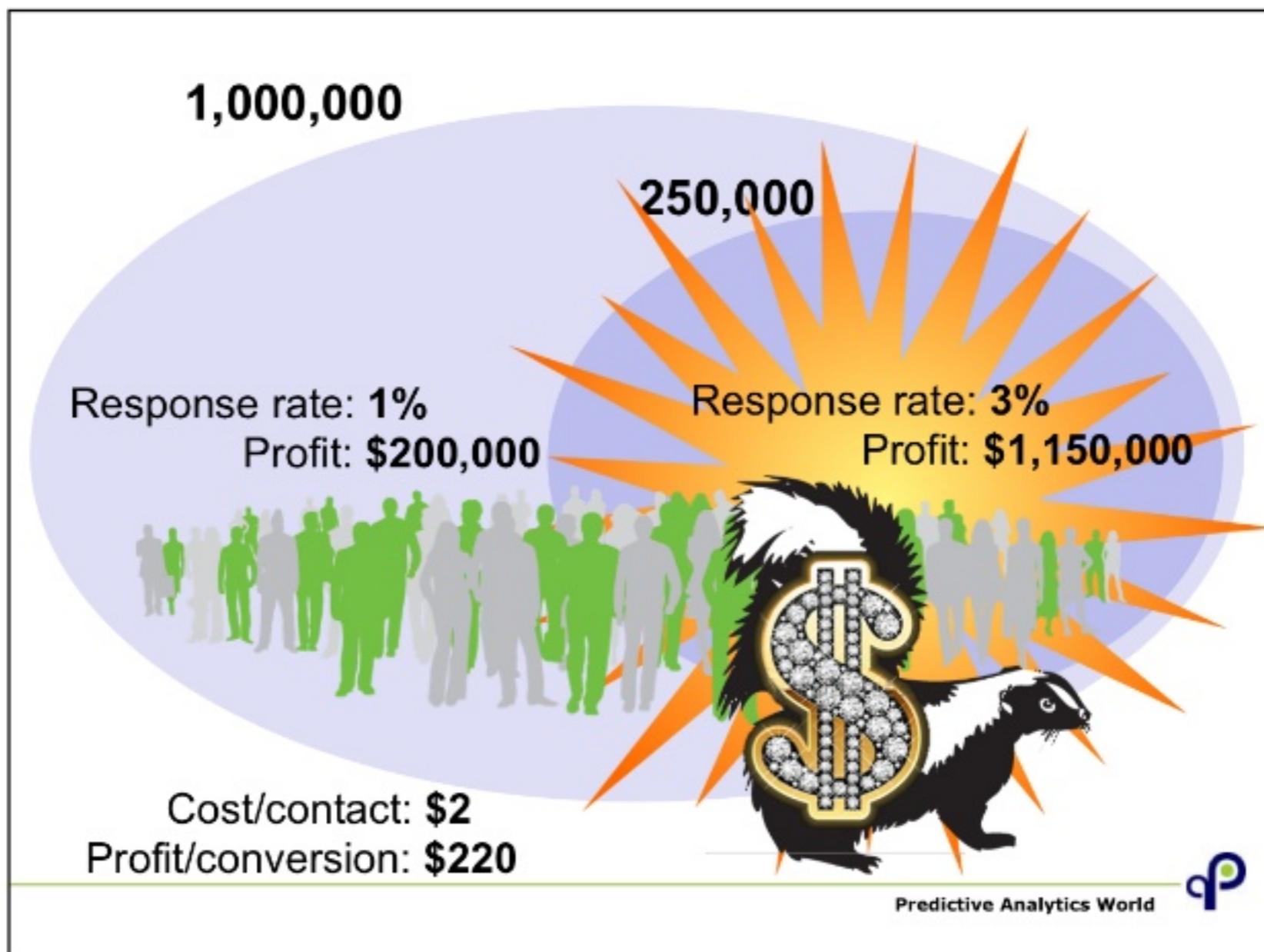


(data)

*Technology that learns from **experience**,
to predict the future behavior of individuals
... **in order to drive better decisions.***







A crummy predictive model delivers big value. It's like a skunk with bling.

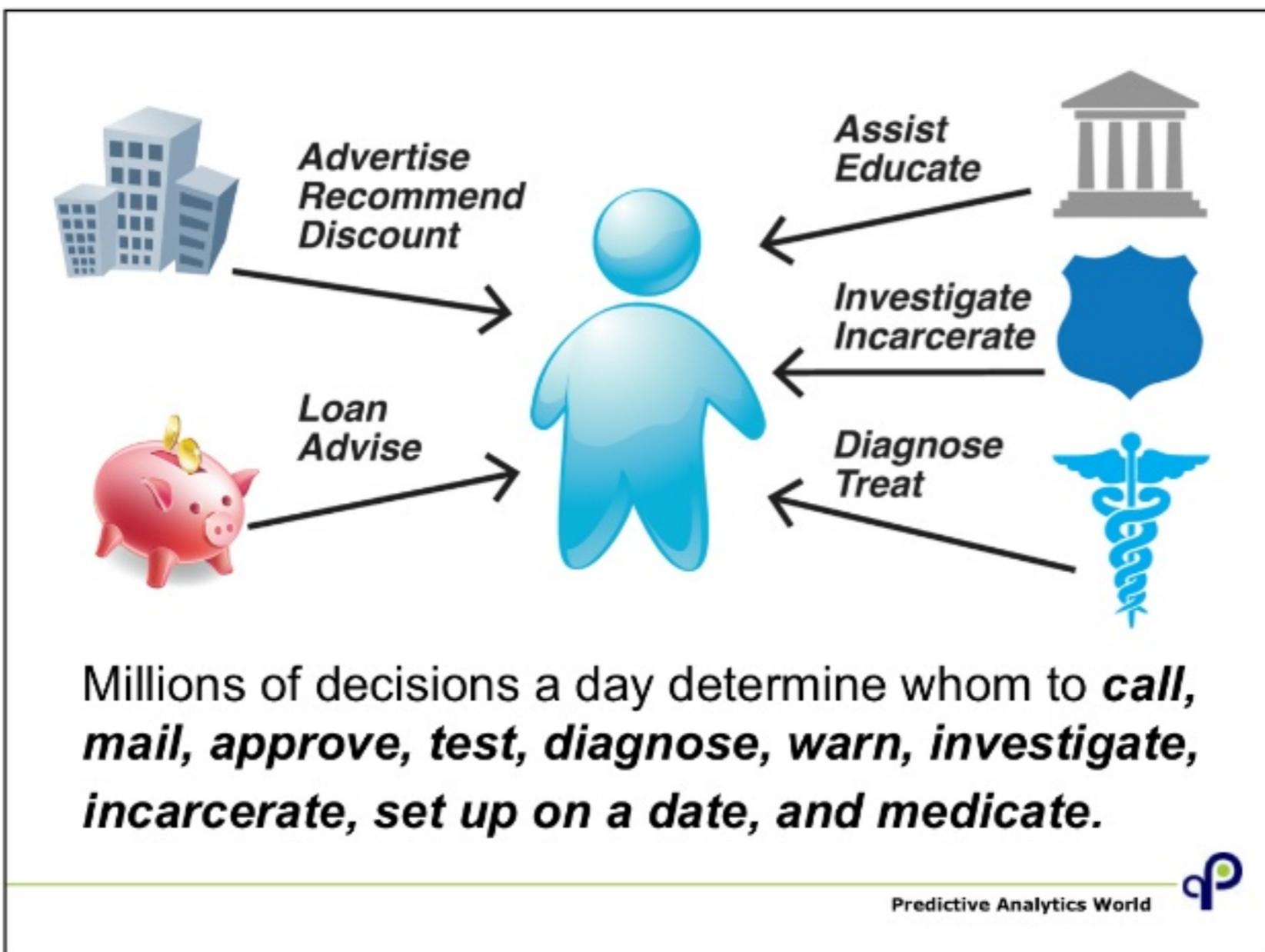
Simple arithmetic shows the bottom line profit of direct mail, both in general and then improved by predictively targeting (and only contacting 25% of the list). The less simple part is how the predictive scores are generated for each individual in order to determine exactly who belongs in that 25%. For details on how this works, see Chapter 1 of the book "Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die" (<http://www.thepredictionbook.com>).

The Prediction Effect:

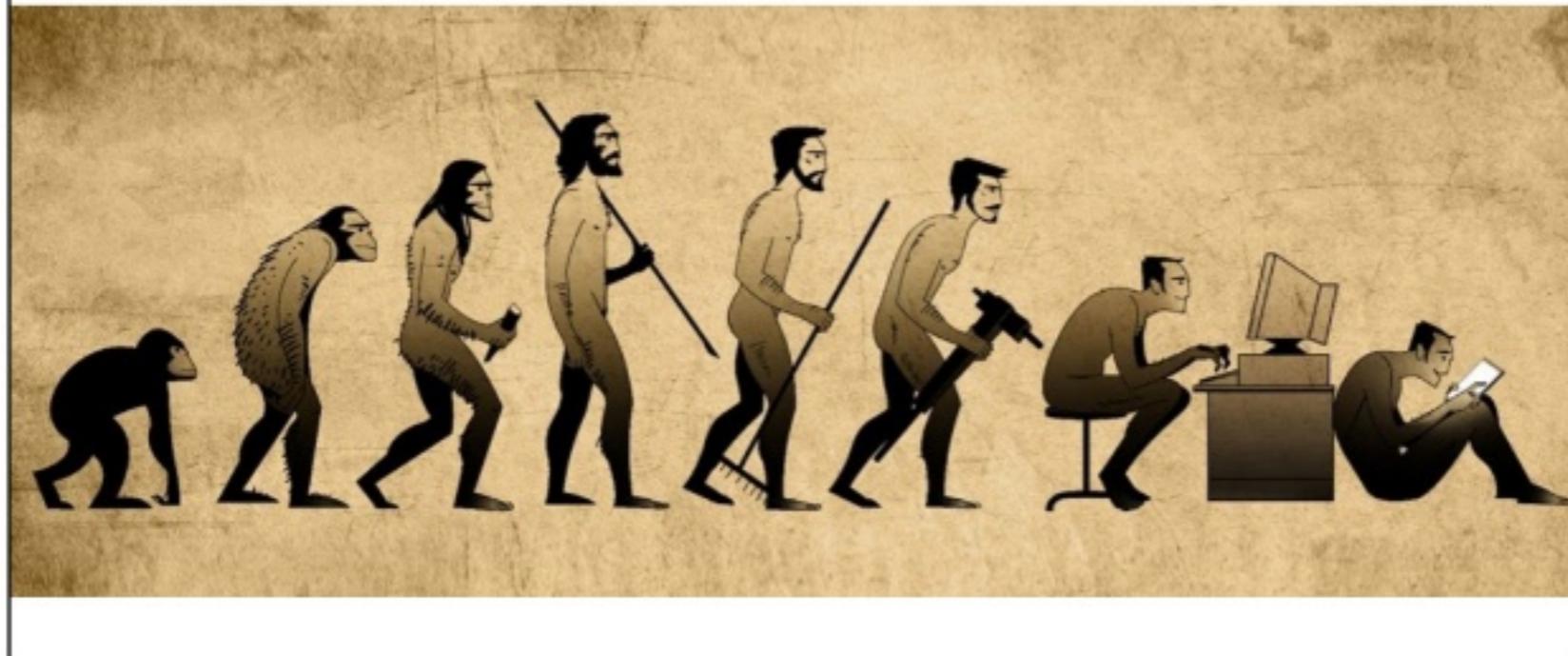
A little prediction goes a long way.

Predictive Analytics World 

Put another way, predicting better than guessing is often sufficient to generate great value by rendering operations more efficient and effective. For details on how this works, see the Introduction and Chapter 1 of the book "Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die" (<http://www.thepredictionbook.com>).



Predictive analytics is the Information Age's latest evolutionary step.



This singular, universally applicable force improves every large-scale thing we do—how we build things, sell things, and prevent bad things from happening—because every function benefits from prediction. As its deployment takes hold across industries, we have moved beyond engineering the infrastructure that manages big data to implementing the science that taps its contents for more intelligent

Has emerged as a commonplace business practice necessary to sustain competitive advantage.

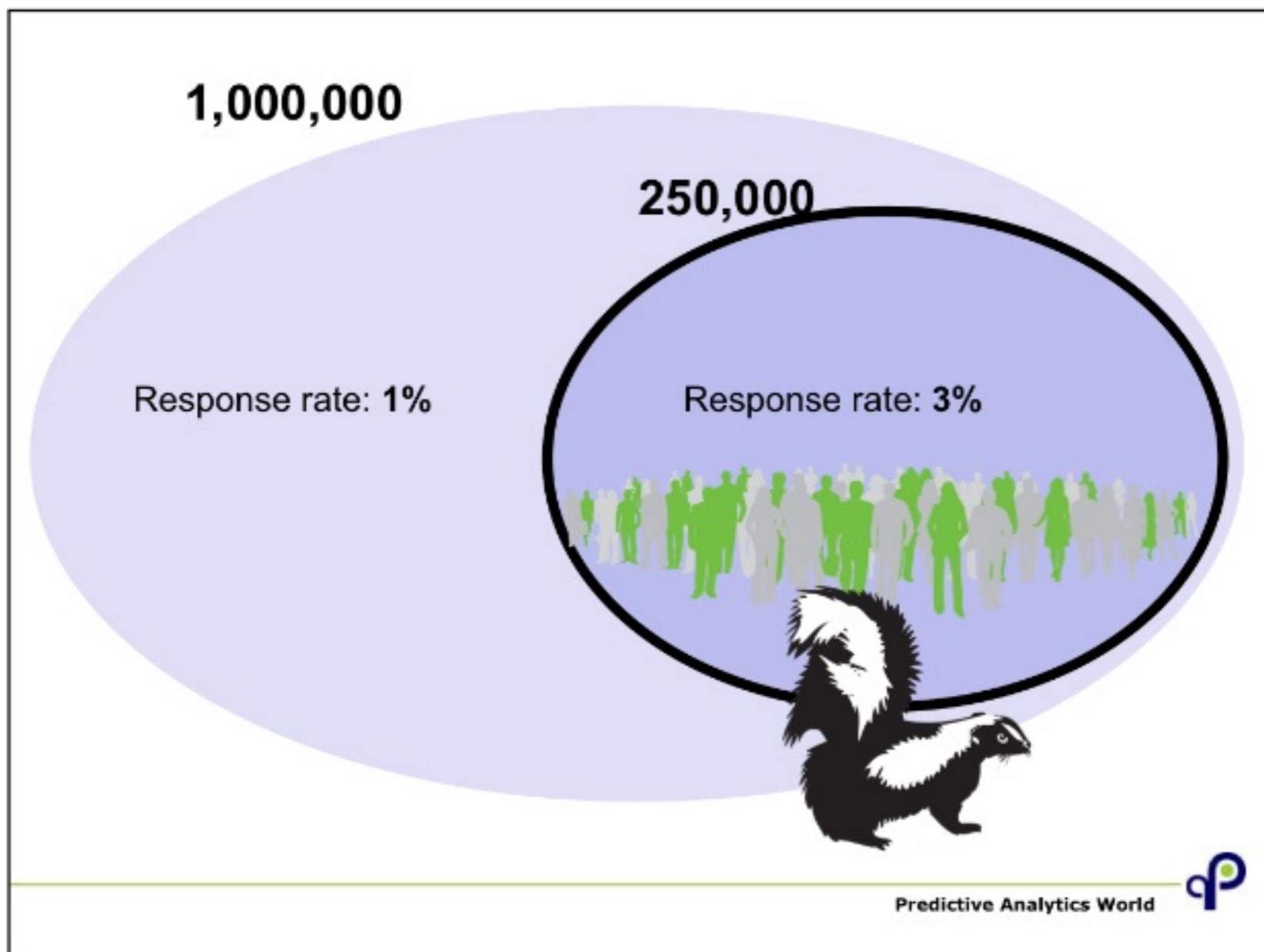
\$6.5 billion within a couple years

84% of marketing leaders intend to increase the role of predictive analytics in marketing over the next 12 months (Forbes)

U.S. shortage of analytics experts at 140,000 in the near-term

LinkedIn's first/second "Hottest Skills That Got People Hired" is "statistical analysis and data mining"

1) Uplift Modeling



A crummy predictive model delivers big value. It's like a skunk with bling.



only
Will the customer buy if contacted?
▲

only
Will the patient become healthy if treated?
▲

Predictive Analytics World 

Does contacting the customer make them more likely to respond?

MEDICAL:

Will the patient survive if treated?

"My headache went away!" Proof of causality by example.

Driving medical decisions with personalized medicine: selecting treatment, e.g., treating heart failure with betablockers

Personalized medicine. Naturally, healthcare is where the term treatment originates. While one medical treatment may deliver better results on average than another, personalized medicine aims to decide which treatment is best suited for each patient, since a treatment that helps one patient could hurt another. For example, to drive beta-blocker treatment decisions for heart failure, researchers "use two independent data sets to construct a systematic, subject-specific treatment selection procedure." (Claggett et al 2011) Certain HIV treatment is shown more effective for younger children. (McKinney et al 1998) Cancer

Uplift modeling:

Analytically modeling to predict the influence on a customer's buying behavior that results from choosing one marketing treatment over another.

"How much more likely is this treatment to generate the **desired outcome** than the alternative treatment?"

Predictive Analytics World 

a.k.a., net lift modeling and persuasion modeling





Resulting improvements with uplift modeling:

- Campaign ROI increased over 5 times
- Costs down 40%
- Lift increased by a factor of 2

Predictive Analytics World



US BANK EXAMPLE

Business case: Direct mail for a home-equity line of credit

Approach: Target campaign with an uplift model

... to existing customers (i.e., a cross-sell marketing campaign)

Resulting improvements over prior conventional analytical approach:

Campaign ROI increased over 5 times previous campaigns (75% to 400%)

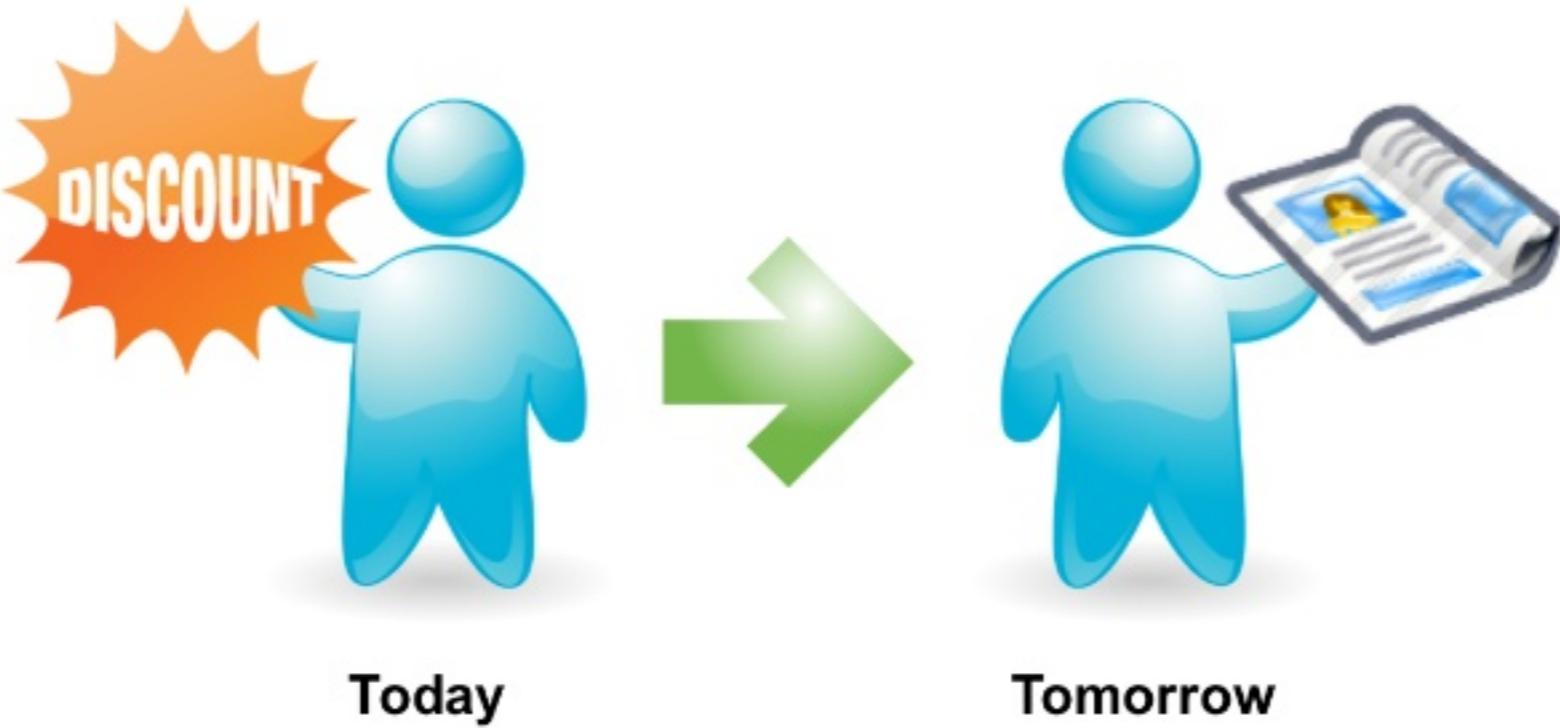
Cut campaign costs by 40%

Increase incremental cross-sell revenue by over 300%

Decrease mailings to customers who would purchase whether contacted or not, and customers who would purchase only if *not contacted*.

Sources: Radcliffe & Surry (2011), Tsai (2010), Patrick Surry (Pitney Bowes Business Insight), Michael Grundhoefer (US Bank)

Retention with Churn Modeling



Predictive Analytics World



**RENEW YOUR
CELL PHONE
CONTRACT...**



**AND
GET A FREE
PHONE!**





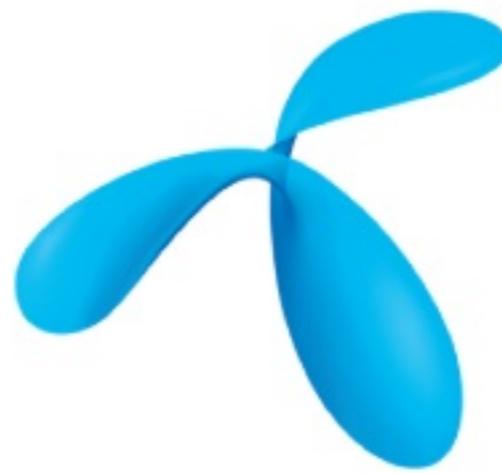
Two articles I wrote provide more details on the Obama campaign's use of predictive analytics (one a reprint of the pertinent section in my book):

<http://www.thefiscaltimes.com/Articles/2013/01/21/The-Real-Story-Behind-Obamas-Election-Victory>

<http://bigthink.com/experts-corner/team-obama-mastered-the-science-of-mass-persuasion-and-won>

Then check out this video of a Predictive Analytics World presentation by one of the hands-on leads on this Obama project, which gets more technical:

<http://www.predictiveanalyticsworld.com/patimes/video-pinpointing-persuadables-convincing-right-customers-right-voters/>



telenor

Resulting improvements with uplift modeling:

- Campaign ROI increased 11-fold
- Reduce churn a further 36% relative to baseline
- Targeted volume reduced by 40%

Telenor: world's 7th largest mobile operator - 159 million subscribers

Improvements are relative to their existing best-practice retention models.

Case study presented at Predictive Analytics World, February 2009, San Francisco.

Case study and graph courtesy of Pitney Bowes Business Insight.

2) Sound Science

The Seattle Times

2012



“

*An orange used car is least likely
to be a lemon.*

”

Predictive Analytics World

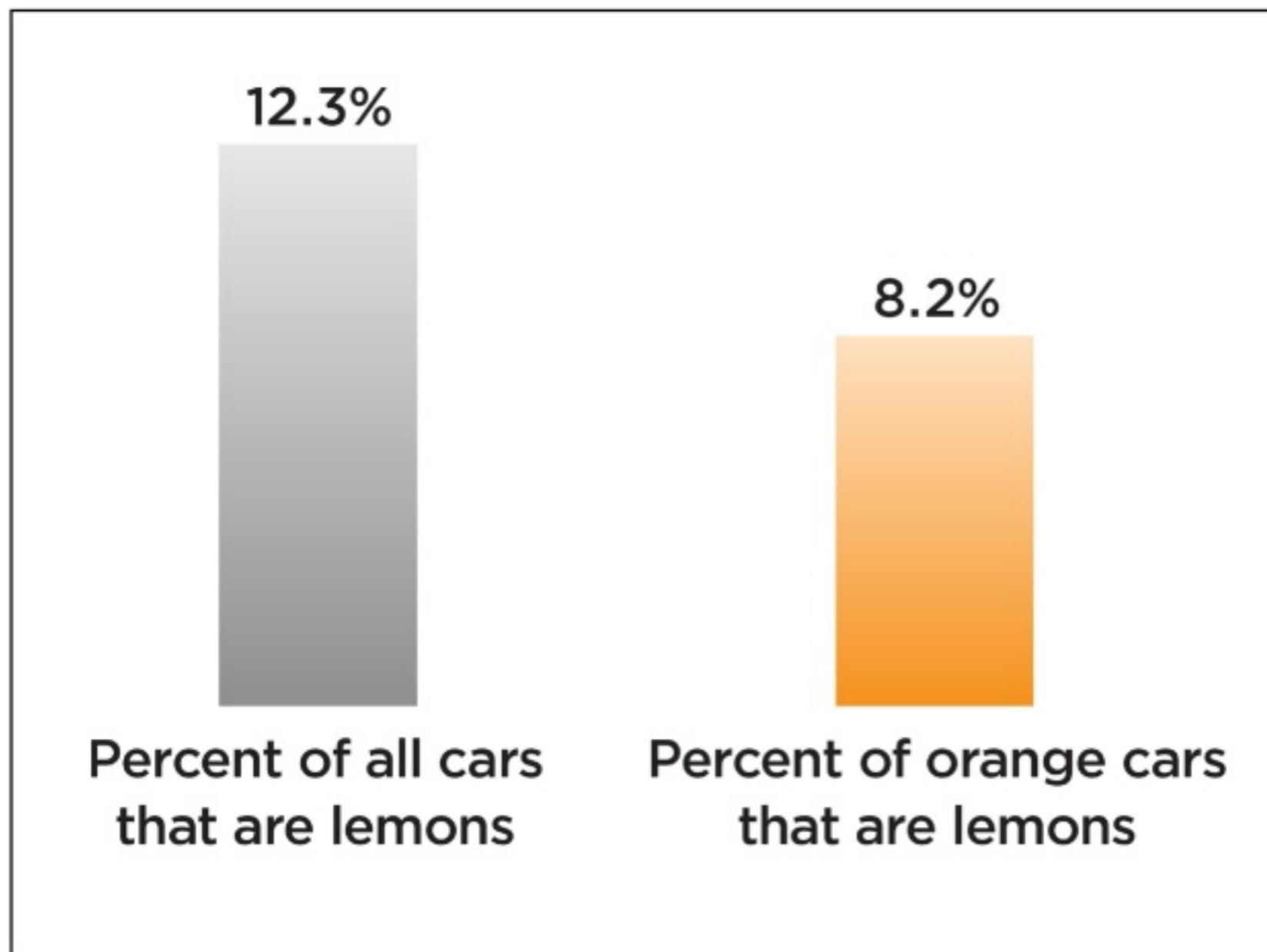


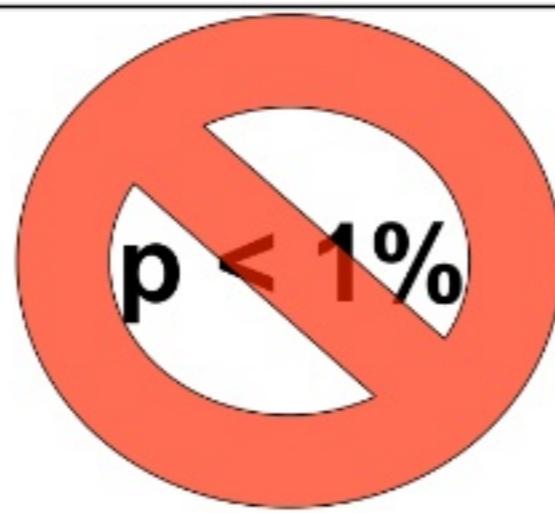
Marcus Wohlsen, “San Francisco startup makes data science a sport.” The Seattle Times, April 14, 2012.







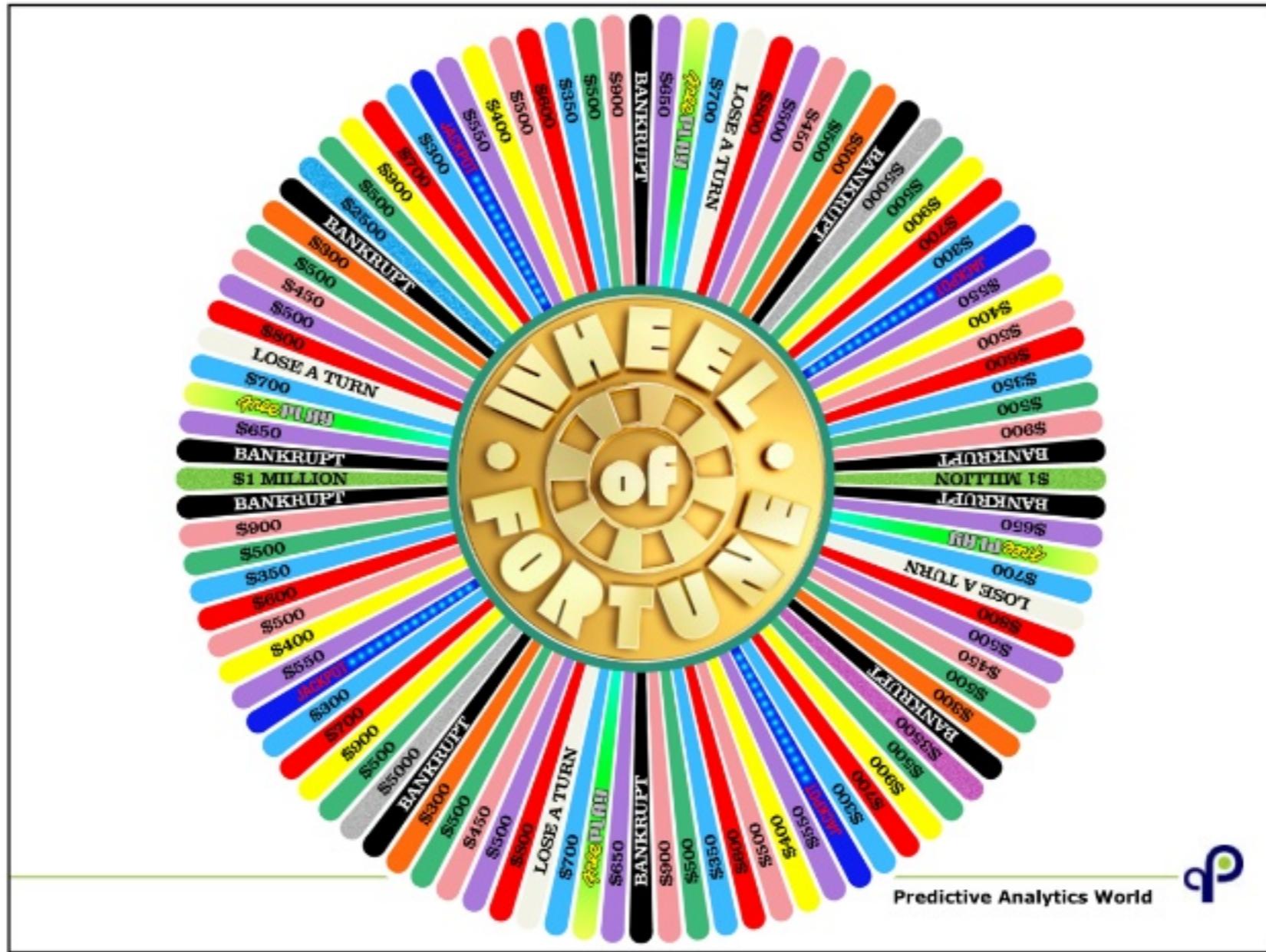


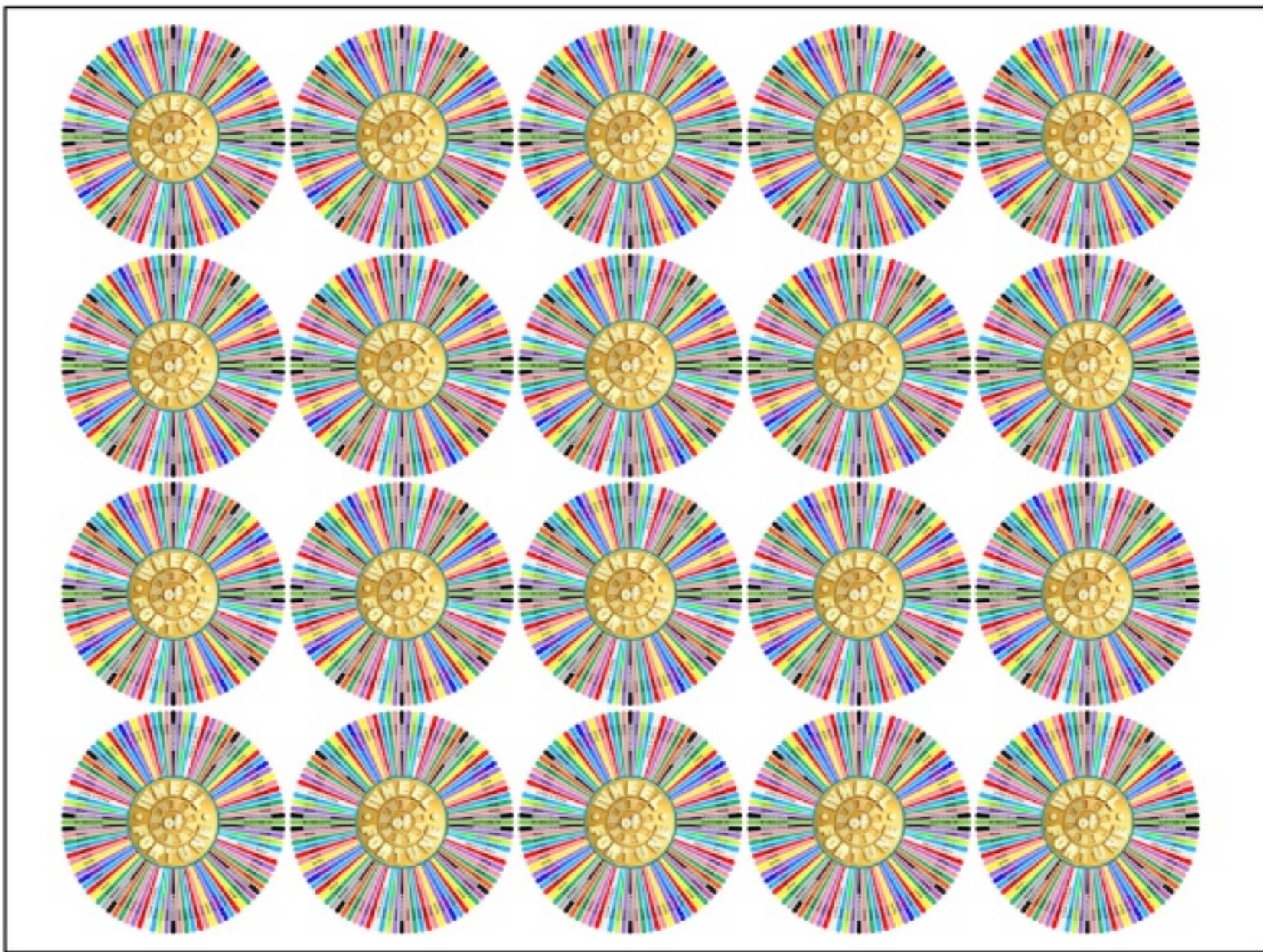


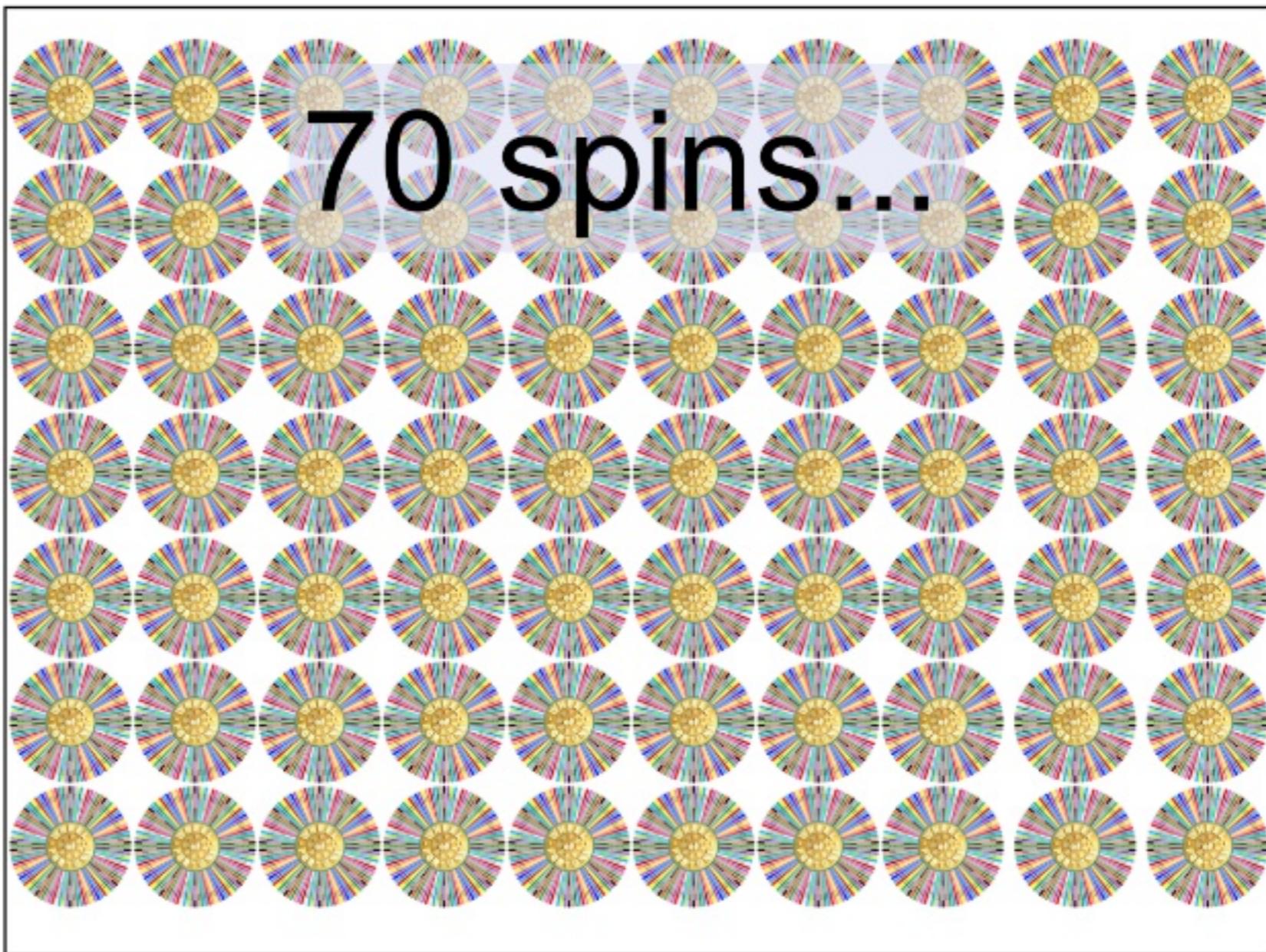
if orange cars were no different,
there'd be a $< 1\%$ chance we'd
see that extreme a result anyway

BAD SCIENCE

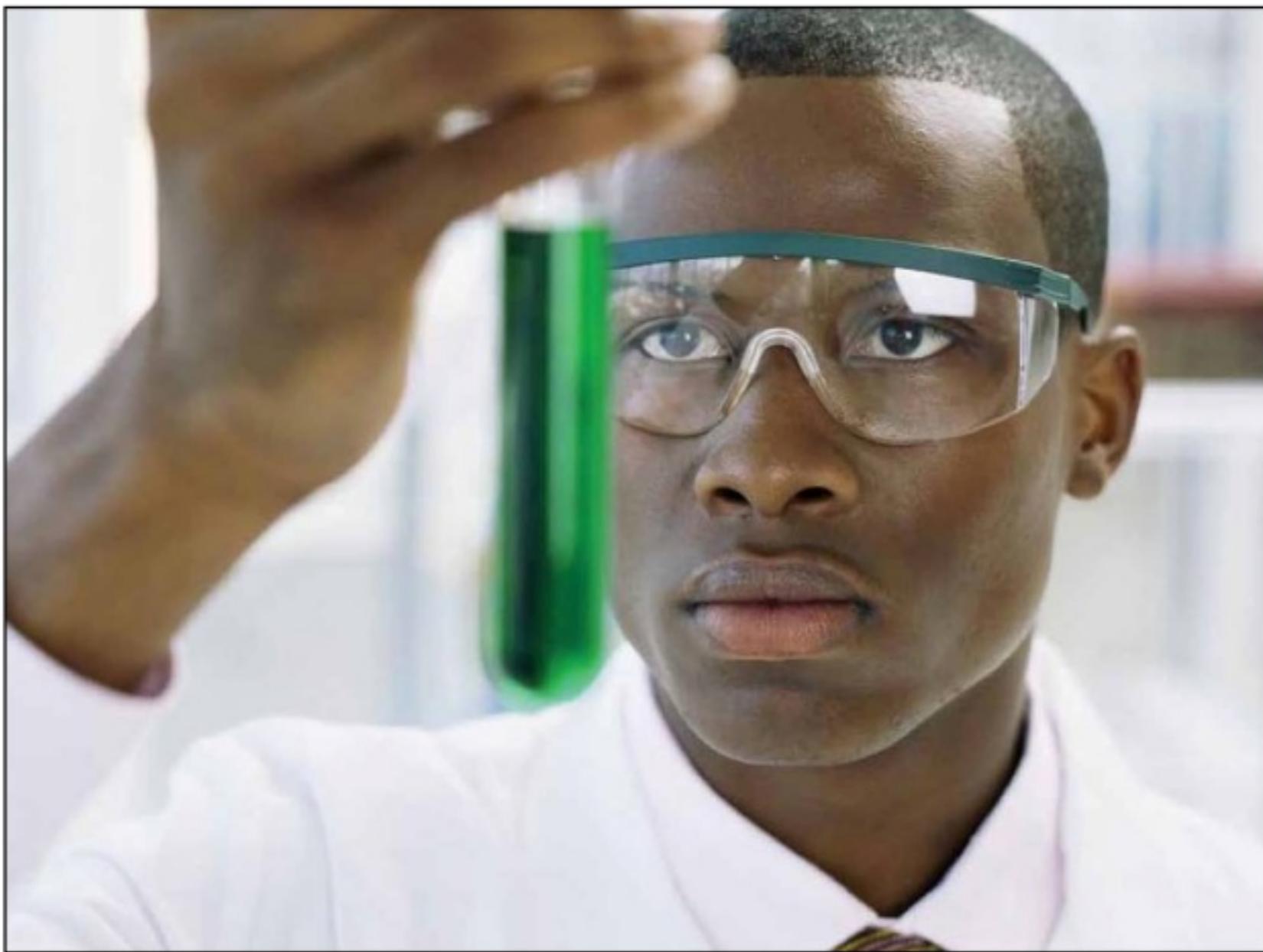


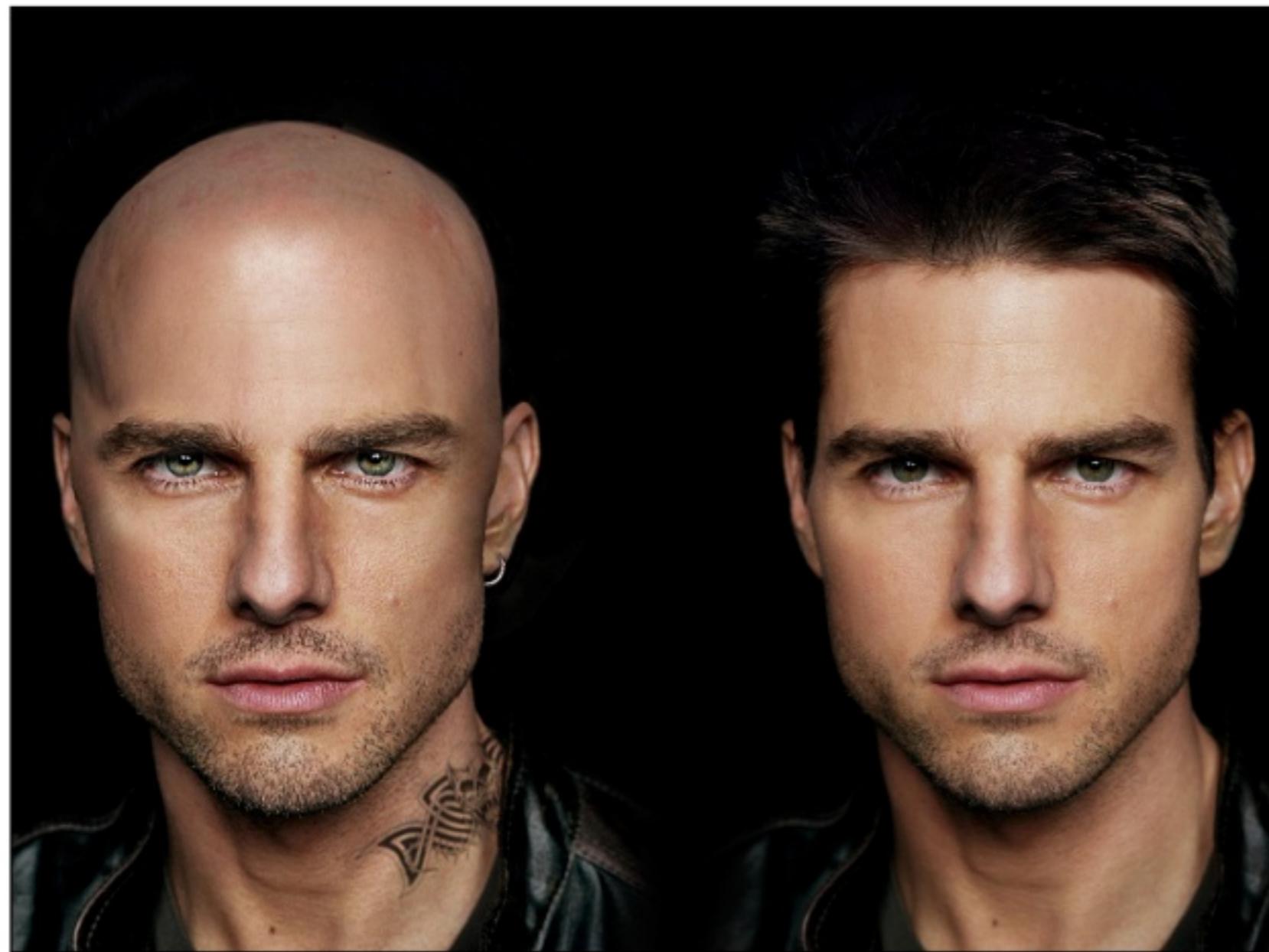


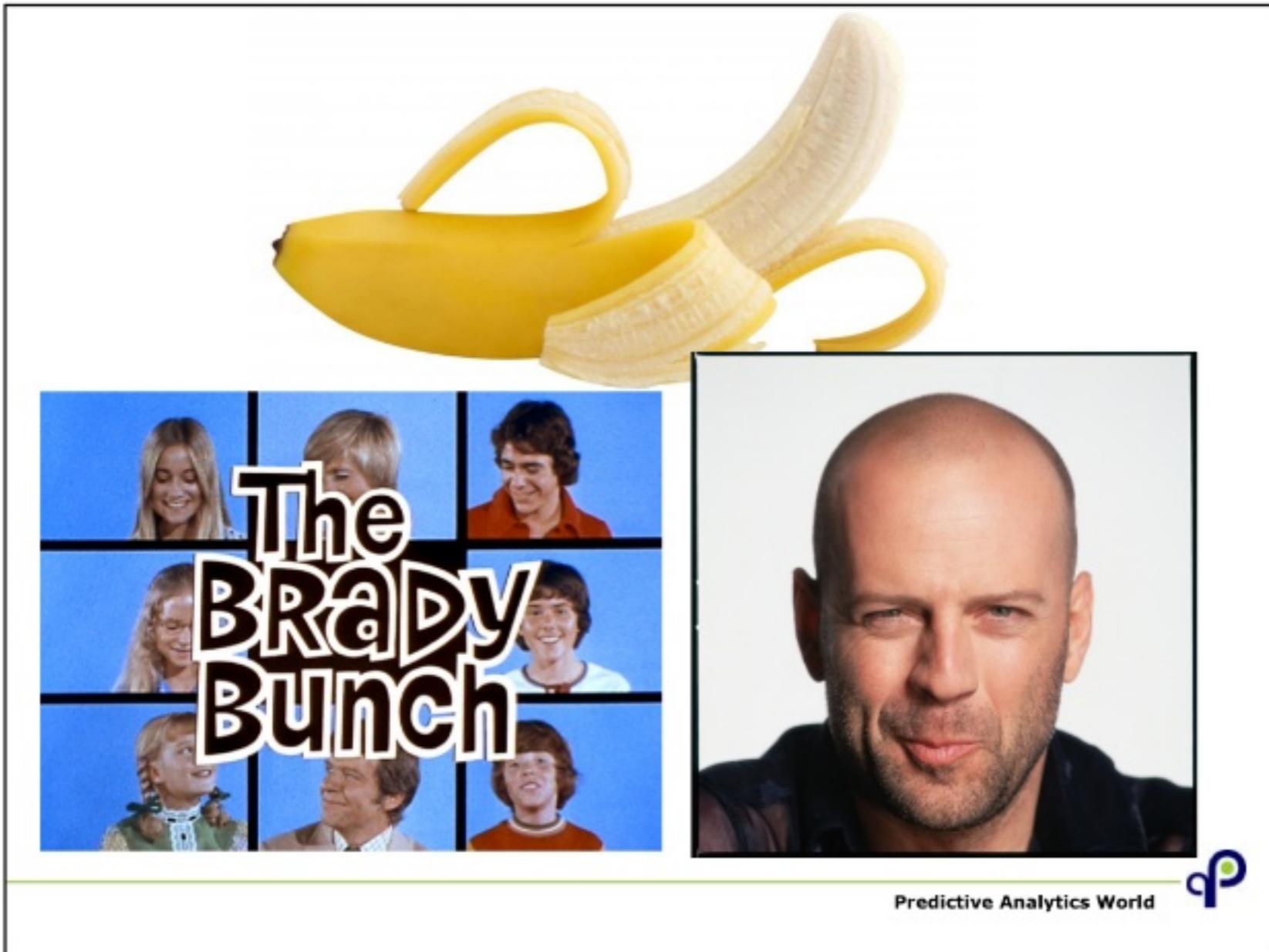




$$\left(\frac{99}{100}\right)^{70} = 0.495$$









$$\left(\frac{99}{100}\right)^{70} = 0.495$$



vast search

p-hacking

wider
~~bigger~~ data is more hazardous



“Last Week Tonight” with John Oliver, May 8, 2016:

<https://youtu.be/0Rnq1NpHdmw>

A Better p-Value

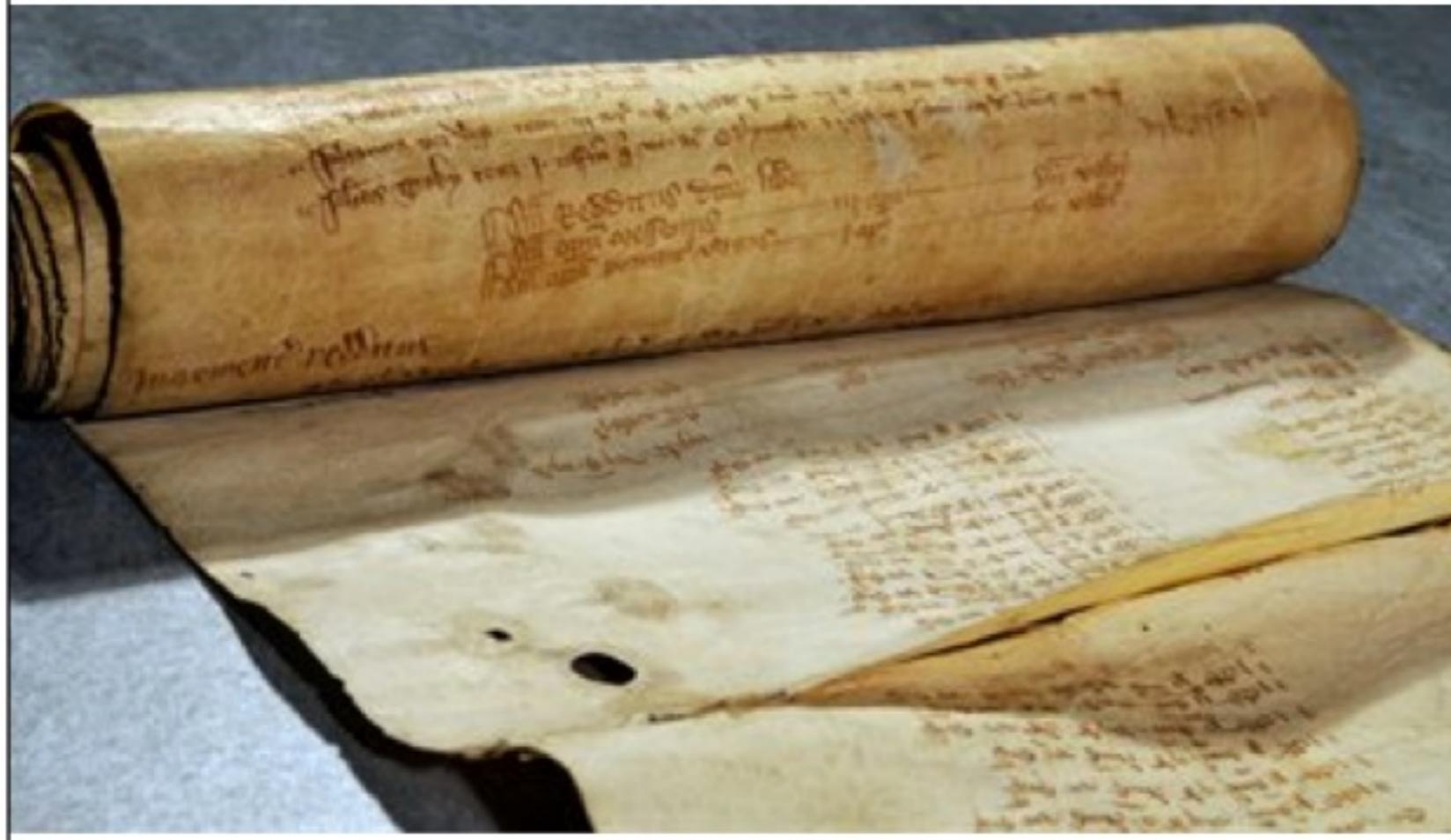


P=7.2%

Predictive Analytics World

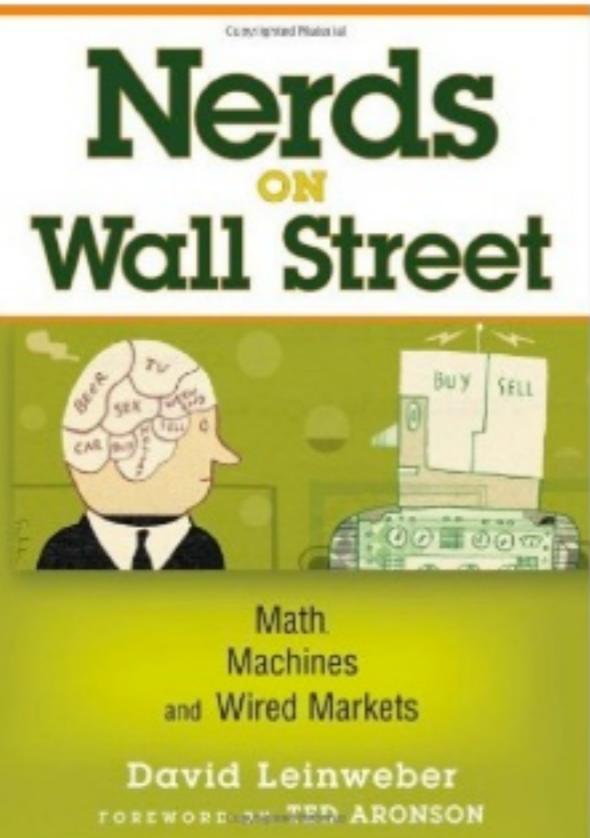


Bigger Longer Data



Employ a Test Set





“An unlimited amount of computational resources is like dynamite: If used properly, it can move mountains. Used improperly, it can blow up your garage or your portfolio.”

—David Leinweber

After 150 Years, the American Statistical Association Says No to p-values:

<http://www.kdnuggets.com/2016/03/asa-says-no-p-values.html>

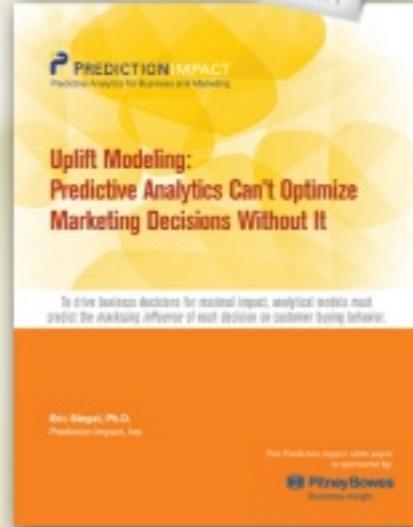
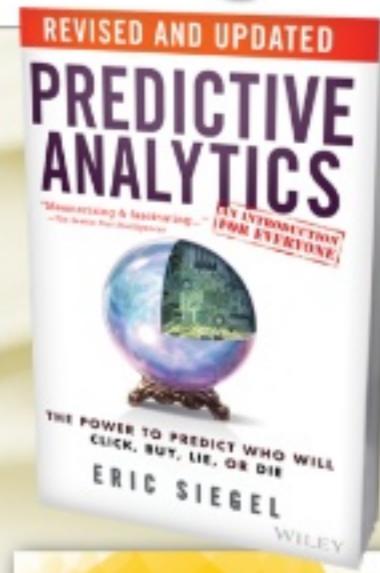
"Even if something is published in the flagship journal of the leading association of research psychologists, there's no reason to believe it. The system of scientific publication is set up to encourage publication of spurious findings."

- Andrew Gelman, professor of statistics and political science at Columbia University

http://www.slate.com/articles/health_and_science/science/2013/07/statistics_and_psychology_multiple_comparisons_give_spurious_results.2.html

Further Info: Uplift Modeling

Chapter 7: Persuasion by the Numbers
How Telenor, U.S. Bank, and the Obama Campaign Engineered Influence



White paper:

Uplift Modeling: Predictive Analytics Can't Optimize Marketing Decisions Without It

www.pawcon.com/signup-uplift-whitepaper.php

Free white paper: www.predictiveanalyticsworld.com/signup-uplift-whitepaper.php

More updated version thereof is Chapter 7 of "Predictive Analytics" (www.thepredictionbook.com)

Or see the (much more) technical papers that chapter cites - see under Chapter 7 of the book's Notes PDF, available online at <http://www.PredictiveNotes.com>.

See also: <http://www.predictiveanalyticsworld.com/patimes/personalization-is-back-how-to-drive-influence-by-crunching-numbers/>

Further Info: Vast Search

Chapter 3: The Data Effect

Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die

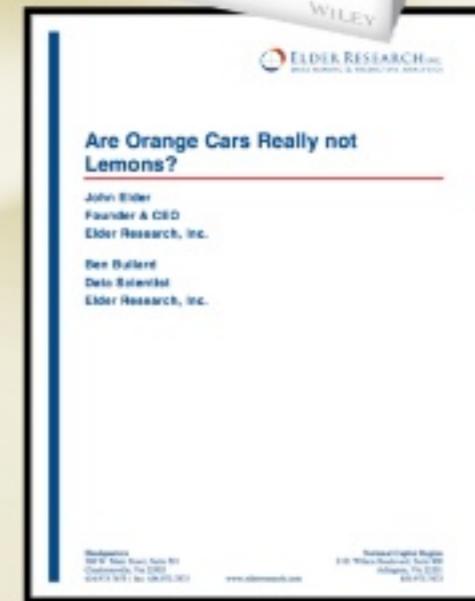
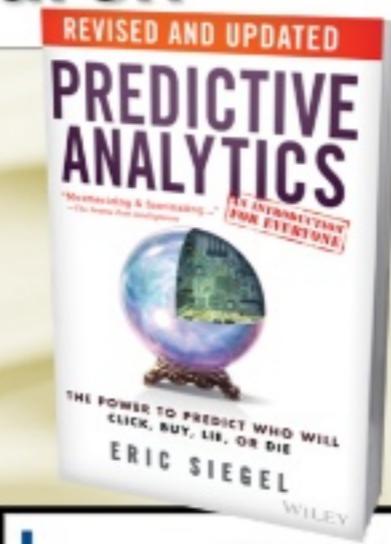
Paper and Video:

Are Orange Cars Really not Lemons?

by John Elder and Ben Bullard, Elder Research, Inc.

www.elderresearch.com/orange-car

www.predictiveanalyticsworld.com/patimes/video-the-power-and-peril-of-predictive-analytics/



For more on vast search, see Chapter 3 of *Predictive Analytics* by Eric Siegel (www.thepredictionbook.com). For further citations on this topic, also see the final portion that chapters Notes (citations), available as a PDF online at www.PredictiveNotes.com.

Also on a broadly-accessible level, see chapter seven, "Dead Fish Don't Read Minds," of "How Not to Be Wrong: The Power of Mathematical Thinking," by Jordan Ellenberg.

Video: www.predictiveanalyticsworld.com/patimes/video-the-power-and-peril-of-predictive-analytics/

See also: <http://www.predictiveanalyticsworld.com/patimes/breakthrough-avert-analytics-treacherous-pitfall/3366/>

See also Elder Research's short, friendly video illustrating target shuffling:
<http://www.elderresearch.com/target-shuffling-video>

For a more technical treatment, in addition to Elder and Bullard's paper cited on the slide above, see "Handbook of Statistical Analysis and Data Mining Applications" by Nisbet et al - second edition only - Chapter 20, "SIGNIFICANCE vs LUCK in the age of MINING." Coincidentally, I used "BS" in earlier versions of this talk before knowledge of similar vocabulary in that new edition of the "Handbook" and elsewhere ☺. Here's a quote from that chapter:

Replication is supposed to be a hallmark of scientific integrity. The concern over this lapse of 'scientific accuracy/integrity' has been growing for some time, especially over the past 3 years, such that the University of Washington (Seattle) is offering a new course for Spring 2017 titled "Calling Bullshit in the Age of Big Data." Among the Learning / Behavioral Objectives for this course is the following: 'After taking this course, you will be able to provide a statistician or fellow scientist with a technical explanation of why a claim or conclusion is 'in error'. Below are three links that go to this new course, including a syllabus:

- <http://callingbullshit.org/syllabus.html#Big>;
- <http://www.recode.net/2017/2/19/14660236/big-data-bullshit-college-course-university-washington>;
- <https://www.statnews.com/2017/02/17/science-fights-alternative-facts/>

One more link:

http://www.slate.com/articles/health_and_science/science/2016/01/amy_cuddy_s_power_pose_research_is_the_latest_example_of_scientific_overreach.html



San Francisco • Boston
Chicago • Toronto
London • Berlin
Washington DC (Government)

Bigger wins!
Strengthen the business impact
delivered by predictive analytics.

www.pawcon.com

Predictive Analytics World Conference

*"Predictive Analytics World
was probably the best analytics
conference I have attended... turned
into my new must-go-to conference."*

Dennis Mortensen
Director of Data Insights
Yahoo!

PAW March 2012 drew 500 attendees.
PAW has included case studies from: AT&T, Aflac,
Amazon.com, Amway, Anheiser-Busch, BBC, Barclays,
Blue Cross, BoA, Citibank, Fidelity, GE Capital,
Google, HP, HSBC, Infinity, Intuit, LinkedIn, Macy's,
Match.com, MetLife, Microsoft, NRA, Netflix, Orbitz,
Overstock.com, PayPal, Reed Elsevier, Target, The
Economist, US Bank, US Bank, Wells Fargo, YMCA,
Yahoo! and many more.

91 % of attendees rated
the instructor
Excellent or Very Good



PREDICTION IMPACT

Predictive Analytics for Business and Marketing

**Online training program:
*Predictive Analytics Applied***

Available on-demand, at your convenience

"Thank you for the perfect
introduction to predictive analytics.
It's just the right blend of a business
and technical angle I was looking for."

—Emiliano Pasqualetti,
CEO, DomainsBot Inc.



businessprediction.com

**To Learn More:
The Predictive Analytics Guide**

www.pawcon.com/guide

Predictive Analytics World





@predictanalytic

Predictive Analytics World 