Analytics 3.0

Big Data in Big Companies

Tom Davenport

Babson/MIT/International Institute for Analytics/Deloitte

Caterpillar

March 27, 2014

Two Types of Data

Name	Sex	Birth Date	Post Code	Complaint
Peter	Male	02/12/1954	SE24 6TY	Pain in left eye
Stewie	Male	05/01/1984	NW1 6XD	Chest pains
Chris	Female	04/08/1978	E17 7WE	Chest pains
Louis	Female	03/10/1960	WC1 7RA	Back pains
Meg	Male	09/09/1990	NW7 5LK	Headaches

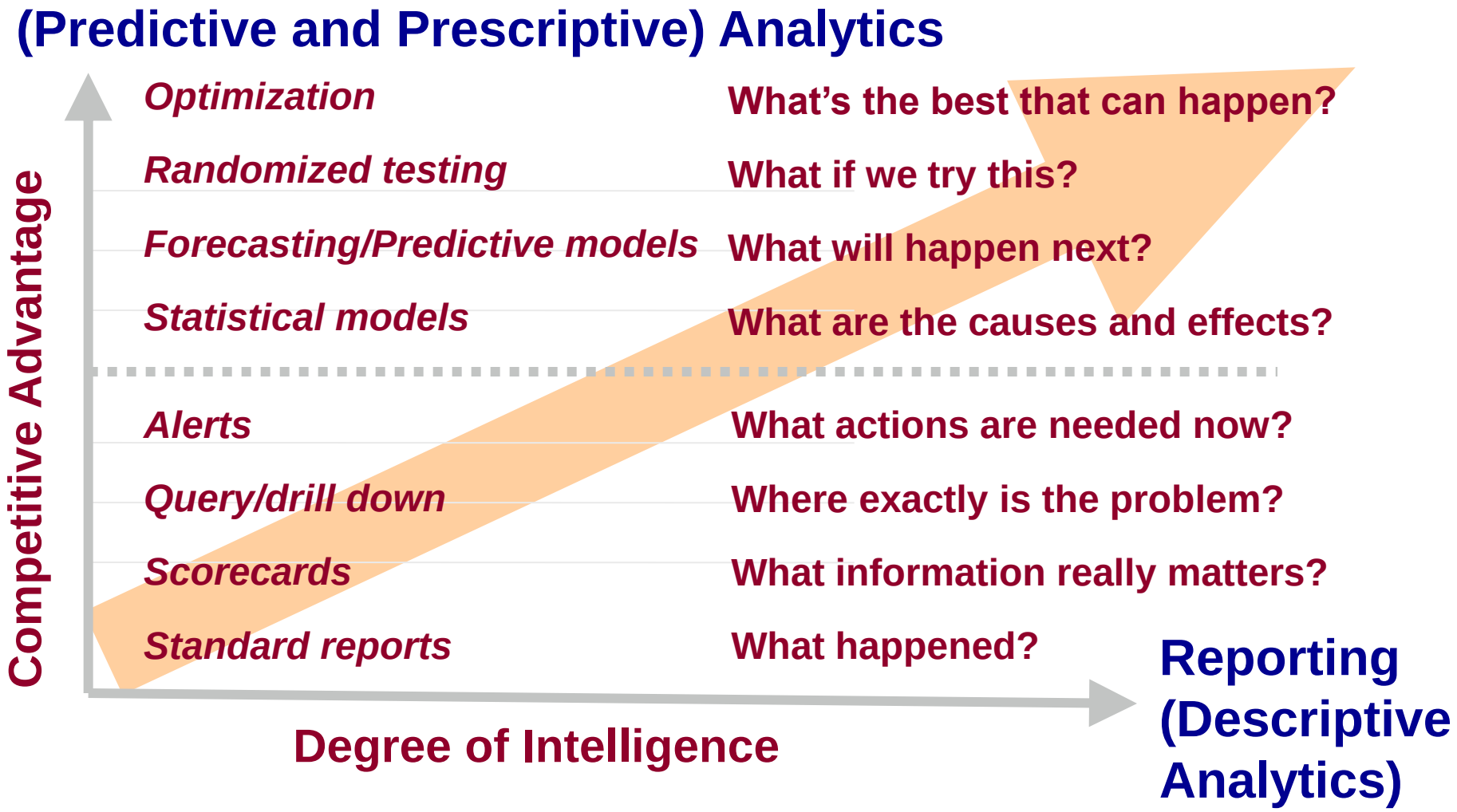
"Small data"

- Small volumes—< 100 TB
- Usually internal
- Structured in rows and columns of numbers
- Slow-moving enough to segregate
- Already suitable for analytics

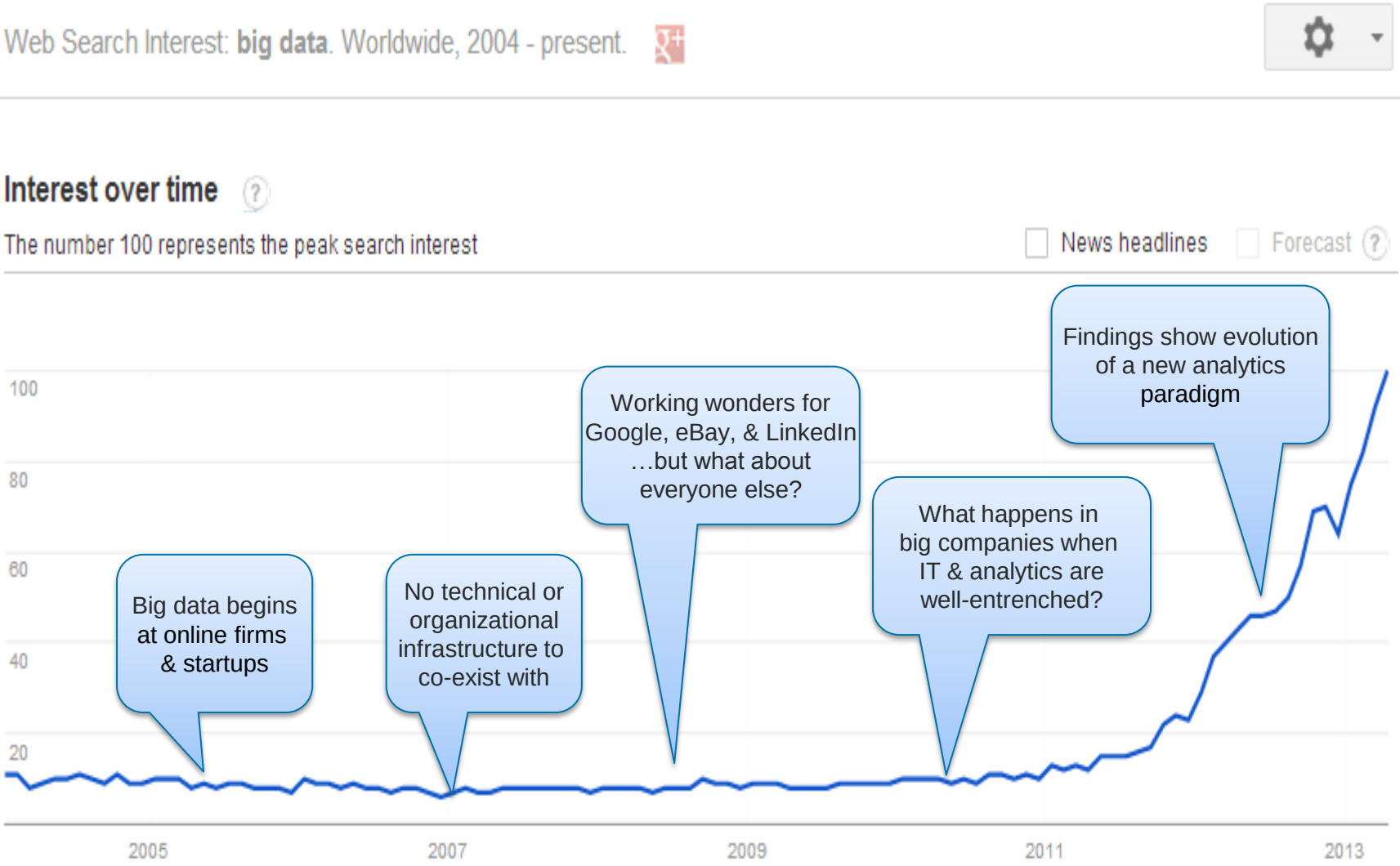
"Big data"

- Large—in petabytes
- Often external
- Unstructured—text, voice, video, image, etc.
- Continually flowing
- Needs to be structured to be analyzed

Three Types of Analytics



The Rise of Big Data



Analytics 1.0 | Traditional Analytics

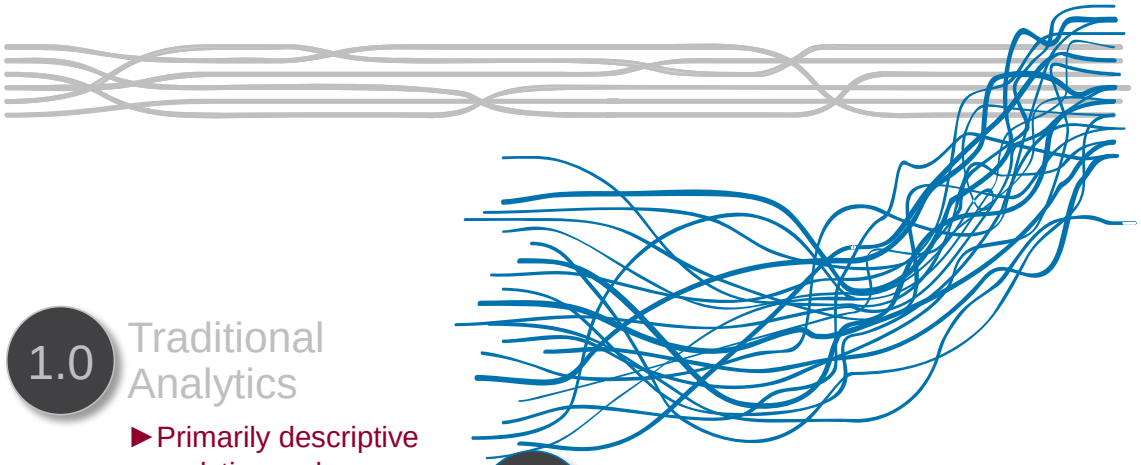
1.0 Traditional Analytics

- Primarily descriptive analytics and reporting
- Internally sourced, relatively small, structured data
- "Back office" teams of analysts
- Internal decision support

Analytics 1.0 | Technologies

- Data warehouses
  - Challenging to get data in
  - Initially for exploring data, now for production
- Standalone spreadsheets
  - Replete with errors, multiple versions of the truth
- BI and analytics “packages”
  - Too many functions
- Pre-digested data cubes
  - Easy but limiting

Analytics 2.0 | The Big Data Era



1.0 Traditional Analytics

- Primarily descriptive analytics and reporting
- Internally sourced, relatively small, structured data
- "Back room" teams of analysts
- Internal decision support

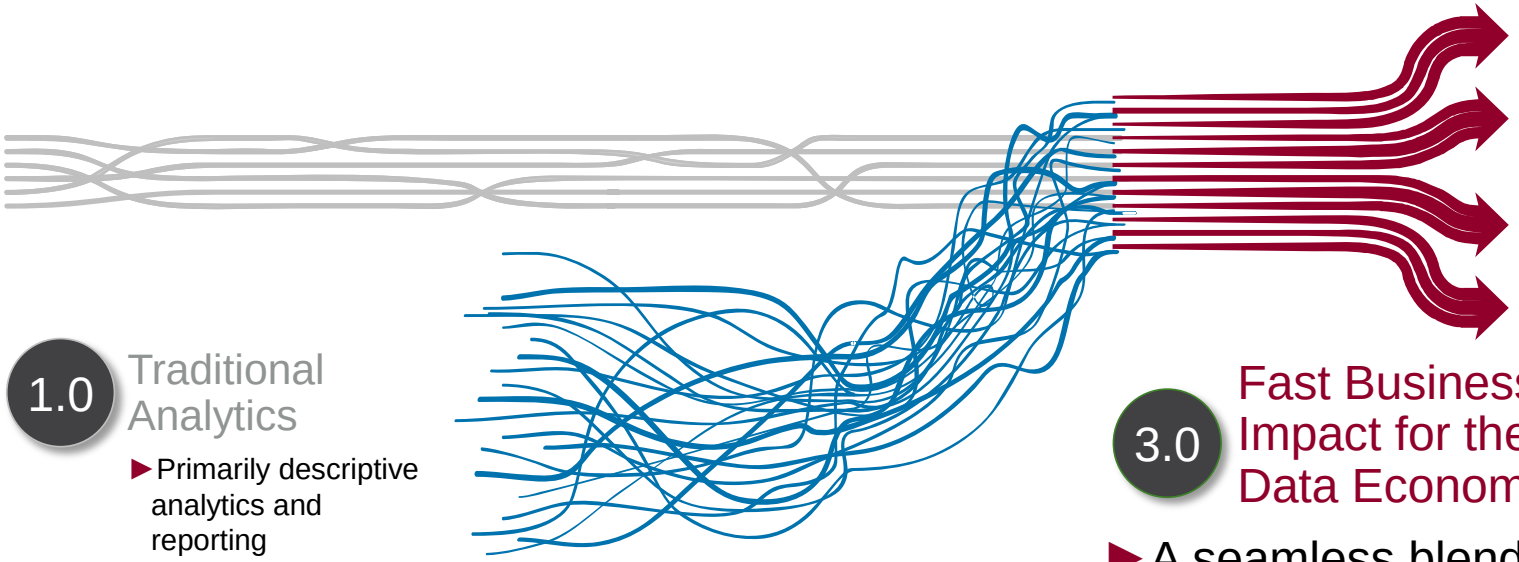
2.0 Big Data

- Complex, large, unstructured data
- New computational capabilities required
- "Data Scientists" emerge
- Online firms create "data products"

Analytics 2.0 | Ethos

- Be “on the bridge” if not in charge of it
- “Agile is too slow”
- “Being a consultant is the dead zone”—develop products, not presentations or reports
- Information (and hardware and software) wants to be free and shared
- “Nobody’s ever done this before!”

Analytics 3.0 | Fast Business Impact for the Data Economy



1.0 Traditional Analytics

- Primarily descriptive analytics and reporting
- Internally sourced, relatively small, structured data
- "Back room" teams of analysts
- Internal decision support

2.0 Big Data

- Complex, large, unstructured data sources
- New analytical and computational capabilities
- "Data Scientists" emerge
- Online firms create data-based products and services

3.0 Fast Business Impact for the Data Economy

- A seamless blend of traditional analytics and big data
- Analytics integral to running the business
- Data and analytics-based products in every business
- Industrialized decision-making at scale

Analytics 3.0 | Goals

Developing products and services based on data and analytics—now available to every industry

- "Precision agriculture" offerings for growers
- Conditional and predictive services for industrial equipment

Data and analytics-based decisions at scale and supporting the front line of organizations

- Real-time routing
- Granular, targeted marketing programs

Analytics 3.0 | Data Types

Purchase *history*
Hosted applications
Social Feeds
Customer *segmentation*
Blogs
Twitter
Customer *value*
Mobile devices
Presentations
Website activity
Purchasing *behavior*
LinkedIn
Articles
Customer *profiles*
Recommendations
Device sensors
Organization *contacts*
Sentiment *analysis*
Cloud
Billing
Clickstream logs
Spatial GPS
Target marketing
Marketing
Documents
Text messages
Contracts/*orders*
Customer experience *management*
Email
Shipping
Service tiers
Images
Claims
RSS
Call center
Customer *service*
Videos
XML

GE 3.0

- $2B initiative in software and analytics
- Primary focus on data-based products and services from “things that spin”
- Will reshape service agreements for locomotives, jet engines, turbines
- Gas blade monitoring in turbines produces 588 gigabytes/day—7 times Twitter daily volume
- Marketing new industrial data platforms and brands like “Predicity” and “Predix”

Monsanto 3.0

MONSANTO

- FieldScripts program uses data from field testing and Monsanto research to recommend what corn hybrids to plant where
- Genotypes and phenotypes of plants add up to tens of petabytes of data for analysis
- Field photographs analyzed to determine correct watering, fertilizer
- Paid almost $1B for The Climate Company, which gathers and analyzes weather data for agriculture
- Embarking on data and analytics education programs for farmer customers

Procter & Gamble 3.0

- Primary focus on improving management decisions
- “Information and Decision Solutions” (IT) embeds over 250 analysts in leadership teams
- Over 50 “Business Spheres” for executive information viewing and decision-making
- “Decision cockpits” on 50K desktops
- Real-time social media sentiment analysis for “Consumer Pulse”
- Financial restatements in seconds versus several days in the past
- P&L’s by brand and retailer on the fly

Schneider National 3.0

- Has invested heavily in sensors to automate data collection on trucks, trailers and intermodal containers
- Quality of decisions has improved as a result of sensor data
- Prescriptive analytics are changing job roles and relationships
- Sensor data related to safety predicts drivers at risk of safety accident for preventative conversations

What Should Organizations Do with Analytics?

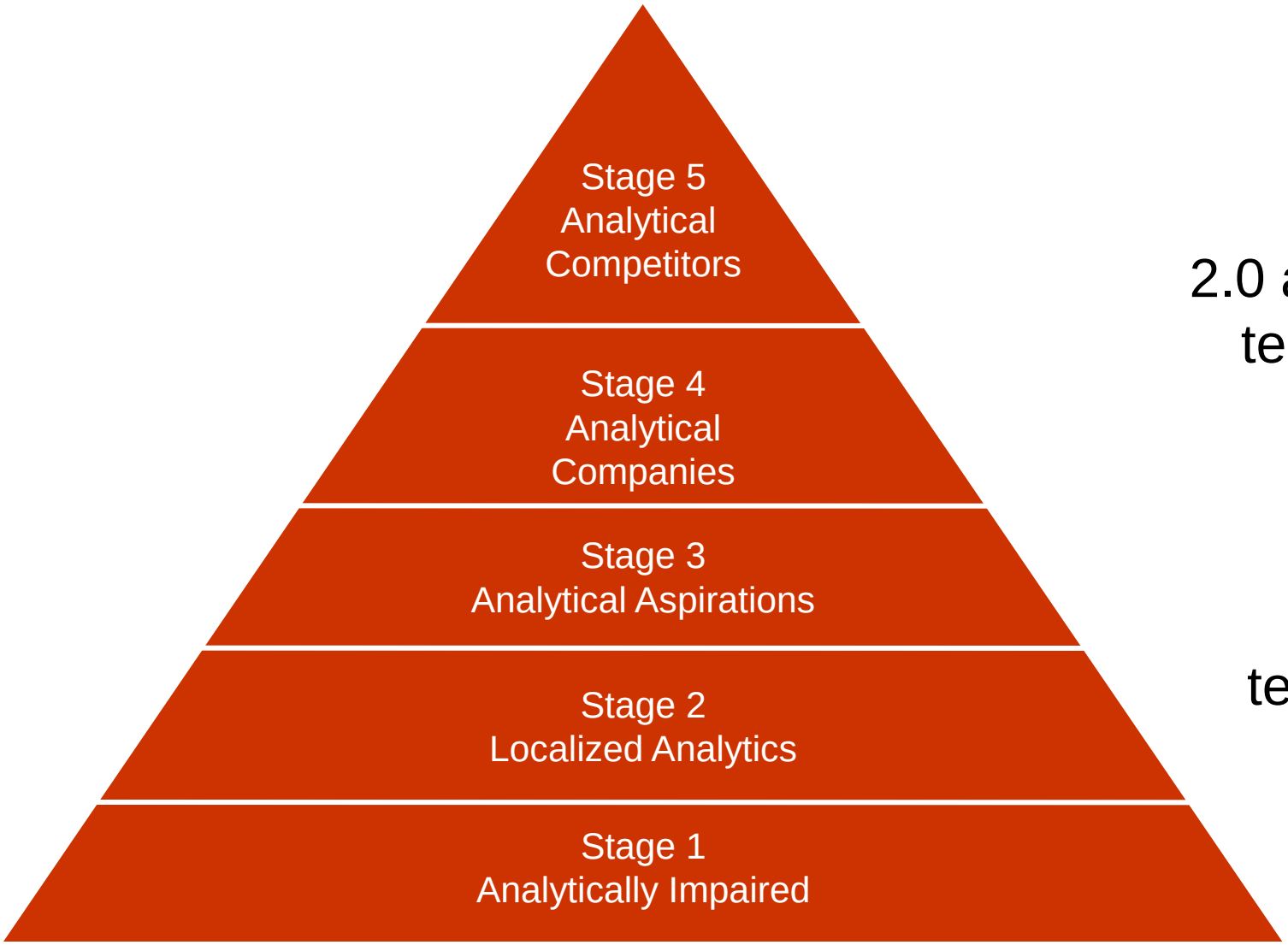
1.0 organizations *use* analytics

- Optimizing aspects of operations
- Finding the best customers, and charging them the right price
- Allocating costs accurately and understanding how financial performance is driven

2.0 and 3.0 organizations *compete on analytics*

- Making analytics and fact-based decisions key elements of products and services, strategy, and competition

Levels of Analytical Capability



Analytical Competitors

Primarily 2.0 firms

- **Netflix—Cinematch, Max, etc.**
- **eBay—testing on a massive scale**
- **Zillow—Zestimates, Rent Zestimates, Underwater Index**
- **Facebook—People You May Know, Custom Audiences, Exchange**

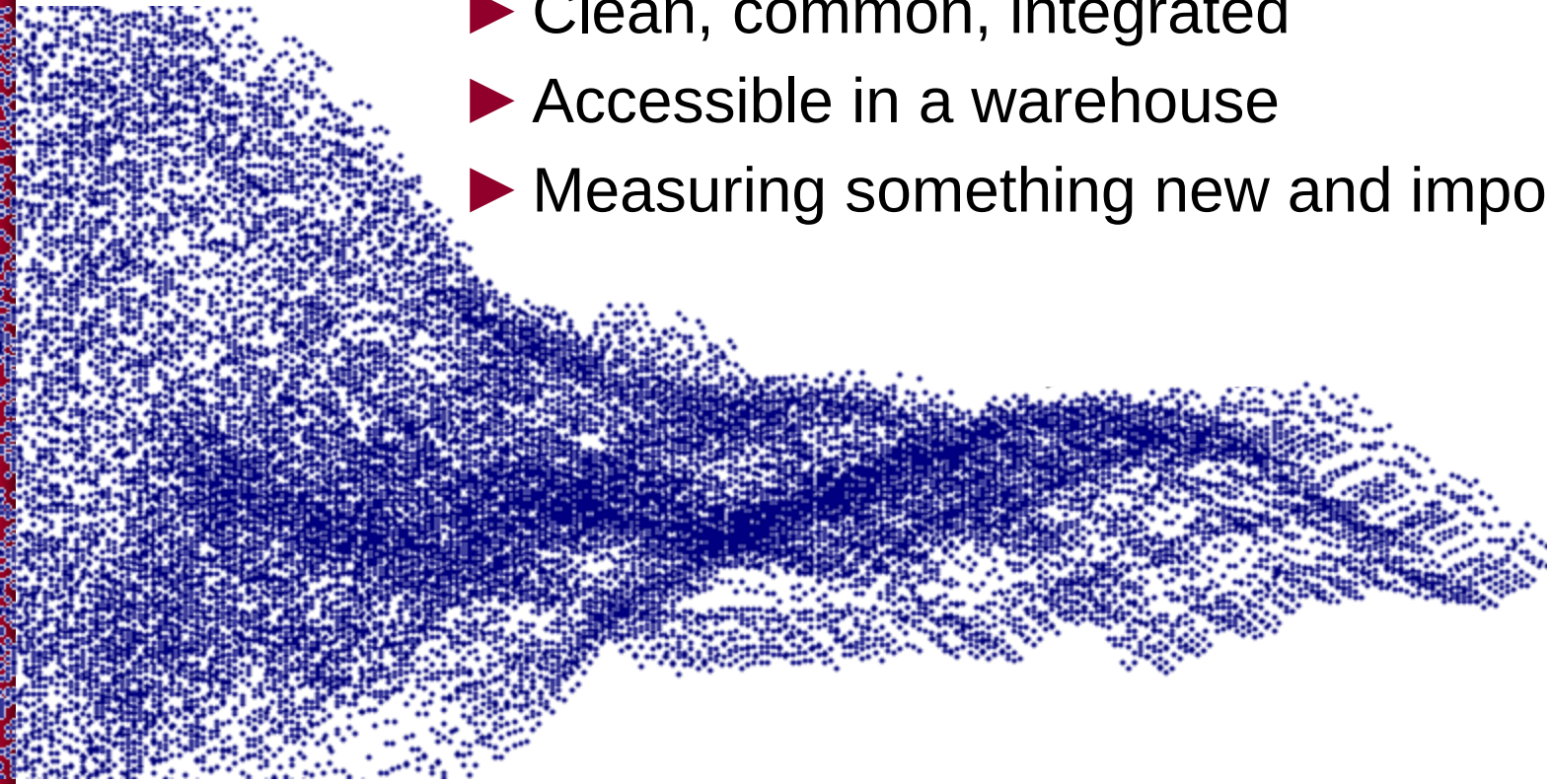
Primarily 3.0 firms

- **Google—page rank, advertising, HR, ventures**
- **LinkedIn—PYMK, Jobs, Groups, and sales**
- **Marriott — Revenue management, web, loyalty**
- **UPS — Real-time routing of 46,000 vehicles**
- **Caesars—Loyalty and service in real time**
- **Tesco—13 million customers, 13 million offers**
- **Capital One—”information-based strategy”**

The Analytical DELTTA

Data

- The prerequisite for everything analytical
- Clean, common, integrated
- Accessible in a warehouse
- Measuring something new and important



New Metrics / Data

Wine Chemistry

Field-Level Weather

Smile Frequency

Enterprise

Enterprise perspectives and coordination on:

- Data
- Analysts
- Technology
- Avoid multiple versions of the truth
- Taking an enterprise perspective can reduce expense

Leadership

Gary Loveman at Caesars

- "Do we think, or do we know?"
- "Three ways to get fired"

Ed Clark at Toronto Dominion

- "Nobody ever accuses us of not running the numbers"

Jeff Bezos at Amazon

- "We never throw away data"

Jeff Inmelt at GE

- From "big iron" to "big data"

"Our CEO is a real data dog"

Sara Lee executive

Targets

With limited analytical resources (or to make an impact), pick a major strategic target, with a minor or two

- Caesars = Loyalty + Service
- AC Milan = Injuries
- UPS = Operations + Customer data

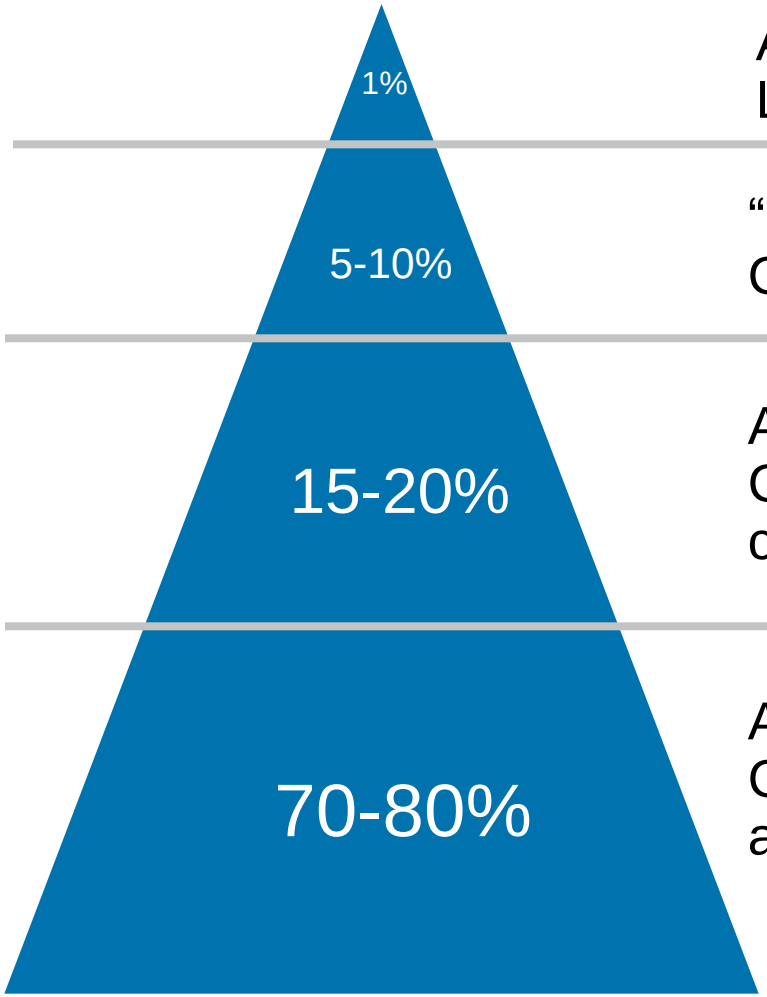
Also need to clarify user targets

- Wal-Mart = Category managers + Suppliers
- Owens & Minor = Logistics + Hospitals

Technologies for Analytics 2.0 and 3.0

- **Hadoop and related technologies for splitting problems across many cheap servers**
- **Cloud processing and analytics**
- **Machine learning**
- **Visual analytics**
- **Mobile “analytical apps”**
- **Sensors and “the Internet of Everything”**

Analysts



Analytical Champions--Own
Lead analytical initiatives

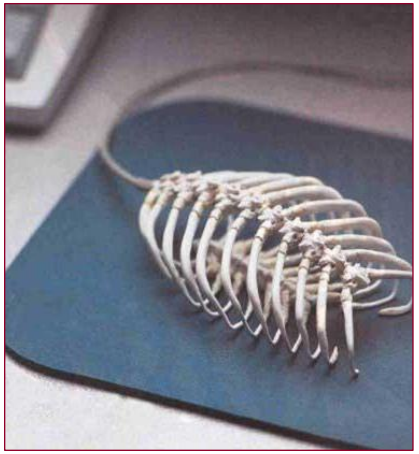
"Data Scientists"—Own/Rent
Can create new algorithms

Analytical Semi-Professionals—Own/Rent
Can use visual and basic statistical tools, create simple models

Analytical Amateurs--Own
Can use spreadsheets, use analytical transactions

* percentages will vary based upon industry and strategy

It Doesn't Happen Overnight — Start Now!



- Takes a while to put data and infrastructure foundation in place, and even longer to develop human capabilities, a fact-based culture, and "success stories"
- Barclay's five-year plan for "Information-Based Customer Management"
- UPS
  - "We've been collecting customer data for six or seven years, but it's only become usable in the last two or three"
  - "We've been installing telematics sensors since 1986."

STRENGTH IN NUMBERS

tdavenport@babson.edu

