Big Data at DST

Bill Nixon, Matt Crouch

DST®

The enclosed materials are highly sensitive, proprietary and confidential. Please use every effort to safeguard the confidentiality of these materials. Please do not copy, distribute, use, share or otherwise provide access to these materials to any person inside or outside DST Systems, Inc. without prior written approval.

This proprietary, confidential presentation is for general informational purposes only and does not constitute an agreement. By making this presentation available to you, we are not granting any express or implied rights or licenses under any intellectual property right.

If we permit your printing, copying or transmitting of content in this presentation, it is under a non-exclusive, non-transferable, limited license, and you must include or refer to the copyright notice contained in this document. You may not create derivative works of this presentation or its content without our prior written permission. Any reference in this presentation to another entity or its products or services is provided for convenience only and does not constitute an offer to sell, or the solicitation of an offer to buy, any products or services offered by such entity, nor does such reference constitute our endorsement, referral, or recommendation.
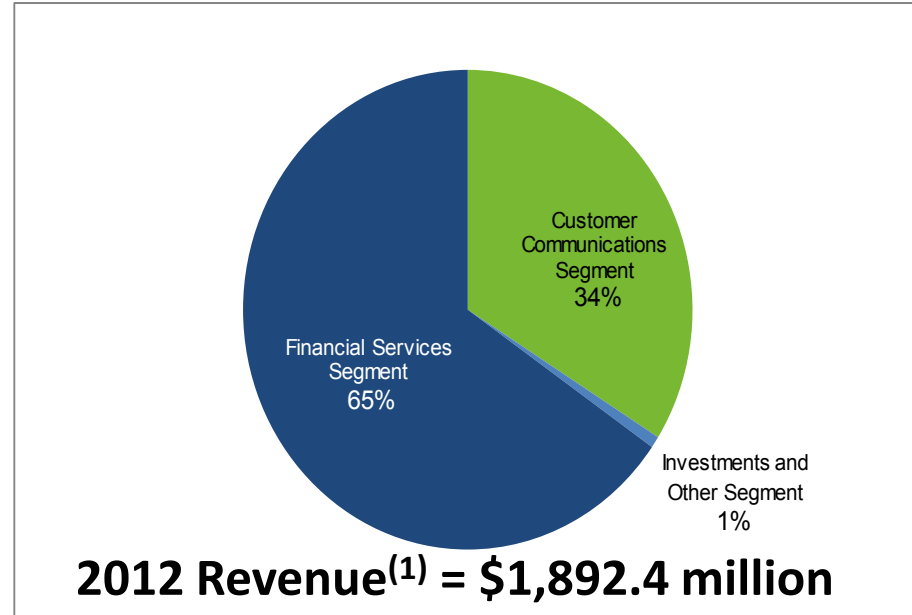
Our trademarks and service marks and those of third parties used in this presentation are the property of their respective owners.

DST Systems Overview

DST

DST Profile

- **Leading provider of technology and processing solutions**

- **Deep penetration in five key verticals**
  - **Mutual fund asset management**
  - **Retirement**
  - **Insurance**
  - **Brokerage**
  - **Healthcare**

- **Global leader in customer communications**

- **Headquartered in Kansas City, MO**

- **13,000 associates in 50 cities worldwide**

- **Listed on the New York Stock Exchange since 1995**

Financial Services Segment 65%

Customer Communications Segment 34%

Investments and Other Segment 1%

2012 Revenue[1] = $1,892.4 million

(1) Excludes out-of-pocket reimbursements and includes intersegment eliminations.

DST

DST's Key Vertical Markets

ASSET MANAGEMENT		**BROKERAGE**		**RETIREMENT**		**INSURANCE**	**HEALTHCARE**	
TA2000	75.6 M accts.	Subaccounting	12.4 M accts	TRAC	4.8 M accts	Percana	Healthcare Lives	22.5 M
Fast/iFast	19.5 M accts						Pharmacy Claims	362 M
		Subserveo		PlanSponsor Web			CareConnect	
Vision	140 K advisors	VisionPro		**myPLANnetwork**			CareAnalyzer	
FAN Mail		**Books and Records**		**RICC**				

Business Process Management	203,500 users
Customer Communications	11.1B images produced, 2.7B items mailed
Global IT Sourcing	Technology development, support, testing, processing and consulting
Hosting Solutions	Web, mainframe, content management and BPM hosting
Data Centers	Robust infrastructure, communications network, business continuity and disaster recovery
Business Process Outsourcing (a.k.a. "Full Service")	Cost-effective processing and infrastructure support

DST

Significant Position in Key Markets

Our size and market position fuel our ability to invest and remain market leaders

Transfer Agency	• Largest third-party provider of US mutual fund shareholder recordkeeping • Largest third-party processor in the UK
Retirement	• Leading provider of defined contribution recordkeeping • Second largest defined recordkeeping platform when current conversions are completed
Brokerage	• Vision/Vision HomeOffice are the most widely used web-based account aggregation sources for advisors and broker-dealers (140,000 active advisors) • Second largest subaccounting platform
Healthcare	• Providing health plan administration solutions and medical claims processing products since 1976 • Providing independent customized pharmacy benefit solutions since 1983
Insurance	• Currently providing services to 21 of the top 25 US life insurance carriers
Communications	• One of the largest First-Class™ mailers in US • World's largest inkjet systems printer with large scale regional facilities in Canada, the US and the UK

DST

Big Data

DST

What is Big Data?

Big Data is an **organic flow** of **structured and unstructured** data that can be organized into **related sets** from which **insight and understanding** can be derived by employing **powerful analytical tools** that **detect, surface, and ultimately visualize patterns**.
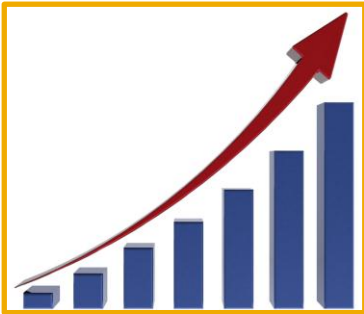
- Jason Thibeault

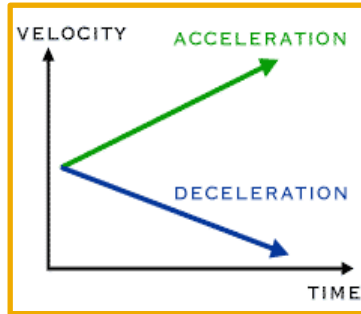**Source: http://www.rethinkeverythingblog.com/are-you-having-conversations-with-your-data/**

DST

Big Data Dimensions – the Four V's

| Volume	Velocity	Variety	Value
• Rows • Columns • Retention • History • Additional sources	• Rapid arrival • Fast processing • Quick Availability	• Structured • Unstructured • Social Feeds • Operational Data • Multimedia	Data is the asset

Realizing the Value with Big Data

Business Goal:

Data is the Asset

- New revenue sources
- Better answers and decisions from facts
- Customer targeting
- Influencing and predicting outcomes
- Determining next best action
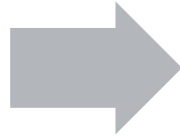
Technology Enabler:

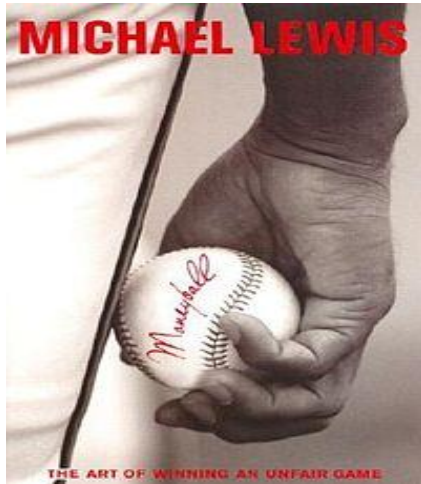Data Platform

- Faster data loading and analysis tools
- Consume information from more and varied sources
- Handling exponential growth
- Integrating into new and existing products

Value from Data Analytics

"**Moneyball is the true story of the Oakland Athletics baseball team that used analytics to assemble a competitive baseball team without the high salaries or revenues of other teams.**"

Value from Existing Unstructured Data

Source: Shahar Solomianik, Crowd Sourcing Sucked, So I Put Instagram API to Work Instead

Value from Visualization of Details

State

County

% of Vote
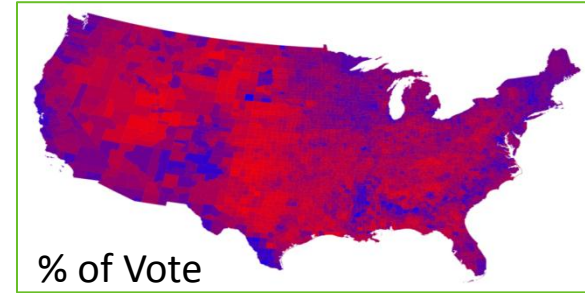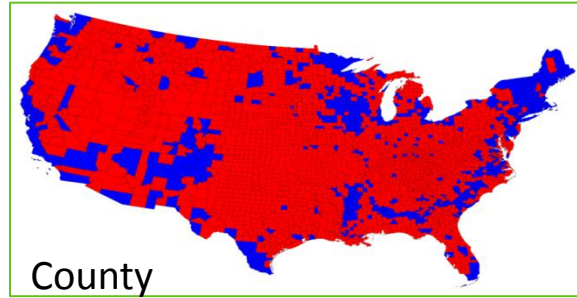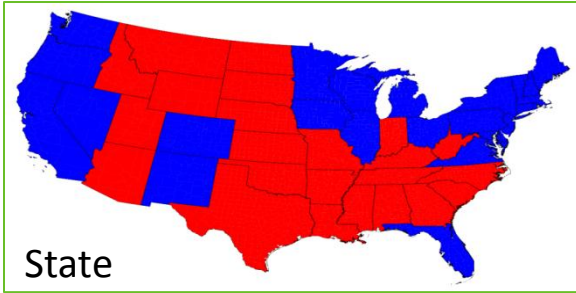
Source: Mark Newman, Department of Physics and Center for the Study of Complex Systems, University of Michigan

State	Obama	Romney	Johnson	Stein	Others	Total
Alabama	795,696	1,255,925	12,328	3,397	6,992	2,074,338
Alaska	122,640	164,676	7,392	2,917	2,870	300,495
Arizona	1,025,232	1,233,654	32,100	7,816	452	2,299,254
Arkansas	394,409	647,744	16,276	9,305	1,734	1,069,468

Value from Social Media
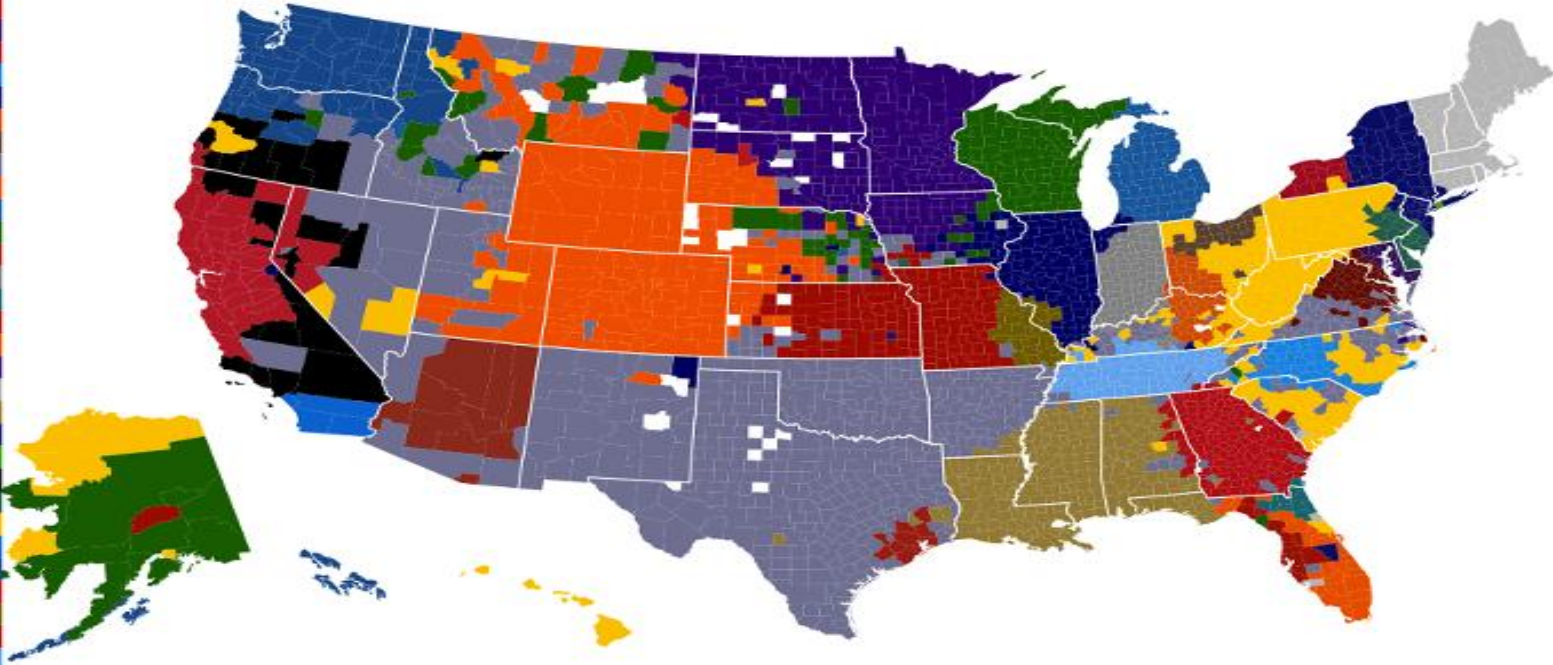
Source: <u>NFL Fans on Facebook</u>

Relationships, mood, sentiment, conversations, …

Value for Next Best Action

Relationship Status change
→ Joint Accounts?

Job Title / Company change →
Rollover IRA, 401k?

Tweet about bad product
→ Customer Service Call

Completed Crib Warranty
→ 529 account?

Big Data Business Model Maturation Index

Measures the degree to which the organization has integrated big data and advanced analytics into their business model

Business Metamorphosis

Data Monetization

Business Optimization

Business Insights

Business Monitoring

Monitoring business performance to flag areas of interest

Integrate insights & recommendations into existing business processes

Embed analytics to optimize select business processes

Leverage insights to identify new revenue opportunities

Transform customer and product insights to move into new markets

EMC²

Source: Bill Schmarzo, CTO, EIM&A Practice, EMC Consulting

|D|S|T|

Big Data Strategy

A platform to leverage data as an asset

Our Principles

1. **Capacity to process large volume & varieties of data**
2. **Cost effective and easy access**
3. **Explore, connect & present with visual & analytical tools**
4. **Enhance data with external sources**
5. **Provide training and skills**
6. **Ensure contract, audit, regulatory & privacy controls**

Challenges

Mix of Quantitative Analytical Skill with
Business Subject Matter Expertise

Compliance, ethics
and regulatory

How do we value our data
assets?

Analytics

Compliance

Business Valuation

Technical Skills
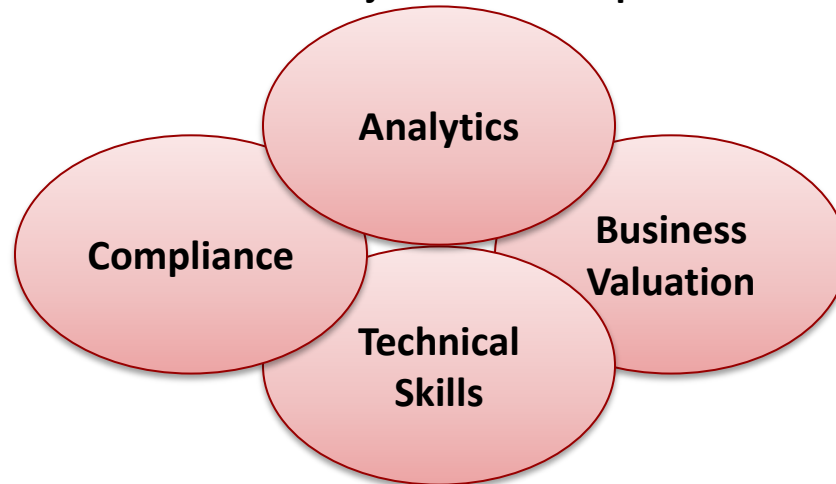
Up-to-date skills on emerging technology
and techniques

DST

Big Data Hype

"Through 2015, more than 85 percent of Fortune 500 organizations will fail to effectively exploit big data for competitive advantage" - Gartner

Who's Really Using Big Data?

High Expectations

- 85% of organizations reported that they have Big Data initiatives planned or in progress.
- 70% report that these initiatives are enterprise-driven.
- 85% of the initiatives are sponsored by a C-level executive or the head of a line of business.
- 75% expect an impact across multiple lines of business.
- 80% believe that initiatives will cross multiple lines of business or functions.

Capabilities Gap

- 15% of respondents ranked their access to data today as adequate or world-class.
- 21% of respondents ranked their analytic capabilities as adequate or world-class.
- 17% of respondents ranked their ability to use data and analytics to transform their business as more than more than adequate or world-class.

Source: HBR Blog, Who's Really Using Big Data, Paul Barth and Randy Bean

DST

Big Data Landscape

Vertical Apps
PREDICTIVE POLICING
bloomreach MYRRIX

Log Data Apps
splunk> loggly sumologic

Ad/Media Apps
rocketfuel collective[i]
bluefin Recorded Future
Media Science TURN LUCKYSORT DataXu

Data As A Service
factual kaggle knoema beta
GNIP DATASIFT Windows Azure Marketplace INRIX LexisNexis SPACE CURVE LOQATE

Business Intelligence
ORACLE | Hyperion
SAP Business Objects RJMetrics
Microsoft | Business Intelligence
IBM COGNOS birst
MicroStrategy
Autonomy bime DOMO
QlikView
Chart.io GoodData

Analytics and Visualization
tableau Palantir
OPERA metaLayer dataspora
METAMARKETS centrifuge
TERADATA ASTER
SAS TIBCO KARMASPHERE
panopticon
Real-Time Visual Data Analysis pentaho
Datameer
platfora ClearStory CIRRO
alteryx visual.ly AYATA

Analytics Infrastructure
Hortonworks VERTICA An HP Company MAPR
cloudera INFOBRIGHT PARACCEL
EMC² GREENPLUM
NETEZZA kognitio EXASOL DATASTAX calpont

Operational Infrastructure
COUCHBASE 10gen MongoDB company
TERADATA HADAPT
TERRACOTTA VoltDB
MarkLogic INFORMATICA

Infrastructure As A Service
amazon web services
Windows Azure
infochimps
Google BigQuery

Structured Databases
ORACLE MySQL
Microsoft SQL Server PostgreSQL
IBM DB2 SYBASE
memsql

Technologies
hadoop hadoop MapReduce mahout APACHE HBASE Cassandra

dave@vcdave.com

blogs.forbes.com/davefeinleib

DST

Challenges
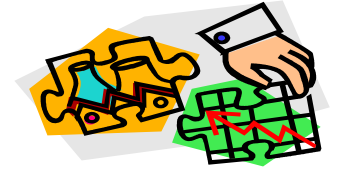
End User Informal Exploration and Analysis

Trusting fact based decisions instead of "gut" intuitive

Finding resources with

- Data mining and statistical experience (Data Scientist)
- A blend of Java, Functional and ETL Programming
- Distributed Cluster Computing experience

Evolving Open Source and Commercial Tools

Different Data Analysis Practices

❮ **Magnetic**

- ▪ Old: Repel new data sources until cleansed and integrated
- ▪ New: Attract new data sources regardless of format and quality niceties

❮ **Agile**

- ▪ Old: Long-Range, careful design and planning of changes
- ▪ New: Continuous Rapid Evolution

❮ **Deep**

- ▪ Old: Traditional Rollups and Drilldowns, Extracts, Samples
- ▪ New: All Data, Embedded statistical methods

Source: http://db.cs.berkeley.edu/jmh/papers/madskills-032009.pdf

Business Intelligence Changes

Expanded Questions

❱ **What are customers saying and feeling? Happy, Sad, Satisfied, Engaged, Confident**

❱ **What are customers doing? Clickstream Analysis, Fraud or other pattern detection**

❱ **How can we characterize and group entities? Market Segmentation, Risk Categories**

❱ **What are customers likely to do? Buy, Sell, Default**

❱ **What could we change to influence customers? Behavior influence**

❱ **What is changing in our environment? Thresholds, Alerts, Monitoring, Comparisons**

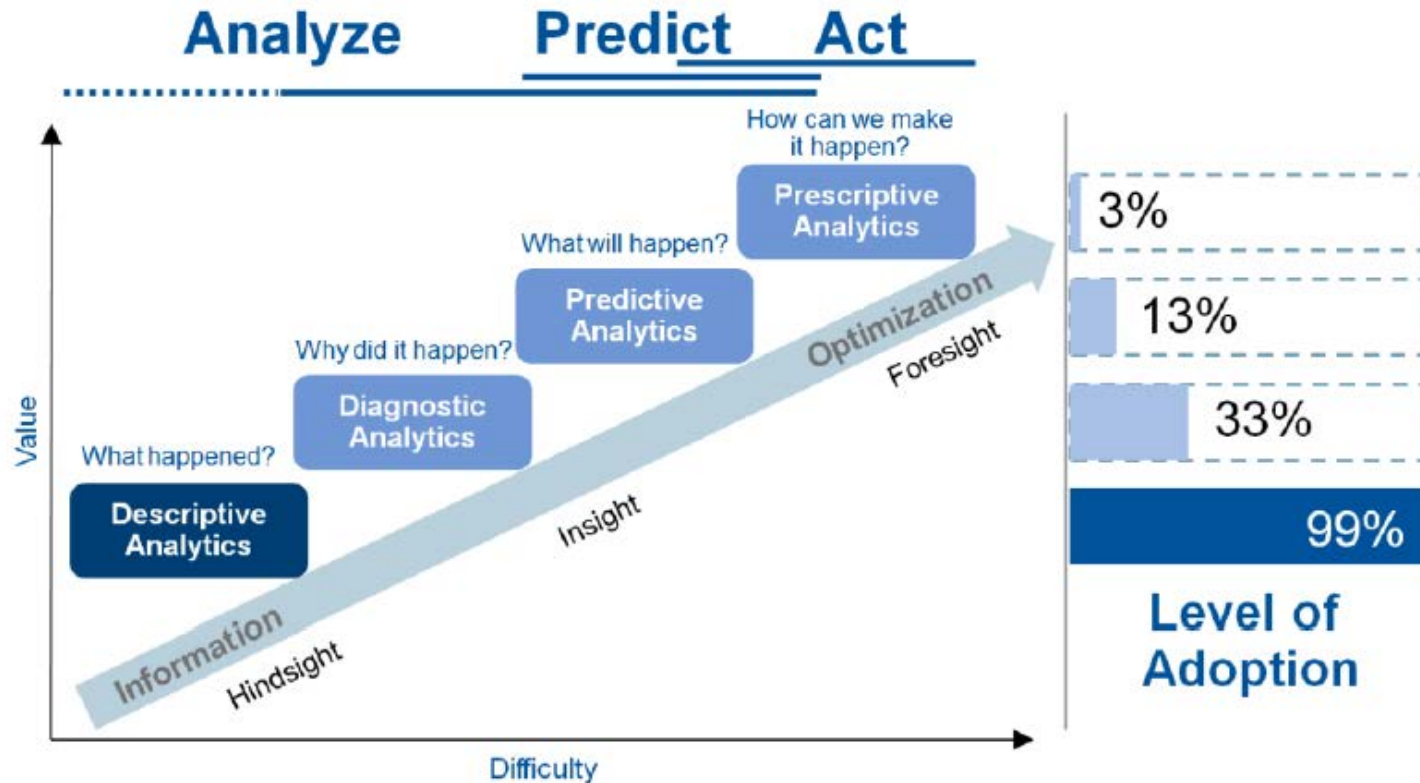❱ **What artifacts and events are most relevant to our requests? Importance of past**

Expanded Requirements

❱ **Massive volumes**

❱ **Mining text and other unstructured sources**

❱ **Specialized processing – natural language, statistical analysis**

Capabilities Allowing Exploration

❱ **Increased processing power via lost cost parallelization allowing exploration of entire, not just sample, data sets**

❱ **Less formal and schema-less data management**

Source: Gartner, Big Data Means Big Changes for Business Intelligence

DST

Adoption of the Analytics Continuum

Source: Magic Quadrant Survey, 2012

25

Prescriptive Analytics Are the "Final Frontier" of Analytic Capabilities

| What Happened, and Why? | What Will Happen? | What Should I Do? |

Information   **Insights**   **Decisions**

For Example

- Who are my most valuable customers?
- Which customers are most likely to respond to my next offer?
- Which offer should I make to each customer?

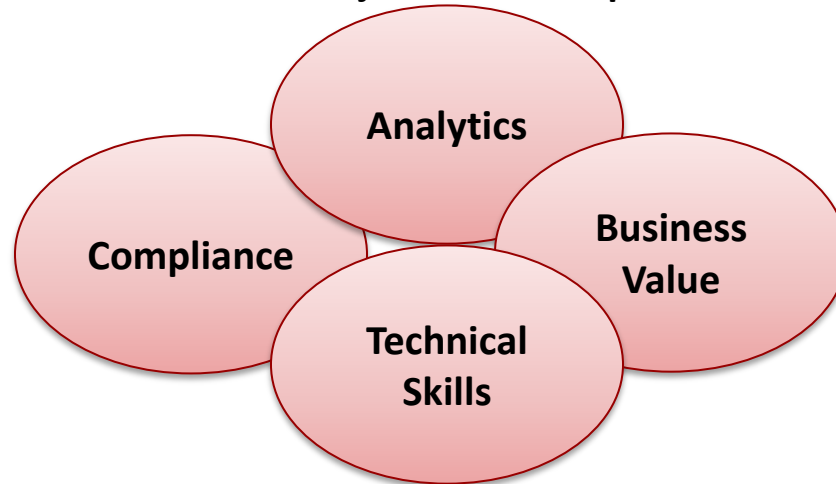Descriptive & Diagnostic Analytics   Predictive Analytics   Prescriptive Analytics

But What Exactly Are They?

Gartner.

Challenges

Mix of Quantitative Analytical Skill with Business Subject Matter Expertise

Compliance, ethics and regulatory

How do we value our data assets?

Analytics

Compliance

Business Value

Technical Skills

Up-to-date skills on emerging technology and new releases