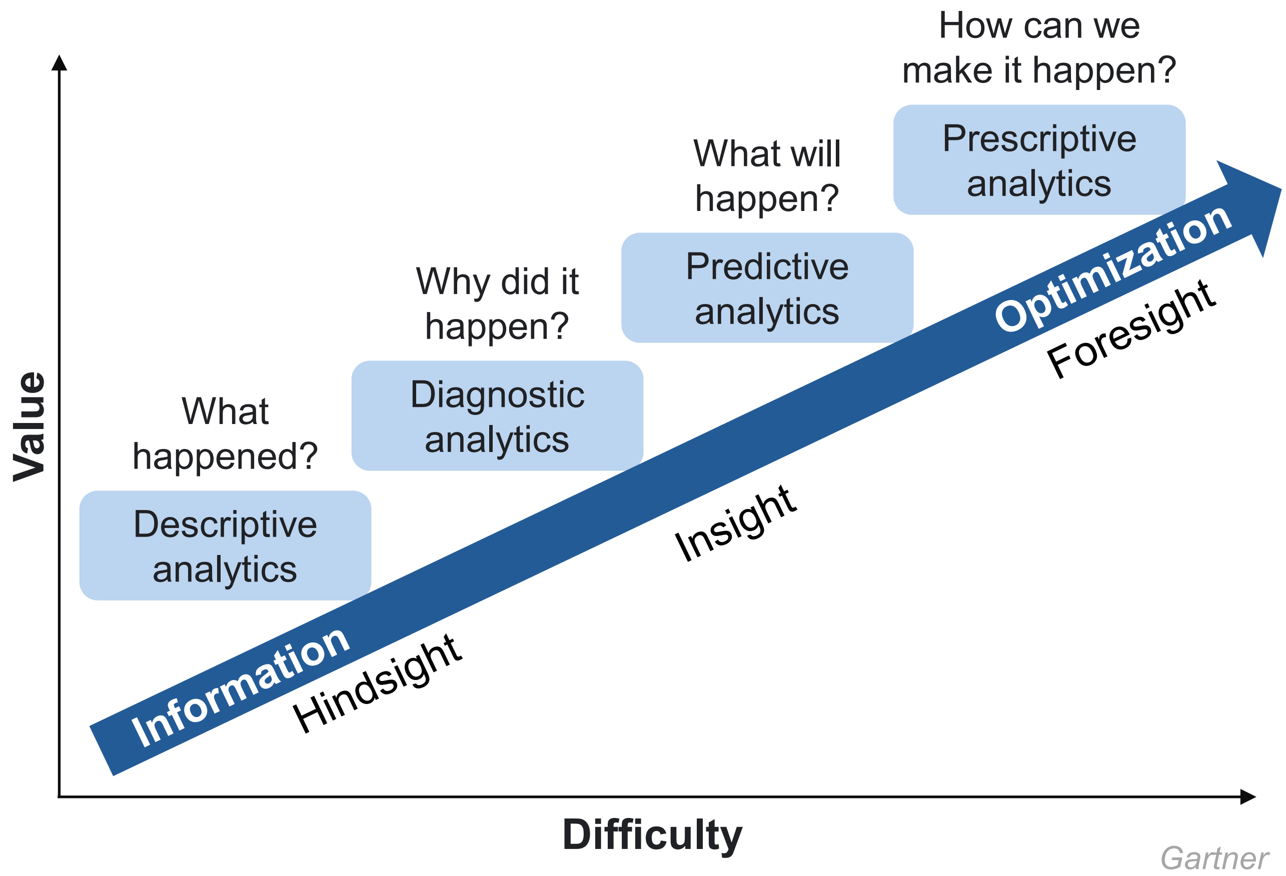
Hierarchy

Gartner and Davenport

Doug Gray

Gartner’s Analytics Hierarchy



Davenport’s Analytics Hierarchy p. 27

DataScience@SMU

Competing on Analytics

Doug Gray

Competing on Analytics: The New Science of Winning 2017 Edition

- ***Part I: Nature of Analytical Competition***
- ***Part II: Building an Analytical Capability***
- Analytics as a critical component of competitive strategy
  - Thinking, behaving, and executing analytically
- Applying analytics methods and techniques
  - People, process, and technology

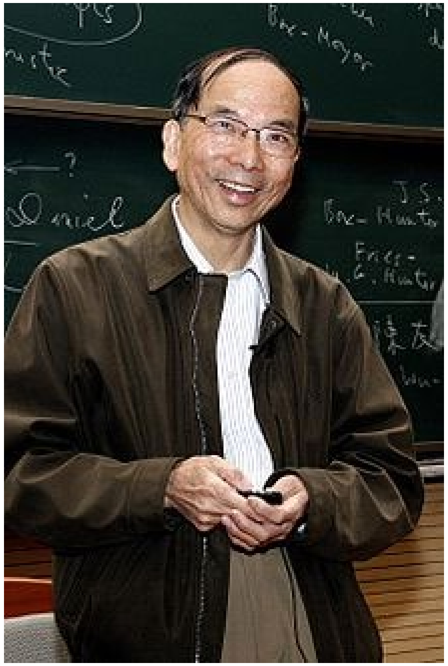
DataScience@SMU

Buzzword Compliance

Doug Gray

Buzzword Compliance

- New school
  - Analytics (De-Di-Pd-Ps) vs. business intelligence (De-Di)
  - Data science (machine learning, deep learning, AI)
  - Big data (unstructured data: text, audio, video, sensor, social media)
- Old school
  - Operations research, management science, industrial engineering
  - Quantitative (business) analysis
  - Mathematics, probability, statistics, forecasting
  - Computer science, machine learning, expert systems, artificial intelligence (AI)

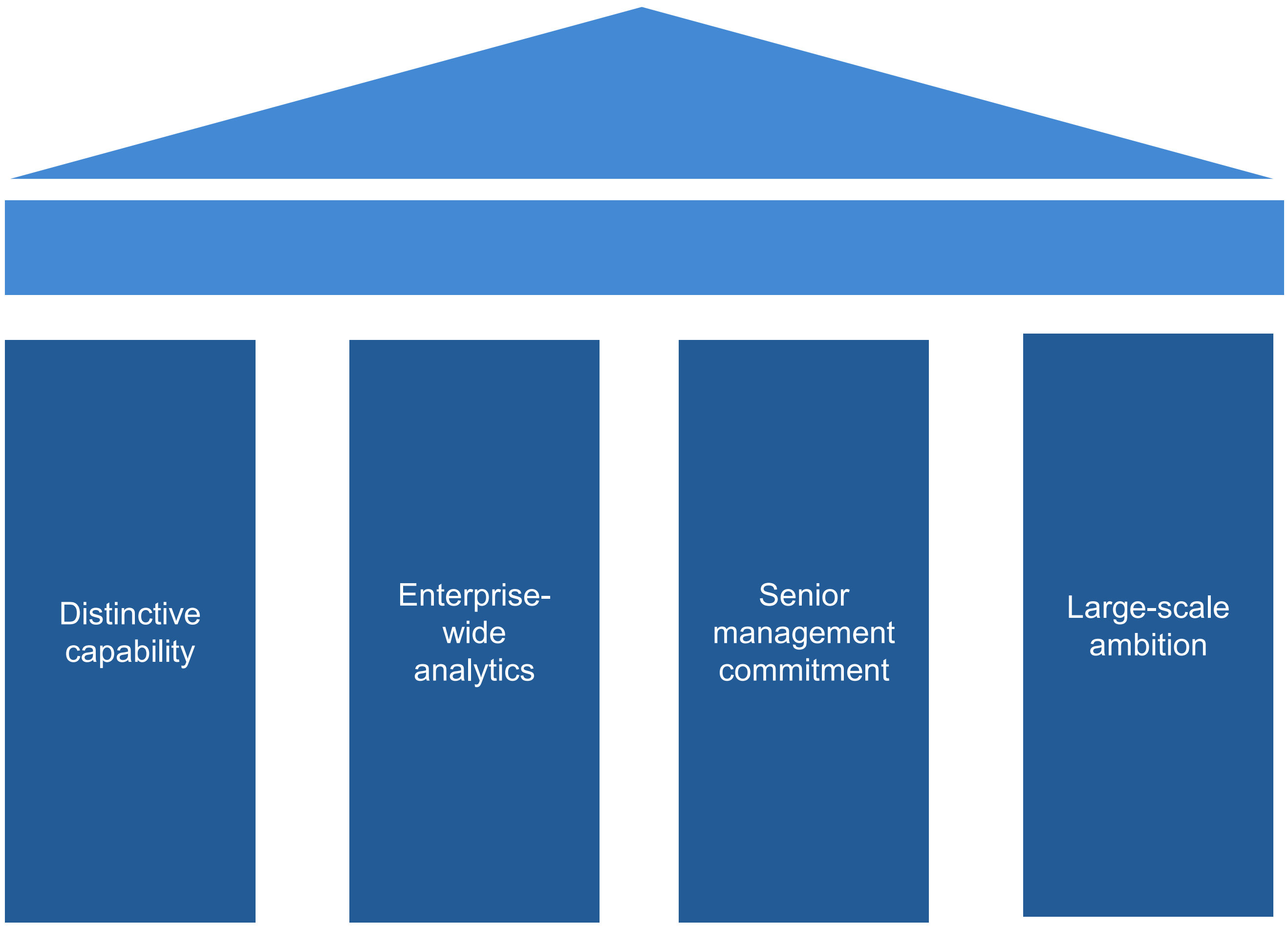


DataScience@SMU

Four Pillars of Analytical Competition

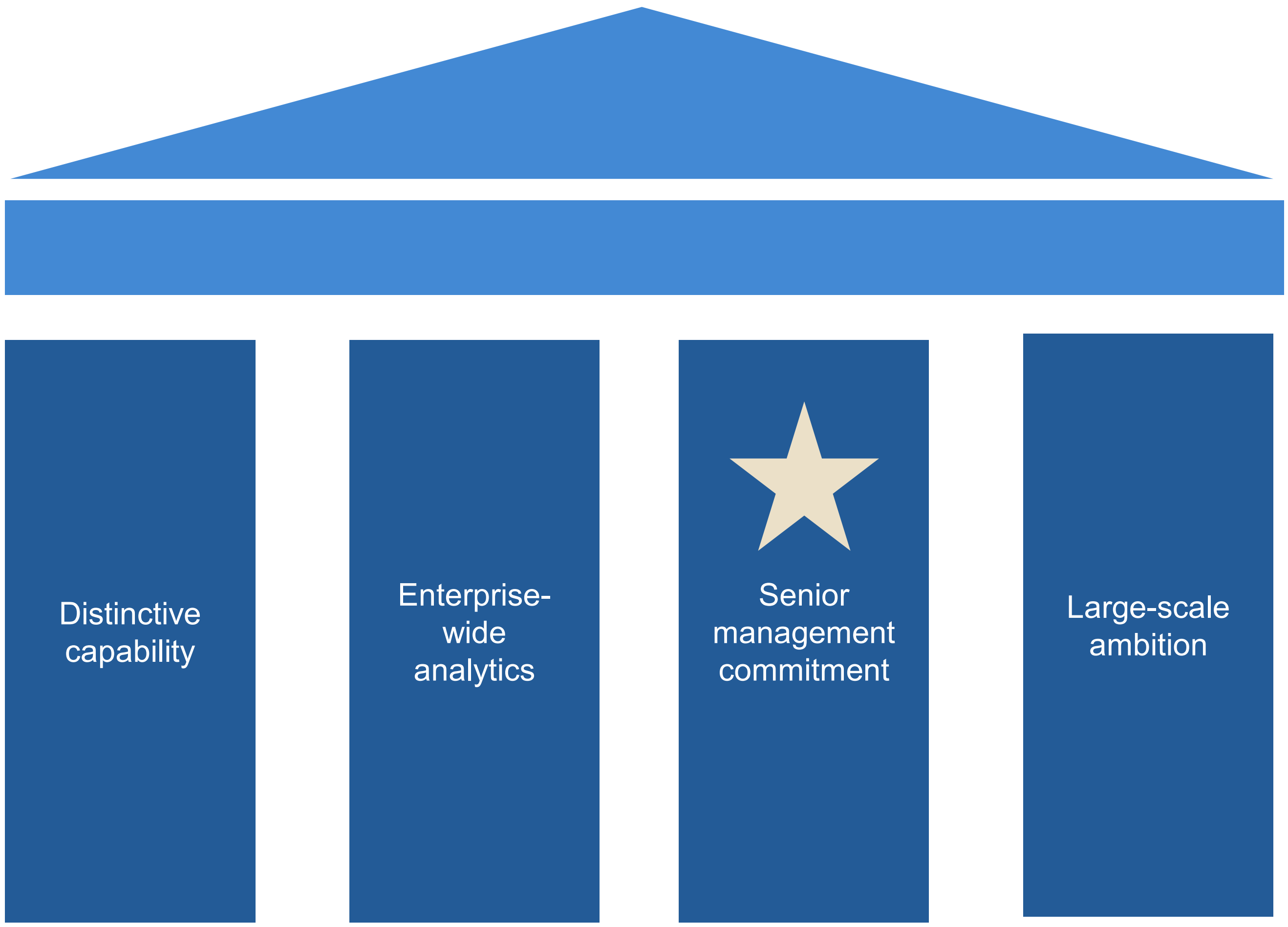
Doug Gray

Four Pillars of Analytical Competition



Source: ***Competing on Analytics: The New Science of Winning***, T. Davenport, J. Harris p. 60

Four Pillars of Analytical Competition



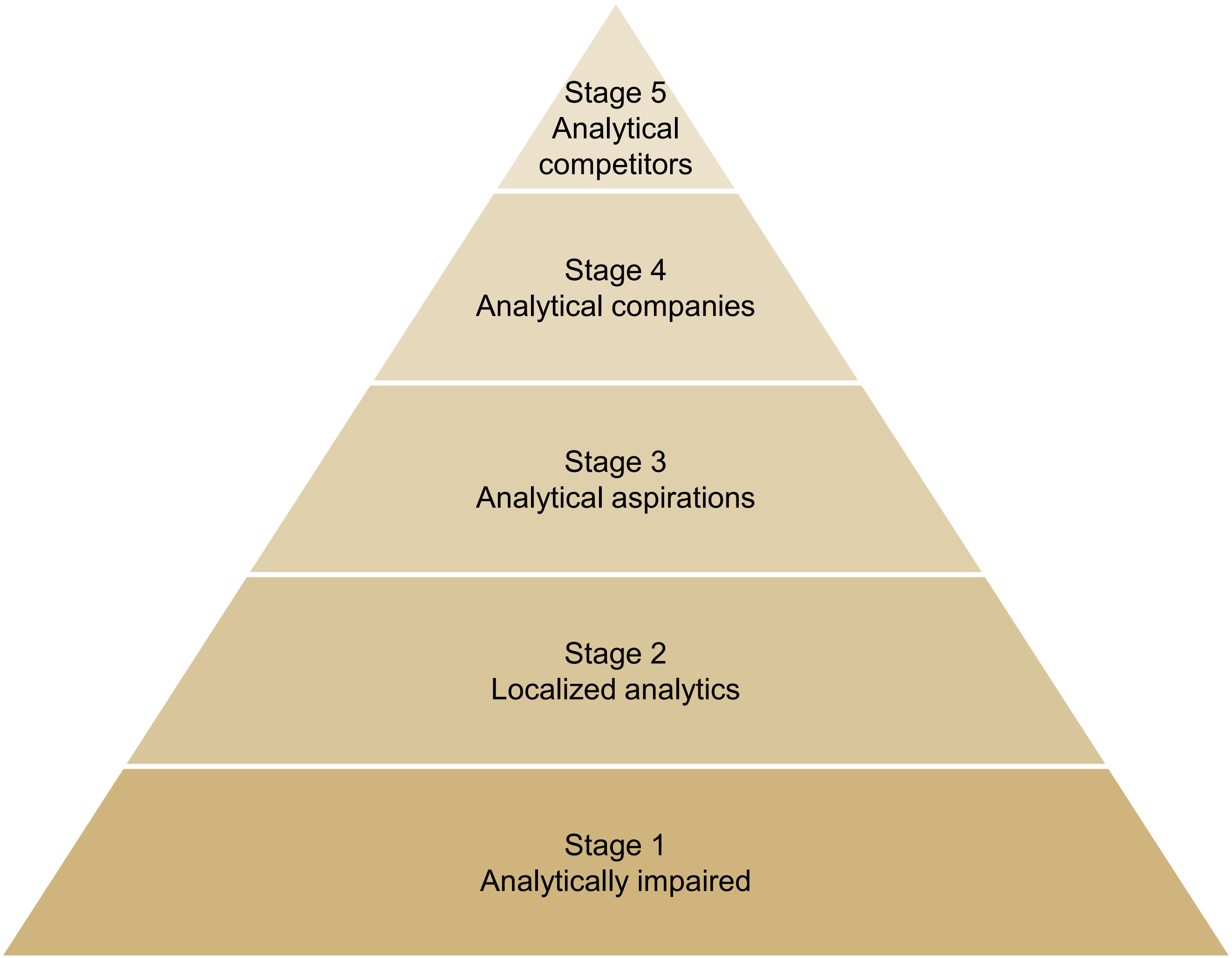
Source: ***Competing on Analytics: The New Science of Winning***, T. Davenport, J. Harris p. 60

DataScience@SMU

Five Stages of Analytical Competition

Doug Gray

Five Stages of Analytical Competition



Source: ***Competing on Analytics: The New Science of Winning***, T. Davenport, J. Harris p. 61, p. 62 (Attributes by Stage)

DataScience@SMU

DELTA

Doug Gray

DataScience@SMU

D-E-L-T-A (Original: 2007)

- **D**ata
- **E**nterprise
- **L**eadership
- **T**argets
- **A**nalysts

Source: Competing on Analytics: The New Science of Winning, T. Davenport, J. Harris pp. 46-47

D-E-L-T-T-A-A (Evolved: 2017)

- **D**ata
- **E**nterprise
- **L**eadership
- **T**argets
- ***Technologies***
- ***Analytical Techniques***
- **A**nalysts

Source: Competing on Analytics: The New Science of Winning, T. Davenport, J. Harris p. 46-47, DELTTAA Model Analytical Capabilities by Stage pp. 182-183

DataScience@SMU

Internal Factors

Doug Gray

DataScience@SMU

Internal Factors

1. Leadership (executives, champions)
2. Data (or lack thereof)
3. Culture, including politics
   - Data-driven (“let the data speak”) vs. gut/experience-driven (“HiPPO-driven”)
     - **Hi**-**P**-**P**-**O** theory: **h**ighest **p**aid **p**erson’s **o**pinion
   - Openness/ability to change, adaptive, challenge the *status quo*
   - Proactive vs. reactive, action-oriented vs. resistant to new ideas/change
   - Incentive and reward structures
   - Continuous improvement (Kaizen), creative destruction
   - Seeing analytics as a means to ***creative competitive advantage***, not “number crunching”

DataScience@SMU

Corporate Strategy

Doug Gray

Corporate Strategy: Enterprise Orientation and Targets (Core Business Model Drivers)

- Industry structure, value chain position, competition
- Strategy, tactics, operations; problems, decisions, opportunities
  - One-time, repeatable-frequency, embedded in business processes
- Performance drivers: revenue drivers, cost drivers, and profit drivers
  - How does the firm make money? How analytics can affect/improve the free cash flow, earnings/EPS, profitable growth
- Customer experience, satisfaction, loyalty drivers (NPS)
- Market share drivers (pricing, quality vs. low cost)

DataScience@SMU

Road Map to Analytical Capabilities

Doug Gray

A Roadmap to Enhanced Analytical Capabilities

- *Prerequisites*: stage one—What is lacking? D-E-L-T-T-A-A
- *Assessing analytical capabilities:* D-E-L-T-T-A-A
  - *Organization—human—technology* (key elements p. 161)
- *Choosing a strategic focus or target* (e.g., Southwest)
- *Choosing a path*
  - *Full steam ahead vs. prove-it* (attributes pp. 164-7) *Stage two: prove-it detour*
- *Stage three*: *triggered by executive sponsorship*
- *Stage four: building world-class analytical capabilities at the enterprise level*
- *Stage five: analytics move from being a very important capability to the key to strategy and competitive advantage* D-E-L-T-T-A-A in-depth pp. 177-183) **(*self-assessment exercise*)**

Source: Competing on Analytics: The New Science of Winning, T. Davenport, J. Harris Chapter 6 - p. 158

A Roadmap to Enhanced Analytical Capabilities

- *Managing for outcomes (pp. 181-184)*
  - *Employee behavior: executives, middle management, front-line*
  - *Processes and programs: incorporating analytics into (automated) processes*
  - *Products and services*
  - *Financial results*
- *Establishing priorities (pp. 184–185)*
- *Avoiding potholes (pp. 185–186)*

Source: Competing on Analytics: The New Science of Winning, T. Davenport, J. Harris Chapter 6 - p. 158

DataScience@SMU

Davenport

Doug Gray

Organizational Model Types (Davenport)*

Organizational Model Type	Description	Pros	Cons	Rating
Decentralized	Analyst groups reside inside business units without any corporate or consolidating structure; most prevalent, but least effective, model which reflects immaturity of most corporations' analytical capabilities	Effective only in the rare case of a large, diversified, multi-business corporation in which business units have little in common, e.g., GE, Samsung, Lockheed-Martin	Difficult to set enterprise priorities and to develop (retain) and deploy staff effectively through borrowing and rotation	**Low**
Consulting	All analysts are part of one organization, but instead of deploying analysts to business unit projects, business units "hire" analysts as consultants to their analytical projects	Market-driven; consolidating analysts enables enterprise-wide view of what's going on; consultants educate and advise customers how to utilize services; makes market demand smarter	***Falters under weak enterprise focus, poor executive leadership, or faulty business value targeting mechanisms; analysts work on whatever project BU chooses to pay for rather than delivering the most business value ("squeaky wheel gets the oil")***	***Medium*** ***Acceptable to Davenport***
Centralized	All analyst groups report to one corporate organization, even if they are assigned to different business units	Easier to deploy analysts on projects with strategic priority	Can create distance between analysts and the business, especially if analysts are all housed in the corporate location	**Low**
Community of Practice or Center of Excellence "Federated"	Analyst groups are decentralized and reside in every major business unit that has an appetite for analytics, but all groups are members of (and perhaps report dotted line to) a corporate COP or COE for analytics	***Builds a community of analysts who can learn from each other by sharing experiences and best practices; COP/COE doubles as a program office looking across analytics initiatives, advising on priorities and staffing***	Limitations of centralization must be overcome by holding regular conferences, knowledge sharing through online portals	***High*** ***Recommended by Davenport***

* *Source: Analytics at Work: Smarter Decisions, Better Results*, T. Davenport, J. Harris, pp. 104-109

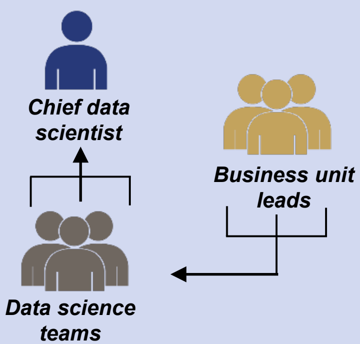
DataScience@SMU

Booz Allen Hamilton

Doug Gray

Organizational Model Types (Booz | Allen | Hamilton)*

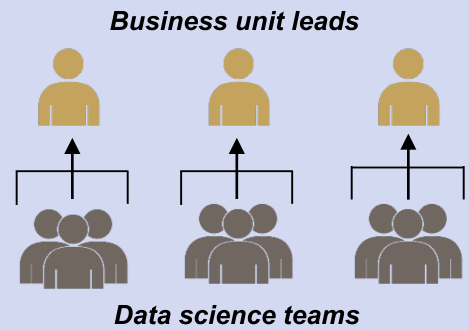
Centralized model



Business units bring their problems to a centralized data science team, overseen by a chief data scientist.

Centralized data science teams serve the entire organization but report to a chief data scientist, who decides which projects the teams will work on, and how to manage the projects. Business units work with the data science teams to solve specific challenges.

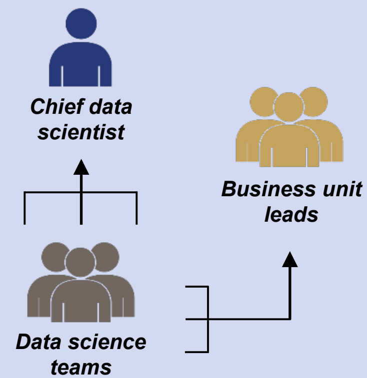
Decentralized model



Data science teams are fully embedded in business units and report to individual business unit leaders.

Diffused, or decentralized, data science teams are fully embedded in business units such as marketing, research and development, operations, and logistics. The teams report to individual business unit leaders and perform work under their leadership.

Deployed model



Data science teams are overseen by a chief data scientist and forward deploy to business units.

As with the diffused model, data science teams are embedded in the business units. The difference is that the embedded teams in the deployed model report to a single chief data scientist as opposed to business unit leaders. In this model, also called the matrixed approach, teams are generally assigned to individual business units, though they are sometimes also assigned to broader product lines, or to mission sets comprised of members from several business units.

* *Source:* ***Tips for Building a Data Science Capability***, Booz | Allen | Hamilton

DataScience@SMU

Leaders

Doug Gray

Analytical Leaders

- Passionate believers in analytical, fact-based, data-driven decision making
- Appreciative of analytical tools and methods
- Willing to act on the results of analyses
- Willing to manage a meritocracy
- Analytical leaders are there from the start (Bezos), enter (Loveman) or emerge (Kraft, Gallo)
- CEO, CFO, CIO, and CDAO

Source: Competing on Analytics: The New Science of Winning, T. Davenport, J. Harris pp. 192-198

DataScience@SMU

Size

Doug Gray

Analytics Organization Size

- Amazon (1,200 supply chain)
- Google (500)
- Ford (900; 300 with PhDs)
- GE (200)
- P&G (200)
- Southwest Airlines (200 in the community; less than 50 doing advanced analytics)
- Analytics professionals (and amateurs) vs. (citizen) data scientists (pp. 205, 206)

Source: *Competing on Analytics: The New Science of Winning*, T. Davenport, J. Harris p. 202

DataScience@SMU

Workers and Decision Makers

Doug Gray

DataScience@SMU

Analytical Information Workers and Decision Makers

- Experimental
- Numerate
- Data literate
- Tools: from MS Excel to Alteryx to autonomous decision-making
- Overrides
- Interactive optimization and analytics: human, computer, and model-algorithm

Source: Competing on Analytics: The New Science of Winning, T. Davenport, J. Harris pp. 210-211

DataScience@SMU

Critical Success Factors

Doug Gray

DataScience@SMU

Critical Success Factors

- Building a sustainable pipeline of projects
- Relationship to IT
- Governance and funding
- Managing politics and change
- Don't get ahead of the end users
- In-house vs. outsourced vs. offshore

Source: *Competing on Analytics: The New Science of Winning*, T. Davenport, J. Harris p. 205

DataScience@SMU

FACE Introduction

Doug Gray

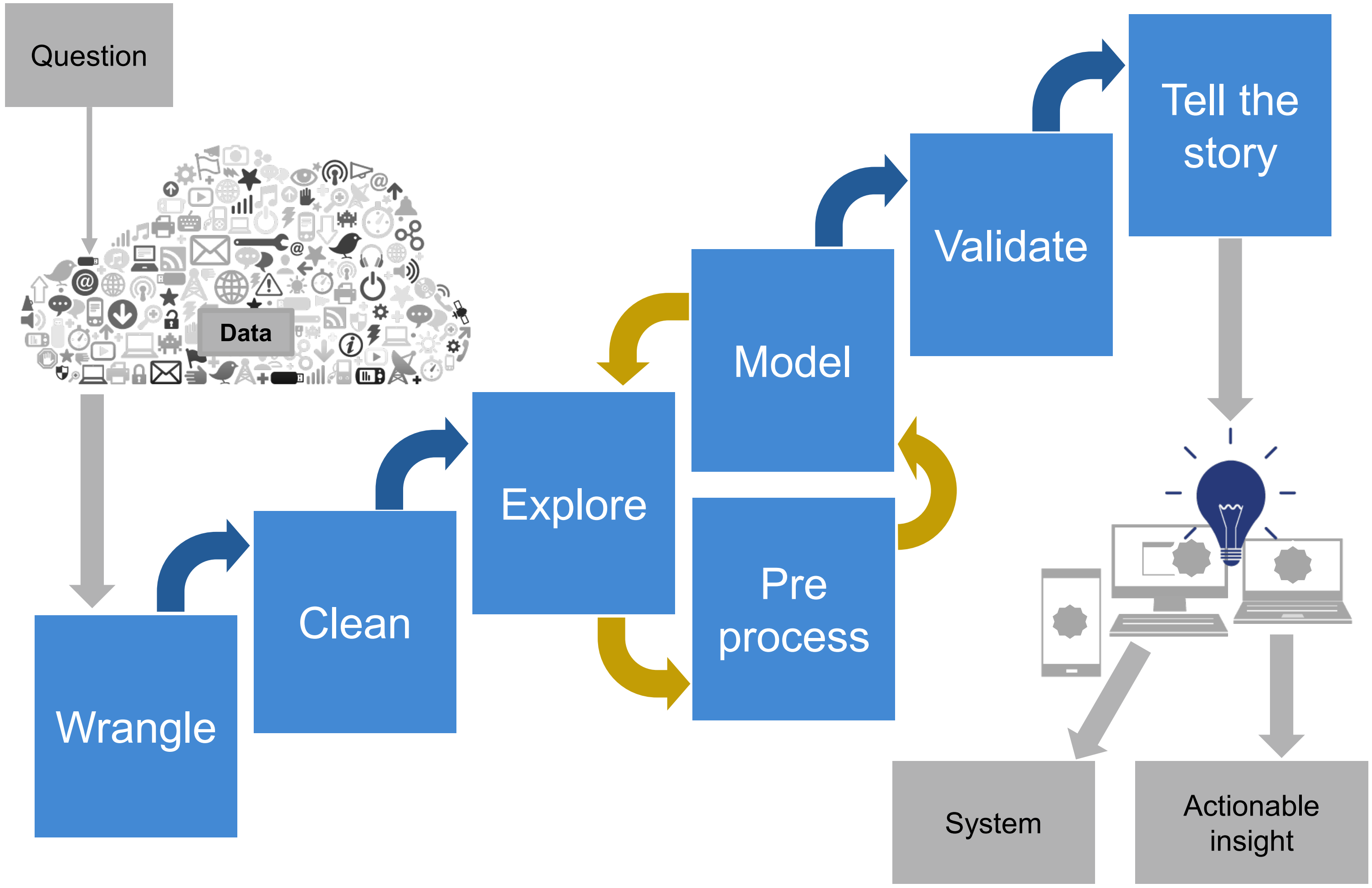
F-A-C-E

- **F**rame the problem/decision/question
- **A**nalytically “model” and solve for the problem/decision/question
- **C**ommunicate and *act* on/implement the results
- ***E****mbed models in enterprise business processes and systems*

FACE

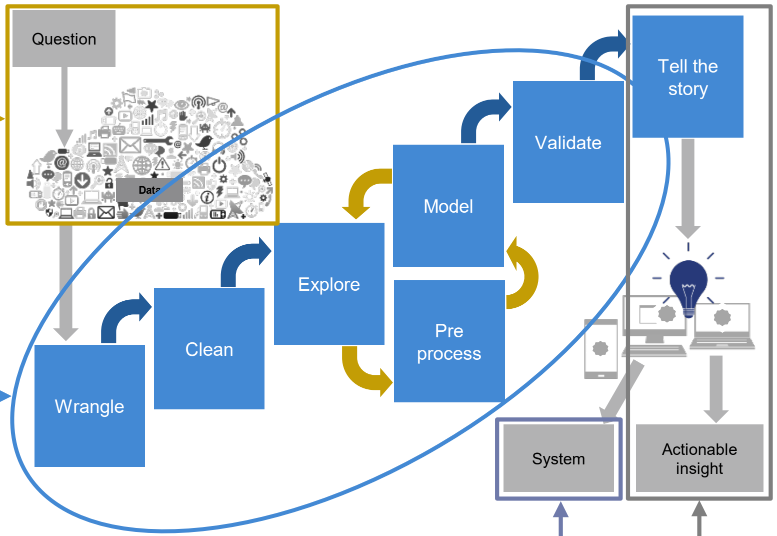
- Framing the problem
  - Problem recognition
  - Review of previous findings
  - Modeling approach
- Analysis/solving the problem
  - Data collection
  - Modeling
  - Data analysis
- Communicating and acting on results
  - Results presentation and action, i.e., “telling a story”
  - Creating impactful visualization
- Embedding final models and methods in enterprise business processes and systems

Data Science Pipeline



FACE

- Framing the problem
  - Problem recognition
  - Review of previous findings
  - Modeling approach
- Analysis/solving the problem
  - Data collection
  - Modeling
  - Data analysis
- Communicating and acting on results
  - Results presentation and action, i.e., "telling a story"
  - Creating impactful visualization
- Embedding final models and methods in enterprise business processes and systems



DataScience@SMU

Frame the Problem

Doug Gray

FACE

Framing the problem

- Problem recognition
  - *Identify stakeholders*
  - *Focus on decisions*
  - *Storytelling*
  - *Deterministic vs. stochastic*

FACE: Deterministic vs. Stochastic (Probabilistic)

Deterministic

- Capacity of an oil/gas pipeline
- Capacity of an aircraft
- Aircraft taxi time (with no delay)
- Other examples?

Stochastic

- Demand for gasoline over time in various geographical markets
- Demand for seats over time in various O/D markets, fare classes
- Aircraft taxi times (with delays)
- Other examples?

FACE

Framing the problem

- Problem recognition
  - *Identify stakeholders*
  - *Focus on decisions*
  - *Storytelling*
  - *Deterministic vs. stochastic*
  - *Normative vs. evaluative*

FACE: Normative vs. Evaluative

Normative

- Optimizing some objective
  - Minimize cost
  - Maximize revenue or profit
  - Maximize flow
- Subject to constraints
  - Capacity
  - Service level
  - Raw material supply
  - Time windows

Evaluative

- *What if*…scenario analysis
- System (discrete-event or continuous) simulation

FACE

- Framing the problem
  - Review of previous findings
    - Google and the Internet are your friends
    - Consult with analysts in your company, and others (noncompetitors)
    - Analytics conferences
    - Consultants (see vendor list)
    - University professors and students

FACE

- Framing the problem
  - Modeling approach
    - What type of variables are involved (explanatory and response)?
      - Categorical
        - Binary (dichotomous) example: {0,1}
        - Multinomial example: {blue, red, purple}
        - Ordinal example: Likert scale—{disagree, neutral, agree}
      - Numerical/quantitative example: length of delay, profit in dollars, etc.

FACE

- Framing the problem
  - Modeling approach
    - What are we using the models for? (Prediction? Interpretation? Both?)
    - Which models do we posses the technological and practical expertise to implement?

DataScience@SMU

Solving the Problem

Data Collection

Doug Gray

DataScience@SMU

F**A**CE

Analysis/solving the problem

- Data collection
  - Primary vs. secondary
  - Structured vs. unstructured (i.e., CSV vs. images vs. text)
  - Format (JSON, XML, CSV, etc.)
  - Source (API, proprietary, third party…next slide!)

F**A**CE: Third Party Consumer Data Sources

- Experian
- Equifax
- ACXIOM
  - Personicx customer segmentation
- FICO
- KBM Group
- Government data, e.g., BLS

DataScience@SMU

Solving the Problem

Modeling

Doug Gray

F**A**CE

Analysis/solving the problem

- Modeling
  - Applying the scientific method
  - Pragmatically
    - *“All models are wrong, but some are useful” - George Box*
    - *“ OR is the art and science of giving bad answers to questions that would otherwise be given far worse answers”*
    - *“The forecast is always wrong; the question is by how much”*
  - ***Type of model/method…next few slides!***

DataScience@SMU

Modeling

Deterministic

Doug Gray

F**A**CE: Deterministic vs. Stochastic (Statistical)

Deterministic

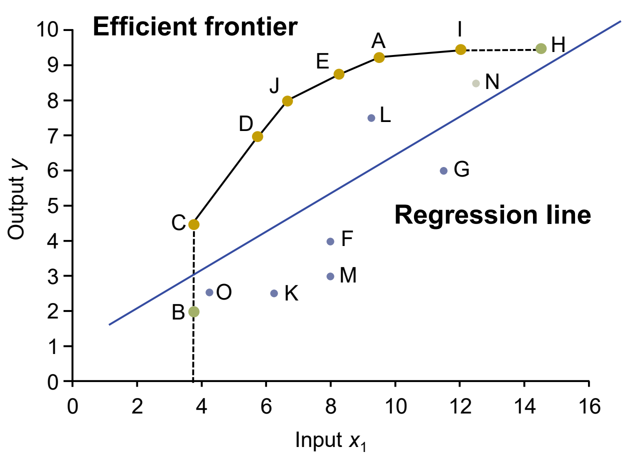
- Mathematical programming
  - Linear programming (LP)
    - DEA (data envelopment analysis)
  - Integer programming (IP)
  - Mixed integer programming (MIP)
  - Nonlinear programming (NLP)
    - Quadratic programming (QP)
  - Constraint programming (CP)

FACE: Linear Programming Model

The “Swiss Army Knife” of deterministic optimization modeling

Maybe data analysis section

$$
\begin{array}{ll}
\text{maximize} & \mathbf{c}^\top \mathbf{x} \\
\text{subject to} & A\mathbf{x} \leq \mathbf{b} \\
\text{and} & \mathbf{x} \geq \mathbf{0}
\end{array}
$$



Decision making units (DMU) = {A,B,C,D,E,F,G,H,I,J,K,L,M,N}

Data envelopment analysis (DEA)

Modeling maximization (or minimization) of an objective function subject to constraints

- Revenue s.t. capacity
- Network flow s.t. capacity
- Cost s.t. customer demand

Solve using MS Excel Solver (Simplex) for linear programming

DataScience@SMU

Modeling

Stochastic/Statistical

Doug Gray

F**A**CE: Deterministic vs. Stochastic (Statistical)

Stochastic/statistical

- Binomial models
- Markov process models
  - Poisson arrivals
  - Exponential service times
  - Queueing models (M/M/n)

Stochastic/statistical

- Classification models
  - Logistic regression
  - Random forest
  - KNN
  - SVM
  - LDA/QDA

F**A**CE: Deterministic vs. Stochastic (Statistical)

Stochastic/statistical

- Probability models
  - Normal distribution
  - Bayesian inference
  - Latent class analysis

Stochastic/statistical

- Continuous response regression analysis
  - Simple/multiple linear
  - Nonlinear
  - Partial least squares
  - Lasso/Ridge/Elastic Net
  - Random forest

F**A**CE: Methods and Models

Stochastic/statistical

- Correlation analysis
- Cluster and segmentation analysis
  - K-means, k-clusters
  - KNN
  - CART/CHAID
- ANOVA/DOE
  - Analysis of variance
  - Design of experiments
  - Factor analysis

Stochastic/statistical forecasting

- Time series analysis
  - Exponential smoothing
    - With seasonal variation
      - Holt-Winters model
  - Box-Jenkins (ARIMA)
    - Auto-regressive
    - Moving average
  - Neural network (MLP, RNN)

DataScience@SMU

Modeling

Commercial Off The Shelf (COTS) Packages and Tools

Doug Gray

F**A**CE: Commercial off the Shelf (COTS) Software Packages and Tools

Solving the problem

- Modeling
  - Applying the scientific method
  - Pragmatically
    - *“All models are wrong, but some are useful,”* George Box
    - *“OR is the art and science of giving bad answers to questions that would otherwise be given far worse answers”*
    - *“The forecast is always wrong; the question is by how much”*
  - *Type of model/method*
  - ***Commercial off the shelf technology***

F**A**CE: COTS Software Packages and Tools

Deterministic

- MS Excel Solver, Lindo What's best?
- R, MATLAB
- IBM ILOG CPLEX, FICO Xpress, Gurobi
- IBM ILOG Solver/Views Suite, FICO Optimization Modeler
- ILOG JRules,
  FICO Blaze, Drools

Stochastic

- MS Excel Statistics
- R, MATLAB, Python
- SAS
- IBM SPSS
- Minitab
- H2O.ai
- Data Robot

F**A**CE: COTS Software Packages and Tools

Big data

- Cloudera, Hortonworks, Pivotal (Hadoop Tools)
- Apache Foundation big data tools, e.g., Mahout for machine learning and data mining
- R tm plug-in (text analytics)

Visual analytics

- Tableau
- Qlikview
- Microstrategy
- Sisense
- TIBCO Spotfire

F**A**CE: COTS Software Packages and Tools

Analytics platforms

- Two tools available at SMU Cox
- **Alteryx** (licensed COTS product)
  - Intuitive, easy to use GUI
  - Data integration, cleansing
  - Visualization
  - Statistical/data science modeling
- **KNIME** (open source product)
- H2O.ai
- Data Robot

DataScience@SMU

Data Analysis

Doug Gray

FACE

Solving the problem

- Data analysis
  - Descriptive statistics, i.e., mean, mode, median, variance, standard deviation
  - Inferential statistics, i.e., correlation, covariance, regression parameters
  - Visualization, i.e., scatter plots, heat maps, probability distribution
  - Confidence/prediction intervals, hypothesis tests, goodness of fit tests
  - Evaluation/cross validation
    - Confusion matrix (classification)
    - RMSE, MAE, MAPE, etc. (numerical regression/prediction)

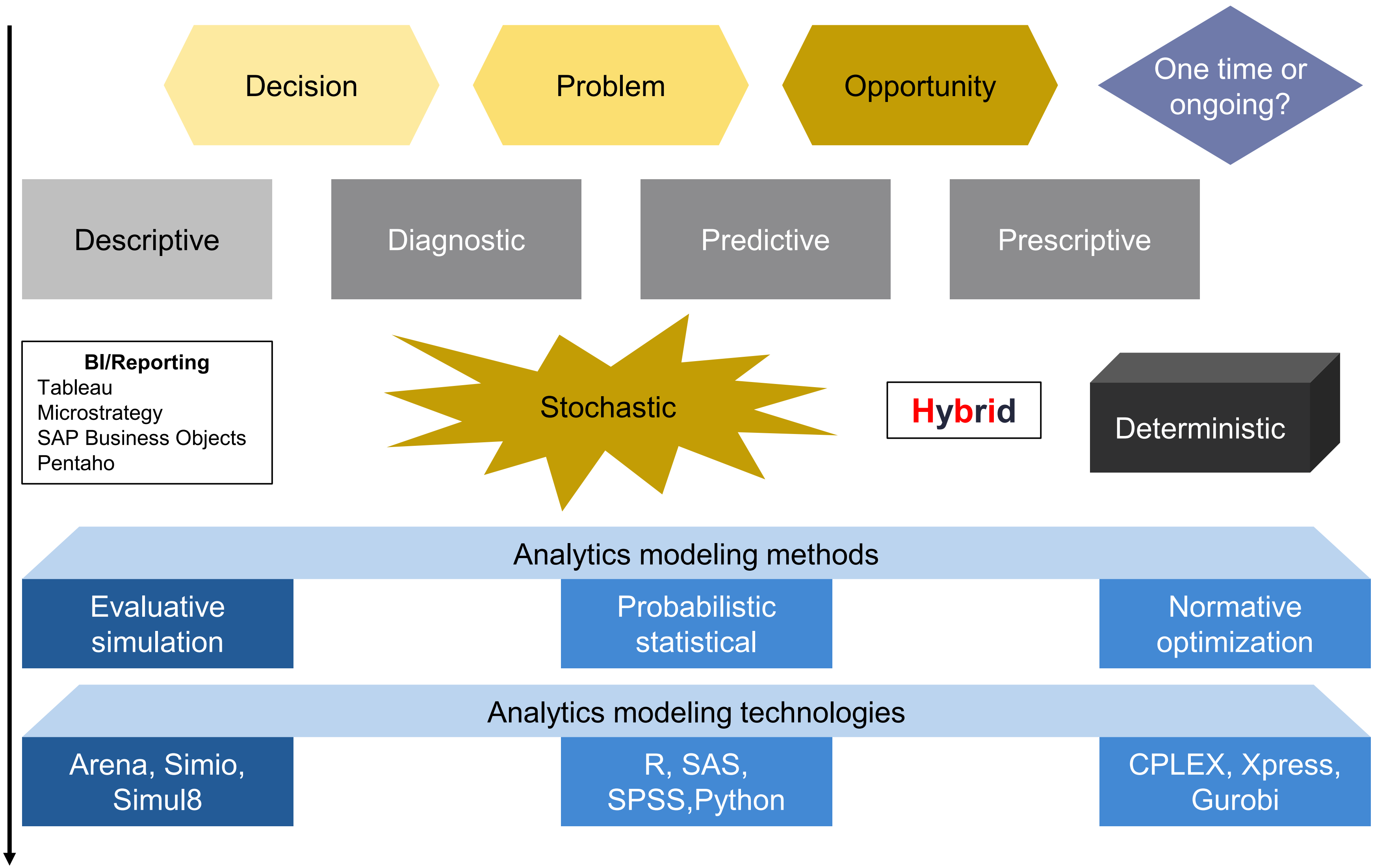
DataScience@SMU

The Pachinko Machine

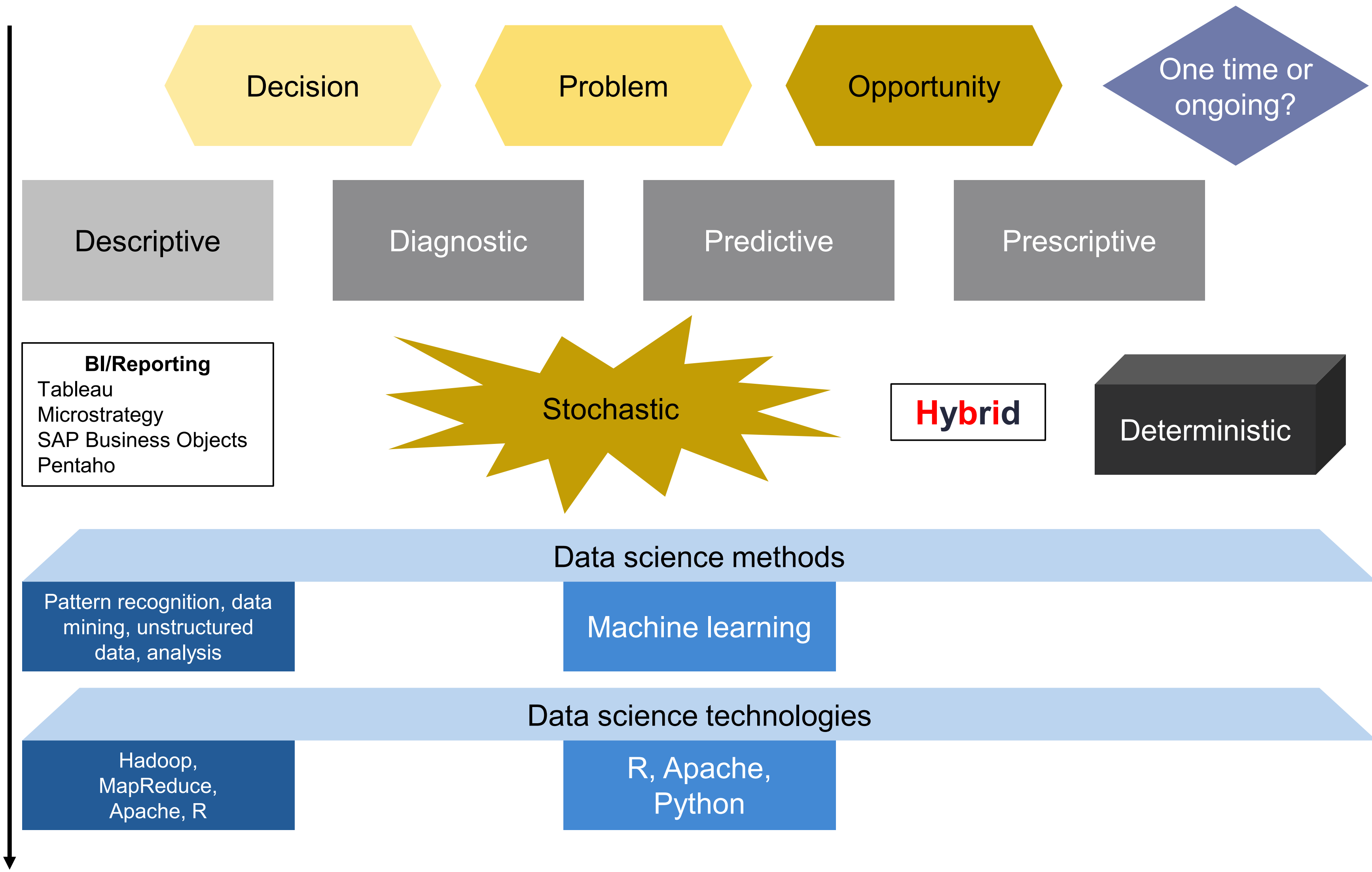
Doug Gray

FACE: The Pachinko Machine!

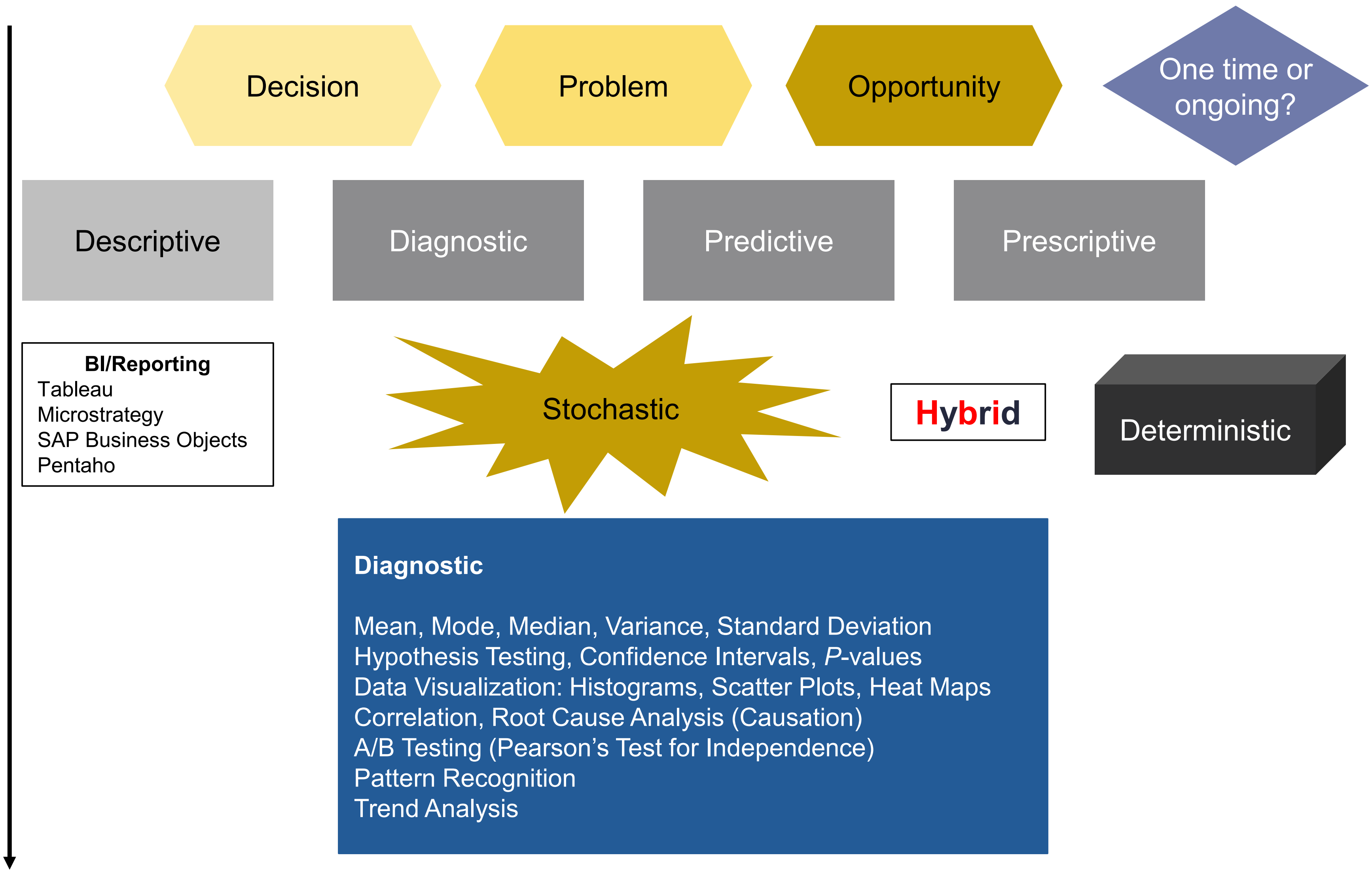
Frame and Solve: Analytics Solution Approach



Frame and Solve: Analytics Solution Approach



Frame and Solve:
Analytics Solution Approach



Frame and Solve:
Analytics Solution Approach

Decision

Problem

Opportunity

One time or ongoing?

Descriptive

Diagnostic

Predictive

Prescriptive

BI/Reporting
Tableau
Microstrategy
SAP Business Objects
Pentaho

Stochastic

Hybrid

Deterministic

If you seek to Predict the (x), then use (y)…(See *The Field Guide to Data Science* for details!)

Probability	Norm(μ,σ), bin(n,p), Pois(λ), Exp(μ), Bayesian (Network, Naïve), Markov
Odds (p / (1-p))	Logistic Regression (1 / (1 + e^{-z}))
Value or Class	(ML, NL, Ridge, LASSO) Regression Time Series (ARIMA), Exponential Smoothing Machine Learning Methods, e.g., Neural Networks, Tree-based Methods, e.g., Gradient Boosting, Random Forest
Relationship	Classification, Clustering, Segmentation, k-means

Frame and Solve:
Analytics Solution Approach

Decision

Problem

Opportunity

One time or ongoing?

Descriptive

Diagnostic

Predictive

Prescriptive

BI/Reporting
Tableau
Microstrategy
SAP Business Objects
Pentaho

Stochastic

Hybrid

Deterministic

If you seek to Prescribe *(Optimize or Advise on)* an outcome (See *The Field Guide to Data Science*)

Mathematical Programming (i.e., Linear, Nonlinear, Integer, Quadratic)

Constraint-based Programming (Feasibility Only)

Heuristic Algorithms (e.g., Genetic, Simulated Annealing)

Discrete-event or Continuous (a.k.a., “Monte Carlo”) Simulation

Systems Dynamics

Artificial Intelligence (AI)

Expert (or Rule-based) System

Fuzzy Logic

Frame and Solve:
Analytics Solution Approach

Example: hospital readmissions

Conditions With the Largest Number of Adult Hospital Readmissions by Payer, 2011

Anika L. Hines, Ph.D., M.P.H., Marguerite L. Barrett, M.S., H. Joanna Jiang, Ph.D., and Claudia A. Steiner, M.D., M.P.H.

Introduction

Health care reform has pinpointed hospital readmissions as a key area for improving care coordination and achieving potential savings.[1] Stakeholders are using data to devise strategies to reduce readmissions. Two criteria for evaluating potential areas of impact include volume and costs. For example, the Centers for Medicare & Medicaid Services (CMS) Hospital Readmissions Reduction Program has selected acute myocardial infarction, heart failure, and pneumonia as target areas for the Medicare population. CMS chose these conditions, in part, because of their high prevalence and their associated high costs for total admissions and readmissions among Medicare beneficiaries.[2] In 2015, CMS will expand their assessment of readmissions to additional conditions that represent high volume and costs.

Identifying conditions that contribute the most to the total number of readmissions and related costs for *all payers* may aid health care stakeholders in deciding which conditions to target to maximize quality improvement and cost-reduction efforts. This Statistical Brief uses readmissions data from the Healthcare Cost and Utilization Project (HCUP) to present the conditions with the largest number of 30-day all-cause readmissions among U.S. hospitals in 2011 and their associated costs. We limited the study population to Medicare beneficiaries aged 65 years and older and to individuals aged 18–64 years who were privately insured, uninsured, or covered by Medicaid. We display the 10 conditions with the largest number of readmissions for each payer.

Readmission was defined as a subsequent hospital admission within 30 days following an original admission (or index stay) that occurred from January through November 2011. Patients were followed across the same and different hospitals. All-cause readmissions were examined; thus, readmissions may or may not include conditions that were listed as the principal diagnosis during the index stay. Some readmissions may be planned or

Frame and Solve: Analytics Solution Approach

- Healthcare industry
- Healthcare provider
- *“Readmissions of orthopedic hip and knee surgery patients have increased 25% YOY”*
- *Readmissions result in a 50% decrease in orthopedic surgery procedure profits”*

Frame and Solve: Analytics Solution Approach

- Healthcare industry
- Healthcare provider
- *“Readmissions of orthopedic hip and knee surgery patients have increased 25% YOY”*
- *Readmissions result in a 50% decrease in orthopedic surgery procedure profits”*

Problem

Frame and Solve: Analytics Solution Approach

- Healthcare industry
- Healthcare provider
- *“Readmissions of orthopedic hip and knee surgery patients have increased 25% YOY”*
- *Readmissions result in a 50% decrease in orthopedic surgery procedure profits”*

Problem

Descriptive What?

Why? Diagnostic

How many readmissions are occurring?

Percentage of the total number of knee and hip surgeries?

What factors cause readmissions?

As a percentage of the total number of readmissions?

Frame and Solve: Analytics Solution Approach

- Healthcare industry
- Healthcare provider
- *"Readmissions of orthopedic hip and knee surgery patients have increased 25% YOY"*
- *Readmissions result in a 50% decrease in orthopedic surgery procedure profits"*

Problem

"How can we prevent readmissions from occurring?"

Process improvements, e.g., more rigorous pre- and post-op sterilization to prevent infection

Predictive

*Identify **(predict)** patients who may be prone to readmit **before** they are discharged to take preventative measures*

Frame and Solve: Analytics Solution Approach

- *"Readmissions of orthopedic hip and knee surgery patients have increased 25% YOY"*
- *"Readmissions result in a 50% decrease in orthopedic surgery procedure profits"*

Problem

"How can we prevent readmissions from occurring?"

Predictive

*Identify **(predict)** patients who may be prone to readmit **before** they are discharged to take preventative measures*

There are dozens of variables that can influence whether or not a patient will readmit

Stochastic

Frame and Solve: Analytics Solution Approach

“How can we prevent readmissions from occurring?”

Predictive

*Identify **(predict)** patients who may be prone to readmit **before** they are discharged to take preventative measures*

There are dozens of variables that can influence whether or not a patient will readmit

Stochastic

Predicting readmission with certainty is not possible; predict the “odds” or “probability” of a patient readmission event

Potential classifiers:

- *Logistic regression*
- *Bayesian inference*
- *Latent class model*
- *Neural network*

Frame and Solve: Analytics Solution Approach

- Airline industry
- Commercial airline
- *“We need a crew schedule to cover our flight schedule (CA, FO, FA, 4K flights/day)”*
- *“The crew schedule needs to comply with contractual and regulatory work rules”*
- *“Crew labor is our number 1 operating expense; crew the schedule at minimum cost”*

Frame and Solve: Analytics Solution Approach

- Airline industry
- Commercial airline
- *“We need a crew schedule to cover our flight schedule (CA, FO, FA, 4K flights/day)”*
- *“The crew schedule needs to comply with contractual and regulatory work rules”*
- *“Crew labor is our number 1 operating expense; crew the schedule at minimum cost”*

Frame and Solve:
Analytics Solution Approach

- Airline industry
- Commercial airline
- *“We need a crew schedule to cover our flight schedule (CA, FO, FA, 4K flights/day)”*
- *“The crew schedule needs to comply with contractual and regulatory work rules”*
- *“Crew labor is our number 1 operating expense; crew the schedule at minimum cost”*

- *“How can we develop a crew schedule that complies with all contractual and regulatory work rules and covers the entire flight schedule at minimum cost?”*

Frame and Solve:
Analytics Solution Approach

- Airline industry
- Commercial airline
- *“We need a crew schedule to cover our flight schedule (CA, FO, FA, 4K flights/day)”*
- *“The crew schedule needs to comply with contractual and regulatory work rules”*
- *“Crew labor is our number 1 operating expense; crew the schedule at minimum cost”*

Decision

Problem

- *“How can we develop a crew schedule that complies with all contractual and regulatory work rules and covers the entire flight schedule at minimum cost?”*

*Rules represent **constraints** that must be satisfied; minimizing cost is an **objective** resulting from crew assignment decisions*

Prescriptive

Frame and Solve: Analytics Solution Approach

- *“We need a crew schedule to cover our flight schedule (CA, FO, FA, 4K flights/day)”*
- *“The crew schedule needs to comply with contractual and regulatory work rules”*
- *“Crew labor is our number 1 operating expense; crew the schedule at minimum cost”*

Decision → Problem

- *“How can we develop a crew schedule that complies with all contractual and regulatory work rules and covers the entire flight schedule at minimum cost?”*

*Rules represent **constraints** that must be satisfied; minimizing cost is an **objective** resulting from crew assignment decisions*

Prescriptive

Deterministic

*The flight schedule is a **plan** that is set/known; the crew schedule is a **plan** that must be set in advance of operation*

Frame and Solve: Analytics Solution Approach

Decision

Problem

- *"How can we develop a crew schedule that complies with all contractual and regulatory work rules and covers the entire flight schedule at minimum cost?"*

*Rules represent **constraints** that must be satisfied; minimizing cost is an **objective** resulting from crew assignment decisions*

Prescriptive

Deterministic

*The flight schedule is a **plan** that is set/known; the crew schedule is a **plan** that must be set in advance of operation*

***Decision variables** to **assign** flights to crew schedules are **integer** {1:Assign, 0: Not}*
*Duty times are **linear** {≥ 0}*

- *Mathematical programming*
  - *Linear and Integer Programming*
  - *Mixed Integer Programming (MIP)*
  - *Set Partitioning Problem Formulation*

DataScience@SMU

Communicating and Acting on Results

Doug Gray

FACE

- Communicating and acting on results
  - Results presentation and action
    - *Tell a story, with your audience in mind*
      - Focus on the business problem/decision, and the business value/impact
      - Interesting and comprehensible formats
      - Avoid boredom and confusion at all costs
      - Explain the experiment, analysis, and results in business terms (NPV, IRR)
      - Save the math for the appendix, unless the audience is highly quantitative
    - The clearer the results presentation, the more likely the quantitative analysis will lead to decisions and actions

FACE

- Communicating and acting on results
  - Results presentation and action
    - Start with the ending: summary, conclusions, and recommendations up front
    - Avoid jargon, technical terminology (R-squared, heteroskedasticity)
    - Use visuals/graphics vs. words/numbers (p.107 KUWTHQ) to tell the story
    - *Before and after, cause and effect*
  - Change management dimension; non-analytically inclined staff
    - "You can't get a computer to ____________?!"
  - Factor in political considerations; not everyone will be happy with the results
    - "What? That can't be right?!"

FA**C**E

- Communicating and acting on results
  - Analytics is often ***very disruptive*** on many levels
    - Threatens the *status quo*
    - Business process and workflow
      - Manual to streamlined, automated, and optimized
      - Difficult to get some decision-makers comfortable with yielding control
    - IT, i.e., data purchasing, collection, storage, governance
      - New types of data, and ***lots*** more data, and technologies
  - The *pros*/*benefits*/*rewards* ***far outweigh*** the *cons*/*costs*/*risks* but changes and transitions need to be managed carefully to unify stakeholders/constituents

DataScience@SMU

Embedding

Doug Gray

FACE

Embedding final models and methods in enterprise business processes and systems

Uber example

GIS locator method

Optimal route detection method

Dynamic pricing model

DataScience@SMU