Data Analytics Tools & Technique

Chapter 1: Introduction to Business Analytics

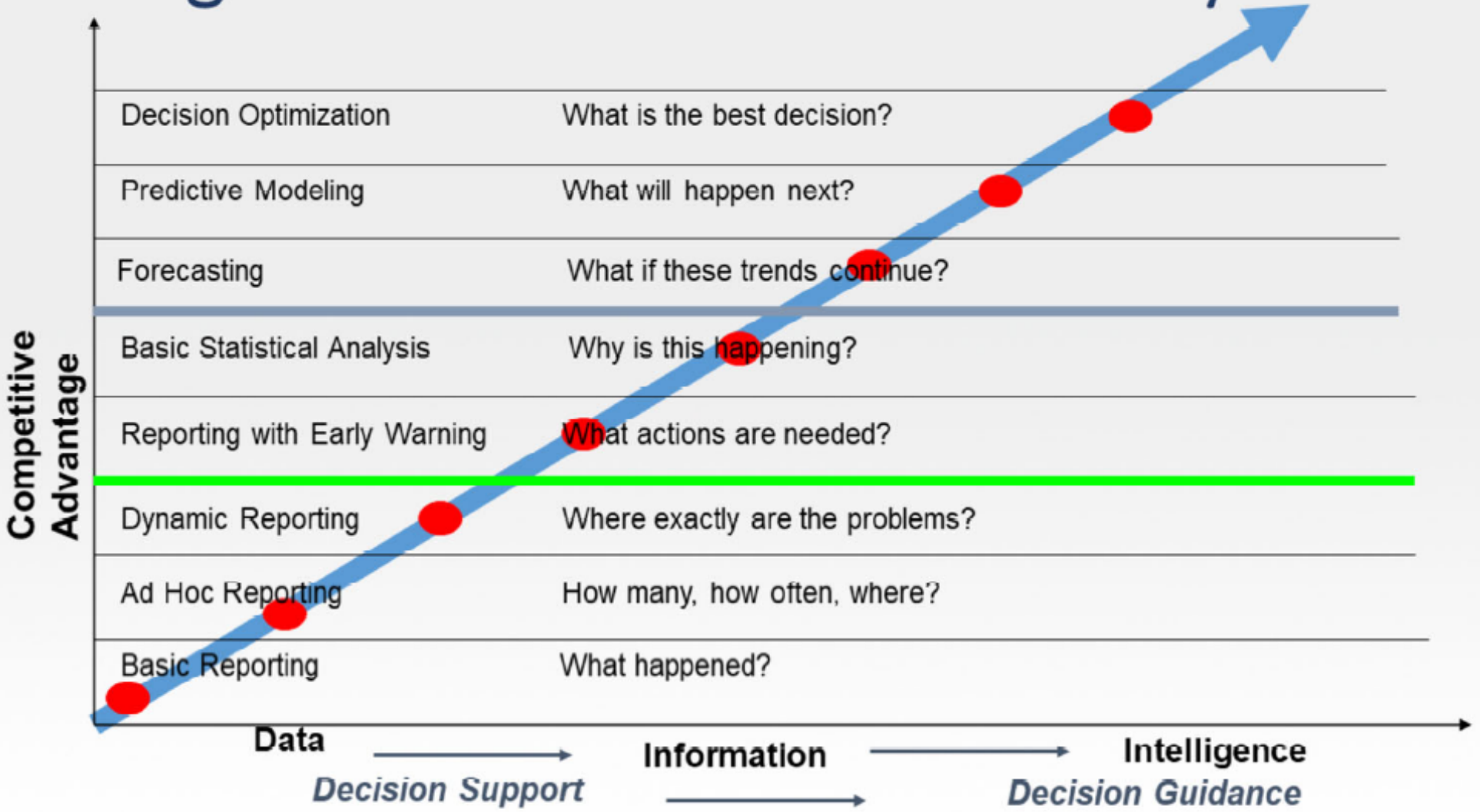
DA 6223

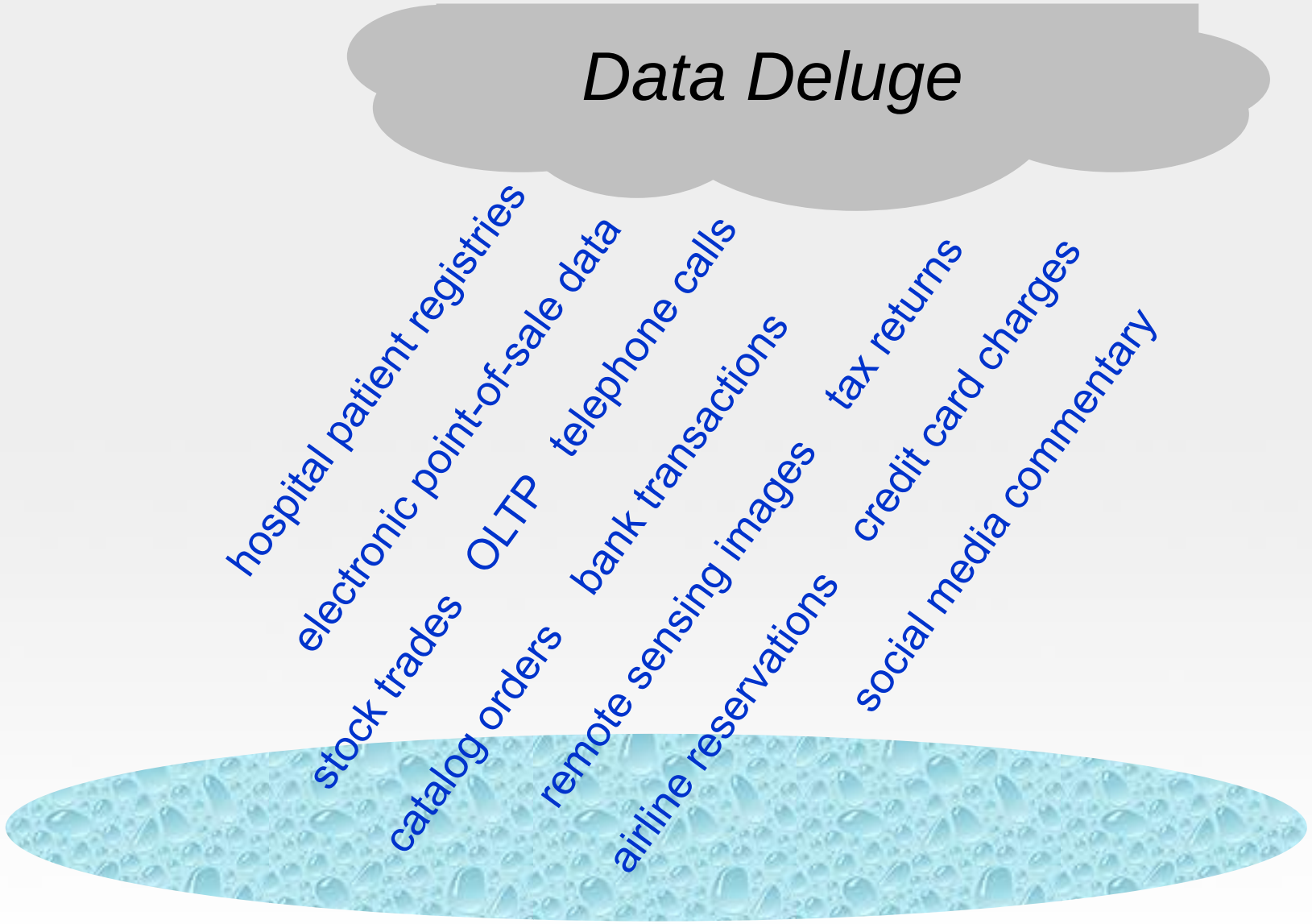
The University of Texas at San Antonio

Business Analytics

- **Anticipate opportunity, take action, and make an impact**
  - Find and seize opportunities for profitability and growth.
  - Take strategic, operational, and tactical actions – in record time.
  - Turn those actions into tangible results across the enterprise.

Achieving Success with Business Analytics




Data Deluge
hospital patient registries
electronic point-of-sale data
telephone calls
OLTP
bank transactions
stock trades
tax returns
catalog orders
credit card charges
remote sensing images
social media commentary
airline reservations

Three Consequences of the Data Deluge

1. Every problem will generate data eventually.

2. Every company will need analytics eventually.

3. Everyone will need analytics eventually.

Three Consequences of the Data Deluge

1. Every problem will generate data eventually.
   Proactively defining a data collection protocol will result in more useful information, leading to more useful analytics.
2. Every company will need analytics eventually.

3. Everyone will need analytics eventually.

Three Consequences of the Data Deluge

1. Every problem will generate data eventually.
   Proactively defining a data collection protocol will result in more useful information, leading to more useful analytics.
2. Every company will need analytics eventually.
   Proactively analytical companies will compete more effectively.

3. Everyone will need analytics eventually.

Three Consequences of the Data Deluge

1. Every problem will generate data eventually.
   Proactively defining a data collection protocol will result in more useful information, leading to more useful analytics.
2. Every company will need analytics eventually.
   Proactively analytical companies will compete more effectively.

3. Everyone will need analytics eventually.
   Proactively analytical people will be more marketable and more successful in their work.

The Business Analytics Challenge

- Getting anything useful out of tons and tons of data

Data Deluge

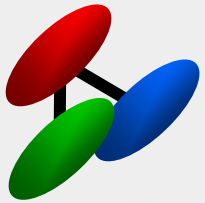
hospital patient registries
electronic point-of-sale data
telephone calls
OLTP
bank transactions
tax returns
credit card charges
social media commentary
stock trades
catalog orders
remote sensing images
airline reservations

+ analytical tools

= actionable knowledge

Changes in the Analytical Landscape

Historically…



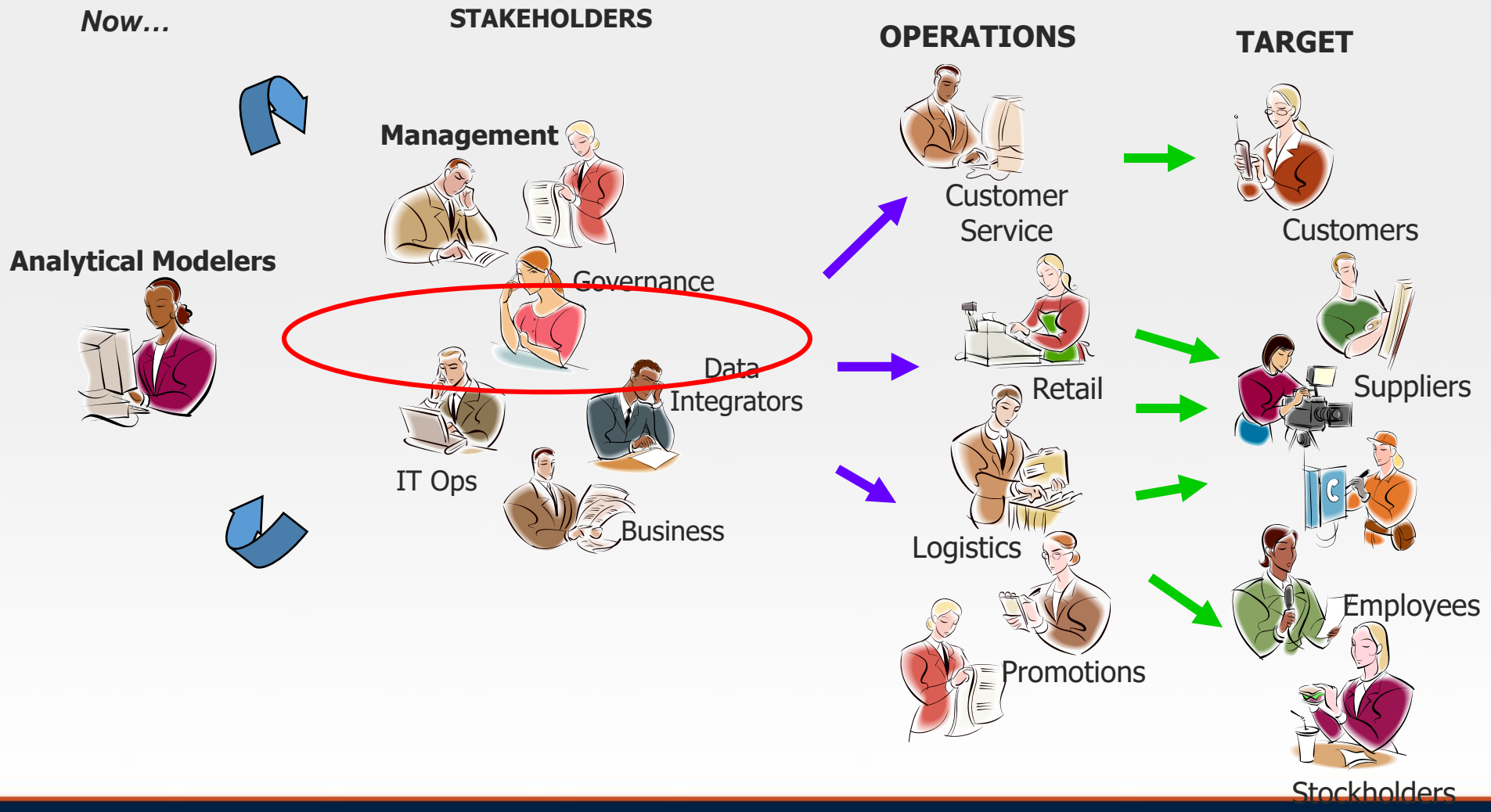
Models

Analytical Modelers

Management

Historically, analytics have typically been handled in the "back office," and information was shared only by a few individuals.

Changes in the Analytical Landscape



Changes in the Analytical Landscape

- Relational databases
- Enterprise resource planning (ERP) systems
- Point of sale (POS) systems
- Data warehousing
- Decision support systems
  - Reporting and ad hoc queries
  - Online analytical processing (OLAP)
- Performance management systems
  - Executive information systems (EIS)
  - Balanced scorecard
  - Dashboard
- Business intelligence

Idiosyncrasies of Business Analytics

1. The Data

- Massive, operational, and opportunistic

2. The Users and Sponsors

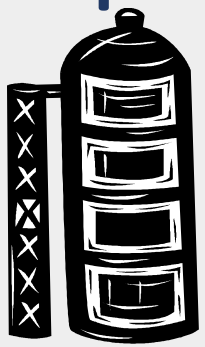
- Business decision support

3. The Methodology

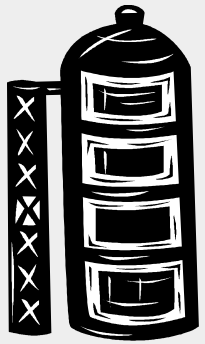
- Computer-intensive ad-hoc
- Multidisciplinary lineage

Data mining can be defined as advanced methods for exploring and modeling relationships in large amounts of data. Data mining is an essential component of business analytics.

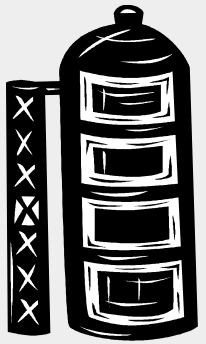
The Data: Disparate Business Units



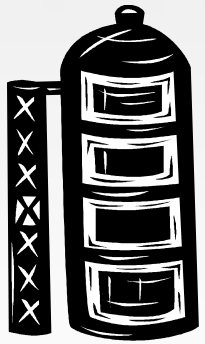
Marketing



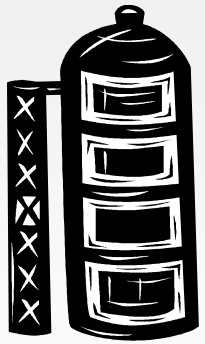
Invoicing



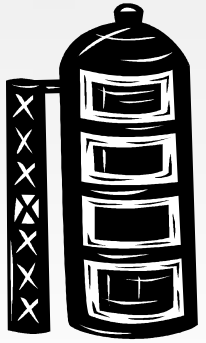
Risk



Acquisitions



Operations



Sales

Opportunistic Data

- Operational data is typically not collected with data analysis in mind.
- Multiple business units produce a silo-based data system.
- This makes business analytics different from experimental statistics and especially challenging.

The Methodology: What We Learned Not to Do

- Prediction is more important than inference.
  - Metrics are used "because they work," not based on theory.
  - *p*-values are rough guides rather than firm decision cutoffs.
  - Interpretation of a model might be irrelevant.
  - The preliminary value of a model is determined by its ability to predict a holdout sample.
  - The long-term value of a model is determined by its ability to continue to perform well on new data over time.
  - Models are retired as customer behavior shifts, market trends emerge, and so on.

Using Analytics Intelligently

- Intelligent use of analytics results in the following:
  - better understanding of how technological, economic, and marketplace shifts affect business performance
  - ability to ***consistently and reliably*** distinguish between effective and ineffective interventions
  - efficient use of assets, reduced waste in supplies, and better management of time and resources
  - risk reduction via ***measurable*** outcomes and ***reproducible*** findings
  - early detection of market trends hidden in massive data
  - continuous improvement in decision making over time

Simple Reporting

- **Examples**: OLAP, RFM, QC, descriptive statistics, extrapolation
- ***Answer questions such as***
  - Where are my key indicators now?
  - Where were my key indicators last week?
  - Is the current process behaving like normal?
  - What is likely to happen tomorrow?

Proactive Analytical Investigation

- **Examples**: inferential statistics, experimentation, empirical validation, forecasting, optimization
- ***Answer questions such as***
  - What does a change in the market mean for my targets?
  - What do other factors tell me about what I can expect from my target?
  - What is the best combination of factors to give me the most efficient use of resources and maximum profitability?
  - What is the highest price the market will tolerate?
  - What will happen in six months if I do nothing?
  - What if I implement an alternative strategy?

- Many companies have data that they do not use or that is used by third parties. These third parties might even resell the data and any derived metrics back to the original company!

- **Example**: retail grocery POS card

Every Little Bit…

- Taking an analytical approach to only a few key business problems with reliable metrics → tangible benefit.
- The benefits and savings derived from early analytical successes → managerial support for further analytical efforts.
- **Everyone has data.**
- **Analytics can connect data to smart decisions.**
- **Proactively analytical companies outpace competition.**

Areas Where Analytics Are Often Used

- New customer acquisition
- Customer loyalty
- Cross-sell / up-sell
- Pricing tolerance
- Supply optimization
- Staffing optimization
- Product placement
- Churn
- Insurance rate setting
- Fraud detection
- …

Which residents in a ZIP code should receive a coupon in the mail for a new store location?

Areas Where Analytics Are Often Used

- New customer acquisition
- Customer loyalty
- Cross-sell / up-sell
- Pricing tolerance
- Supply optimization
- Staffing optimization
- Product placement
- Churn
- Insurance rate setting
- Fraud detection
- …

What advertising strategy best elicits positive sentiment toward the brand?

Areas Where Analytics Are Often Used

- New customer acquisition
- Customer loyalty
- Cross-sell / up-sell
- Pricing tolerance
- Supply optimization
- Staffing optimization
- Product placement
- Churn
- Insurance rate setting
- Fraud detection
- …

What is the best next product for this customer? What other product is this customer likely to purchase?

Areas Where Analytics Are Often Used

- New customer acquisition
- Customer loyalty
- Cross-sell / up-sell
- Pricing tolerance
- Supply optimization
- Staffing optimization
- Product placement
- Churn
- Insurance rate setting
- Fraud detection
- …

What is the highest price that the market will bear without substantial loss of demand?

Areas Where Analytics Are Often Used

- New customer acquisition
- Customer loyalty
- Cross-sell / up-sell
- Pricing tolerance
- Supply optimization
- Staffing optimization
- Product placement
- Churn
- Insurance rate setting
- Fraud detection
- …

How many 60-inch HDTVs should be in stock? (Too many is expensive; too few is lost revenue.)

Areas Where Analytics Are Often Used

- New customer acquisition
- Customer loyalty
- Cross-sell / up-sell
- Pricing tolerance
- Supply optimization
- Staffing optimization
- Product placement
- Churn
- Insurance rate setting
- Fraud detection
- …

What are the best times and best days to have technical experts on the showroom floor?

Areas Where Analytics Are Often Used

- New customer acquisition
- Customer loyalty
- Cross-sell / up-sell
- Pricing tolerance
- Supply optimization
- Staffing optimization
- Financial forecasting
- Product placement
- Churn
- Insurance rate setting
- Fraud detection
- …

What weekly revenue increase can be expected after the Mother's Day sale?

Areas Where Analytics Are Often Used

- New customer acquisition
- Customer loyalty
- Cross-sell / up-sell
- Pricing tolerance
- Supply optimization
- Staffing optimization
- Product placement
- Churn
- Insurance rate setting
- Fraud detection
- …

Will oatmeal sell better near granola bars or near baby food?

Areas Where Analytics Are Often Used

- New customer acquisition
- Customer loyalty
- Cross-sell / up-sell
- Pricing tolerance
- Supply optimization
- Staffing optimization
- Product placement
- Churn
- Insurance rate setting
- Fraud detection
- …

Which customers are most likely to switch to a different wireless provider in the next six months?

Areas Where Analytics Are Often Used

- New customer acquisition
- Customer loyalty
- Cross-sell / up-sell
- Pricing tolerance
- Supply optimization
- Staffing optimization
- Product placement
- Churn
- Insurance rate setting
- Fraud detection
- …

How likely is it that this individual will have a claim?

Areas Where Analytics Are Often Used

- New customer acquisition
- Customer loyalty
- Cross-sell / up-sell
- Pricing tolerance
- Supply optimization
- Staffing optimization
- Product placement
- Churn
- Insurance rate setting

- Fraud detection
- …

How can I identify a fraudulent purchase?

When Analytics Are *Not* Helpful

- Snap decisions required

Deciding when to run from danger

- Novel approach (no previous data possible)
- Most salient factors are rare (making decisions to work around unlikely obstacles or miracles)
- Metrics are inappropriate
- Naïve implementation of analytics
- Confirming what you already know

When Analytics Are *Not* Helpful

- Snap decisions required
- Novel approach (no previous data possible)
- Most salient factors are rare (making decisions to work around unlikely obstacles or miracles)
- Metrics are inappropriate
- Naïve implementation of analytics
- Confirming what you already know

Predicting the adoption of a new technology

When Analytics Are *Not* Helpful

- Most salient factors are rare (making decisions to work around unlikely obstacles or miracles)

Planning contingencies for employees winning the lottery

When Analytics Are *Not* Helpful

- Expert analysis suggests a particular path

The seasoned art critic can recognize a fake

When Analytics Are *Not* Helpful

- Metrics are inappropriate

Predicting athletes' salaries or quantifying love

When Analytics Are *Not* Helpful

- Snap decisions required
- Novel approach (no previous data possible)
- Most salient factors are rare (making decisions to work around unlikely obstacles or miracles)
- Metrics are inappropriate
- Naïve implementation of analytics
- Confirming what you already know

Only looking at one variable at a time

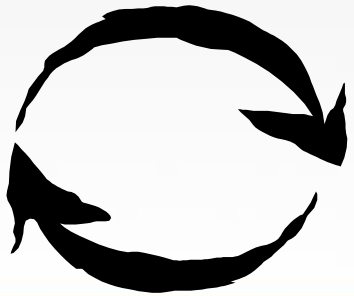
When Analytics Are *Not* Helpful

- Confirming what you already know

Ignoring variables that might be important

Expectations Leading the Analysis

- Even sophisticated analytics are not immune to personal bias such as the following:
  - selectively fitting models with variables because they place someone's opinion or agenda in a positive light
  - ignoring information that might disprove a hypothesis.
- Personal bias in model fitting, whether intentional or otherwise, can diminish the usefulness of your analytical efforts.



Trustworthy Analytics

- Let the data guide your conclusions.
- Ask the following questions:
  - Are my assumptions about the causes of my data patterns warranted?
  - Should I try something different?
- Assign a cynic to the analytical team whose purpose is to question the assumptions.
  - What would my critic say is the flaw with my analysis?
  - Investigate the data in such a way that a critic's concerns can be ruled out.

Data Analytics Tools & Technique

Chapter 1: Introduction to Business Analytics

DA 6223

The University of Texas at San Antonio