
Previews of TDWI course books are provided as an opportunity to see the quality of our material and help you to select the courses that best fit your needs. The previews can not be printed.

TDWI strives to provide course books that are content-rich and that serve as useful reference documents after a class has ended.

This preview shows selected pages that are representative of the entire course book. The pages shown are not consecutive. The page numbers as they appear in the actual course material are shown at the bottom of each page. All table-of-contents pages are included to illustrate all of the topics covered by a course.


# TDWI Introduction to Business Analytics

# TABLE OF CONTENTS




# Module 1

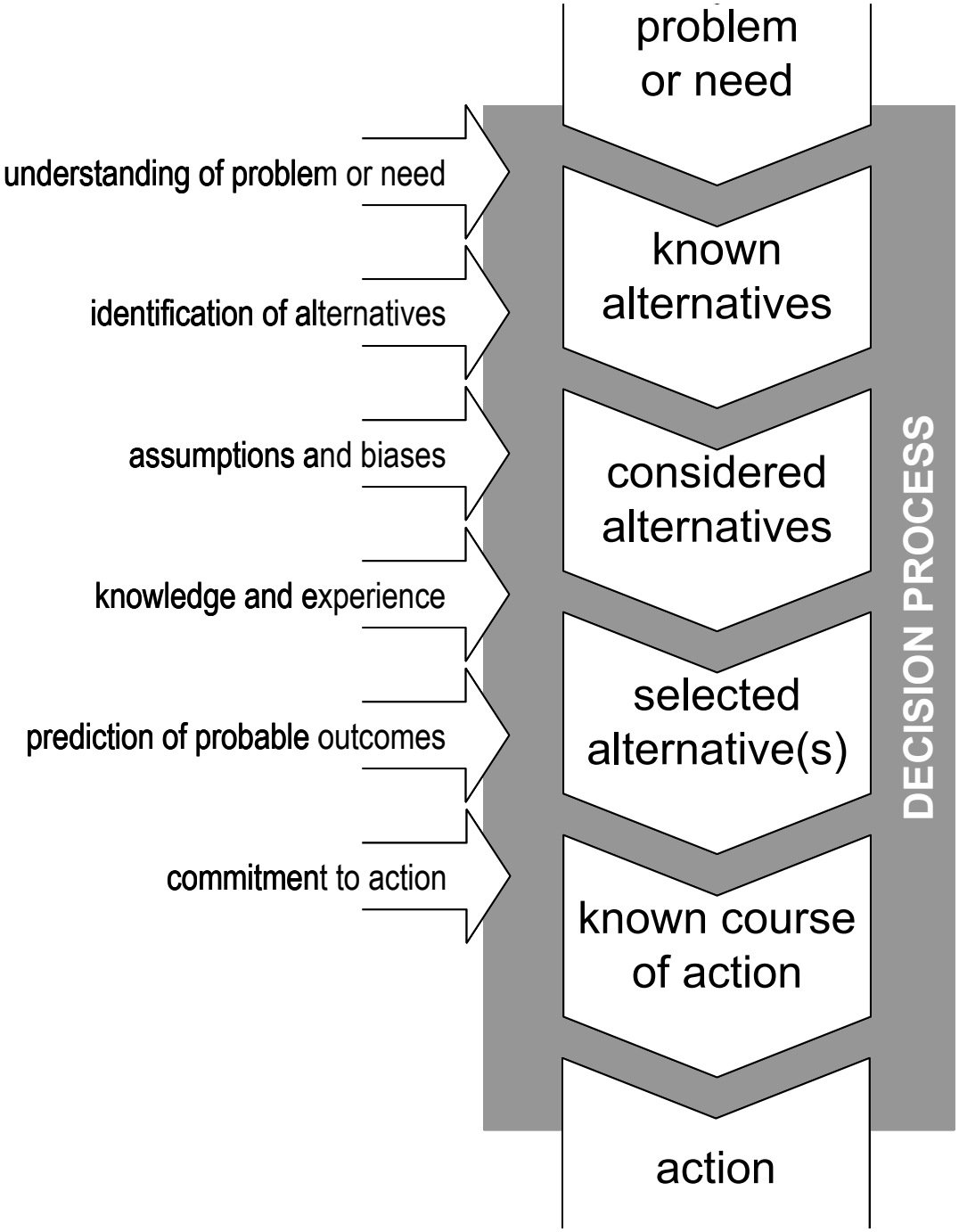
## The Business Analytics Landscape



# Analytic Concepts

## Decision Concepts



# Analytic Concepts

## Decision Concepts

### OVERVIEW

The Decision Making process shown in the diagram on the facing page provides a guideline and a set of specific steps to help managers and knowledge workers achieve consistency in the quality of their decisions. The process outlines a prescribed set of steps that help to eliminate "gut feel" and "emotion" from decisions. These "human instinct" components can be replaced by business analytics components at key locations along the process.

### STEPS IN THE DECISION MAKING PROCESS

A series of general steps are used to define a decision making process. Although there are different methodologies and approaches used in decision making, most of them have the following core steps.

1. Identify and define the problem or need
2. Generate a broad set of possible alternatives
3. Reduce the alternatives to those considered for serious evaluation
4. Select and recommend the alternative(s)
5. Define and clarify the course of action
6. Execute the recommendation

### ENABLING THE DECISION MAKING PROCESS WITH ANALYTICS

Within each step in the Decision Making process, there are several activities that take place. Many of these steps either depend on or benefit from business analytics to clarify our understanding about the problem being solved. Analytics can be used to identify new problems or opportunities that enter the decision making funnel. Business analytics can enable the Decision Making process in the following activities.

1. Monitoring the business to identify Problems or Opportunities
2. Clarifying the definition and formulation of the Problem
3. Validating assumptions and constraints
4. Mitigating human biases within the problem formation
5. Establishing the feasibility of generated alternatives
6. Predicting and evaluating the outcomes of selected alternatives

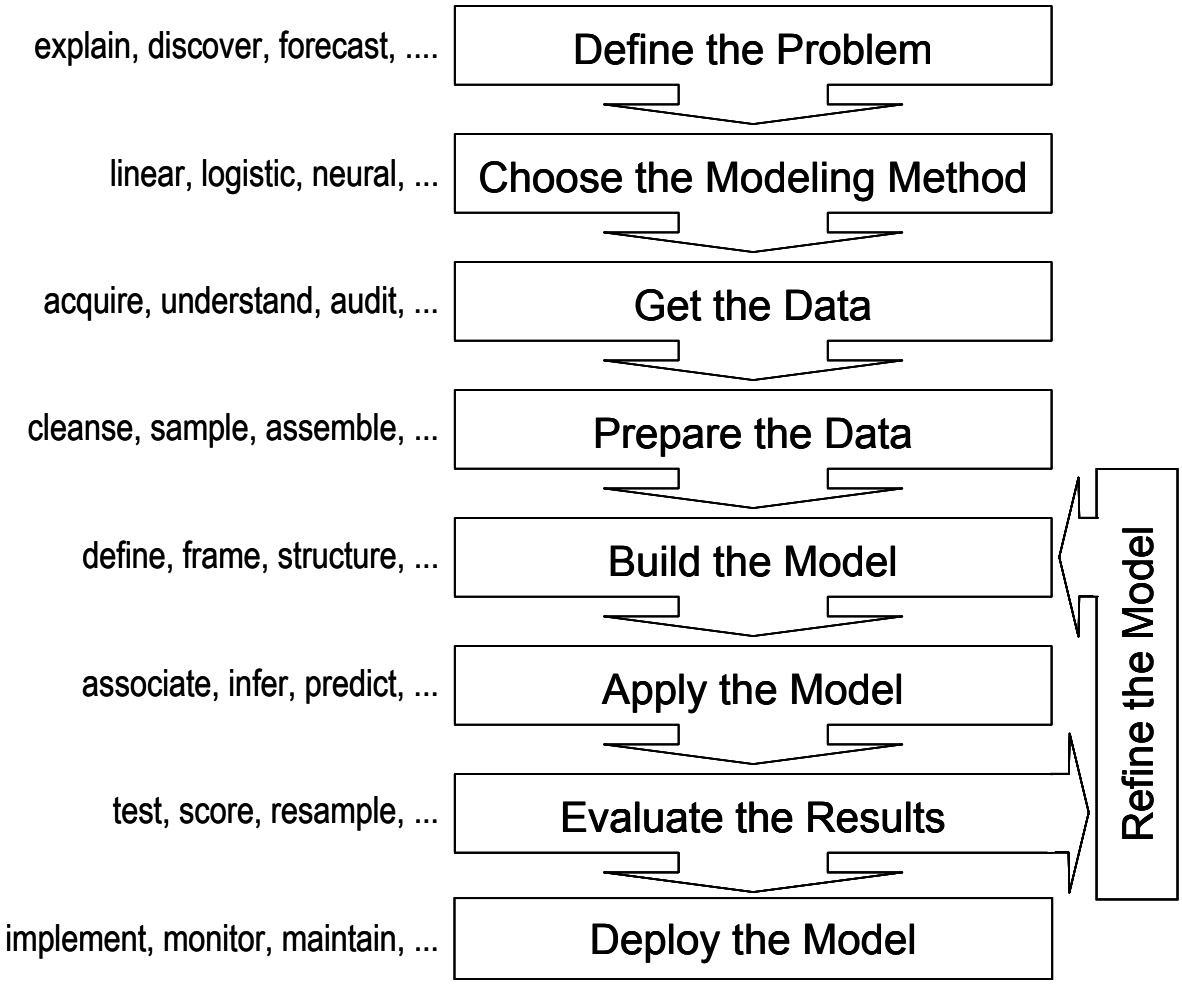

# Module 2

## Analytic Skills

This page intentionally left blank.

# Modeling

## Developing Analytic Models



# Modeling

## Developing Analytic Models

### APPROACH

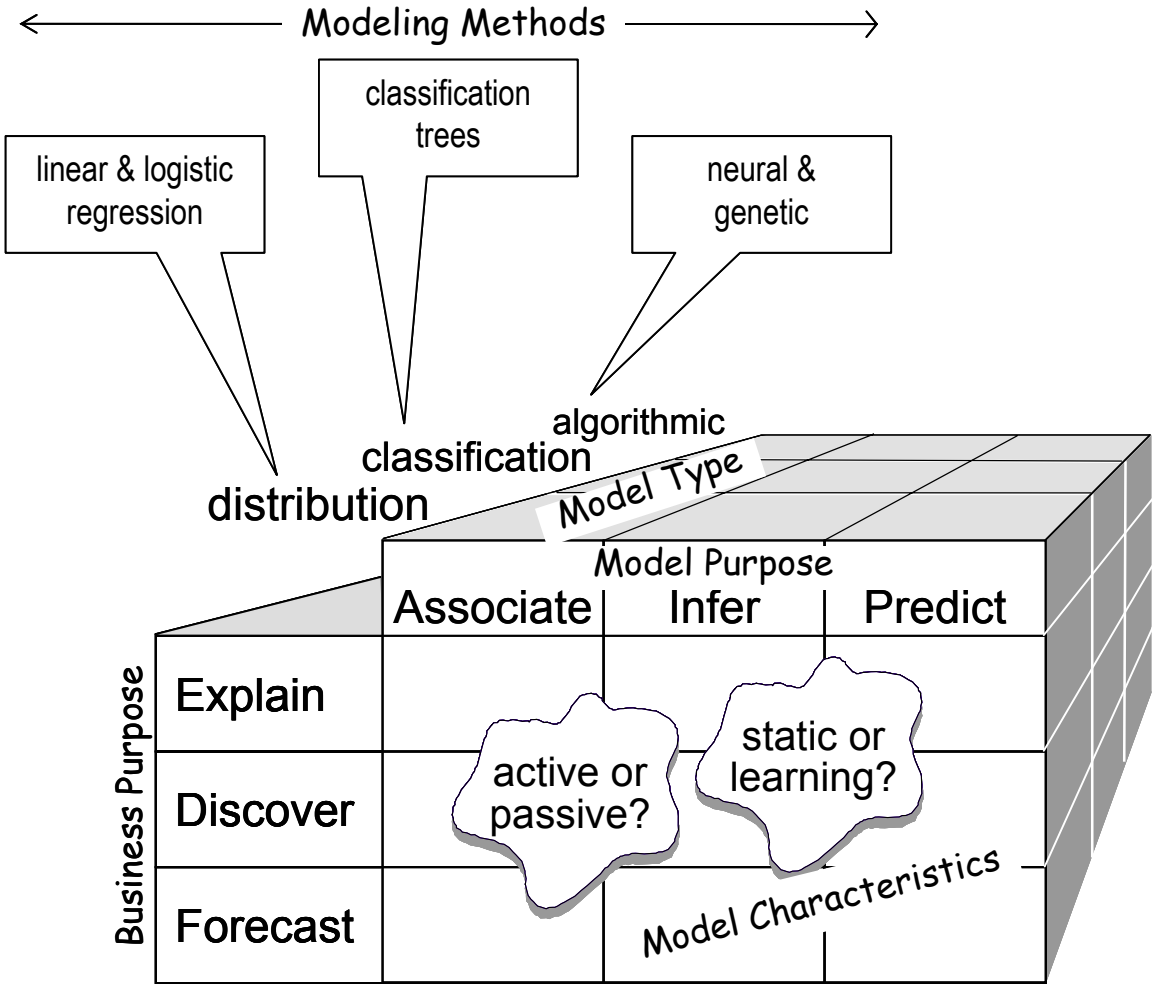
The following approach is recommended for developing analytical models. It is important to use business concepts, terminology, semantics and naming conventions whenever possible.

1. Define the Problem in Business Terms
    a. Explain goals and success factors of the business
    b. Discover key variables used by the business
    c. Forecast possible business benefits and model limitations
2. Choose the Modeling Technique
    a. Linear Regression
    b. Logistic Regression
    c. Neural Network, etc
3. Get the Data
    a. Acquire data sources
    b. Understand the true nature of each potential variable
    c. Audit the data quality in terms of modeling requirements
4. Prepare the Date
    a. Cleanse the data to meet the quality requirements
    b. Sample sets of data to evaluate it for quality
    c. Structure the data into a suitable data store
5. Build the Model
    a. Define a model to meet the requirements
    b. Frame the limitations, constraints and assumptions.
    c. Structure the model to meet the business requirements
6. Apply the Model
    a. Associate variables together
    b. Infer new information the model is telling you
    c. Predict new results that can be validated with test data
7. Evaluate the Results and Iterate back to 5
    a. Test the model against test data sets
    b. Score the model results
    c. Re-Sample the data to improve its scoring
8. Deploy the Model
    a. Implement in a software application
    b. Monitor results of the model and the decision quality
    c. Maintain, tune and calibrate the model

# Modeling

## Using Analytic Models




Understand the business purpose.
Determine the modeling purpose.
Identify the model type.
Determine the modeling method.
Define the model characteristics.
Prepare the data.
Build and test the model.
Deploy the model.
Apply for decision making and action taking.

# Modeling

## Using Analytic Models

### MEETING BUSINESS REQUIREMENTS

It is important that all analytic models be applied to business problems that were intended to address. Every model has a set of capabilities, limitations, constraints and ranges of applicability. A good model applied to an inappropriate situation or problem will produce disappointing results. The framework shown on the facing page shows how the following business and modeling concepts are related to each other.

- Business purpose defines what the business wants to accomplish
  - Explain
  - Discover
  - Forecast
- Modeling purpose defines the capability of the model
  - Associate
  - Infer
  - Predict
- Model characteristics defines the performance of the model
  - Responsive
  - Adaptive
  - Etc.
- Model type defines the category of models, for example, use existing formulae, develop new formulae and develop the appropriate structure.
  - Statistical distributions
  - Classification
  - Algorithmic
    - Deterministic – does not consider randomness
    - Stochastic – considers randomness
- Modeling methods are the fundamental tools available to create a variety of modeling solutions
  - Regression
  - Neural Nets
  - Genetic Algorithms
  - Etc.

It is important that alignment and understanding is achieved at both the business and technical levels to ensure that the "right" model is developed using an appropriate set of techniques and that is used to address the "right" business problem or opportunity.

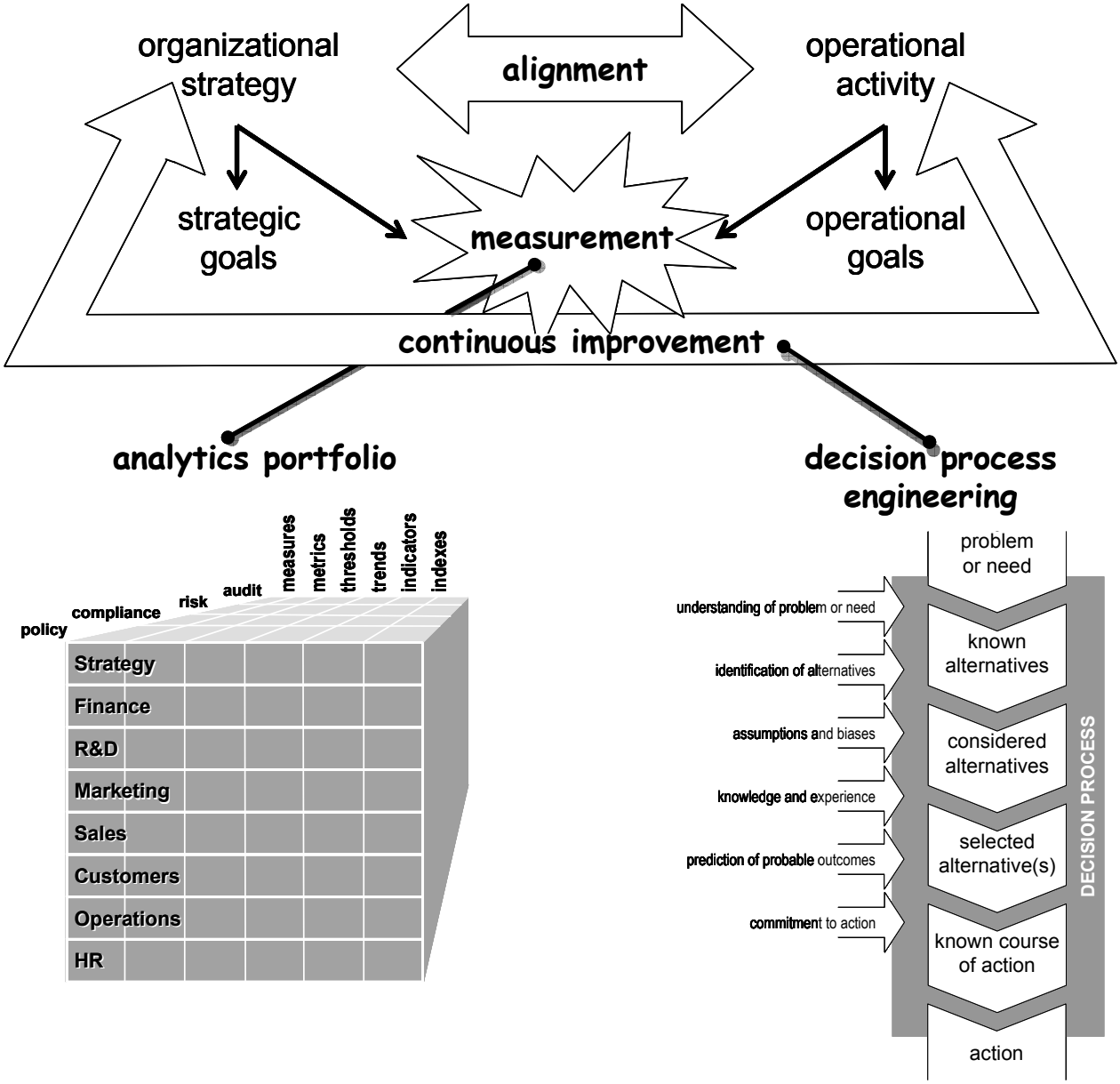

# Module 3

## The Analytics Enabled Business

This page intentionally left blank.

# Business Environment

## Analytics Culture



# Business Environment

## Analytics Culture

**BACKGROUND**

In his recent paper *Competing on Analytics*, published in the Harvard Business Review, Thomas Davenport identified some common traits found in organizations that have successfully used analytics to their competitive advantage.

- widespread use of modeling and optimization
- experiments used to evaluate "lift" from strategies
- followed an enterprise approach for information management
- advocated by senior executive
- CEO's lead the charge
- found and maintained the right focus
- built the right culture
- hired the right people to make use of the data
- developed a strategy based on analytics
- implemented the right technology (hardware & software)
- developed and executed a data strategy

These traits can be grouped into the five categories described below.

**ALIGNMENT**

Alignment is the organizational characteristic that results when all investments, resources, initiatives, people, processes and strategies are focused on achieving common goals. Success in analytics heavily depends on achieving levels of alignment between organizational strategy and operational activity.

**MEASUREMENT CULTURE**

A culture based on using measurements to monitor and improve business performance implies that managers and staff know what to measure, they know why they are measuring it, they trust the measurements, and they know what actions they should take in response to the measurements.

**CONTINUOUS IMPROVEMENT CULTURE**

A culture based on Continuous Improvement exists when all employees take ownership for success. There is an open and collaborative atmosphere to support ideas and compensation rewards innovation.

**ANALYTICS PORTFOLIO**

Managing analytics as a portfolio means that a diverse set of approaches and techniques are applied to a broad set of business opportunities.

**DECISION PROCESSES**

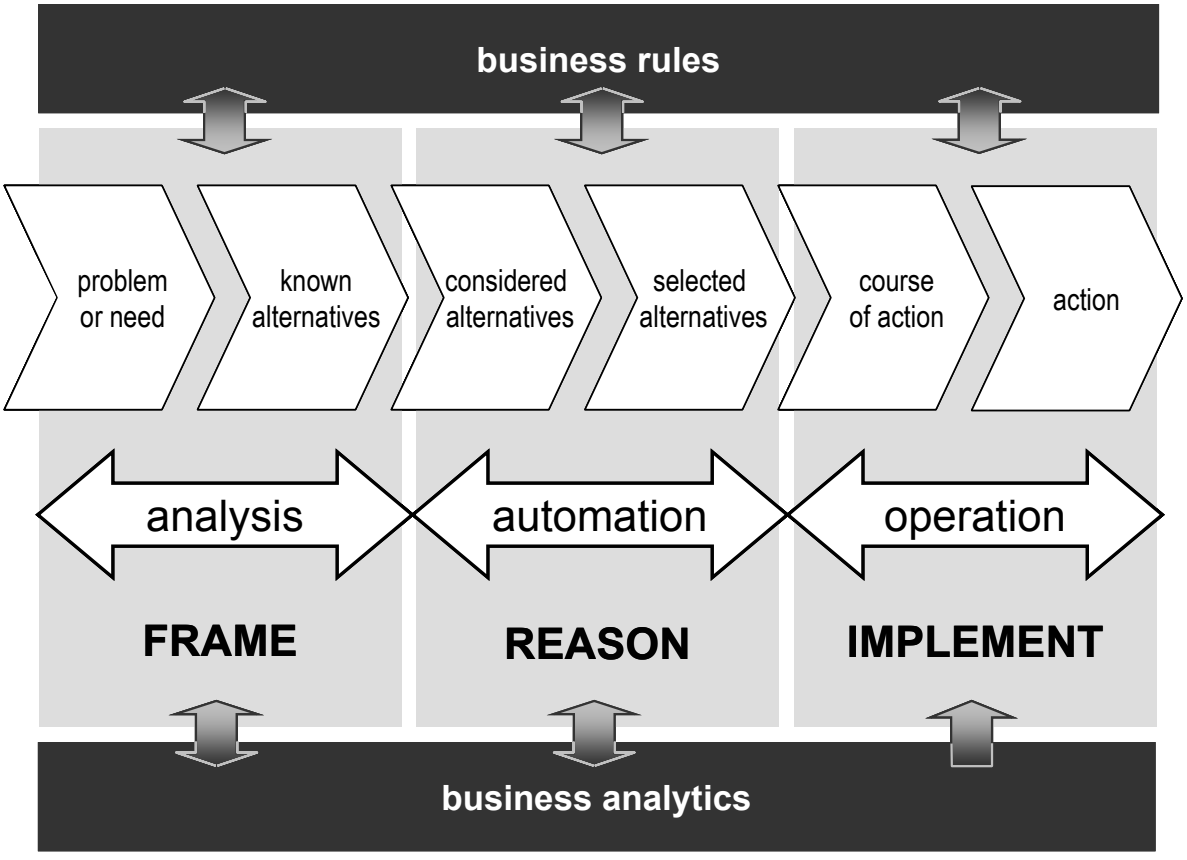
Following well defined and consistent decision processes at an enterprise level is a foundation and basis for achieving long term success and a competitive advantage with analytics.

# Business Environment

## Decision Processes


strategic decisions vs. **operational decisions**
one-time decisions vs. **recurring decisions**
conflict decisions vs. **routine decisions**



# Business Environment

## Decision Processes

### AUTOMATING DECISION PROCESSES

The drive for organizations to automate their processes is based on the competitive needs to increase productivity, enhance quality, maintain consistency and implement best practices. Decision processes are also candidates for automation for the same reasons. We need to understand the opportunities, benefits, costs, limitations and risks that full or partial automation of the decision processes would introduce into our organization. The following two questions need to be answered.

1. Can a specific decision be automated?
2. Should the specific decision be automated?

The answer to the first question is based on the characteristics of the specific decision. As discussed earlier, decisions that are operationally oriented, carried out on a frequent basis and have a routine level of complexity are quite often candidates for automation. Decisions in this category usually have the following attributes.

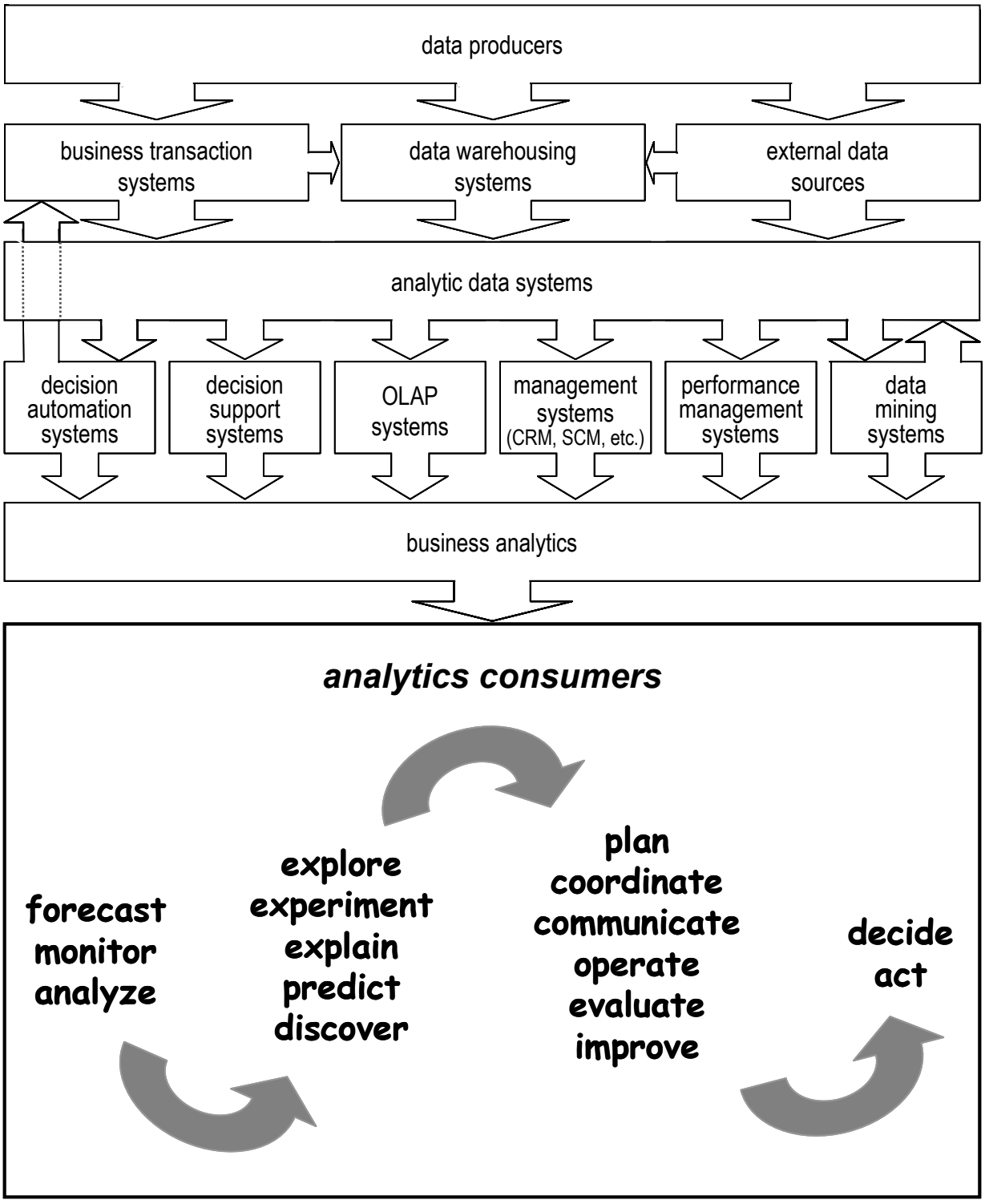
- the situation or problem is clearly understand by all stakeholders
- information is available that can identify the need for a decision
- there are only a few possible alternatives
- measures are available that can estimate the value of each alternative
- the evaluation of alternatives can be clearly made based on available measures
- all input information and outcome measures are quantitative

Other issues related to the first question are related to the degree of automation being considered. The full decision process includes the Framing, Reasoning and Implementation phases. All three of these phases can potentially be automated. The key issues that must be assessed are related to decision attributes, the automated availability of sufficient data to drive each phase and whether the action resulting from the process can also be automated and provide measurements to evaluate its performance. The second question should be answered based on a business case to evaluate the costs and benefits of decision automation.

A common implementation is to automate the reasoning phase of the process. Many of the barriers for automating the framing and implementation phases are related to the actual business processes that are being analyzed within the Decision Processes.

# Capabilities

## Analytic Usage



# Capabilities

## Analytic Usage

### APPLICATIONS OF ANALYTICS

At the final stage of the analytics supply chain, the analytics consumers combine their own knowledge and experience with the information provided to them to make a decision, take an action and impact the business to some degree. The types of activities that the analytics Consumers carry out to improve some aspect of the business are related to answering questions about past, present of future behavior of the organization.

Senior managers or executives carry out the following tasks as analytics consumers.

- Forecast to estimate what future activities will be
- Monitor to ensure that current activities are in control
- Analyze to determine why certain results occurred in the past

At a more detailed level, analysts may carry out the following activities

- Explore data to find new cause/effect relationships
- Experiment to assess the quality of previous decisions
- Explain why specific and surprising results occurred
- Predict future results based on actions to be taken today
- Discover a pattern or relationship that was previously unknown

Managers will use the results of the analysts work to carry out their own responsibilities at a process level. They will do the following activities as analytics consumers

- Plan
- Coordinate
- Communicate
- Operate
- Evaluate
- Improve

Managers and their staff will also use business analytics to make many informed decisions and take actions within the business as a result. An integrated approach to business analytics has a wide and diverse range of application potential across the management and operations spectrum.


# Module 4

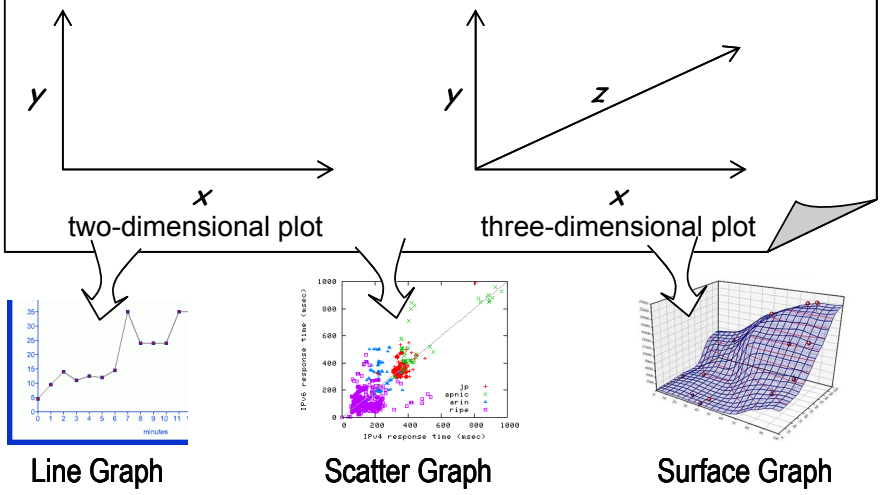
## Analytics Delivery

This page intentionally left blank.

# Visual Delivery Methods

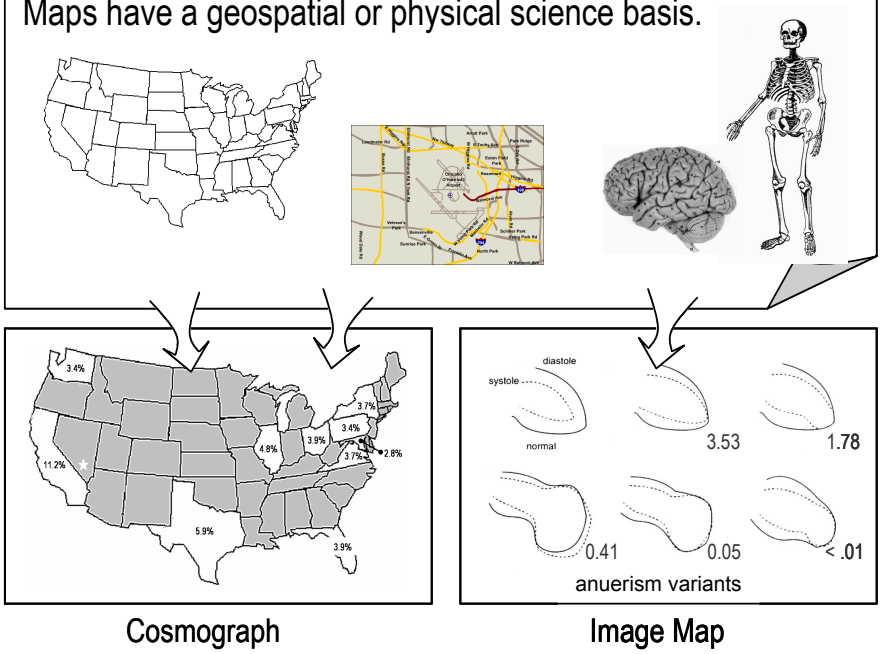
## Plots and Maps

**Plot:** Shows relationships among *variables* using the values as coordinates for placement along two or more axes. Plots are based on Cartesian mathematics.




Line Graph Scatter Graph Surface Graph

**Map:** Shows the values of *variables* relative to their location in two or three dimensional space of something physical. Maps have a geospatial or physical science basis.




Cosmograph Image Map

# Visual Delivery Methods

## Plots and Maps

**DESCRIPTION**

Plots and maps are two different graphical techniques useful for presenting and communicating information that was created from a business analytics study or investigation.

A Plot shows relationships among variables by using the data values as coordinates for placement along two or more axes. Most business analytics plots are based on Cartesian coordinates. Each axis is at right angles to the other(s). Data can be plotted as two-dimensional or three-dimensional plots. The types of plots are diverse. Three examples plots shown on the facing page are a Line Graph, Scatter Graph and Surface Graph.

A Map shows the values of variables relative to their location in two or three dimensional space of something physical. Maps have a geospatial or physical science basis. Examples of maps are shown on the facing page.

**PURPOSE**

The purpose of graphical representations of data is to present the shapes contained in the data values and in the relationship between data values.

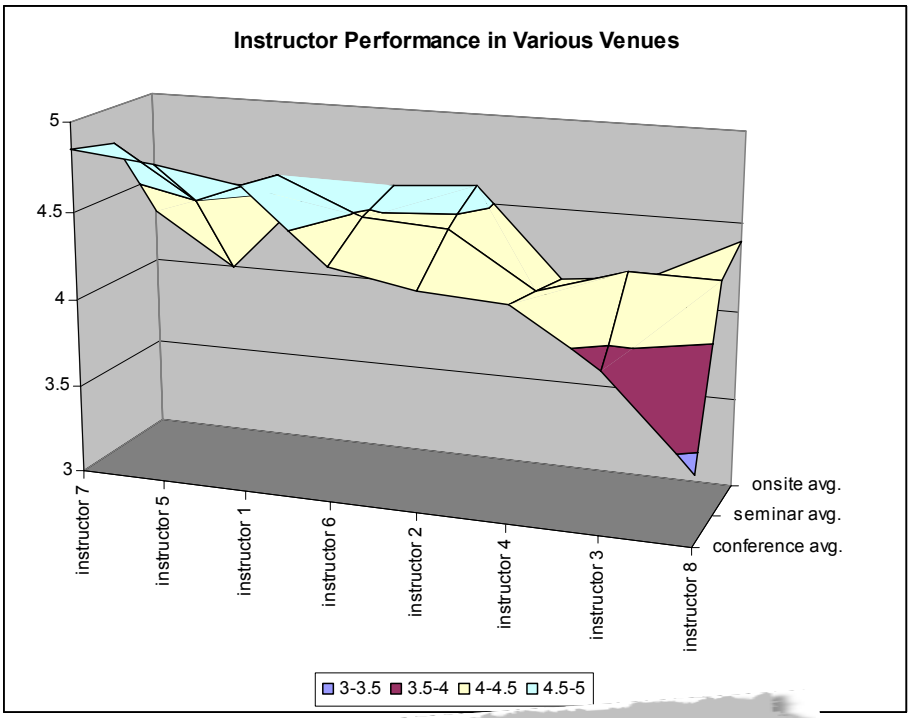
**APPLICATIONS**

Graphical representations of data using plots or maps are used when there is a requirement to analyze and identify relationships among many data values.

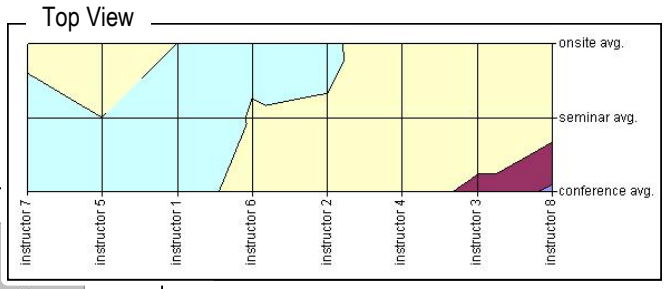
 

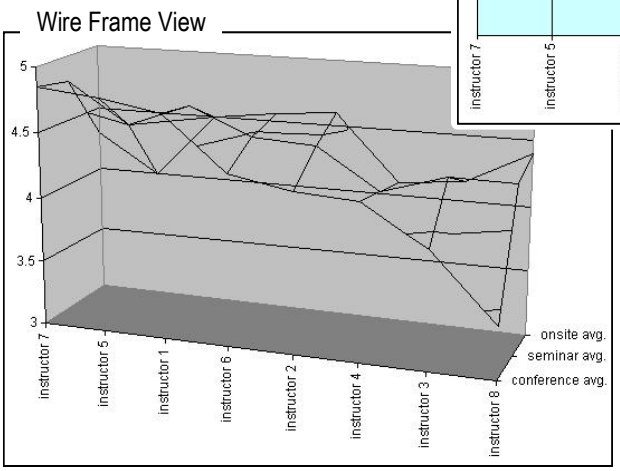
# Charts and Graphs

## Surface Graphs




three-dimensional view of values and trends shown in a continuous curve.



# Charts and Graphs

## Surface Graphs

**DESCRIPTION**

Surface graphs provide a visual presentation of the relationships that exist between three variables at a time.

This type of graph presents a more complex picture than the 2-dimensional styles discussed earlier. They are more difficult to interpret but they convey more information.

Surface graphs can be presented in two different styles. The first one is called a wire frame view. An example of this is shown on the facing page as the first graph. It presents a 3-dimensional perspective of the surface plat representing the relationships among the three variables. A second style is to project a 2-dimensional view of the surface from different perspectives. In the example, the project view is from the Top perspective. Variations of data in the Z-axis are shown as shaded areas enclosed by contour lines. All of the values within a shaded area have similar values. The number of contour lines chosen to represent data values in the Z direction is based on the granularity of the data and the choice of scales.

**PURPOSE AND USAGE**

Surface graphs use be used sparingly to communicate a message that requires the full perspective of a 3-dimensional perspective. Because of the extra complexity that results from communicating the additional information, this type of graph is more difficult to interpret. The trade off between information density and ease of interpretation must be carefully considered.

Another important point to consider when creating a surface graph is the impact that the choice of sorting and scaling has on the message being communicated. Distorted views and miscommunication of the message are possible due to poor choice of these factors. This is also a concern with other graph types as well, but becomes more important as the graphs become more complex.


# Module 5

## Analytics Design

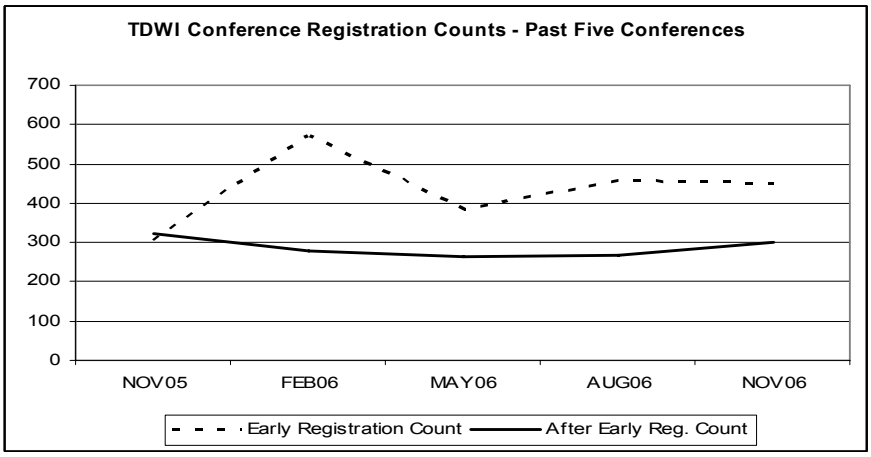
This page intentionally left blank.

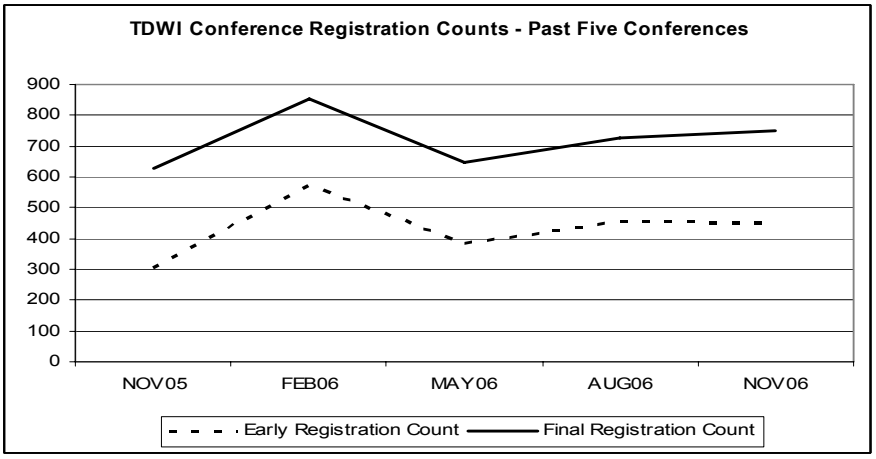
# Presentation and Audience

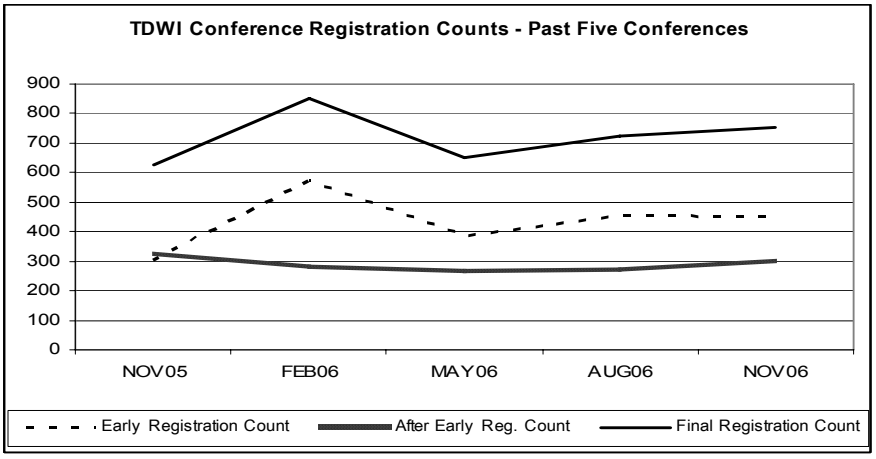
## Comprehension, Interpretation, and Misinterpretation

**Which graph best represents the data from the table?**

| | NOV05 | FEB06 | MAY06 | AUG06 | NOV06 | TOTAL |
|---|---|---|---|---|---|---|
| Early Registration Count | 305 | 573 | 385 | 457 | 451 | 2171 |
| After Early Reg. Count | 323 | 278 | 264 | 268 | 300 | 1433 |
| Final Registration Count | 628 | 851 | 649 | 725 | 751 | 3604 |



# Presentation and Audience

## Comprehension, Interpretation, and Misinterpretation

### MENTAL MODELS

People have mental models stored in their minds that have evolved and developed throughout their lifetime. Individual models are based on a person's experience, observations, education, values, beliefs, relationships and environment. It is through this mental model that reality is perceived, filtered and interpreted. Mental models influence and govern how different individuals might observe the same physical situation at the same time and make separate and contradictory conclusions about what was observed. Our mental models shape how information, graphics, situations and events are interpreted, filtered and perceived.

Designers of graphical reports and presentations must have some understanding of the mental models embedded in the minds of the target audience for the presentation. The graphical technique used for presenting information must align with the expected mental models of the audience.

### EXAMPLE

The example on the facing page helps to explain some of the design issues facing a report or graphics designer. The information that needs to be presented in some graphical format includes three separate counts of conference registration for the Past Five TDWI Conferences. The registration counts are classified as "Early", "After Early" and "Final". The designer must decide which graphical format should be used to communicate the information to the expected target audience.

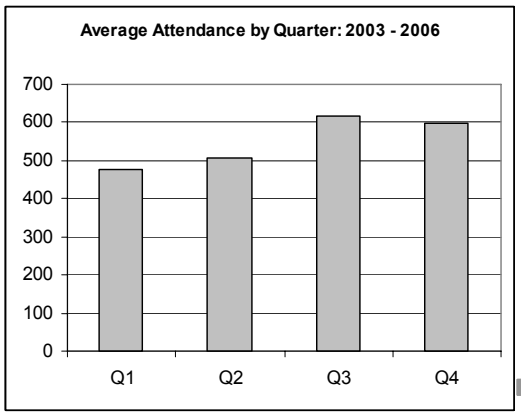
All three graphs in the example accurately reflect the raw data. Which graph is less likely to be misinterpreted? The first graph shows "Early" and "After Early" counts. The second one shows "Early" and "Final" counts. The third one shows all three variables, "Early", "After Early" and "Final". Based on the expected mental models and purpose of the graph, the designer could select any of these graphs.

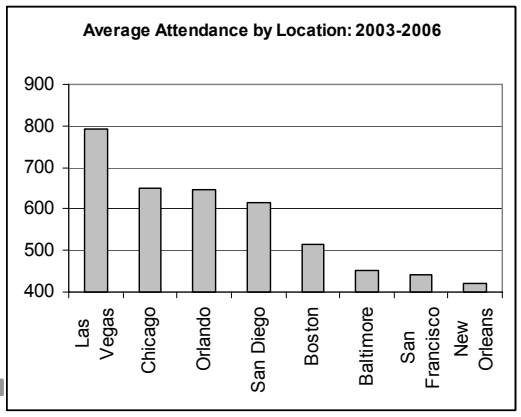
If the purpose is to show the Whole-Part relationship between the "Final" counts and the two component parts of "Early" and "After Early", the middle graph shows this breakdown clearly. However, if it is important to directly identify from the graph the value of "After Early" counts, then the first or third graphs work better.

# Presentation and Usage

## Anticipating Risk



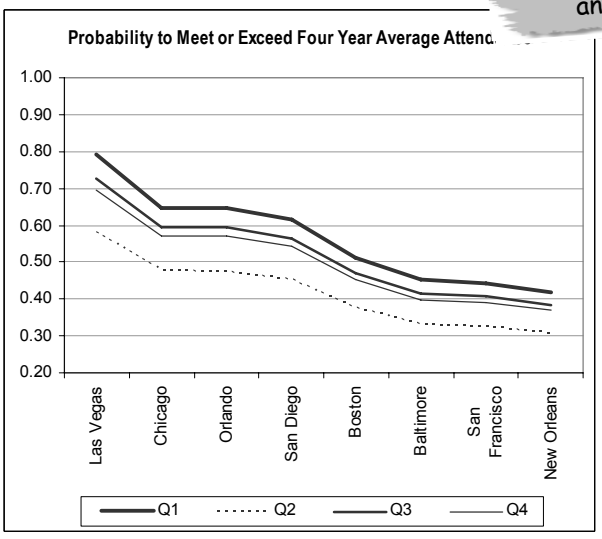



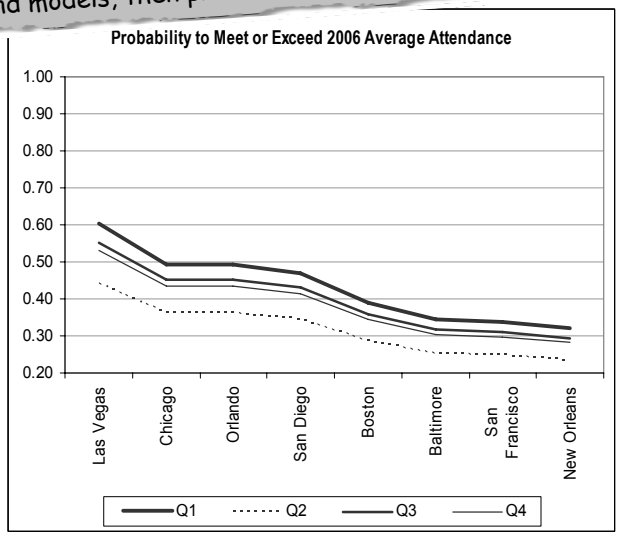

Using what we know from the data to estimate and illustrate probability of success or failure to meet goals …

| 2006 | Q1 | Q2 |
|---|---|---|
| Las Vegas | 0.60 | 0. |
| Chicago | 0.49 | 0.36 |

| 2003-2006 | Q1 | Q2 | Q3 | Q4 | |
|---|---|---|---|---|---|
| Las Vegas | 0.79 | 0.58 | 0.73 | 0.70 | .41 |
| Chicago | 0.65 | 0.48 | 0.59 | 0.57 | .34 |
| Orlando | 0.65 | 0.48 | 0.59 | 0.57 | .30 |
| San Diego | 0.62 | 0.45 | 0.57 | 0.54 | .30 |
| Boston | 0.51 | 0.38 | 0.47 | 0.45 | .28 |
| Baltimore | 0.45 | 0.33 | 0.41 | 0.40 | |
| San Francisco | 0.44 | 0.32 | 0.41 | 0.39 | |
| New Orleans | 0.42 | 0.31 | 0.38 | 0.37 | |

… where estimates are realized as derived data based upon assumptions and models, then presented visually.



# Presentation and Usage

## Anticipating Risk

**OVERVIEW**

Risk is a measure of uncertainty. It is a combination of two concepts. The first one is an estimate of the probability of some negative event happening. The second one is an estimate of the impact if the event does occur. Impact may be expressed in financial terms, such as dollars or it may be expressed in operational terms such as inventory count or attendance counts. When the probability of the event is multiplied by the estimated impact of the event the risk exposure is determined. This is the expected impact if the event happens. Typically this is calculated in dollars but it could also be calculated in operational terms.

**INFORMATION NEEDS**

Carrying out a Risk Assessment helps business managers understand where their significant exposures are. Well run organizations develop risk mitigation plans to help them reduce exposures in selected areas of the business. The information needed to monitor a Risk Management program is a list of significant events, the probability of the event occurring and an estimated impact if the event does happen.

Events can be grouped by business activity, process, schedules, projects and programs. Probabilities are estimated from historical data or from sources such as actuarial tables.

**EXAMPLE**

The example on the facing page shows an analysis carried out for TDWI attendance data at their conferences. The risk that is analyzed is related to achieving specific attendance goals at a future conference. There are a number of factors that historical data has indicated will impact meeting attendance goals. It is related to the city and the season of the conference. Different cities and seasons have historically shown that are different probabilities of meeting their targets.

The probabilities and estimated impact data can then be applied to planning of future conferences. Based on the future city and season of the conference, different risk mitigation strategies can be put in place. Examples of this might be related to more aggressive marketing or more flexible contractual terms with the conference venue company.

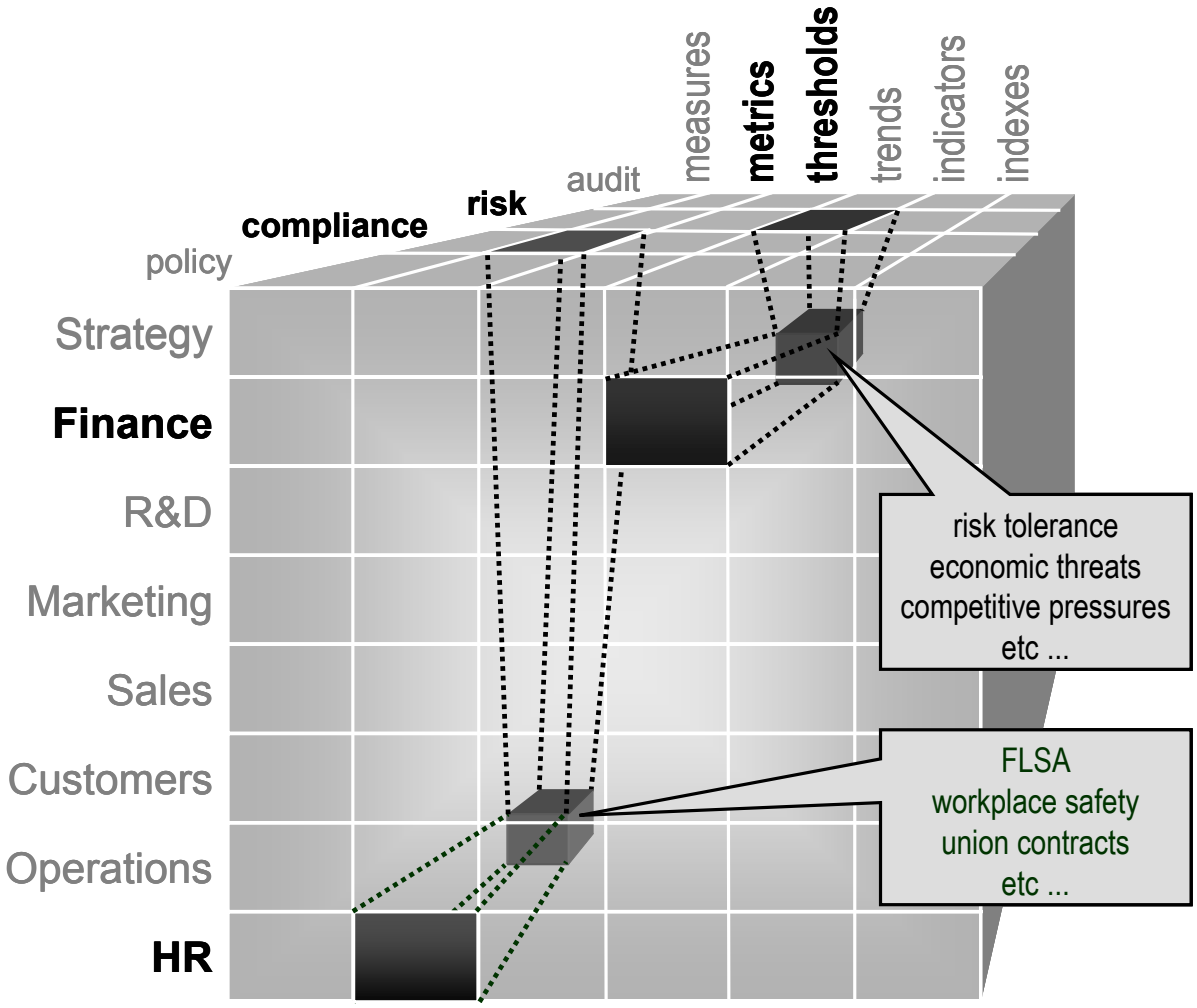

# Module 6

## The Analytic Portfolio

This page intentionally left blank.

# Analytic Content

## Matched to Decision Processes



# Analytic Content

## Matched to Decision Processes

### OVERVIEW

The decision processes that are carried out in an organization have been discussed in an earlier section. Depending on the culture and management style, there are different styles and approaches used for making decisions. Some decisions are made frequently about well understood options and constraints. Other decisions are made infrequently about a subject or its options that are poorly understood. Another perspective about decision processes is the amount of detail or rigor that must be followed when identifying or evaluating alternatives. Some organizations may be very conservative and want detailed analysis before moving forward. Others may find higher risk levels to be acceptable, and make decisions with a high degree of uncertainty and after somewhat high level analysis. All of these decision making styles and approaches are found in most organizations.

### ALIGNMENT TO DECISION PROCESSES

A key challenge for adopting business analytics in an organization is to ensure that the category of decision and the style or existing process used to make that category of decision is aligned to the applications contained in the analytics portfolio.

This area of alignment management is important because if the results from an analytic solution are not used, believed or acted upon by the business decision makers, then the investment in analytics will not return value. The decision categories, decision processes and decision styles must be understood and aligned with the analytic capabilities.

Another issue that must be resolved is management commitment to consistent use of analytics in decision making. Consistency is the key factor. There may be situations where the analytics support the manager's intuition and "gut feel," and others where the analytics are contrary to a manager's mental model and intuition. This is a difficult situation to deal with. One approach is to implement a series of business experiments in parallel to the decision processes that will help managers to gain confidence in analytic models, and to also understand their limitations to know when manual override may be the appropriate solution. The key message is that there needs to be a reasonable degree of alignment between decision process needs and constraints and analytic portfolio capabilities.


# Module 7

## Summary and Conclusion

# Beyond Fundamentals

## What Else Do You Need to Know?

### BEYOND THE BASICS

The content of this course introduced and described the main topics and subjects that make up the business analytics discipline. However, as mentioned earlier, this is a complex topic that requires further study to develop deeper understanding of the major areas within business analytics.

Some of the key areas for further study and skill development are:

- Analytic Requirements
  - Requirements Gathering Techniques
  - Balanced Scorecard and Other Management Approaches
  - Applied Analytics Techniques such as Planning, Monitoring and Forecasting
- Statistics
  - Concepts and Terminology
  - Statistical Methods
  - Statistical Modeling
  - Forecasting with Statistics
- Visualization
  - Visualization Techniques
  - Design of Charts and Graphs
  - Dashboard and Scorecard Design
  - Testing Visual Designs
- Measurement and Metrics
  - Metrics Definition and Design
  - Measurement System Concepts
  - Integrating Business Metrics
  - Data Modeling for Metrics
- Analytics Technology
  - Technology Architecture
  - Technology Administration
  - Evaluating and Selecting Tools
  - Data Mining Technology