

# Predictive Analytics: Low-Risk / High-Impact Implementation Course Outline

## INTRODUCTION

- Predictive Analytics Definition & Core Concepts
- Terms Used in Today's Analytics Environment
- Statistics vs. Predictive Analytics: Complimentary Technologies
- Goal-Driven Analytics
- "The Main Thing is to Keep the Main Thing the Main Thing."
- What is the Goal of the Analysis Project?
- What are the Performance Metrics for Evaluating Success of the Decision Process?
- What is the Behavior that Impacts Performance?
- Is There Sufficient Data for the Target Behavior to Develop an Adequate Model?
- The Modeling Practice Framework™
- The Analytic Project Team
- Analytic Opportunity Identification

## BUILDING THE GOAL-CENTERED ANALYTICS OPERATION

- Attracting and Hiring the Right Analytic Talent
- The Roles and Functions of the Fully-Formed Analytic Project Team
- Specialization in Analytic Project Teams
- Analytic Opportunity Identification, Qualification and Prioritization
- Organizational Resistance and Developing a Culture for Change
- Project Failure is Not the Worst Outcome
- Staging the Organizational Mind Shift to Data-Driven Decisioning
- Motivating Adoption by Domain Experts, End Users and Leadership
- Recording Ongoing Organizational Changes
- Monitoring and Advancing Organizational Analytic Performance
- "Democratizing" Analytics: Advantages and Risks of "Self-Service"
  - Tableau
  - Watson Analytics
  - Establishing Performance Dashboards
- Standing Up an Agile Analytic Modeling Factory
- Knowledge Retention and Skill Reinforcement

## Phase 1: ASSESS

- Comprehensive Project Assessment
- Organizational Objectives
- Motivation and Alignment of Leadership
- Behavior(s) of Interest
- Environmental Constraints
- Operational Requirements

- Identification of Scarce Resources
- Threats to Project or Process
- Defining Baselines and Evaluating Project Potential

#### Phase 2: PLAN

- Project Definition: The Blueprints for Actionable Analytics
- The Three Steps of Model Development
  - Train
    - Construct Candidate Models
    - Sample Size Requirements
    - Matching Modeling Methods to Project Type
  - Test
    - Decision Cycle Identification
    - Sample Size Requirements
    - Performance Evaluation of Candidate Models
  - Validate
    - Operational Decision Consistency
    - Strategy Specification
    - Validation Study Requirements

#### Phase 3: PREPARE

- Know Your Data and How it Was Generated
  - Importance of Face-to-Face Interviews with those Close to Data Collection
  - Difficulty of Obtaining Appropriate Data
  - Data is Never Presented on a Silver Platter
  - What Data Should I Include in My Analytic Sandbox?
  - Some Data is Not Math-Compatible
  - What Does the Outcome or Target Variable Look Like?
  - What Data Representations Should I Use?
  - What Data Transformations Should Apply?
  - How Do I Select Variables for My Model?
    - Beware of Dependent Variables Masquerading as Input Variables
    - Example: Response to Credit Card Solicitation vs. Number of Plastics Used
  - How do I construct the Train / Test / Validate data sets?
  - Structuring Data for Modeling

#### Phase 4: MODEL

- Process Objectives and Goals
- Experimental Design: TRAIN Revisited
- Selecting Condition Attributes
- Analytic Model Assessment
  - Statistics
  - Tables

- Graphs
- Resampling / Bootstrapping
  - Ensemble Modeling Conceptualization
  - Bias – Variance Tradeoff
  - Classification Models
- Logistic Regression
- Decision Trees / Boosted Trees / Random Forests
- K-Nearest Neighbor
- Neural Networks
  - Forecasting Models
- Linear Regression
- Bayesian Regression
- Neural Networks
- Multiple Models are Usually Needed
- Perfect Correlation is Not a Good Thing
- and No Correlation is a Waste of Time

#### Phase 5: VALIDATE

- Does Our Math Make Business Sense?
- Organizational Performance is the Only Priority
- Analytic Metrics Do Not Equal Organizational Performance Metrics
- Establish a Model Competition
- How to Pick a Challenger
- Confirming That a Valid Challenger Was Selected

#### Phase 6: DEPLOY

- Evaluating the Expected Performance of our Challenger
- Adoption by Domain Experts
- Adoption by the Operational Environment or End Users
- Adoption by Leadership and Stakeholders
- Project Failure is Not Our Worst Outcome...

#### Phase 7: MONITOR

- Adapting to a Changing Environment
- The Environment Always Changes
- Our Organizational Goals Also Change
- Measuring Primary Model Performance Degradation
- Determine When to Install A Hot-Spare Challenger Model
- Determine When to Refresh the Full 7-Phase Model Development Cycle

#### SPECIAL TOPICS

- The Complexity of Large-Scale Analytics
- Start with the Low-Hanging fruit: Structured Data

- Unstructured Data May be 90% of Overall Content, But Usually Holds Only 10% of the Value
- Specialization in Project Teams
- The Power of Adapting Core Analysis Skills
- The Even Greater Power of Honing Soft Skills
- Where to Go from Here
- Resources to Get You on Your Way

#### RESOURCES

- Analytic Glossary
- Recommended Books
- LinkedIn Groups
- Data Repositories
- Predictive Analytics Across Social Media
- Webinars, Courses, Conferences