PROBLEM SOLVING USING ADVANCED ANALYTICS - PRACTICAL CONSIDERATIONS

Alper Murat, Ph.D.

Industrial & Systems Engineering Department
Wayne State University

alpmur@gmail.com
Tel: 313.443.4429
Outline

- Overview of Business Analytics
  - BI, Data Science, Analytics (Business)
  - Descriptive, Diagnostic, Predictive, **Prescriptive**

- Case Examples
  - Near Real-time Decision Support System in Healthcare Delivery Systems
  - Streamlining Patient Flow in Emergency Department
  - Assortment Planning for Configurable Products in Automotive

- Key Takeaways
What is Business Analytics?

- **Business Analytics is:**
  - Understanding and learning from past relationships,
  - Predicting and controlling future outcomes,
  - Making decisions to generate business value.

  **using:**
  - data
  - information technology
  - statistical analysis
  - quantitative methods, and
  - mathematical or computer-based models

- **Challenges:**
  - Massive amounts of data (Big Data) and accuracy
  - Rapid decision making (latency)
  - Multi-disciplinary decision making (siloed data and conflicting goals)
  - Efficient Balancing of \( \rightarrow \) Data+Analytical Approach+Business Value
**BI, Data Science, Analytics (Business)**

- **Business Intelligence (BI):** Forrester
  
  “Methodologies, processes, architectures, and technologies for analysis, reporting, performance management, and information delivery.”

- **Data Science:** Dataversity
  
  “Process of deriving understanding, significance, and form from the myriads of variety of structured and unstructured Data that Big Data can encompass.”

- **Analytics:** Gartner
  
  “Statistical and mathematical data analysis that clusters, segments, scores and predicts what scenarios are most likely to happen.”

- **Business Analytics:**
  
  “Tailored analytics and BI for non-technical and business users to create business value.”
<table>
<thead>
<tr>
<th><strong>Business Intelligence</strong></th>
<th><strong>Data Science</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Perspective</strong></td>
<td>Looking backwards</td>
</tr>
<tr>
<td><strong>Actions</strong></td>
<td>Slice and Dice</td>
</tr>
<tr>
<td><strong>Expertise</strong></td>
<td>Business User</td>
</tr>
<tr>
<td><strong>Data</strong></td>
<td>Warehoused, Siloed</td>
</tr>
<tr>
<td><strong>Scope</strong></td>
<td>Unlimited</td>
</tr>
<tr>
<td><strong>Questions</strong></td>
<td>What happened?</td>
</tr>
<tr>
<td><strong>Output</strong></td>
<td>Table</td>
</tr>
<tr>
<td><strong>Applicability</strong></td>
<td>Historic, possible confounding factors</td>
</tr>
<tr>
<td><strong>Tools</strong></td>
<td>SAP, Cognos, Microstrategy, SAS</td>
</tr>
<tr>
<td><strong>Hot or not?</strong></td>
<td>So 1997</td>
</tr>
</tbody>
</table>
Analytics Spectrum

Business Intelligence to Advanced Analytics

Data Science

Advanced Business Analytics

Business Intelligence

Michigan Spark Users Group Meeting – 14 May 2015
BI Capabilities

- Structured and Unstructured Data Modeling
- Dashboarding: Standard & Ad-hoc Reporting and Dashboards
- Exploration: Analysis/Query/Drill-down/Discovery
- Statistics: Trends and Pattern Discovery
- Visualization
Diagnostic and Predictive Analytics

Diagnostic: Why it happened?
Predictive: What will happen (if/when...)?

- Basic and Advanced Statistics
  - Single vs Multivariate Statistics; Bayesian and Monte Carlo Statistics
- Descriptive Data Mining:
  - Clustering
  - Association Analysis
  - Feature Extraction
- Linear / Logistic Regression
- Time-series analysis

- Predictive Data Mining:
  - Classification
  - Decision Trees
  - Parametric/Nonparametric Regression
  - Neural Networks
- Simulation
- Survival Analysis
- Text analytics
Operations Research (OR) is the application of advanced analytical methods (optimization) to help make better decisions:

- Integrates computer science, mathematics, statistics and probability theory
- Management Science (MS) vs Operations Research (OR)

OR Model: Decisions, Objective(s), Constraints/Restrictions
Discrete vs Continuous Optimization

- **What Matters:** Decisions that are ‘truly’ discrete in nature.
  - Ex: Number of TL shipments Detroit-NYC vs binary decision to launch the production of a new product line.

- **Business problems with discrete decisions are more difficult to optimize.**

  ![Graph showing discrete vs continuous optimization](image)

  Function to maximize: \( f(x, y) = 6x + 5y \)
  - Optimum LP solution: \((x, y) = (2.4, 3.4)\)
  - Pareto optima: \((0, 4), (2, 3), (3, 2), (4, 1)\)
  - Optimum ILP solution: \((x, y) = (4, 1)\)
Local vs Global Optimization

- Local optimum is the best decision in the vicinity.
- Global optimum is the best decision among all admissible decisions.

- Local vs Global optimality is a characteristic of the decision problem
  - Trade-offs, characteristics of decisions, objectives of decision maker.

- Heuristic vs Exact Global Opt.
  - Eg.: Genetic Algorithm vs Divide & Conquer
### Single vs. Multiobjective (MO) Optim.

- Single Objective optimization seeks to find the best decision to optimize a single criteria (e.g., profit).
- MO aims to optimize conflicting objectives that cannot be traded-off a-priori.
- Goal in MO is a set of non-dominated (Pareto optimal) solutions.

- Dynamic systems results in shifting Pareto frontier.
- Uncertainty in MO.
Select balanced, feasible combinations of subsystem options that:

1. Maximizes Performance Score

\[ \sum \text{Performance} \]

- Speed \( w_1 \)
- Acceleration \( w_2 \)
- Fuel Efficiency \( w_3 \)
- Ride Quality \( w_4 \)
- Rate of Fire \( w_5 \)
- SIGMAN \( w_6 \)
- Survivability \( w_7 \)
- Maint. Ratio \( w_8 \)

2. Minimizes Burdens
- Life Cycle Cost
- Risk
- Power
- Thermal
- Weight
- MTBSA etc.

Non-dominated Combinations

Options per Subsystem vs. Theoretical Combinations

<table>
<thead>
<tr>
<th>Options per Subsystem</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theoretical Combinations</td>
<td>2,097,152</td>
<td>10,460,353,203</td>
</tr>
</tbody>
</table>
Simulation: Imitation of real world system

Discrete Event Simulation (DES)
- DES is process oriented, top-down, single control thread, passive entities.
- ABS entity oriented, bottom-up, multiple agent control thread, active entities.

Agent Based Simulation (ABS)

System Dynamics (SD)
- Adoption rate
- Adoption from advertising
- Adoption from word of mouth
- Total population
- Adoption fraction
- Contract rate
Case Examples

Joint work with Dr. Ratna Babu Chinnam, PhD Students and Sponsor Staff:

- Near Real-time Decision Support System (NRT-DSS) in Healthcare Delivery Systems
- Streamlining Patient Flow in Emergency Departments
- Assortment Planning for Configurable Products in Automotive
Near Real-time Decision Support System (NRT-DSS) in Healthcare Delivery Systems
Project with Veteran’s Affairs

Healthcare delivery systems feature an interplay between multiple services, providers/resources, and entities.

Key performance indicators are:

- Access to Care: Waiting times
- Quality of Care: Health outcomes
- Efficiency: Total cost of healthcare services provided

KPI Issues

- Underutilized resources (e.g., Operating Rooms, PACU/ICU, Surgeon, staff, equipment, …)
- Waiting times (patient waiting for surgery, bed, diagnostic test, …)
- Quality of care (poor quality of care due to waiting, rush/chaos, staff satisfaction, …)
Objective: Developing advanced analytics solution to address KPI issues.

Analytics:
- Descriptive: Historical processing rates, process start delays...
- Diagnostic: Delay reasoning
- Predictive: SoS Simulation (operational for proactive management); Text Mining, Classification, Clustering,...
- Prescriptive: Simulation based Optimization for job sequencing, cancellation, task assignment, rescheduling recommendations,...

Users
- Schedulers, frontline and clinics staff, services, patients, management...

Enterprise level BI deployment: In progress
- Dashboards, reports, what-if analysis,...

Experience based co-design: Users in the loop for analytics and BI development
System of Systems Approach
NRT-DSS in Healthcare Delivery Systems

NRT-DSS

- Clinics (GI, ENT)
- Logistics
- Dermatology
- Wards
- CENSITRAC/RTLS
- OR
- Clinics (Eye, Dental, Podiatry)
- SPS / IVN
- Radiology & Labs
- ICU/PACU
- ED
**Motivation:**

- **Quality** and **Effective Sterilization** of Re-Useable Medical Equipment (RME) is critical for **Patient Safety** and **Access** of care in many VAMC departments
  - **Sterile Processing Service (SPS):** Expired Staff Certification, Rushed Processing & Contamination Risks
  - **Operating Room (OR):** Surgery Delays & Cancellations
  - **Dental, Eye, Podiatry, GI:** Appointment Cancellations & Delays
  - **Emergency Department (ED):** Delay in Care & Increased Wait Times

**Objective:**

- Improve the **RME Availability** and **Sterilization Quality**

**Deliverable:** RME Near Real-Time (NRT) Decision Support System for

- **Increased RME Availability** at Clinics and Emergency Department
- **Improved Surgery Schedules** with in-advance RME Availability Info.
- **Effective Scheduling** and **Certification** Tracking of SPS Staff
- **Reduced** Sterilization **Cycle Time** and **Lean RME Inventory**
OR-RME NRT DSS Framework

- Prioritization of RME Sterilization
- RME Delay & Availability Reports
- Appointment & Surgery Case Cancellation Reports
- OR, ED, Clinic, Lab, Radiology, BMS, Utilization Reports
- Expedited RME Orders
- Advisory Schedule Changes
-...

Simulation & Optimization Tools

NRT DSS (Modular Architecture)

SPS/RTLS Input

SPS Output

User/Supplier Input

User/Supplier Output

Sterile Processing Service

GI
ENT
Dermatology

OR

Radiology

ED

Labs

Clinics

Inpatient Wards
ICU/PACU

Logistics

Users/Suppliers

SPS

RTLS & IVN

CHEDISTRAC

SPS/RTLS Input

User/Supplier Output

Users/Suppliers
NRT DSS integrates RME and OR patient flows through multiple input-output relations between SPS and OR models.
Data Flow and Analytical Process

**Database**
- OR Case Schedules
- Clinical Schedules
- ...

**Data Pull**
- Frequency – Multiple times a day

**RME Input**
- Dynamic and on-demand

**Input Resources**
- OR Case Schedules
- Clinical Schedules
- ...

**Input Controls**
- Staff Competency & Prioritization

**Output Reports**
- Generated with every simulation run
- Data stored for future report generation

**SIMULATION & OPTIMIZATION**
- STAFF SCHEDULING
- TASK SEQUENCING
- DELAY REPORTS

**Input Resources**
- SPS Resources:
  - 1. Staffing
  - 2. Equipment

**Output Reports**
- RME Task Sequence
- Staff Schedules
- RME Assignment for Staff Competency
- OR Delay Reports
- Utilization Reports
- ...

**RTLS TagID-RME Mapping**

**OR Schedules**

**User Front Ends**

**Censitrac**

**RME-OR Mapping**

**OR RME Demand**

**RME Inventory**

**NRT DSS**

**Generated with every simulation run**

**Data stored for future report generation**

**To be Completed**
Surgery Duration Prediction

Predictive Analytics

Text-based Source Documents

PreOp Diagnosis
Principal Procedure Description

Text Mining Analytics

Parse
Filter
Extract
Detect

Knowledge

Medical Dictionary / Thesaurus
CPT Code/Description Libraries

Text Features

Requested Durations
Attending ID / Facility

Non-text Features

Assigned CPT Codes

Machine Learning Analytics

Classification Trees
Multi-class SVMs

Training & Prediction

Historical Surgery Duration – Attending / Site

Text Mining & Ensemble Machine Learning

Surgery Duration Prediction Tool

Michigan Spark Users Group Meeting – 14 May 2015
Streamlining Patient Flow in Emergency Departments
**Motivation: Patient Crowding/Boarding**

- **Motivation:** Hospitals with average occupancy levels above 85% routinely experience bed shortages and crises and long patient waiting times
  - *Shortages* result in Excessive Waiting in Emergency Rooms, Ambulance Diversions ...
  - *Missed Opportunities* such as Cancellations of Elective Surgeries ...

- **Industry Interest:**
  - *Singapore General Hospital* – Reduced Bed Turnaround, Waiting, Search, and Hiding
  - *Mount Sinai Medical Center - GE Healthcare:* Exploits near real-time information to attain efficiencies not seen before
    
    “*You can run a hospital at 95% Occupancy, and do it in a way that's Safer, more Sustainable & Reliable - puts less strain on staff than running at 85% - by Truly Optimizing how things work*”
    
    - Jeff Terry, Principal of Clinical Excellence at GE Healthcare
Modeling & Decision-Support Models Allow Proactive Planning

Reduction in Wait Time
Decision-Support Model

Proactive Coordination with Wards

NOTATION:

\( \text{LOS} \): Predicted Patient’s Length-of-Stay in ED  \( p \): Predicted Probability of Patient Admission

\( T_L \): Lead-time for Ward-Bed  \( T_R^* \): Optimal Reservation Slot Time  \( T_B \): Time when Bed is Ready

OPTIMAL ACTIONS: Select Actions that Leads to Least Expected Cost!

Reserve Ward-Bed: Yes or No

If Reserving, What Time (\( T_R^* \))?

Cost of Patient Waiting

Cost of Bed Wastage & Another Patient Waiting

\( T_B > \text{LOS} \)

No Cost: Reservation Cancelled Early

\( T_B < \text{LOS} \)

Cost of Bed Wastage & Another Patient Waiting

\( T_B > \text{LOS} \)

Cost of Patient Waiting

\( T_B < \text{LOS} \)

No Cost

Identify \( T_R^* \) to Min Expected Cost

\( T_B = \max\{T_R^*, T_L\} \)
Two-Stage Approach: Admission -> Target Ward

- Admission Prediction: Binary Classifier
  - Need probabilistic output ($\rho$)
  - Class imbalance issues

- Target Ward Prediction: Multi-class Classifier
  - Admission probability threshold is necessary
  - Class imbalance potential

One-Stage: Ward Admission

- Multi-class Classifier
  - More complex separating hyper-surface (higher VC dimension model)
  - Increased computational burden (Hsu & Lin, 2002)
  - Need probabilistic output ($\rho$)
Classifiers: Support Vector Machine

Attempts to Maximize the “Separation Margin”

- **QP Formulation:**

  \[
  \text{minimize} \quad \frac{1}{2} ||w||^2
  \]
  
  such that \( y_i(w^T x_i + b) \geq 1 \)

- **Strengths:**
  - QP formulation leads to global optimal solution
  - Can handle nonlinear and non-separable classes
  - Very effective in practice

- **Limitations:**
  - Intrinsically handles two classes
  - Large # of kernels (even under sparsity)
  - No probabilistic output

---

**Linear Model:**

\[
y(x, w) = \sum_{m=0}^{M} w_m \phi_m(x) = w^T \phi
\]
Classifiers: Relevance Vector Machine
Bayesian Alternative to SVM

**Offers Probabilistic Output**

Output Probability:

\[ p(t|x, w) = (\sigma(y)^t(1 - \sigma(y)))^{1-t} \]

- **Strengths:**
  - Can be naturally extended to multi-class classification
  - Compact: Yields fewer kernels
  - Probabilistic outputs
    - Indicate ‘confidence of predictions’

- **Limitations:**
  - Non-convex Optimization
  - Slower to train than SVM

(Bishop, 2004)
Iteration #1: Detroit, VAMC

Admission: Naïve Bayes Classifier

Too many false positives!

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Home</th>
<th>Admitted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home</td>
<td>71%</td>
<td>29%</td>
</tr>
<tr>
<td>Admitted</td>
<td>12%</td>
<td>88%</td>
</tr>
</tbody>
</table>

Destination Ward: Support Vector Machines

Too many misclassifications!
Overall Accuracy: 42.7%

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Medicine</th>
<th>Surgery</th>
<th>Psychiatry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicine</td>
<td>26%</td>
<td>22%</td>
<td>52%</td>
</tr>
<tr>
<td>Surgery</td>
<td>16%</td>
<td>57%</td>
<td>27%</td>
</tr>
<tr>
<td>Psychiatry</td>
<td>25%</td>
<td>22%</td>
<td>53%</td>
</tr>
</tbody>
</table>
Investigate Underlying Causes

Data & Model

- **Data Inaccuracy**
  - Ex.: Triage Staff “Patient Complaint” Data Entry Errors

- **System Design/Capacity Induced Failures:**
  - Surgery Ward Staff Leave @ 5PM; Patients are admitted to Medicine Ward
  - Overflow between wards not recorded (e.g., Medicine & Surgery wards)

- **Methodology Related**
  - Text mining not effective: usually based on word/query frequency and similarity
    - Very Short Complaints (2 letters to 3 words)
    - Confusion: call back & back pain
  - Classifier inadequacies
    - Ex.: Lack of prediction confidence
Admission: Naïve Bayes Classifier

**Improved accuracy!**

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Home</th>
<th>Admitted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home</td>
<td>86.9%</td>
<td>13.1%</td>
</tr>
<tr>
<td>Admitted</td>
<td>17.1%</td>
<td>82.9%</td>
</tr>
</tbody>
</table>

Actual Outcome

**Destination Ward: Support Vector Machine**

**Psychiatry accuracy quite low. Small sample size!**

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Medicine</th>
<th>Surgery</th>
<th>Psychiatry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Outcome</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td>93.4%</td>
<td>2.5%</td>
<td>4.1%</td>
</tr>
<tr>
<td>Surgery</td>
<td>21.7%</td>
<td>78.3%</td>
<td>0%</td>
</tr>
<tr>
<td>Psychiatry</td>
<td>78.8%</td>
<td>0%</td>
<td>20.7%</td>
</tr>
</tbody>
</table>
Iteration #3: Relevance Vector Machine

**One-Stage Approach**

Acceptance Threshold
Probability: 0.6

<table>
<thead>
<tr>
<th>Total Volume:</th>
<th>1788</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold Volume:</td>
<td>1672</td>
</tr>
</tbody>
</table>

### Prediction

<table>
<thead>
<tr>
<th>Actual Outcome</th>
<th>Home</th>
<th>Medicine</th>
<th>Surgery</th>
<th>Psychiatry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home</td>
<td>94.6%</td>
<td>4.4%</td>
<td>0.7%</td>
<td>0.3%</td>
</tr>
<tr>
<td>Medicine</td>
<td>13.7%</td>
<td>76.3%</td>
<td>10.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Surgery</td>
<td>0.0%</td>
<td>40%</td>
<td>60%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Psychiatry</td>
<td>14.3%</td>
<td>14.3%</td>
<td>0.0%</td>
<td>71.4%</td>
</tr>
</tbody>
</table>

**Very Good Overall Performance**

*Medicine & Surgery Patient Separation Difficulties*
### Impact of Acceptance Threshold Probability

<table>
<thead>
<tr>
<th>Prob. Threshold:</th>
<th>0.2</th>
<th>0.4</th>
<th>0.6</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Base Volume:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Accuracy:</strong></td>
<td>91.3%</td>
<td>91.3%</td>
<td>92.2%</td>
<td>93.9%</td>
<td>94.8%</td>
</tr>
<tr>
<td><strong>Home Volume:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Accuracy:</strong></td>
<td>93.8%</td>
<td>93.8%</td>
<td>94.6%</td>
<td>95.5%</td>
<td>96.0%</td>
</tr>
<tr>
<td><strong>Medicine Volume:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Accuracy:</strong></td>
<td>75.6%</td>
<td>75.6%</td>
<td>76.3%</td>
<td>80.4%</td>
<td>83.3%</td>
</tr>
<tr>
<td><strong>Surgery Volume:</strong></td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td><strong>Accuracy:</strong></td>
<td>66.7%</td>
<td>66.7%</td>
<td>60.0%</td>
<td>50.0%</td>
<td>33.3%</td>
</tr>
<tr>
<td><strong>Psychiatry Volume:</strong></td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Accuracy:</strong></td>
<td>62.5%</td>
<td>62.5%</td>
<td>71.4%</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>
Assortment Planning for Configurable Products in Automotive
Assortment Planning

- **Overly Complex Assortments**: Until recent years, automotive companies steadily increased the size of product assortments/diversity.

- **Questionable Impact**: Correlation between sales and assortment diversity is questionable.

<table>
<thead>
<tr>
<th>European Model</th>
<th>Available Configurations</th>
<th>European Sales (2002)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VW Golf</td>
<td>1,999,813,504</td>
<td>595,465</td>
</tr>
<tr>
<td>Ford Focus</td>
<td>366,901,933</td>
<td>523,356</td>
</tr>
<tr>
<td>GM Astra</td>
<td>27,088,176</td>
<td>440,567</td>
</tr>
<tr>
<td>Renault Megane</td>
<td>3,451,968</td>
<td>261,383</td>
</tr>
<tr>
<td>Fiat Stilo</td>
<td>10,854,698,500</td>
<td>173,453</td>
</tr>
<tr>
<td>Toyota Corolla</td>
<td>162,752</td>
<td>139,837</td>
</tr>
<tr>
<td>Nissan Almera</td>
<td>3,036</td>
<td>87,474</td>
</tr>
</tbody>
</table>

Source: Pil and Holweg, 2004

- **Recent Cutbacks**: Examples from Ford
  - Reduced ordering complexity of 2009 F-150 truck by more than 90%
  - Focus 2010 was planned with only 150 major combinations, a drop of 95% from 2008 model

Source: Automotive News, August 18, 2008
Overly Complex Assortments

Complexity with Little Choice Accessible to Avg. Customer

Mfg. & Supply Chain

Realized Customer Choice

Engineering Complexity
### Terminology

- **Product Content: Models, Packages, Exterior/Interior, Accessories**
  - Differentiates product configurations
    - **Parts:** Engine type, Transmission type, Seat type ...
    - **Options:** Navigation system, Moon-roof, Satellite radio, Floor mat ...

- **Product “Configuration”**
  - Particular combination of parts and options that can be ordered
    - Example: Red sedan with manual transmission, V6 engine, black heated leather seats, moon-roof ...

- **“Core Entity” Configuration**
  - Mostly focuses on primary parts (e.g., engines, transmissions, leather bucket seats, xenon headlights)
Product “Assortment” (Variety)
- Set of configurations offered for the product (orderable)
- Product Assortment Structure:
  - Model Nests vs. Catalog Listing

Product “Substitution”
- Customer’s “ideal” configuration not part of the assortment or out of stock
- Customer either substitutes to similar configurations or “walks”
Objective and Analytics

- **Objective:** Develop decision support models to facilitate objective assortment planning of configurable automotive products
  - Manage overall complexity while maximizing profitability
  - Tiered Decisions: Component Facilities->Assembly Facilities->Markets

- **Analytics:**
  - Descriptive: Engineering and integration costs, manufacturing costs, economies of scale,…
  - Predictive: Customer choice preferences, substitution patterns (Feature Extraction, Classification, Neural Networks, Logistic Regression)…
  - Prescriptive: Optimization of assortment, production and distribution allocation,…

- **Users:**
  - Product Planning, Marketing & Sales, Engineering, Manufacturing, SC, Purchasing
### Assortment Optimization
- Considerations for Commonality
- Manufacturing and Supply Network
- Take-Rates & Production Capacities
- Sensitivity Analysis

### Powertrain Technology Selection
- CAFE Implications, Emissions, Profitability

### Bundling Options/Features (i.e., Options Packaging)
Mid-Segment Program: Aggregate Results
Markets: U.S. & Canada

**Profit & Loss Projections Summary**

- Sales: 300k
- Revenue: $7 billion
- Costs: $6.25 billion
- Profit: $750 million

---

**Cost Breakdown**

All costs in millions

- Administrative, Marketing, and Common Parts Engineering: $731 million
- Design, Engineering, and Development: $83 million
- Facilities: $20 million
- Tooling: $174 million
- Manufacturing (OEM): $435 million
- Manufacturing (Supplier): $28 million
- Transportation: $3,284 million
- Complexity: $1,500 million

---

**Illustrative Example**
# Mid-Segment Program: Aggregate Results

## Markets: U.S. & Canada

<table>
<thead>
<tr>
<th>REVENUES (in millions)</th>
<th>$7,003</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated Program Wholesale Unit Volume (in thousands)</td>
<td>300</td>
</tr>
<tr>
<td>Ford U.S. Total</td>
<td>270</td>
</tr>
<tr>
<td>S-Series</td>
<td>25</td>
</tr>
<tr>
<td>SE-Series</td>
<td>106</td>
</tr>
<tr>
<td>Standard</td>
<td>106</td>
</tr>
<tr>
<td>Electrified</td>
<td>0</td>
</tr>
<tr>
<td>SEL-Series</td>
<td>88</td>
</tr>
<tr>
<td>TI-Series</td>
<td>51</td>
</tr>
<tr>
<td>Ford Canada Total</td>
<td>30</td>
</tr>
<tr>
<td>S-Series</td>
<td>3</td>
</tr>
<tr>
<td>SE-Series</td>
<td>12</td>
</tr>
<tr>
<td>Standard</td>
<td>12</td>
</tr>
<tr>
<td>Electrified</td>
<td>0</td>
</tr>
<tr>
<td>SEL-Series</td>
<td>10</td>
</tr>
<tr>
<td>TI-Series</td>
<td>6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>COSTS (in millions)</th>
<th>$6,256</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administrative, Marketing, &amp; Common Parts Engineering - Program share of annualized costs</td>
<td>$1,500</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Design, Engineering, &amp; Development Costs - Program share of annualized costs</th>
<th>Powertrains</th>
<th>$144</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>Electrified</td>
<td>$0</td>
</tr>
<tr>
<td>Radios</td>
<td>Seats</td>
<td>$13</td>
</tr>
<tr>
<td>Moon Roof</td>
<td>Navigation System</td>
<td>$5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Facility Costs - Program share of annualized costs</th>
<th>OEM Plants (Body, Paint, Final Assembly)</th>
<th>$250</th>
</tr>
</thead>
<tbody>
<tr>
<td>Powertrain Plants</td>
<td>$110</td>
<td></td>
</tr>
<tr>
<td>Radio Plants</td>
<td>$6</td>
<td></td>
</tr>
<tr>
<td>Seat Plants</td>
<td>$17</td>
<td></td>
</tr>
<tr>
<td>Moon Roof Plants</td>
<td>$28</td>
<td></td>
</tr>
<tr>
<td>Navigation System Plants</td>
<td>$25</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tooling Cost - Program share of annualized costs</th>
<th>Powertrains</th>
<th>Seats</th>
<th>Radios</th>
<th>$17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Powertrains</td>
<td>$2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seats</td>
<td>$10</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Manufacturing Cost - Labor, Materials, Energy, Consumables</th>
<th>OEM</th>
<th>$3,284</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suppliers</td>
<td>$731</td>
<td></td>
</tr>
<tr>
<td>Powertrains</td>
<td>$500</td>
<td></td>
</tr>
<tr>
<td>Standard</td>
<td>$500</td>
<td></td>
</tr>
<tr>
<td>Electrified</td>
<td>$0</td>
<td></td>
</tr>
<tr>
<td>Radios</td>
<td>$43</td>
<td></td>
</tr>
<tr>
<td>Seats</td>
<td>$110</td>
<td></td>
</tr>
<tr>
<td>Moon Roof</td>
<td>$31</td>
<td></td>
</tr>
<tr>
<td>Navigation System</td>
<td>$55</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Transportation costs</th>
<th>Finished Product Distribution</th>
<th>$74</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component/Sub-Assembly Shipping</td>
<td>$2</td>
<td></td>
</tr>
<tr>
<td>Component Duties</td>
<td>$7</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Complexity Cost - Additional costs from variant complexity</th>
<th>Powertrains</th>
<th>$11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radios</td>
<td>$1</td>
<td></td>
</tr>
<tr>
<td>Seats</td>
<td>$9</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NET PROFIT BEFORE TAXES (in millions)</th>
<th>$747</th>
</tr>
</thead>
</table>

---

**Illustrative Example**
Conclusions and Key Takeaways

- Descriptive, Diagnostic, and Predictive analytics efforts are critical to the success of Prescriptive analytics.
- Need to be able to efficiently and effectively try many alternative models and methods.
- Data does not drive analytical models/methods; nor fall in love with models/methods.
- Don’t make assumptions regarding data availability, completeness, fidelity and accuracy.
- Start small and iteratively build more advanced models/methods and try new approaches (baseline).
- Experience based co-development is key.
  - Stakeholders and users are great source of information for choosing and implementing the right analytical approach.
  - Also adoption and user acceptance relies on their involvement.