Pricing Analytics & Optimization Solutions
Workshop | October 22, 2014

Greg Thomas, M.Sc
V.P. Pricing Research & Analytics
gthomas@pricingsolutions.com

Frederic Puech, PhD
Director of Analytics
fpuech@pricingsolutions.com

pricingsolutions.com
Pricing Analytics & Optimization

• AGENDA
  – 9:00-10:00  Introduction: Analytics & Optimization
  – 10:00-10:30  BREAK
  – 10:30-12:00  Step 1: The Data
  – 12:00-1:00  LUNCH
  – 1:00-2:30  Step 2: The Statistical Model
  – 2:30-3:00  BREAK
  – 3:00-4:00  Implementation
9:00 – 10:00 Introduction
Defining Pricing Analytics & Optimization

• ...analytics... the extensive use of data, statistical and quantitative analysis, explanatory and predictive models, and fact-based management to drive decisions and actions

Competing on Analytics: The New Science of Winning. by Thomas H. Davenport & Jeanne Harris
Defining Pricing Analytics & Optimization

Discovery

Explanation

Predictive

Transformational
Defining Pricing Analytics & Optimization

• …optimization…
Why create pricing analytical & optimization processes?

<table>
<thead>
<tr>
<th>Pricing Without Analytics</th>
<th>Pricing With Analytics</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Image of pricing without analytics]</td>
<td>![Image of pricing with analytics]</td>
</tr>
</tbody>
</table>
## How to create a pricing analytics & optimization process

<table>
<thead>
<tr>
<th>Higher Order Thinking Questioning Matrix</th>
<th>Event</th>
<th>Choice</th>
<th>Reason</th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Remembering</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Possibility</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Analysing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predictability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Imagination</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Application
- **Past**
- **Possibility**
- **Probability**
- **Predictability**
- **Imagination**

### Evaluating & Creating
# How to create a pricing analytics & optimization process

## Analytical question framework

<table>
<thead>
<tr>
<th>Past</th>
<th>Present</th>
<th>Future</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Information</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What happened? (Reporting)</td>
<td>What is happening now? (Alerts)</td>
<td>What will happen? (Extrapolation)</td>
</tr>
<tr>
<td><strong>Insight</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How &amp; why did it happen? (Modeling, experimental design)</td>
<td>What is the next best action? (Recommendation)</td>
<td>What is the best / worst that can happen? (Prediction, optimization, simulation)</td>
</tr>
</tbody>
</table>

Source: Analytics at Work: Smarter Decisions, Better Results, Thomas H. Davenport, Jeanne G. Harris and Robert Morison
How to create a pricing analytics & optimization process

Theoretical foundations of the model – how the customer's behavior is represented

• Extract
• Cleanup

• Identify trends, data gaps and issues
• Measure Frequency of observations,

Calculation of sensitivity coefficients (to price, segments, etc.)

“Engine” of the model, where the coefficients and data are put together so we can run scenarios

What the end user will see and use:
• Simulation dashboard
• Results dashboard
How to create a pricing analytics & optimization process

• Appreciative Inquiry:
  – a method for analysis, decision-making & the creation of strategic change
  – 4D’s:
    • Discover – identification of organizational process that work well
    • Dream – envisioning of process that would work well in the future
    • Design – planning & prioritizing processes that would work well
    • Deploy – implementation of the proposed design
  – Goal:
    • build/re-build around what works, rather than trying to fix what doesn’t
How to create a pricing analytics & optimization process

• Appreciative Inquiry:

| What works well? |  
| What do you wish could be better? |  
| What needs to be done to get better? |  

10:00 – 10:30 BREAK
10:30 – 12:00 The Data and The Model
Step 1: The Data
What do you need to build a pricing tool?
How to create a pricing analytics & optimization process

Data Collection and Treatment

- Extract
- Cleanup

Theoretical foundations of the model – how the customer’s behavior is represented

Data Mining

- Identify trends, data gaps and issues
- Measure Frequency of observations,

Model Design

Calculation of sensitivity coefficients (to price, segments, etc.)

Model Back-End

“Engine” of the model, where the coefficients and data are put together so we can run scenarios

Model Front-End

What the end user will see and use:
- Simulation dashboard
- Results dashboard
Assessing and gathering data

- Internal
  - Pricing
  - Transactions
  - Marketing
  - Costs
  - Customers

- External
  - Competitors
  - Economic conditions
  - Other factors
Assess what is available (Internal data)

- **Pricing – MUST HAVE**
  - List price
  - Net price
  - Discounts
  - Rebates
  - Other incentives

- **Transactions – MUST HAVE**
  - Frequency of observations (daily, weekly, monthly, etc.)
  - Time period available
  - How transactions can be classified (by account/customer, by location, etc.)

- **Costs – MUST HAVE IF MARGIN OPTIMIZATION**
  - COGS and/or margins

- **Customers – NICE TO HAVE**
  - Demographic info: age, gender, etc. (B2C), company size, industry (B2B), location (all)
  - If loyalty program in place, data collected on customer can prove very useful
Assess what is available (External data)

- Competitors
  - List of products
  - Price (net) of those products
  - Units sold
  - Promotional activities

- Economic conditions (by location/state/region)
  - Economic growth
  - Unemployment rate
  - Level of income
  - ...
  - Mostly applicable for B2C

- Other factors
  - Weather
  - Other events that may be of relevance (public holidays, sports events, etc.)
Do not underestimate the impact of weather!

- What good weather can do to beer sales in the UK:

![Graph showing model's accuracy with an exceptional event highlighted.](image)

Exceptional event: a sunny week in the middle of summer in the UK
Where to find the data you need?

• Internal data:
  – The main challenge is that data is usually scattered between IT, finance, marketing and sales
  – It then becomes critical to clearly define your needs when sending a request

  *Prepare a template of your (foreseen) database and send it to each department you need data from. It will save you time and ease the process.*

• External data:
  – Data on competitors is usually hard to get, especially in B2B
  – In some industries (CPG, retail, pharma), market-level data may be available (e.g. Nielsen, IMS, CEESA)
  – Economic and weather data are usually easy to find, especially with the rise of open data
Cleaning your database

• Dealing with errors
  – Entry errors
  – Classification errors
  – Missing values

• Dealing with bad data
  – Outliers
  – “Flat” data
  – Skewed data
Top 3 errors found in a database

• Entry errors (forgotten decimals, extra zeros, etc.)
  – How to spot them? → Simple descriptive statistics (avg., std. dev., min and max) and/or scatterplot by relevant group (product, region, customer)
  – How to handle them? → Remove them, unless you are 100% sure of what the “true” number should be

• Classification errors (wrong product code)
  – Typically not “errors” per se but rather change in nomenclature
  – Can become a nightmare if need to match two different sources based on that code
  – How to spot them? → Pivot table of product name vs. code
  – How to handle them? → Manual cleaning

• Missing values
  – How to spot them? → Count # of records by variable + pivot tables
  – How to handle them? → If not too many (<5% of total), records can be removed. Otherwise, need to look for additional data and/or change approach
  – Main issue: if missing records are not random, results can be biased
    • E.g. records missing for an entire region or sales rep
An example of the problem posed by missing values

Price on Done Deals

Entry errors (to be removed)

- 1% Missing
- 99% Present

Competitor Price on Done Deals

- 100% Present
- Unusable

Price on Lost Deals

Check if there is a pattern (e.g. 1 sales rep never reporting price on lost deals)

- 68% Present

Competitor Price on Lost Deals

- 40% Missing
- 60% Present
What makes a “good” data?

• No outliers
• Enough variability to provide meaningful results – “flat” data does not tell anything
• Not too skewed toward a specific sub-category
• Not too many (or too long) gaps in the time series
  – Issues may arise when dealing with long-tail products with very few observations and/or products taken on/off the market periodically

• Best way to assess quality of data is to draw the distribution of the series
• “Bad” data can still be used (unlike data with errors) but needs to be handled with more care
What is an outlier?

http://public.tableausoftware.com/views/NHLCareerPointsLeaders/AnatomyofanOutlier?:embed=yes&Toolbar=yes&:tabs=no&:showVizHome=no
How flat is too flat?

BAD - UNUSABLE

OKAY

THERE WE GO!
Two examples of “rotating” products

12-unit package replaced by 10-unit package – combined into a single SKU in the analysis

6-unit package of 500mL replaced by 6-unit package of 450 mL, then replaced by 4-unit package – combined into a single SKU in the analysis
Good vs. bad data

Needs cleanup

Much better!
Resources needed to complete Data collection and treatment

- Availability assessment:
  - Time and effort: typically a couple of days of phone calls and online searches

- Data collection:
  - Time and effort: can vary widely depending on data readiness
  - Cost: virtually nothing for internal data, assuming a data warehouse is already in place; external data can be expensive (especially market-level data)
  - This is the stage where you have the least control over time and budget. Think ahead!

- Data cleaning:
  - Time and effort: typically 1-2 weeks, up to 3-4 if lots of issues
  - Cost: depends on tools used (see next page)

- Best executed by someone with at least minimum knowledge of stats and “quant” analysis
Tools/Software needed

• Storing and managing the data:
  – Up to 500k records: Excel or Access are typically sufficient
  – Larger datasets are typically stored in relational databases (SQL, Oracle, etc.)

• Data cleaning and mining: our recommendation is to combine:
  – 1 tool for statistical analysis and data mining:
    • Free (and powerful): R
    • Affordable: SPSS, Stata, Minitab
    • Less affordable: SAS
  – 1 tool for data visualization:
    • Free: Gephi
    • Affordable: Tableau
    • A bit less affordable: QlikView
  – We currently use Stata+Tableau but are investing in R
How to create a pricing analytics & optimization process

Data Collection and Treatment
- Extract
- Cleanup
- Identify trends, data gaps and issues
- Measure Frequency of observations

Data Mining
- Explore your data
- Gather preliminary insights

Model Design
- Theoretical foundations of the model – how the customer’s behavior is represented

Coefficients Calculation
- Calculation of sensitivity coefficients (to price, segments, etc.)

Model Back-End
- “Engine” of the model, where the coefficients and data are put together so we can run scenarios

Model Front-End
- What the end user will see and use:
  - Simulation dashboard
  - Results dashboard
Data Mining Techniques

• Cluster analysis
  – Useful to identify customer segments based on continuous and categorical variables
  – Can be used for price tiering

• Decision trees
  – Useful to identify customer segments based on categorical variables
  – Can be used to identify profitable/unprofitable customers and manage discounts

• Market Basket Analysis
  – Useful to learn about products associations
  – Can be used for bundling

• Other techniques are available (neural network, discriminant analysis, etc.) but are usually less relevant to pricing
Cluster Analysis 101

- Objective: identify customer segments that could be easily identified/targeted for price tiering but also marketing, advertising, etc.
- Data used:
  - Demographic data collected on each customer
    - Personal information (gender, age, state)
    - Socio-economic information: income bracket, marital status, employment status, level of education
  - Behavioral data if available (usually collected via a survey)
- Statistical technique used: K-means clusters
  - Customers grouped into clusters, based on their “distance” to the center of the cluster
  - Clusters and their centers built based on data (center: profile of each customer)
  - Number of clusters can vary: trial-and-error process to find the number of clusters that best fits the data
An example of cluster analysis

- 5 customer segments identified, and overlapping with US states
An example of Classification Tree: Identify Current Margin Drivers

- **Account Size**
  - High: 4170
  - Low: 10419
  - Big
    - High: 1010
    - Low: 6330
    - Premium Machines
      - No
        - High: 490
        - Low: 4572
      - Yes
        - High: 520
        - Low: 1758
    - Small
      - High: 3160
      - Low: 4089
      - Region
        - Region 1
          - High: 2814
          - Low: 2839
        - Region 2
          - High: 346
          - Low: 1250
Market Basket Analysis 101

• Data mining technique used to collect insights on customers’ purchasing behavior
  – Analysis of the “baskets” of goods purchased simultaneously by customers
  – Seeks to uncover affinities between products and/or brands
  – Widely used in retail as well as in web analytics but can also be used in B2B
  – Very useful to uncover potential bundling opportunities

• Key outputs:
  – Sets of items, based on their frequency of joint purchase
  – Sets of “association rules”, indicating affinity between a Basket A and a Product B
  – Various metrics to measure how “interesting” a rule is:
    • Frequency, i.e. how often is the rule observed in the actual data
    • Probability: given the purchase of Basket A, what is the probability of Product B to be purchased as well
    • Other metrics measuring the degree of correlation between A and B

• Example of a rule:
  – (Diapers, Beer) → Milk
    • Frequency: 14%, i.e. 14% of the transactions included diapers, beer and milk
    • Probability: 94%, i.e. if a customer purchased diapers and beer together, there is a 94% probability s/he will purchase milk as well
An example of Market Basket Analysis

- The network graph shows all the brands leading to the purchase of Brand 5.
- The thickness of the line shows the frequency of the relationship in the data.
- The color shows the probability of Brand 5 being purchased in Brand X is purcl.
Case Study: How to use Market Basket Analysis to drive decisions on bundling?

<table>
<thead>
<tr>
<th>Profile</th>
<th>Challenge</th>
<th>Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>• B2B</td>
<td>• Leader on the market with broad portfolio, being challenged by smaller, more specialized competitors</td>
<td>• To grow profit by better managing portfolio of products</td>
</tr>
<tr>
<td>• Healthcare</td>
<td>• Margin growth driven only by price increases – needed new source of growth</td>
<td>• Identify bundling opportunities in portfolio</td>
</tr>
<tr>
<td>• ~200 SKUs</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Data analyzed

- All the invoices sent to customers over a 3-year period

- Information provided:
  - Invoice number
  - Date
  - Customer code
  - Product information: SKU code, brand, product category
  - Quantity purchased
  - Customer tier

- ~400,000 records in total
Demo of the tool built for the case study
Resources needed to complete Data Mining

• Time and effort:
  – A couple of weeks if you are familiar with the technique and software you are using
  – 4-6 weeks if learning of new technique/software is required

• Software:
  – R provides all the algorithms needed to mine your data thoroughly…and is free
    • Downside: beware the learning curve!
  – SPSS (especially SPSS Modeller) and SAS provide good (though costly) alternatives to R that are more user-friendly
  – If familiar with SQL Server, Excel has a nice add-on for data mining

• Best executed by someone with solid knowledge of stats and “quant” analysis
Step 2: The Modeling
What do your pricing data have to say?
How to create a pricing analytics & optimization process

Data Collection and Treatment
- Extract
- Cleanup
- Identify trends, data gaps and issues
- Measure Frequency of observations

Data Mining
- Explore your data
- Gather preliminary insights

Model Design
- Theoretical foundations of the model – how the customer’s behavior is represented

Coefficients Calculation
- Calculation of sensitivity coefficients (to price, segments, etc.)

Model Back-End
- “Engine” of the model, where the coefficients and data are put together so we can run scenarios

Model Front-End
- What the end user will see and use:
  - Simulation dashboard
  - Results dashboard
What do we call “modeling”? 

• Bottom line of a pricing tool: trying to get an understanding and accurate quantification of the relationship between your pricing decisions and your KPIs (sales, revenue & margin)

• Modeling: part of the process where you actually represent and measure that relationship, using maths and data

• KPI = f(Pricing)
Which model is right for you?

• It all comes down to the micro-economic theory of the consumer and her “demand curve”
  – Through this exercise, you are trying to actually draw that curve and measure the price-sensitivity of your customers and products

• So what is the best way to represent the behavior of your customers?
  – The answer to this question eventually influences the tool you will end up building via the “functional form” of your model

• Functional form: mathematical equation you will use to represent the relation between pricing and Demand (the “f” in Demand = f(Price) )
The choices available

• There are many functional forms available but they can be classified as follows:

• **Discrete choice models**: when demand is a Yes/No type
  – Any industry working with contracts or subscriptions
  – Key requirement: data on lost deals or lost customers
    • If you only have data on “Yes”, you cannot model the impact of price on Yes/No
  – The model gives the impact of price on the probability to win a deal

• **Continuous (linear) models**: when demand is a quantity
  – The model gives the impact of price on volume (or market share if market-level data is available)
How to choose?

• Theory is one thing…the reality of your business is another. Leverage your data!
How to create a pricing analytics & optimization process

Data Collection and Treatment
- Extract
- Cleanup
- Identify trends, data gaps and issues
- Measure Frequency of observations

Data Mining
- Explore your data
- Gather preliminary insights

Model Design
- Theoretical foundations of the model – how the customer’s behavior is represented

Coefficients Calculation
- Calculation of sensitivity coefficients (to price, segments, etc.)

Model Back-End
- “Engine” of the model, where the coefficients and data are put together so we can run scenarios

Model Front-End
- What the end user will see and use:
  - Simulation dashboard
  - Results dashboard

Model Front-End
What is a coefficient?

• In the mathematical equation that is your model, a coefficient is what links your **dependent variable** (what you are trying to explain: unit sales, volume, market share) to your **independent variables** (the factors explaining it: price, promos, weather, etc.)
  – It is usually referred to by fancy Greek characters such as \( \alpha, \beta, \epsilon, \lambda \), etc.
• The coefficients measure the sensitivity of your dependent variable to the factors explaining it
How do you compute a coefficient?

• Coefficients are computed via statistical algorithms
  – You may have heard of OLS (ordinary least squares), ML (Maximum Likelihood), MCMC (Monte Carlo Markov Chains), etc.

• What you need to compute your coefficients:
  – A statistical software with those algorithms (and more)
  – A trained statistician/econometrician who knows:
    • Which algorithm to choose to run your model
    • How to read the results and find the best fit for your data

• Computing coefficients is:
  – The most technical stage of the entire process
  – The most painstaking one: lots of trial and error to find the best fit for the data
  – The most critical one: if not done properly, your end product (the pricing tool you are building) will not be accurate
### Case Study: Estimating the probability to close a deal

<table>
<thead>
<tr>
<th>Profile</th>
<th>Challenge</th>
<th>Goal</th>
</tr>
</thead>
</table>
| • B2B  
  • Utilities | • Highly competitive market where even a few cents can make a difference  
  • Need surgical pricing | • To grow profit by better managing probability to win a deal |
Data analyzed

• Done and lost deals, over a 4-month period (proof of concept)

• Information provided:
  – Price, and margin rate
  – Quantity purchased
  – Term length
  – Transaction type (New/Renewal)
  – For lost deals: name and price of the winning competitor (when available)
  – Some demographic information (SIC, area code)

• ~10,000 records in total
Methodology

• Logistic regression: method used to predict the probability of an event to occur, based on observed factors
  – Event: winning a deal (discrete variable: Lost/Won)
  – Factors: price offered, transaction type, etc.

• Typical trial-and-error process:
  – Start with simplest model: \( \text{prob}(\text{winning}) = a + b \times \text{Price} \)
  – Add other relevant factors: \( \text{prob}(\text{winning}) = a + b \times \text{Price} + c \times \text{Factors} \)
  – Keep only the model that best fits the data and yields highest accuracy
The process we have followed

- SIC (1385 segments)
- Area code (291 segments)
- Term length (6 segments)
- Transaction type (2 segments)
- Price only (no segment)
The simplest model – decent accuracy

$80 gives a 50% probability to win a deal

Accuracy*: 65%
The odds of winning a renewal are 360% higher than the odds of winning a new customer.

\[
\text{Prob} = f(\text{Price, Transaction Type})
\]

Accuracy*: 70%

*: Accuracy measured by the percentage of actual deals predicted correctly.
The longer the term, the better the odds of winning a deal (renewal or new)

\[ \text{Prob} = f(\text{Price, Transaction Type, Term Length}) \]

Accuracy*: 70%

Accuracy measured by the percentage of actual deals predicted correctly.
Adding Area Code (90 different codes) shows the value of granular customer segmentation...

Prob = f(Price, Transaction Type, Term Length, Rate Code)

Accuracy*: 77%

*: Accuracy measured by the percentage of actual deals predicted correctly
A snapshot of the end product

Deal ID: 00000343  
Deal Type: RFP  
Start Date: 02-Sep-08  
Win/Loss: Pending  
Approval: Pending - Sales Director  
Approval Urgency: 2 days  
Sales ID: 0734 - Scott Miller  

Customer Details

- Customer ID: 1234567  
- Customer Name: Oster Wolflan  
- City: St. Paul  
- State: Minnesota  
- Zip Code: 55102  
- Office Type: Office Tower  
- Industry: Legal  
- Office Size: 1,000  
- Service Freq: 2 times per week  
- Location: Metro  

Customer Score: 6

Offering/Bid Details

<table>
<thead>
<tr>
<th>Product ID</th>
<th>Description</th>
<th>Unit Qty</th>
<th>Unit Measure</th>
<th>Unit Bid Price</th>
<th>Unit Cost</th>
<th>GP%</th>
<th>REVENUE</th>
<th>GP</th>
</tr>
</thead>
<tbody>
<tr>
<td>125</td>
<td>Coffee - Hazelnut</td>
<td>25</td>
<td>Box</td>
<td>$17.50</td>
<td>$7.50</td>
<td>57%</td>
<td>$12,600.00</td>
<td>$7,200.00</td>
</tr>
<tr>
<td>126</td>
<td>Coffee - French Vanilla</td>
<td>25</td>
<td>Box</td>
<td>$17.50</td>
<td>$8.50</td>
<td>51%</td>
<td>$12,600.00</td>
<td>$6,480.00</td>
</tr>
<tr>
<td>127</td>
<td>Coffee - French Roast</td>
<td>25</td>
<td>Box</td>
<td>$17.50</td>
<td>$7.00</td>
<td>60%</td>
<td>$12,600.00</td>
<td>$7,560.00</td>
</tr>
<tr>
<td>128</td>
<td>Coffee - French Roast</td>
<td>25</td>
<td>Box</td>
<td>$17.50</td>
<td>$7.00</td>
<td>71%</td>
<td>$12,600.00</td>
<td>$9,090.00</td>
</tr>
<tr>
<td>1456</td>
<td>Napkins - 80 pack</td>
<td>80</td>
<td>Pack</td>
<td>$7.50</td>
<td>$2.80</td>
<td>63%</td>
<td>$3,750.00</td>
<td>$2,350.00</td>
</tr>
<tr>
<td>305</td>
<td>Single Cup Brewing Sys</td>
<td>5</td>
<td></td>
<td>$0.00</td>
<td>$175.00</td>
<td>0%</td>
<td>$0.00</td>
<td>-$875.00</td>
</tr>
<tr>
<td>99922</td>
<td>Installation - Single Cup</td>
<td>1</td>
<td></td>
<td>$0.00</td>
<td>$135.00</td>
<td>0%</td>
<td>$0.00</td>
<td>-$135.00</td>
</tr>
<tr>
<td>99934</td>
<td>Servicing Basic - Single Cup</td>
<td>96</td>
<td></td>
<td>$0.00</td>
<td>$35.00</td>
<td>0%</td>
<td>$0.00</td>
<td>-$3,360.00</td>
</tr>
</tbody>
</table>

Total Revenue: $54,150  
Total GP: $28,220  

Payment Terms: Net 60  
Floor Margin: 45.0%

Comments/Competitive Situation

- Competitor#1: Imperial  
- Competitor#2: Imperial  
- Competitor#3: Imperial

Notes: Coming off 3-year contract with Imperial  
Date: 02-Aug-08  
Empl#: 0734

DEAL SCORE

Max Score: 65  
Actual Score: 56  
Approval Level: Sales Director

Floor Margin: 45.0%
Resources needed to complete Step 2

• Time and effort:
  – Typically 1 week for the design stage and 3-4 weeks for the coefficients calculation
  – Up to 6 weeks if difficulties occur (low accuracy, coefficients persistently with the “wrong” sign)
  – Main pitfall: the trial-and-error process can go on forever – need to define what is considered as “good enough” to move on

• Software:
  – R provides the most comprehensive catalog of algorithms for econometric modeling
    • Again: beware the learning curve!
  – Other proprietary software (Stata, SPSS Modeller, SAS, Matlab, etc.) have all their strengths and weaknesses
  – Excel is NOT an option at this stage: its built-in algorithms are not powerful enough to handle such complex calculations
12:00 – 1:00 LUNCH
1:00 – 2:30 The Tool
Step 3: The Predictive Pricing Tool
How to build your pricing “Minority Report”? 
How to create a pricing analytics & optimization process

Data Collection and Treatment
- Extract
- Cleanup
- Identify trends, data gaps and issues
- Measure Frequency of observations

Data Mining
- Explore your data
- Gather preliminary insights

Model Design
- Theoretical foundations of the model – how the customer’s behavior is represented

Coefficients Calculation
- Calculation of sensitivity coefficients (to price, segments, etc.)

Model Back-End
- “Engine” of the model, where the coefficients and data are put together so we can run scenarios

Model Front-End
- What the end user will see and use:
  - Simulation dashboard
  - Results dashboard
What is a pricing tool?

• The tool is where you put it all together:
  – The data
  – The coefficients

• The main goal of the tool is to enable its user to test price changes:
  – Run “what-if” scenarios and read the results provided by the model

• Depending on the needs (and budget invested in building it), the tool can either be:
  – A standalone solution, typically built in Excel
  – Or integrated into the company’s overarching IT system
What makes a good tool?

• Ease of use, even (and especially) for people without a background in stats/maths/econ/science
  – The tool will eventually be used by someone else!

• Adequacy with business needs
  – Does it provide the right answers to the right questions?

• Functionality
  – Does it work?
  – Is it user-friendly?

• Accuracy
  – Vital to get buy-in from key stakeholders…as well as actual results!
Typical Tool Development Workflow

Iterative process reflects best practices in software development:

- Ensures efficient workflow
- Flexible enough to allow modification of tool, even late in the development process
- Minimizes the risk of errors
- Guarantees the client/end user to receive a fully operational deliverable
# Case Study: Building a Price & Promo Planner

<table>
<thead>
<tr>
<th>Profile</th>
<th>Challenge</th>
<th>Goal</th>
</tr>
</thead>
</table>
| • B2C         | • Highly competitive market with complex “go to market models” that differ by province/state  
| • CPG         | • Significant growth in the Value Segment, but cannibalizing volume from the Premium Segment | • To improve the management of go-to-market strategies by becoming more data-driven  
| • ~250 SKUs   |                                                                           | • Measure the impact of price changes  
|               |                                                                           | • Model the expected return on promotional investments in the planning calendar |
Predictive Methodology

**Inputs**

1. Weekly data by SKU & Display Area for all products on the market
   - Sales
   - Promotional activity
   - Pricing
   - Other factors (# of stores, region, etc.)

**Consumer Choice Analysis**

2. Sensitivity to Price
   - Sensitivity to Promotions
   - Sensitivity to Seasonality
   - Sensitivity to Nests & Sub-nests
   - Sensitivity to Holidays
   - Sensitivity to Display Area

**Outputs**

3. Analysis
   - Total volume
   - Market share, Revenue and Margin
   - Change in volume, Revenue and Margin based on price and promotional changes
   - Elasticity, cross-elasticity, etc…
Key Feature: Customer Product Switching Analysis (Nesting)
The tool’s interface

1) Choose Geography
2) Choose Simulation Week
3) Promotions per SKU
4) Set Simulated Price per SKU

Simulation Results: Volume & Margin by SKU, by Brand, by Segment, by Brewer
Model Accuracy: Testing

Average Prediction Error of Total Sales Volume Per Week: 4%
Model Output: Learning About the Customer’s Behavior

Which brands’ sales are more seasonal?

How do holidays impact sales by brand?

How sensitive each brand is to various types of promotions?
Scenarios Output: Impact of a Price Change

Impact of a 5% price increase on volumes...

...and on revenue
### Promotion Sensitivity Impact

<table>
<thead>
<tr>
<th></th>
<th>Volume</th>
<th>Revenue</th>
<th>Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display</td>
<td>0.46%</td>
<td>0.46%</td>
<td>0.46%</td>
</tr>
<tr>
<td>TPR</td>
<td>7.76%</td>
<td>7.76%</td>
<td>7.76%</td>
</tr>
<tr>
<td>Multibuy</td>
<td>3.59%</td>
<td>3.59%</td>
<td>3.59%</td>
</tr>
<tr>
<td>MAD</td>
<td>0.47%</td>
<td>0.64%</td>
<td>0.71%</td>
</tr>
<tr>
<td>In-case</td>
<td>1.49%</td>
<td>0.76%</td>
<td>0.07%</td>
</tr>
</tbody>
</table>
## Case Study: Building a portfolio optimizer

### Profile
- B2B
- Healthcare
- ~200 SKUs

### Challenge
- Leader on the market with broad portfolio, being challenged by smaller, more specialized competitors
- Margin growth driven only by price increases – needed new source of growth

### Goal
- To grow profit by better managing portfolio of products
- Identify bundling opportunities in portfolio
Dual approach for coefficients calculation

Market Basket Analysis
Quantifies interactions between products within portfolio

Econometric Regressions
Quantifies price-sensitivity of products in portfolio

Portfolio Optimizer
Quantifies the impact of a price change on a single product to the entire portfolio
Demo of the tool built for the case study

<table>
<thead>
<tr>
<th>Therapeutic Group</th>
<th>Brand</th>
<th>Item</th>
<th>Current Price</th>
<th>Simulated Price</th>
<th>% change</th>
<th>Current Revenue</th>
<th>Simulated Revenue</th>
<th>Absolute Change</th>
<th>% Change</th>
<th>Current Margins</th>
<th>Simulated Margins</th>
<th>Absolute Change</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-infective</td>
<td></td>
<td></td>
<td>$50,162,389</td>
<td>$50,619,771</td>
<td>0.9%</td>
<td>$457,302</td>
<td>$457,302</td>
<td>$16,019</td>
<td>0.8%</td>
<td>$361,246</td>
<td>$361,246</td>
<td>$0</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>Brand 1</td>
<td></td>
<td>$17,592,906</td>
<td>$17,962,627</td>
<td>2.1%</td>
<td>$369,721</td>
<td>$369,721</td>
<td>$38,090</td>
<td>11.1%</td>
<td>$16,623,635</td>
<td>$16,972,986</td>
<td>$349,351</td>
<td>0.21%</td>
</tr>
<tr>
<td></td>
<td>Brand 2</td>
<td></td>
<td>$21,322,794</td>
<td>$21,290,431</td>
<td>-0.15%</td>
<td>$32,363</td>
<td>$32,363</td>
<td>$-323</td>
<td>-1.0%</td>
<td>$18,173,599</td>
<td>$18,081,844</td>
<td>$91,755</td>
<td>0.50%</td>
</tr>
<tr>
<td></td>
<td>Brand 3</td>
<td></td>
<td>$6,938,513</td>
<td>$7,058,537</td>
<td>1.73%</td>
<td>$120,024</td>
<td>$120,024</td>
<td>$1,371</td>
<td>0.11%</td>
<td>$5,991,916</td>
<td>$6,095,565</td>
<td>$103,649</td>
<td>1.73%</td>
</tr>
<tr>
<td></td>
<td>Brand 4</td>
<td></td>
<td>$286,914</td>
<td>$286,914</td>
<td>0.00%</td>
<td>$195,348</td>
<td>$195,348</td>
<td>$0</td>
<td>0.00%</td>
<td>$195,348</td>
<td>$195,348</td>
<td>$0</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>Brand 5</td>
<td></td>
<td>$25,443</td>
<td>$25,443</td>
<td>0.00%</td>
<td>$16,788</td>
<td>$16,788</td>
<td>$0</td>
<td>0.00%</td>
<td>$16,788</td>
<td>$16,788</td>
<td>$0</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>Brand 6</td>
<td></td>
<td>$3,995,820</td>
<td>$3,995,820</td>
<td>0.00%</td>
<td>$3,202,178</td>
<td>$3,202,178</td>
<td>$0</td>
<td>0.00%</td>
<td>$3,202,178</td>
<td>$3,202,178</td>
<td>$0</td>
<td>0.00%</td>
</tr>
<tr>
<td>Anti-inflammatory</td>
<td></td>
<td></td>
<td>$1,556,986</td>
<td>$1,556,986</td>
<td>0.00%</td>
<td>$1,008,209</td>
<td>$1,008,209</td>
<td>$0</td>
<td>0.00%</td>
<td>$1,008,209</td>
<td>$1,008,209</td>
<td>$0</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>Brand 7</td>
<td></td>
<td>$1,556,986</td>
<td>$1,556,986</td>
<td>0.00%</td>
<td>$1,008,209</td>
<td>$1,008,209</td>
<td>$0</td>
<td>0.00%</td>
<td>$1,008,209</td>
<td>$1,008,209</td>
<td>$0</td>
<td>0.00%</td>
</tr>
</tbody>
</table>
Going from Predictive to Prescriptive

Prescriptive is built upon predictive, i.e. prescriptive analytics cannot be done without building predictive model first.
# Case Study: Building a Margin Optimizer

<table>
<thead>
<tr>
<th>Profile</th>
<th>Challenge</th>
<th>Goal</th>
<th>Methodology</th>
</tr>
</thead>
<tbody>
<tr>
<td>• B2B</td>
<td>• Client facing revenue leakages in their vending business (negative volumes, prices very low or very high, etc.)</td>
<td>• To develop new pricing strategies that were appropriate for different customer segments and regions</td>
<td>• Model built on hold-out data over 3 years (50M+ records)</td>
</tr>
<tr>
<td>• Food services provider for business, schools, government buildings, etc.</td>
<td></td>
<td></td>
<td>• Seasonality and trends by SKU</td>
</tr>
<tr>
<td>• Project dealt with their vending machines business</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
For each individual machine, the model provides:
- Optimal prices
- Optimal facings
Output: Summary

Margin Optimization

- SM + LG Qty
- LG - $1.10
- LG - $1.20
- LG - $1.30
Detailed Output for Daily Users

| M/C Num | Customer Num | Machine Num | Current Factings (A) | Model Factings (B) | Min | Max | Weighted Price | Weighted Cost (Net) | Unit Margin (Net) | Historical Units | ModSales (Rel) | ModSales (Rel) | Revenue | Margin | Factings Change | Price Change | Qty Change |
|---------|--------------|-------------|---------------------|-------------------|-----|-----|----------------|---------------------|------------------|----------------|--------------|--------------|--------------|---------|--------|----------------|-------------|------------|
| 6006    | 0            | 4055        | 0                   | 0                 | 1   | 6   | 0.8            | 0.29                | 0.51             | 36            | 33.6419      | 26.8195      | 17.1574      | 0      | 0      | 0.05           | -2          |            |
| 6006    | 0            | 4055        | 1                   | 1                 | 0   | 6   | 0.8            | 0.29                | 0.51             | 53            | 48.6184      | 38.6947      | 24.7954      | 0      | 0      | 0.05           | -4          |            |
| 6006    | 0            | 4055        | 0                   | 0                 | 0   | 6   | 0.8            | 0.28                | 0.54             | 28            | 25.1623      | 20.1298      | 13.5876      | 0      | 0      | 0.05           | -3          |            |
| 6006    | 0            | 4055        | 1                   | 1                 | 0   | 6   | 0.8            | 0.33                | 0.47             | 10            | 9.47496      | 7.57997      | 4.45323      | 0      | 0      | 0.05           | -1          |            |
| 6006    | 0            | 4055        | 0                   | 0                 | 4   | 6   | 0.8            | 0.33                | 0.47             | 10            | 9.47496      | 7.57997      | 4.45323      | 0      | 0      | 0.05           | -1          |            |
| 6006    | 0            | 4055        | 0                   | 0                 | 4   | 6   | 0.8            | 0.33                | 0.47             | 10            | 9.47496      | 7.57997      | 4.45323      | 0      | 0      | 0.05           | -1          |            |
### Which model is right for you?

<table>
<thead>
<tr>
<th>Do you sell to your customers via contracts?</th>
<th>Portfolio</th>
<th>Single Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>You have data on the deals you lost</td>
<td>Contract pricing optimizer (CPO) and/or Portfolio Pricing Optimizer (PPO) B2B, utilities, cable &amp; phone, publications</td>
</tr>
<tr>
<td></td>
<td>You don’t have data on the deals you lost</td>
<td>MO and/or PPO B2B, utilities, cable &amp; phone, publications</td>
</tr>
<tr>
<td>No</td>
<td></td>
<td>PPO and/or Pricing &amp; Promo Planner (PPP) B2C, CPG, retail, restaurants, pharma</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Are you looking at a portfolio of products or a single product?</th>
<th>Portfolio</th>
<th>Single Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>You have data on the deals you lost</td>
<td>Contract pricing optimizer (CPO) and/or Portfolio Pricing Optimizer (PPO) B2B, utilities, cable &amp; phone, publications</td>
<td>CPO and/or Margin optimizer (MO) B2B (heavy industries), medical devices</td>
</tr>
<tr>
<td>You don’t have data on the deals you lost</td>
<td>MO and/or PPO B2B, utilities, cable &amp; phone, publications</td>
<td>MO B2B (heavy industries), medical devices</td>
</tr>
<tr>
<td>No</td>
<td>PPO and/or Pricing &amp; Promo Planner (PPP) B2C, CPG, retail, restaurants, pharma</td>
<td>MO and/or PPP Pharma</td>
</tr>
</tbody>
</table>
Resources needed to complete Step 3

• Time and effort:
  – Typically 3-4 weeks
  – Up to 6 weeks if difficulties occur or if complex features need to be programmed
  – Building the tool requires a minimum of skills and experience in programming
    (Excel/VBA, SQL, Oracle, etc. depending on the platform on which the tool is built)

• Software:
  – Excel is typically the platform of choice, especially when building a tool for the first time
  – More advanced companies typically integrate their pricing tool into their existing BI platform
2:30 – 3:00 BREAK
3:00 – 4:00 Implementation
What are 5 common road blocks to implementing advanced pricing analytics?

...and the answer is...
Business Intelligence & Analytics Platforms

Dashboards
Magic Quadrant for Business Intelligence & Analytics Platforms
Gartner February 2014
# Business Intelligence & Analytics Platforms

<table>
<thead>
<tr>
<th>Platform</th>
<th>Overview</th>
<th>Strengths</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tableau</td>
<td>Tableau's highly intuitive, visual-based data discovery, dashboarding, and data mashup capabilities have transformed business users’ expectations about what they can discover in data and share without extensive skills or training with a BI platform.</td>
<td>Ease of use, breadth of use and enabling business users to perform more complex types of analysis without extensive skills or IT assistance. Low cost to implement.</td>
</tr>
<tr>
<td>Qlik - QlikView</td>
<td>QlikView is a self-contained BI platform, based on an in-memory associative search engine and a growing set of information access and query connectors, with a set of tightly integrated BI capabilities.</td>
<td>Business User: Ease of use, intuitive interactive experience, dashboards. IT’s requirements: enterprise features, reusability, data governance &amp; control, scalability.</td>
</tr>
<tr>
<td>MicroStrategy</td>
<td>MicroStrategy offers an enterprise-grade and organically grown end-to-end BI platform that is well suited to large and complex requirements.</td>
<td>Supports the largest data volumes &amp; end users. Mobile BI, reporting, dashboards, geospatial &amp; location analytics. Largest breadth and depth of BI tools.</td>
</tr>
</tbody>
</table>

Source: Gartner February 2014
Pricing Infrastructure

People

Systems

Structure

Process

Objectives
What makes a great analytical professional?

- Technical skills & education a starting point not an end point
- Commitment to learning
- Creativity
- Curiosity
- Business savvy
- Communication
- Solution focus
- Performance vs. Perfectionist mentality
What are 5 common challenges faced by pricing analysts?

...and the answer is...
## Analytics & IT challenges

<table>
<thead>
<tr>
<th>Analytics need to…</th>
<th>IT must…</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heavily utilize system resources</td>
<td></td>
</tr>
<tr>
<td>Create tables and use a lot of space</td>
<td></td>
</tr>
<tr>
<td>Run complex ad hoc queries</td>
<td></td>
</tr>
<tr>
<td>Go outside the box</td>
<td></td>
</tr>
<tr>
<td>Experiment with new approaches</td>
<td></td>
</tr>
<tr>
<td>Work with limited rules and restrictions</td>
<td></td>
</tr>
</tbody>
</table>
Reading List

Competing on Analytics
The New Science of Winning
by Thomas H. Davenport & Jeanne Harris

Nudge
Improving Decisions about Health, Wealth & Happiness
by Richard Thaler & Cass Sunstein

Taming the Big Data Tidal Wave
by Bill Franks

Analytics at Work
Smarter Decisions Better Results
by Thomas H. Davenport, Jeanne G. Harris and Robert Morison

Predictive Analytics
The Power to Predict who will click, buy, lie or die
by Eric Siegel
What are 5 business analytics trends for 2015?

...and the answer is...
Thank you!
To learn more, please contact:

**Greg Thomas**  
V.P. Pricing Research & Analytics  
gthomas@pricingsolutions.com

**Frederic Puech**  
Director of Analytics  
fpuech@pricingsolutions.com