PROFITABLE ANALYTICS

Presented by David Haertzen First Place Learning

This document is the property of First Place Software, Inc. Trademarks, products and images are properties of their respective owners.



About the Author

- > Enterprise and data architect
- Provided services to organizations such as: Allianz Life, 3M, Mayo Clinic, IBM, Fluor Daniel, Procter & Gamble and Synchrono – from start up to multinational.
- Experienced author
- Frequent presenter in the areas of:
 - Data modeling
 - Data warehousing
 - Enterprise architecture
 - Analytics and Business Intelligence
 - ✤ SQL
- Instructor for First Place Learning and eLearningCurve
- University of Minnesota MBA, University of St Thomas
- Visit: <u>http://www.firstplacelearning.com/</u>
- Visit: <u>http://www.linkedin.com/davidhaertzen</u>
- Visit: <u>http://ecm.elearningcurve.com/David_Haertzen_s/89.htm</u>



David Haertzen







Topic Objectives

Upon finishing this presentation you will:

- Understand what Analytics is, its goals, and its components
- Know the top Analytics terms
- Understand how to benefit from Analytics
- Be able to discuss Analytical methods and tools
- Be prepared to learn more about Analytics



Session Structure

•Topic 1: Introduction to Analytics

- > Decisions Impact the Bottom Line
- > Evolution of Data Based Decisions
- > Analytics and Data Mining

•Topic 2: Analytics Use Cases

- Customer Profitability Analysis
- Market Basket Analysis
- Next Best Action
- Law Enforcement
- Sports

•Topic 3: Analytics Methods and Tools

- Analytical Methodology
- > Design of Experiments and Random Trials
- Data Mining
- > Analytical Development Analytics Modeling and Programming
- Visual Analytics
- Tricks and Traps

Topic 4: Learning About Analytics

- Self Study -> Do Analytics -> Compete
- Master of Analytics



PROFITABLE ANALYTICS

Topic I: Introduction to Analytics



Copyright(c) 2013, First Place Software, Inc.

I. Introduction to Analytics

- > What is Analytics?
- Decisions Impact the Bottom Line
- Evolution of Data Based Decisions
- > Business Intelligence and Reporting
- > Data Warehouse



Analytics is a Management Tool

Analytics is an approach that translates data into insights and supports effective decision making through a combination of people, processes and technology.

Descriptive Analytics:

The use of analytics methods such as data mining and statistics to better understand data. For example, clustering and affinity analysis can help retailers to better understand customers and the products they buy.

Predictive Analytics:

The use of analytic methods such as data mining and statistics to anticipate future outcomes. For example, predictive analytics may provide insights into future demand for a product or the buying habits of a customer.

Prescriptive Analytics:

The use of analytics methods such as data mining and statistics to make recommendations. For example, a recommendation engine may advise a banking service representative to offer a specific rate of interest a customer.

Decisions Impact The Bottom Line

- Costs may be reduced by:
 - Negotiating improvements in supply
 - Dropping unprofitable products and customers
 - Reducing waste due to low quality
- Risks may be addressed by:
 - Avoiding problems such as poor credit risks
- Revenues may be increased by:
 - Understanding and better serving customers
 - Focusing on the most profitable products and customers
 - Cross selling to customers
 - Capitalizing on trends
 - Growing marketing opportunities

Data Scientists In Demand



Data Scientist: The Sexiest Job of the 21st Century By Thomas H. Davenport and D.J. Patil

How challenging is it to source analytical skills in general?



SOURCE DATA FROM A SURVEY BY NEWVANTAGE PARTNERS

HBR.ORG

(http://blogs.hbr.org/cs/assets_c/2012/11/analyticalskills-2832.html)



There's No Panacea for the Big Data Talent Gap

By Paul Barth and Randy Bean

How challenging is it to source data scientists?





Analytics Results



CapitalOne becomes a top 5 bank using analytics.



Olive Garden, the Italian restaurant chain, has reduced unplanned staff hours by 40% and wasted food by 10%.



The Royal Shakespeare Company increased the number of regulars by 70%.



3M is saving \$10 million in maintenance and has boosted sales force productivity 10%.



AIG failures in the use of analytics, including assumptions about the valuation and risk of loans result in losses that shake the nation.



Increasing Value of Analytics



Wayne Eckerson - A Practical Guide to Analytics - EBook

Data Warehousing and Analytics

Data Warehousing integrates and stores data needed for analysis.



Business Intelligence / Analytics analyzes the data provided through data warehousing.





PROFITABLE ANALYTICS

Topic II: Analytics Use Cases



Copyright(c) 2013, First Place Software, Inc.

II. Analytics Use Cases

- Customer Profitability Analysis
- Next Best Action
- Law Enforcement
- Market Basket Analysis
- Sports



Customer Profitability Analysis (CPA)

Customer Profitability Analysis is an analytic approach that determines the profitability of individual customers or segments of customers by identifying revenue and cost patterns associated with those customers. This includes identifying the most profitable customers (angels) as well as unprofitable customers (devils).



- > Customer behavior and profitability identified by analysis of company databases
- > 20% of customers are angels and result in bulk of profit
- > 20% of customers are devils and reduce profits by 20%
- > Profiles built of profitable and unprofitable customers
- Attracts most profitable customers by promotions, stocking desired products and providing best service
- Avoids unprofitable customers by dropping them from promotion lists, stopping loss-leader promotions and charging fees for restocking

Best Buy Decides Not All Are Welcome – Wall Street Journal - 2004



Watch out for the "Whale Curve"

.. The most profitable 20% of customers generate between 150 – 300% of total profits. The middle 70% of customers about break even, and the least profitable customers lose 50 – 200% of total profits, leaving the company with its 100% of total profits. Often some of the largest customers turnout to be the most unprofitable."

Dr. Robert Kaplan

- "Customer Profitability Measurement and Management"
 - Harvard Business School



Customer Profitability Factors



The Profitable Customer:

- Orders standard products
- Orders standard handling
- Orders via web or ecommerce
- Makes short service calls
- Almost never returns goods
- Orders large volume
- Pays on time
- Praises company on social media

The Unprofitable Customer:

- Orders exception products
- Orders via call centers
- Often makes lengthy service calls
- Frequently returns goods
- Orders special handling
- Orders small volume
- Requires low cost price match
- Pays late requires collection
- Complains on social media
- Only buys on sale



Next Best Action (NBA)

Next Best Action is an immediate action recommended by rules discovered through data mining or statistics that is intended to produce optimal results. This often includes providing service or making an offer to a customer.



- Felco analyzes customer profitability and behavior
- Builds a decision model based on customer profitability and responsiveness
- Customer calls Customer Service, requests lower rate or termination of account
- Telco service rep knows what to offer to customer discounts or other accommodations



Law Enforcement – Richmond, VA

Law Enforcement Analytics utilizes information coupled with predictive models and visual displays to effectively dispatch officers and to combat high priority crime.



- Patterns found in 911 and other police systems
 multiple data warehouses
- > Statistical model and business intelligence used
- > Crime and crime locations predicted
- > Police and special units effectively dispatch
- Models tied to color coded maps sent to policeofficers
- Violent crime headed off through proactive policing using property crimes as a leading indicator.

Results:

- > Dangerous city ranking down from 5th to 99th
- Homicides down 32% *
 - Rapes down 19% *
 - Robberies down 3%
 - Aggravated Assaults down 17%
- Police accomplish more with the same resources
 * From 2006 to 2007

IBM Smart Planet 2010



Copyright(c) 2013, First Place Software, Inc.

 \geq

Market Basket Analysis (MBA)

Market Basket Analysis is an analytical method that identifies product and service combinations that customers tend to buy. It is typically based on records of customer purchases. Market Basket Analysis is also known as Product Affinity Analysis or Association Rule Learning.

amazon.com

- > Amazon's revenue in 2009: \$24.5B
- ~\$5B came from product recommended using Market Basket Analysis (affinity analysis)





How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did

Walmart 🔀

SKU Rationalization Demands Market Basket Analysis (aka Customer Buying Patterns)

http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/ http://emcien.com/sku-rationalization-demands-market-basket-analysis-aka-customer-buying-patterns/



MBA Based Actions

- Locate items in stores and websites
- Improve cross-sell and up-sell
- Offer attractive incentives
- > Target offers to individuals or segments
- Attract traffic to stores and websites
- Obtain inventory to support promotions a sale on one item can lead to increased sales on other items

Avoid:

- Harming sales by dropping products
- Harming sales by increasing prices on related goods



Sports Analytics

Sports Analytics utilizes information coupled with predictive models to both to improve competitive results and to improve marketing results.



- Moved from gut feel approach to \geq statistical approach - Sabermetrics
- Built a record setting team with \geq players not recruited by other teams
- First team in 100 years to win 20 consecutive games (2006)
- Swept Twins in 2006 play offs. \geq



- Built a holistic view of fans correlated purchase records with web activity and 3rd party data
- Evaluated fan value including both \geq direct and indirect
- Took action to improve customer experience
- Generated profits via premium \geq advertising products.

Moneyball: The Art of Winning an Unfair Game is a book by Michael Lewis, published in 2003 Case Study: ESPN Drives Fan Value Through Customer Intelligence. Forrester Research, Inc. 2011





PROFITABLE ANALYTICS

Topic III: Analytical Tools and Methods



Copyright(c) 2013, First Place Software, Inc.

III. Analytical Methods and Tools

- Analytic Steps (CRISP-DM)
- Design of Experiments and Random Trials
- Big Data / Little Data
- Data Mining / Statistics
- Analytical Development Analytics Modeling and Programming
- Visual Analytics
- Tricks and Traps



CRISP-DM Flow

1.0 Business Understanding	2.0 Data Understanding	3.0 Data Preparation	4.0 Modeling	5.0 Evaluation	6.0 Deployment
Step 1.1 Determine Business Objectives	Step 2.1 Collect Initial Data	Step 3.1 Select Data	Step 4.1 Select Modeling Technique	Step 5.1 Evaluate Results	Step 6.1 Plan Deployment
Step 1.2 Assess Situation	Step 2.2 Describe Data	Step 3.2 Clean Data	Step 4.2 Generate Test Design	Step 5.2 Review Process	Step 6.2 Plan Monitoring and Maintenance
Step 1.3 Determine Data Mining Goals	Step 2.3 Explore Data	Step 3.3 Construct Data	Step 4.3 Build Model	Step 5.3 Determine Next Steps	Step 6.3 Produce Final Report
Step 1.4 Produce	Step 2.4 Verify Data Quality	Step 3.4 Integrate Data	Step 4.4 Assess Model		Step 6.4 Review Project
S	ource: www.crisp-dm.org	Step 3.5 Format Data	CRISP-DM (CRos A methodology funded by the E	ss Industry Standard Process developed in late 1990s, par uropean Commission under i	for Data Mining) ially he ESPRIT Program.

I S

Design of Experiments

Experiment or Random Trial is a set of actions designed to test the effectiveness of a treatment followed by observation of the results. Participants randomly receive treatment. This may be a simple A-B comparison or a more complex multifactor comparison.



Big Data / Sample Data Plus (+)

Minus (-)

Big Data is data that is so voluminous that it cannot be managed using traditional databases such as relational databases. This data is typically unstructured and consists of text, images, video, and audio	 May reveal outliers and exception opportunities May reveal new trends Analysis of customer sentiment requires big data Analysis of physical processes requires big data More accurate than a sample 		Requires more time to gather Costs more to store Takes longer to analyze Doesn't fit into memory
Sample Data is a portion of a population selected for statistical analysis.	 Costs less to gather Costs less to store Can have a high confidence level Faster results can be quickly applied 	A A	Requires discrete facts May miss outliers and exceptions



Data Mining / Statistics

Data Mining is the application of analytical methods to large volumes (terabytes) of data. This can include finding patterns, clustering data, prescribing actions and predicting outcomes.

Statistics uses the same techniques using data samples.

Data Mining and Statistics Categories

- Descriptive Statistics (average, standard deviation)
- Affinity Analysis / Clusters
- Regression
- Decision Trees
- Neural Networks



Data Mining - Regression

Regression is a statistical method that predicts the value of one variable based on the value(s) of one or more numeric variables. For example, the variable of wine price might be predicted based on winter rainfall, average growing season temperature and harvest rainfall.





Copyright(c) 2013, First Place Software, Inc.

Regression Example



Ayres, Ian. Super Crunchers: Why Thinking-By-Numbers is the New Way to be Smart. New York NY: Bantam Books, 2007.



Data Mining - Cluster Analysis

Cluster Analysis is an analytic method based on grouping data points with a large degree of affinity. The data points have much in common with data points in the cluster and differ from data points in other clusters.



Applications of clustering include:

- Customer analysis / segmentation
- Cataloging of astronomical objects

FIRST PLACE LEARNING

Grouping books and music into genres.

Data Mining - Decision Tree

A Decision Tree is a structure that enables large collections of inputs to be classified into homogeneous groups through a series of choices called nodes. The tree is processed from left to right or top to bottom, with the first node called the root node, nodes secondary to the root node called child nodes, and nodes at the bottom called leaf nodes.





Copyright(c) 2013, First Place Software, Inc.

Data Mining – Neural Net

Neural Nets are a flexible predictive analytics tool that mimics the learning of the human brain. A neural net model accepts a large collection of known inputs and produces an output that may be continuous-valued. Neural nets include machine learning and so can improve with use.



Data Mining Models Tips and Traps Traps (Don't Do This) Tips (Do This)

- □Clearly define goals and objectives
- Explain the model to establish credibility

 \Box Cleanse the data.

- Allow outliers to skew results.
- □ Include a synonym of the expected result in the training data
- Include noise data that is not relevant to the model.







Analytical Development Patterns

Analytical Development Patterns is the over all approach to implementing the logic of an analytical model. The two approaches are: code the model using programming languages or build the model through a visual drag and drop process.

Code It	Diagram It				
•Solution is developed using a text editor	•Solution is visualized as a flow of data				
•Algorithms are built using included functions	•Algorithms are nodes in the flow selected from a pallet				
•Examples:	•Examples:				
R (open source, implements S)	RapidMiner from Rapid-I GmbH				
> Python	SAS Enterprise Miner from SAS				
> SQL	SPSS Miner from IBM				
> SAS					
≻ C++					
> Java					

Visual Analytics

Visual Analytics is an approach to understanding data through graphical means – charts, graphics, maps, etc.

1854 Broad Street Cholera Outbreak

> Cholera was killing hundreds of people in London's Soho district

John Snow used visualization and statistics to show that the Broad Street pump was the source of the infection

Nearby monks did not contract cholera – they drank only home brewed beer

Similar techniques are used today



John Snow

Analytical Data Structures



Normalized and Unstructured Data Enable

Transaction Processing:

Create, Read, Update and Delete

Graphs Enable Scheduling and Affinity Analysis:

Critical Path, MBA, Logistics, Factories, Work Flows



Cubes and Star Schemas Enable Human Analytic Data Exploration:

Roll Up, Drill Down, Slice and Dice

Flattened Data Supports Analytical Algorithms: > Regression, Decision Tree, Neural Net, MBA



Analytical Tool Examples





THE POWER TO KNOW











Google Analytics



Analytics Projects Tips and TrapsTips (Do This)Traps (Don't Do This)

- Start by identifying project goals, objectives and assumptions
- □Work with key users
- □ Validate data sources
- Allow for iterative improvements to the model
- Staff with experienced project manager, data warehouse architect and data scientist.

- Use waterfall approach
- Start without basic pieces in place
- □ Wait for perfect conditions
- Let outsiders set large budget
- □ "Build it and they will come"
- Set unrealistic expectations with executives.







Copyright(c) 2013, First Place Software, Inc.



PROFITABLE ANALYTICS

Topic IV: Learning About Analytics



Copyright(c) 2013, First Place Software, Inc.

IV. Learning About Analytics

- Competencies
- Self Study
- Do Analytics
- Compete
- Master of Analytics



Data and Analytics Competencies

Data Competencies

- Data Profiling
- Data Cleansing and Data Quality
- Data Integration
- Data Governance
- Technical Architecture
- Relational Data Modeling
- Graph Data Modeling
- Text Data Modeling
- Dimensional Data Modeling
- Flat Data Modeling
- Relational Databases
- Data Query
- BI Operations
- Big Data (Hadoop, Map Reduce)

Analytics Competencies

- Basic Statistics
- Data Mining / Text Mining
- Geospatial Data Analytics
- Linear and Advanced Regression
- Planning Scheduling
- > Machine Learning / Artificial Intelligence
- Linear and Matrix Algebra
- > Time Series and Forecasting
- Cluster Analysis
- Survival Data Analysis
- > Design of Experiments
- Data Visualization
- > Marketing and Customer Analytics
- Financial Analytics
- Risk Analytics
- Web Analytics
- > Tools (R, Python, SAS, SPSS, etc.)

Self Study















Do Analytics

Get Tools:



http://cran.r-project.org



Databases Spreadsheets

Get Data:



FAKE NAME GENERATOR™

Book Downloads

Case Studies:









Compete

kaggle⁻

We're making data science a sport."

Participate in competitions

Kaggle is an arena where you can match your data science skills against a global cadre of experts in statistics, mathematics, and machine learning. Whether you're a worldclass algorithm wizard competing for prize money or a novice looking to learn from the best, here's your chance to jump in and geek out, for fame, fortune, or fun.

Create a competition

Kaggle is a platform for data prediction competitions that allows organizations to post their data and have it scrutinized by the world's best data scientists. In exchange for a prize, winning competitors provide the algorithms that beat all other methods of solving a data crunching problem. Most data problems can be framed as a competition.

Join as a participant

Learn more about hosting





Graduate Analytics Programs

North Carolina State University http://analytics.ncsu.edu/ Michael Rappa

Northwestern University http://www.analytics.northwestern.edu/ Diego Klabjan

Stanford University

University of Minnesota

University of Saint Thomas

University of Wisconsin



David Haertzen – Contact Information



David Haertzen Author and Instructor



http://www.first-place-learning.com/ http://www.linkedin.com/davidhaertzen http://ecm.elearningcurve.com/David_Haertzen_s/89.htm Twitter: #BigHeart7 david at davidhaertzen dotCom

