

Brad Hill – Senior Technical Sales Consultant



Data Mining & Text Analytics

Predicting outcomes with IBM SPSS Modeler



Business Analytics software

© 2011 IBM Corporation

Business Analytics Software



What if..


- Reduce current customer attrition by 89%
- Increase the numbers of student applications by 7%
- Increase profits by 300% with better cross sell offers
- Lower crime rates by 19% over 4 years
- Reduce marketing costs by 40% while increasing profit


© 2011 IBM Corporation

Business Analytics Software 



© 2011 IBM Corporation

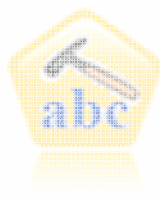
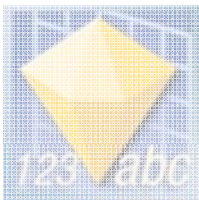

Business Analytics Software 



Predictive analytics **Data mining** **Text analytics**

© 2011 IBM Corporation


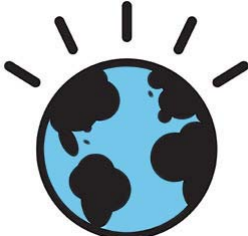
Business Analytics Software for a smarter planet




Predictive analytics Data mining Text analytics

© 2011 IBM Corporation


Business Analytics Software for a smarter planet



Instrumented

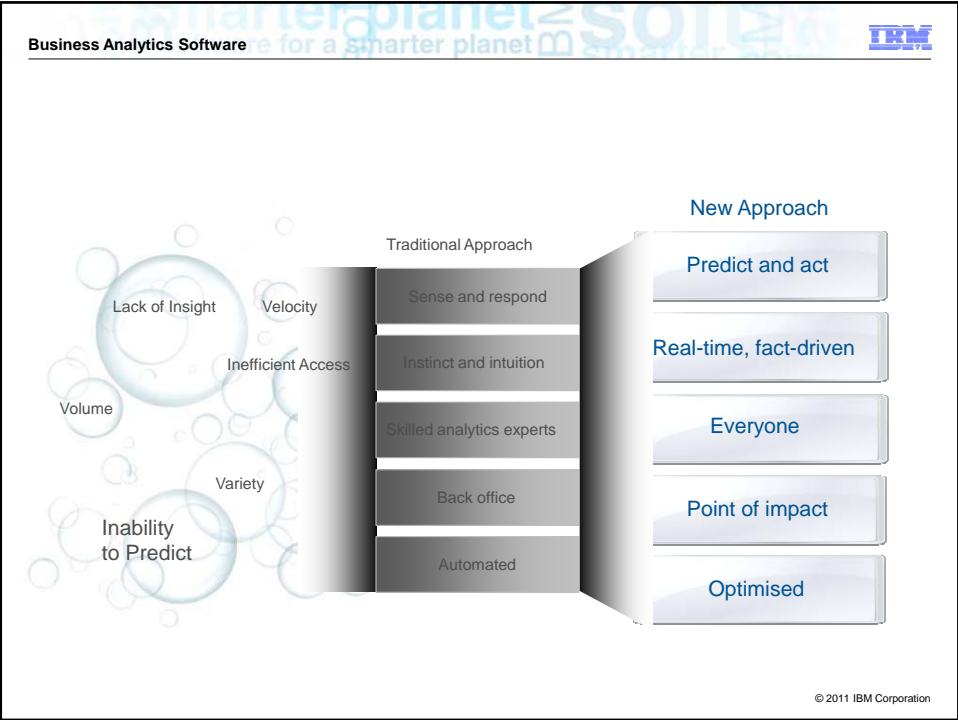


Interconnected




Intelligent

© 2011 IBM Corporation



Business Analytics Software IBM

What is predictive analytics?



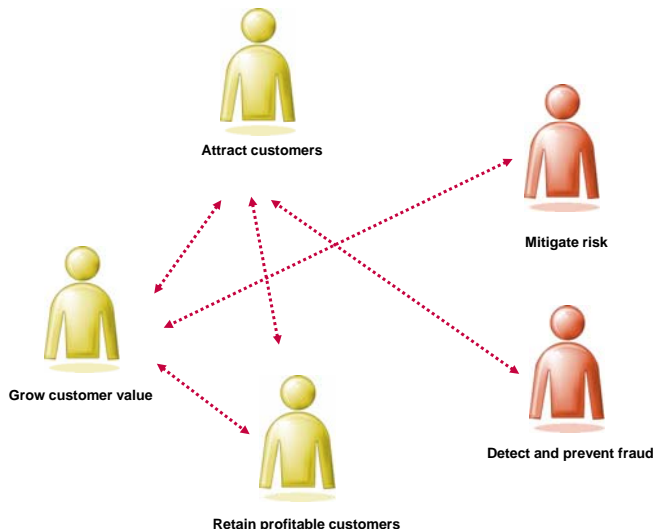
Predictive Analytics helps connect **data** to effective **action** by drawing reliable conclusions about current conditions and **future** events

Gareth Herschel, Research Director, Gartner Group

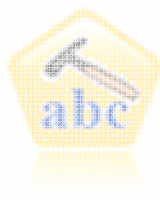
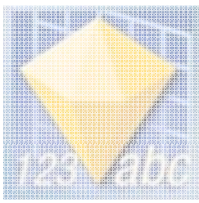

© 2011 IBM Corporation

Predictive analytics in action

- Customer relationship management “analytical CRM”
 - Who are our best customers?
 - Can we get more like that?
 - What/why do they buy?
 - Why do they leave?
- Human capital management
 - Who are our best employees?
 - How do we keep our best employees from leaving?
 - Which prospects should we recruit?
- Science
 - Genetics
 - Drug discovery
 - Medical research
 - Food authentication
- Fraud detection
 - Money laundering
 - Network intrusion
 - Tax audits & collection
- Crime analysis
- Industrial process optimisation & QA and many more...



Business Analytics Software for a smarter planet



Predictive analytics

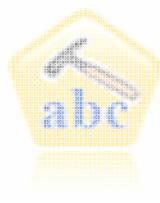


- Faster, better-informed decisions
- Improve business processes
- Better understanding customers/constituents
- Predict and act

Data mining

Text analytics

© 2011 IBM Corporation

Business Analytics Software for a smarter planet



Predictive analytics

Data mining

Text analytics

© 2011 IBM Corporation

Data Mining

- Discover key relationships between variables
- Use advanced analytical techniques on data
- Model effect of variables on outcomes
- Determine influence on outcomes
- Predict outcomes
- Apply models to new data

IBM SPSS Modeler

- High performance data mining and text analytics workbench
- Ability to create and operationalise predictive intelligence
- Used for the proactive and repeated...
 - Identification of revenue opportunities
 - Reduction of costs
 - Increase in productivity



Business Analytics Software **Complete workbench**

The screenshot displays the IBM SPSS Modeler 14.0 interface. At the top, it says "Business Analytics Software" and "Complete workbench". The main workspace shows a workflow diagram with the following steps: 1. Data sources: Comments.xls, Customers, and Churn. 2. Merge: A Merge node combines the data. 3. Comments: A node to handle the Comments field. 4. CHURN + [25 Fields]: A node for the target variable. 5. Type: A node for variable type conversion. 6. CHURN: A node for the target variable. 7. Churn Text: A node for text-based churn. 8. CHURN: A second node for the target variable. 9. Compare the models: A node for model comparison. 10. Analysis: The final analysis node. The interface includes a menu bar (File, Edit, Insert, View, Tools, Superfoster, Window, Help), a toolbar, and a bottom toolbar with various modeling tools like Database, Var, File, Auto Data Prep, Select, Sample, Aggregate, Derive, Type, Filter, Graphboard, Auto Classifier, Auto Numeric, Auto Cluster, Table, Flat File, and Database. A right-hand pane shows "Streams" and "Models" sections.

Business Analytics Software **Complete workbench**

This screenshot is similar to the one above, showing the IBM SPSS Modeler 14.0 interface with the same workflow diagram. However, the bottom toolbar is different, featuring icons for Favorites, Sources, Record Ops, Field Ops, Graphs, Modeling, Database Modeling, Output, Export, SPSS Statistics, and SPSS Test Analytics. The rest of the interface, including the menu bar and right-hand pane, is identical to the first screenshot.



Business Analytics Software IBM

Complete workbench

Hands on 2d - PASWB Modeler 14

The screenshot shows the IBM PASWB Modeler 14 interface. The main workspace contains a workflow diagram with the following nodes and connections:

- Inputs: Comments, Customers, Churn.
- Process: Merge (receives input from Comments and Customers).
- Process: CHURN x [25 Fields] (receives input from Merge).
- Process: Comments (receives input from Merge).
- Process: Type (receives input from CHURN x [25 Fields]).
- Process: CHURN Test (receives input from Type).
- Process: CHURN (receives input from Type).
- Process: Analysis (receives input from CHURN Test and CHURN).

The bottom toolbar includes: Favorites, Sources, Record Ops, Field Ops, Graphs, Modeling, Database Modeling, Output, Export, PASWB Statistics, and PASWB Test Analytics. The bottom status bar shows 'Server: Local Server' and '252MB / 420MB'.

Business Analytics Software IBM

Complete workbench

Hands on 2d - PASWB Modeler 14

This screenshot is similar to the one above but includes annotations:

- A yellow box above the Merge node says "Join the datasets based on ID".
- A yellow box above the CHURN Test node says "Different models with and without using test".
- A yellow box above the Analysis node says "Compare the models".

The workflow diagram and bottom toolbar are identical to the previous screenshot.



Business Analytics Software IBM

Complete workbench

Business Analytics Software IBM

Data mining techniques

Technique	Algorithms	Usage
Classification (or prediction)	Auto Classifiers, Decision Trees, Logistic, SVM, Time Series, etc	Used to predict group membership (ie will this employee leave?) or a number (ie how many widgets will I sell?)

© 2011 IBM Corporation



Data mining techniques

Technique	Algorithms	Usage
Classification (or prediction)	Auto Classifiers, Decision Trees, Logistic, SVM, Time Series, etc	Used to predict group membership (ie will this employee leave?) or a number (ie how many widgets will I sell?)
Segmentation	Auto Clustering, K-means, etc.	Used to classify data points into groups that are internally homogenous and externally heterogeneous.
	Anomaly detection	Identify cases that are unusual

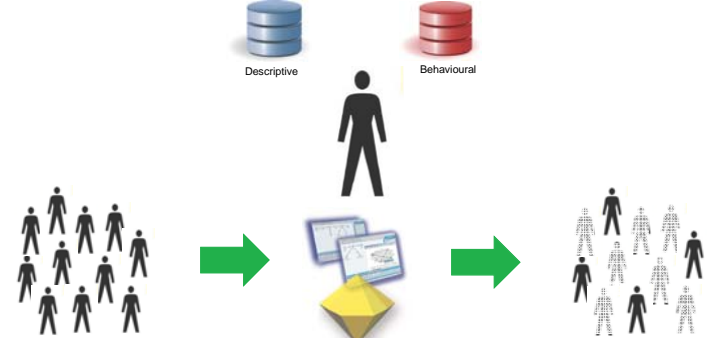


Data mining techniques

Technique	Algorithms	Usage
Classification (or prediction)	Auto Classifiers, Decision Trees, Logistic, SVM, Time Series, etc	Used to predict group membership (ie will this employee leave?) or a number (ie how many widgets will I sell?)
Segmentation	Auto Clustering, K-means, etc.	Used to classify data points into groups that are internally homogenous and externally heterogeneous.
	Anomaly detection	Identify cases that are unusual
Association	APRIORI, Carma, Sequence	Used to find events that occur together or in a sequence (ie market basket).

Scenario

- Customer and product data
- Explore and understand data
- Build a model to identify customers likely to respond
- Generate a list for marketing



© 2011 IBM Corporation

Demonstration



Business Analytics software

© 2011 IBM Corporation



Business Analytics Software

How predictive intelligence gets deployed

The screenshot shows a call center agent's interface. At the top, there are navigation tabs: 'my activities', 'team activities', 'products', 'charts', and 'introductions'. Below this is a header with 'Welcome, John Palmer' and 'log out help'. The main area is divided into several sections:

- Customer Details:** A form with fields for Last name (Wes), Gender (M), First name (Frank), Address (Crossgroun), Age (43), City (Cuburg), Profession (Manager), and Zipcode (6893 OK). A 'get info' button is at the bottom.
- Products:** A table with columns ID, Description, and Group. It lists:

12	Teen Visa Card	Banking
13	Home Equity Loan	Banking
14	Easy Access Account	Banking
- Contact history:** A table with columns Description, Date, and Result.
- Details current call:** A section with a dropdown for 'Type of contact' (set to 'to be determined') and a 'submit' button.
- Recommendation:** A section with 'Interaction' (Prevent Churn-IV (Single)) and 'Offer' (Retention - Racins). A dropdown menu for 'Action' is open, showing options: 'Select option', 'Accept offer', 'F 011: Conversation took too long', 'F 012: Customer not in the mood', and 'F 013: Already on target; list in 5'.
- Message:** A text input field at the bottom.

A call center agent submits customer information during an interaction

Based on the predictive model, a single offer is presented to the customer

The reaction to the offer is tracked and used to refine the model

© 2011 IBM Corporation

Business Analytics Software

Predictive analytics

Data mining

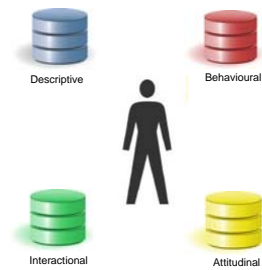
- Access a variety of data
- Easy to use graphical interface
- Automatic data preparation and modelling
- Add value to BI
- Create predictive intelligence which can be operationalised

Text analytics

© 2011 IBM Corporation

Text analytics

- Extract, analyse and create structure from unstructured data
- Integrate analysis results into operational systems
- Integrate analysis results into Business Intelligence applications
- Integrate analysis results with structured data and use as input for Data mining
- Improves model accuracy.





Business Analytics Software IBM

Text analytics within IBM SPSS Modeler

Interactive Workbench - Q4carcomments

File Edit View Generate Categories Tools Help

Category	Descriptors	Docs
All Documents	-	200
Uncategorized	-	48
No concepts extracted	-	1
Pos. General Satisfaction	26	63
Pos. Product Functioning	18	45
Neg. Product Functioning	15	28
Neg. Service Accessibility	38	15
Neg. Product Design/Features	3	12
Other: Don't Know	14	11
Pos. Pricing and Billing	3	8
Neg. Product Availability/Variant/Size	12	6
Neg. General Dissatisfaction	14	6
Cont: Pricing and Billing	5	6
Cont: Company Public Image/Reputation	3	6

Category	Bar	Selection %	Docs
Pos. General Satisfac...	[Bar]	53.8	
Pos. Product Functi...	[Bar]	38.5	
Neg. Product Functi...	[Bar]	23.9	
Neg. Service Acces...	[Bar]	6.0	
Other: Don't Know	[Bar]	4.3	
Neg. Product Desig...	[Bar]	3.4	
Cont: Company Pu...	[Bar]	3.4	
Pos. Pricing and Bil...	[Bar]	2.6	
Neg. General Dissat...	[Bar]	2.6	
Cont: Pricing and Bi...	[Bar]	2.6	
Pos. Service Attitud...	[Bar]	1.7	
Pos. Product Usabil...	[Bar]	1.7	
Neg. Product Avail...	[Bar]	1.7	
Pos. Service: Genera...	[Bar]	0.9	
Neg. Pricing and Bil...	[Bar]	0.9	
Pos. Service: Acces...	[Bar]	0.9	

Concept	In	Global	Docs	UType
car	fx	106 (14%)	93 (47%)	DU +Products+
good	fx	57 (7%)	47 (24%)	DU +Positive+
clean	fx	37 (5%)	37 (19%)	DU +PositiveFeeling+
excellent	fx	19 (2%)	19 (9%)	DU +Positive+
well	fx	12 (2%)	12 (6%)	DU +Positive+
dirty	fx	10 (1%)	10 (5%)	DU -NegativeFeeling-
like	fx	9 (1%)	9 (5%)	DU +Positive+
dislike	fx	9 (1%)	9 (5%)	DU -Negative-
small	fx	8 (1%)	8 (4%)	DU -Contentful-
bad	fx	8 (1%)	8 (4%)	DU -Negative+
no problem	fx	7 (1%)	7 (4%)	DU +Positive+
satisfied	fx	6 (1%)	6 (3%)	DU +Positive+
price	fx	6 (1%)	6 (3%)	DU -Budget-
fun	fx	6 (1%)	6 (3%)	DU +Positive+
brand name	fx	6 (1%)	6 (3%)	DU +Positive+

Q4carcomments (117)	Categories
4 car was [redacted] and [redacted] on miles.	Pos. Product Functioning
5 car was [redacted] but the [redacted] were very dirty on the inside.	Neg. Product Design/Feat. Neg. Service Accessibility Pos. Product Functioning
6 car was [redacted]. it would be helpful if [redacted] contained a [redacted] instruction sheet illustrating where things are (i.e. gas cap, [redacted] [redacted]) as well as how to operate the security keychain.	Pos. Product Functioning Pos. Product Information Pos. Service: Knowledge
7 car was economy. [redacted] [redacted] gave us trouble. We had to change it.	Neg. Product Functioning
8 car was [redacted] [redacted] worked, it was [redacted], and [redacted] drove really fast.	Pos. Product Functioning Pos. General Satisfaction
9 car was [redacted], but it had an [redacted] [redacted] or something	Pos. General Satisfaction

Business Analytics Software IBM

Text analytics within IBM SPSS Modeler

Interactive Workbench - Q4carcomments

File Edit View Generate Categories Tools Help


Global	In	Type1	Type2
1		<CustomerSupport>	-Negative+
1	fx	<Documentation>	-Negative+
1	fx	<Budget>	-NegativeBudget-
1	fx	<Unknown>	-NegativeBudget-
7	fx	<Products>	-NegativeFeeling-
6	fx	<Messages>	-NegativeFunction-
2	fx	<Unknown>	-NegativeFunction-
1	fx	<Characteristics>	-NegativeFunction-
59		<Products>	+Positive+
28		<Unknown>	+Positive+
4	fx	<Characteristics>	+Positive+
3	fx	<Budget>	+Positive+
2		<Performance>	+Positive+
1		<Location>	+Positive+
1	fx	<Documentation>	+Positive+
1		<Products>	+Positive+

Global	Docs	In	Concept	Concept2
3	3		condition	good
1	1	fx	speakers	blown
1	1	fx	windshield wiper	not working
1	1	fx	red car	fading
1	1	fx	air conditioning	not working
1	1	fx	speedometer	not working
1	1	fx	paint	fading
1	1	fx	mid-range sedan	disconnected
1	1		warning	light
1	1		map	light
1	1		sun	well maintain
1	1		aspects	satisfied
1	1		idea to rent a	excellent
1	1		drive	good
1	1		ford taurus	roomy
1	1		vacation	fun

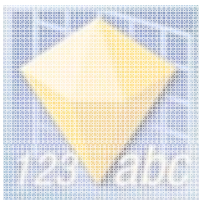
ConceptWeb Type Web

To populate data pane:
Make a selection in a table and click Display


Business Analytics Software IBM



Predictive analytics



Data mining




Text analytics

- 360° view of customer
- Improve model accuracy
- Achieve better insight
- Do something with qualitative data


© 2011 IBM Corporation

Business Analytics Software IBM




Predictive analytics

- Faster, better-informed decisions
- Improve business processes
- Better understanding customers/constituents
- Predict and act



Data Mining

- Access a variety of data
- Easy to use graphical interface
- Automation data preparation and modelling
- Add value to BI
- Create predictive intelligence which can be operationalised



Text analytics

- 360° view of customer
- Improve model accuracy
- Achieve better insight
- Do something with qualitative data

© 2011 IBM Corporation

Business Analytics Software IBM


Instrumented

Interconnected

Intelligent

Grow, Attract, Retain, Risk, Fraud

Business Analytics Software IBM



Increase public safety

Challenge

- Find innovative ways to fight escalating crime
- Find a cost-efficient way to analyse crime data, assess public safety risks, make intelligent decisions about personnel

Solution

- Analysts and officers use IBM SPSS Modeler to pore through data and find crime patterns and predict outcomes
- Forecast strategic positions for personnel and deployed "hot spot" maps to officers
- Used to identify key crime patterns to develop proactive policing strategies


Results


- Dramatic reduction in crime between 2006 and 2007 despite economic conditions
- New Year's Eve test saw 246% increase in weapon seizure, 49% decrease in gunfire, and \$15,000 savings in overtime
- Gives even rookie officers veteran-like insight into crime data

"IBM SPSS Modeler and data mining represent a revolution in our ability to access previously unobtainable data, and pull meaning and value from it. This is as close to a crystal ball as we are ever going to get."
 — Colleen McCue, program manager for the Department's Crime Analysis Unit

© 2011 IBM Corporation



Business Analytics Software 

 **INFINITY**
PROPERTY & CASUALTY CORPORATION

Claims identification

Challenge

- Reduce payments on fraudulent claims
- Improve ability to collect payments from other insurance companies

Solution


- Used IBM SPSS Modeler to develop models of fraudulent claims
- Leveraged text analytics to interpret and analyse handwritten notes for use in investigation
- Extended use of predictive analytics beyond claims to customer retention and pricing analysis


Results

- 403% ROI in first 3 months
- Realised \$5 Million in benefit in the first year post-implementation
- Reduced cost of claims payment by enabling earlier, more targeted investigations
- Models deployed within call center to streamline claims process and gather the right data

"The relationship we have with our customers is put to the test when they file a claim, as they want a resolution so their lives can return to normal as quickly as possible. With SPSS, we can fast track valid claims or flag possible counterfeit claims for further review, saving our customers time and money."
↳ Bill Dibble, SVP of Claims

© 2011 IBM Corporation

Business Analytics Software 



Maximises revenue from targeted email marketing

Challenge

- What factors drive direct business through their e-commerce channels
- How to use wealth of customer data to tailor each marketing communication to a customer's unique needs

Solution

- Used IBM SPSS Modeler to develop customer profiles
- Used IBM SPSS Modeler to develop more accurate segmentation models
- Applied predictive intelligence to e-mail marketing campaign to target the right communication to the right customer


Results

- Cost of e-mail marketing as a percentage of revenue (CPR) cut by 42% in 2009 vs. 2008
- Increased insight into customer activity drives loyalty
- Models and customer segmentation revealed where to target marketing spend

"The Customer Segmentation project allows us to keep in touch with our large database using cost-effective e-mail, but with all the benefits of a one-to-one relationship because we now have a clearly defined picture of each customer."
↳ Chris Parker, direct analytics specialist at Avis Europe

© 2011 IBM Corporation



Business Analytics Software 

Baruch COLLEGE

Gain and retain the right students

Challenge

- Access data held in multiple silos (admissions office, registrar, accounts receivable, etc.)
- Increase market visibility and target specific segments of prospective students

Solution

- Used IBM SPSS Modeler to access and consolidate multiple data stores to create a single view
- Created models for at-risk students, course placement, and student retention, and more
- Applied predictive intelligence across the student lifecycle


Results

- In a declining business school market, saw 7.1% increased applications to business school
- 21% annual increase in transfer students
- Decreased dropouts significantly by using predictive analytics to improve freshman placement

“These days, no meeting to make policy changes takes place without analysis based on predictive analytics”

— Jimmy Jung, Assistant VP for Enrollment Management


© 2011 IBM Corporation

Business Analytics Software 

Contact us

Phone:
132 426

Web:
www.ibm.com/spss/au



© 2011 IBM Corporation

Seminar feedback

WIN an Apple iPad2!!
Just complete the online survey
via email link after the event.



© 2011 IBM Corporation



© 2011 IBM Corporation