



Getting Started With Predictive Analytics

Rob Risany, IBM Business Analytics

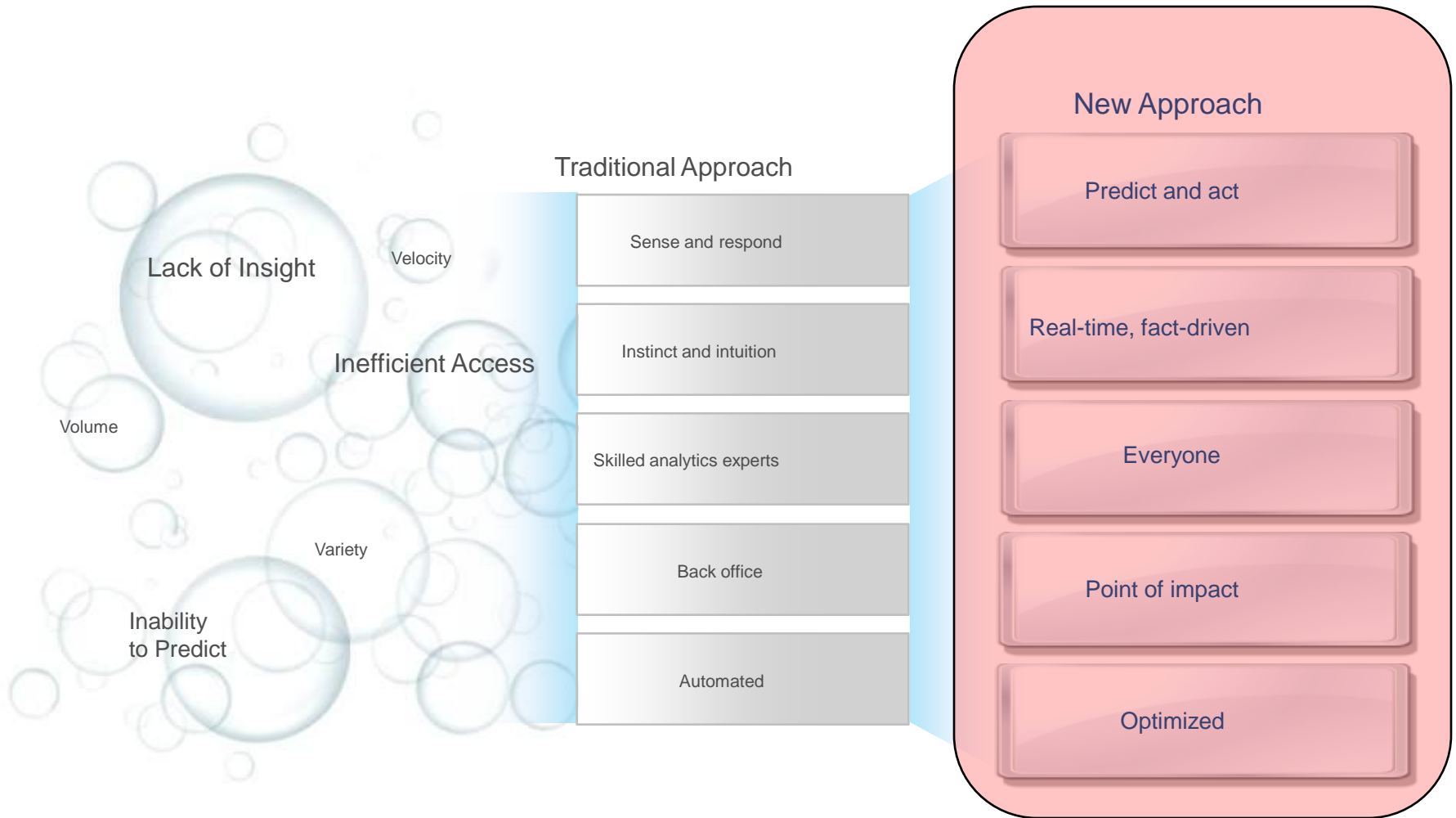
Information On Demand 2010

The Premier Forum for Information & Analytics

Gain Insight. Optimize Results.



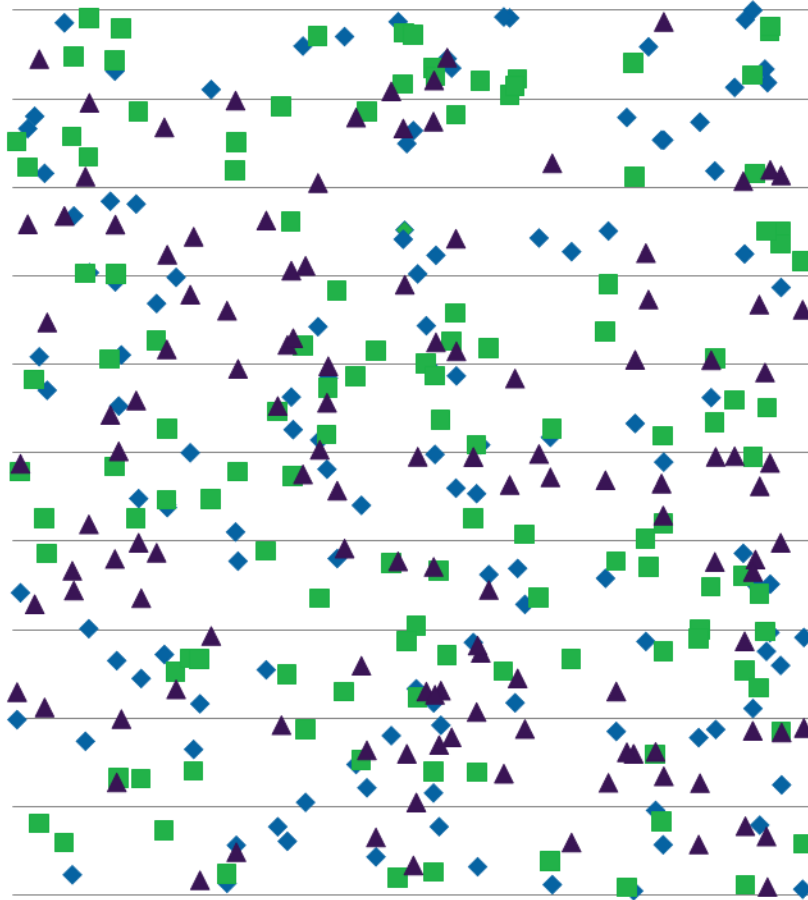
Transformation – The goal of predictive analytics



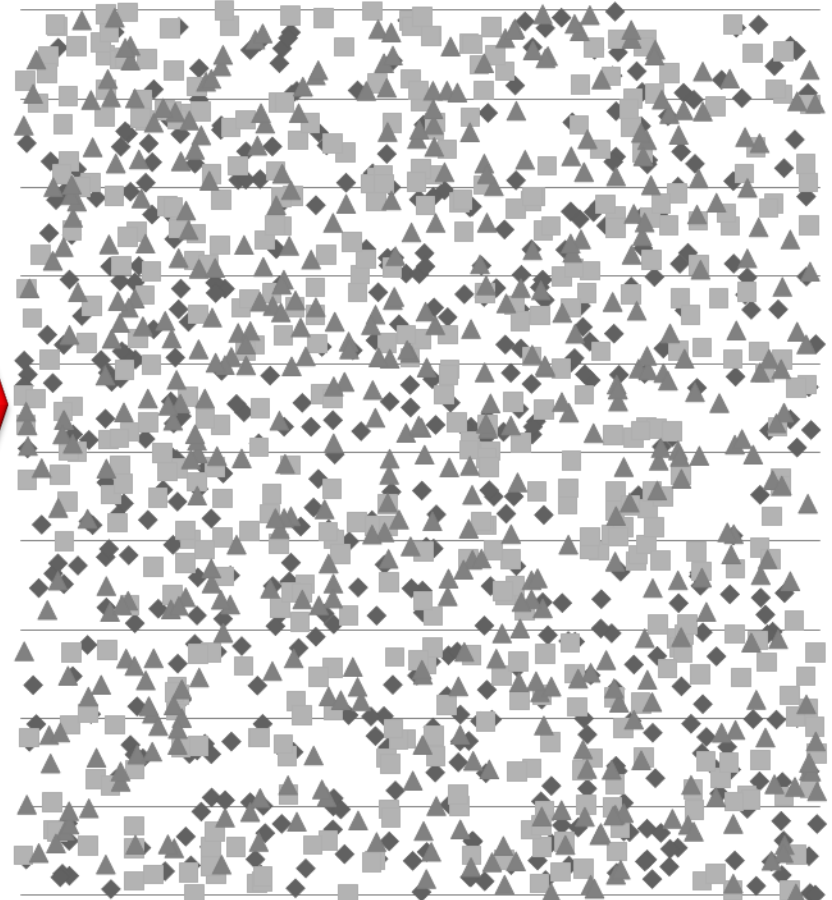


Predictive Analytics in a Nutshell: *Statistics*

A Sample of Data



A Universe of Data

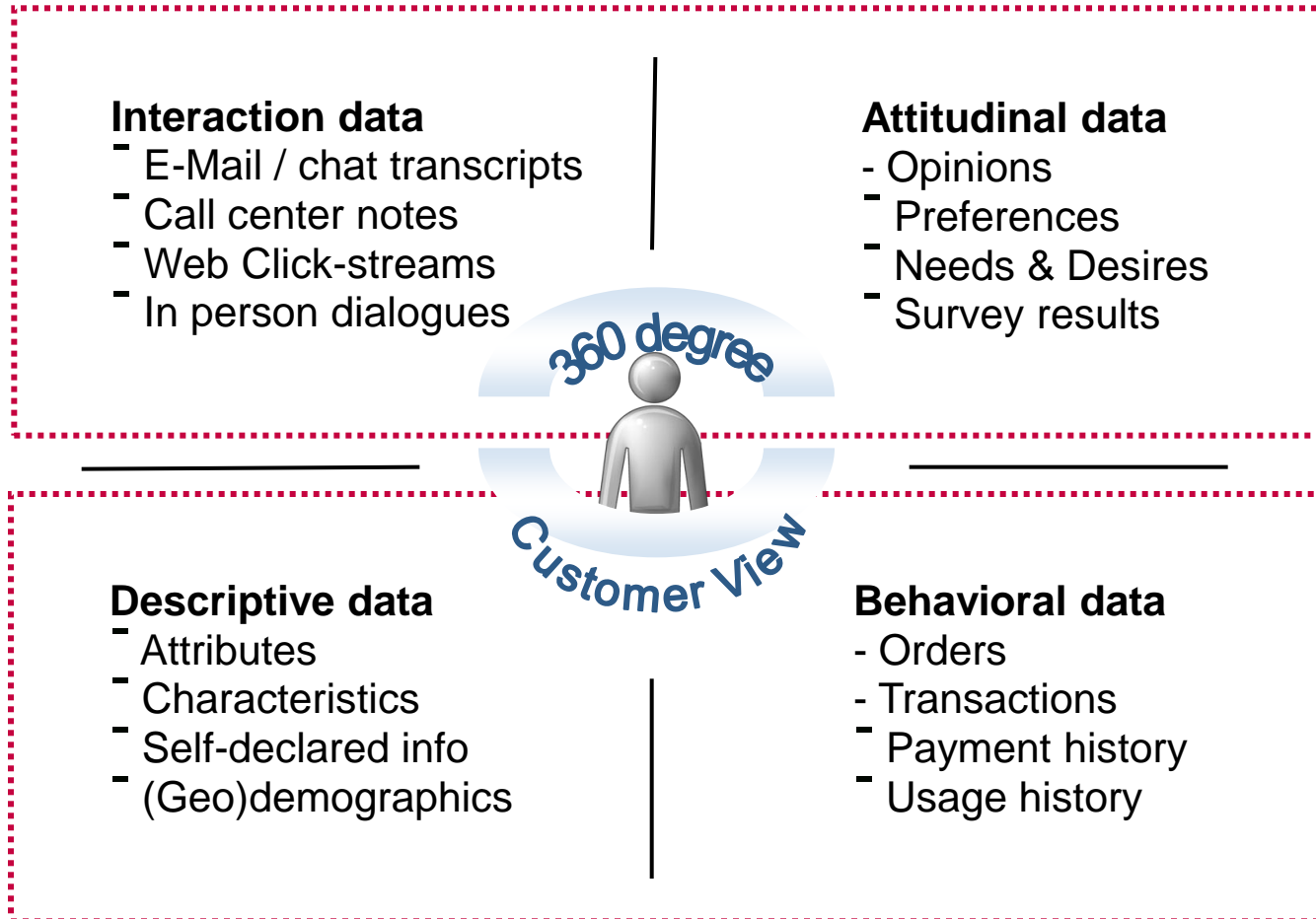


Getting a More Accurate Picture

All the Data Matters



High-value, dynamic
- source of competitive differentiation

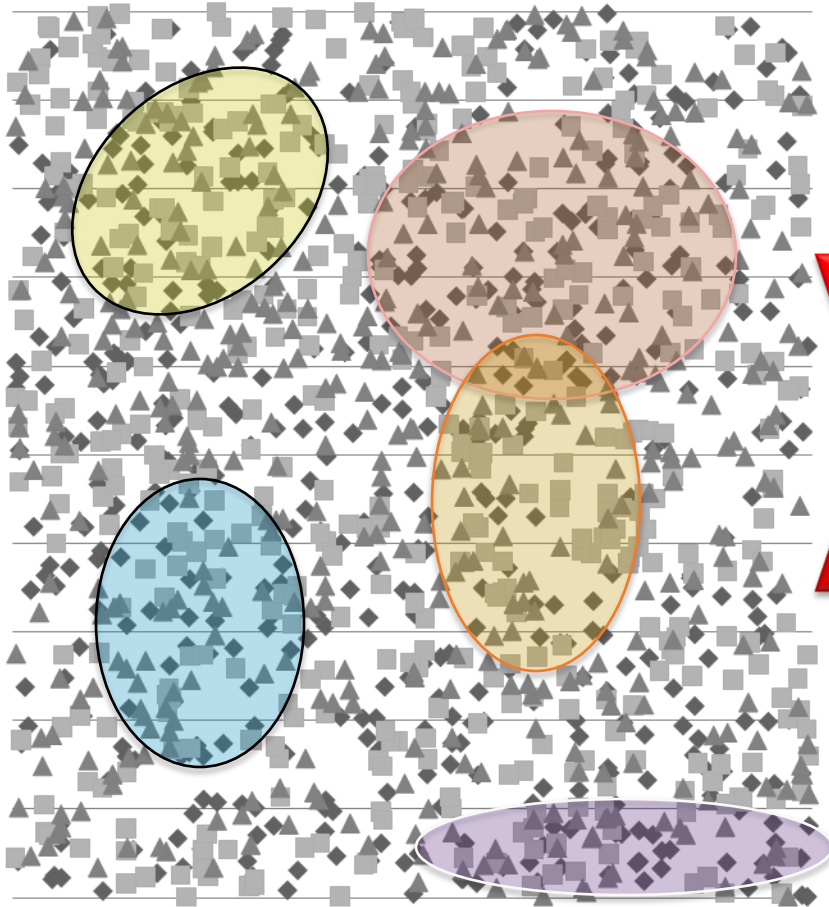


“Traditional”



Predictive Analytics in a Nutshell: *Data Mining*

A Universe of Data



A Predictive Model

Attributes:

- Married, 1 kid
- Lives in Suburbs of Chicago
- Owns two Cars
- 38 years old
- Drinks Scotch

Predicted Attributes

- Likes Guns n Roses
- Likes Johnnie Walker Black
- Works long hours
- Commutes
- Upper Middle Income

Predicted Behavior

- Wants to Buy a Sportscar!
- Buys Car Washes!
- Buys Chardonnay
- Vacations where it's warm

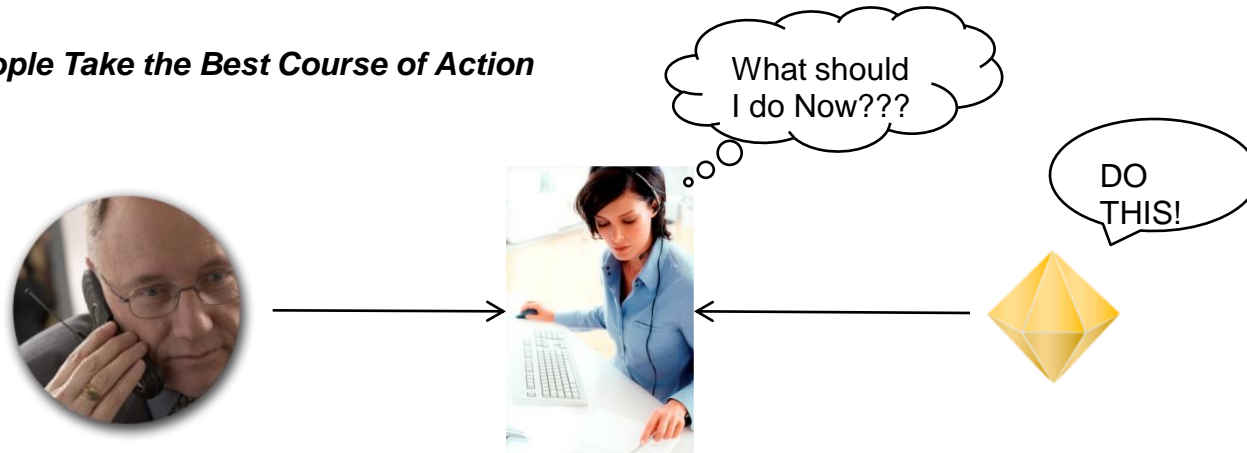


The End Game of Predictive Analytics

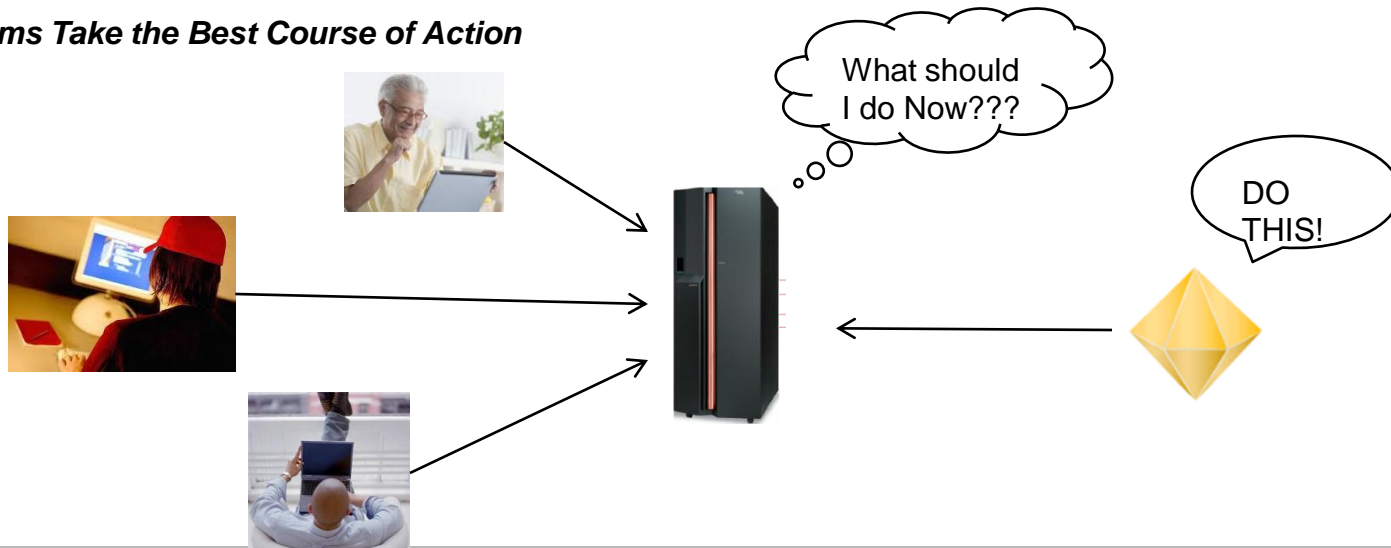
Deployment: Apply Predictions at the Point of Interaction



Help People Take the Best Course of Action

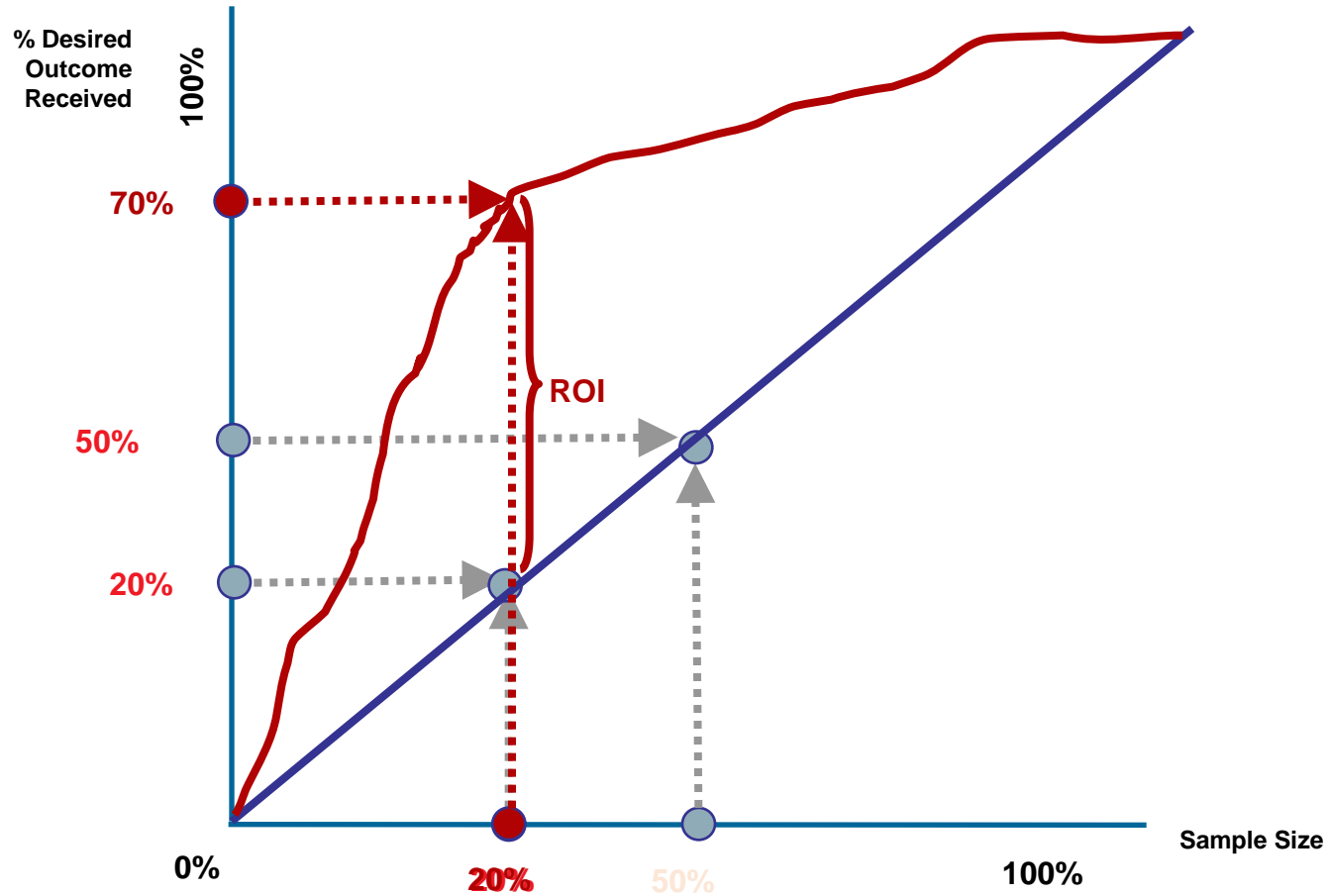


Help systems Take the Best Course of Action





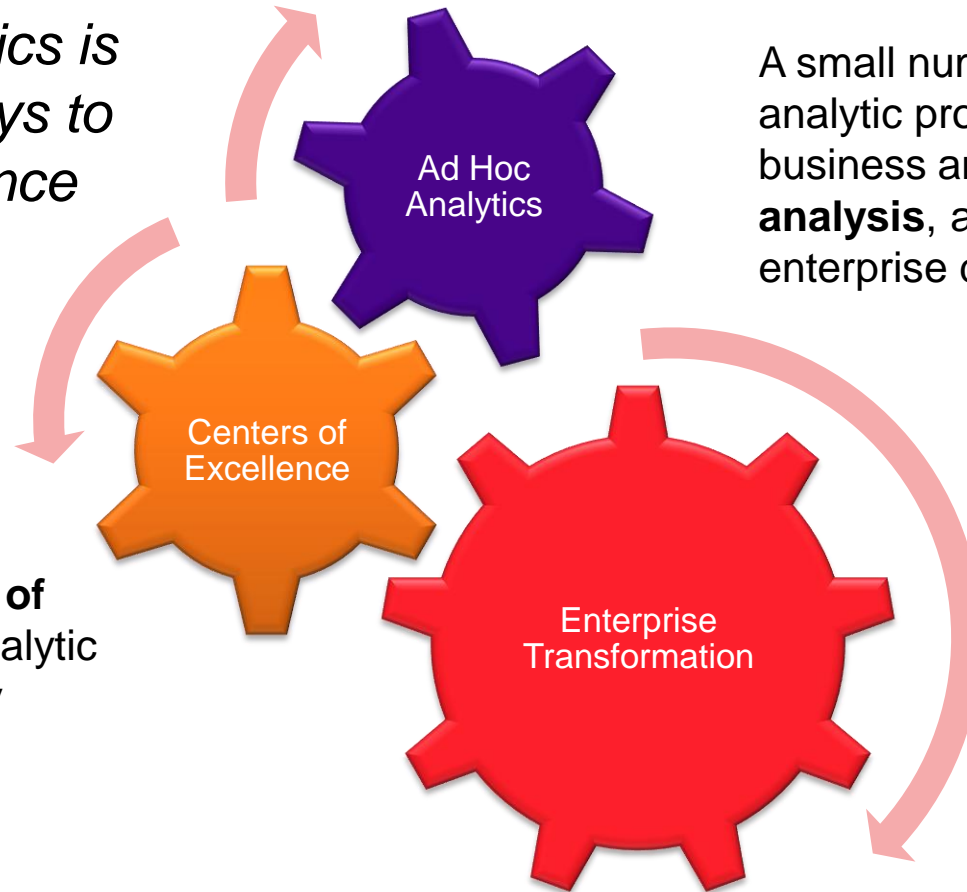
Predictive Analytics in a Nutshell: *Better Outcomes*





How do analytics get into the DNA of an organization?

Predictive analytics is used in three ways to deliver performance



A small number of individual analytic professionals to use business analytics for **ad-hoc analysis**, and understanding of enterprise data





Departmental **centers of excellence** provide analytic foresight for use in key business decisions

Operational outcomes in one or more **domain specific points of interaction** use the results of business analytics in the context of business process





Different Approaches All Provide Value

Sector	Company	Need	Results
Retail		Market Mix Analysis	<u>Analytic Center of Excellence</u> Improved inventory turns, and optimized product availability to move SKUs at stores
Insurance		Insurance Abrogation	<u>Enterprise Transformation</u> Found overpayments and collected millions
Consumer Goods		Product Design & Customer Sensitivity	<u>Ad Hoc Analytics</u> Researched key product features to drive increased differentiation
Telecommunications		Customer Intimacy	<u>Enterprise Transformation</u> Reduced customer churn from 19% to 2%



How you'll know if it's time to move beyond ad-hoc analytics



- Tough Analysis
 - My analysis is really complicated, and IT won't get me a new laptop...
 - I've got other things to do other than this analysis....
- Big Data
 - I've got large data sets in my data warehouse.... ..
 - My data is messy and needs a lot of work...
 - I've got too much unstructured content
- Governance
 - Sometimes our analysts work from home...
 - My analysts don't seem to work together effectively...
 - We waste a lot of time doing grunt work





Business Considerations

- What are you trying to accomplish
 - The Goal
 - “Improve claims processing via fast tracking of claims, automate fraud analysis and improve subrogation process.”
 - “Reliably predict the areas with the highest likelihood for violent crime, thereby enabling the Department to make informed deployment decisions and proactively prevent crime. Automatically incorporate new patterns of crime.”





Data Considerations

- What data is important to the problem and do you have it?
 - Understanding of existing data
 - Structured vs. Unstructured data
 - Number of sources
 - Centralized vs. Geographically dispersed
 - Warehouse vs. Operational
 - What other data will help improve the analysis?
 - Centralize data for performance
 - During model deployment, will I have the relevant data available?
 - Enterprise View for consistent semantics and view of data across uses





Design Considerations

What techniques are available and useful for your analysis?

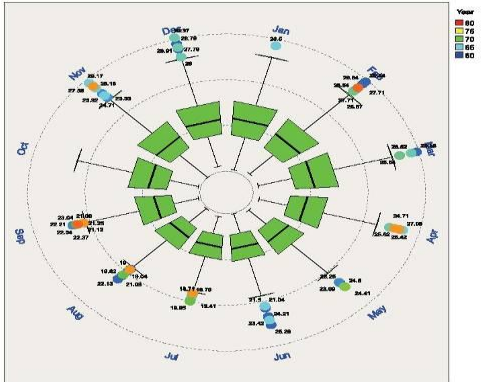
- Analytical Considerations
 - The analysis designed around available data to accomplish your goal
 - Designing the process / model so it's repeatable
 - Updating analysis to reflect new patterns
 - Automating all or part of the analysis process
- Technical Considerations
 - Dependencies between processes
 - Development vs. Production environment
 - Integrating results into operations: real-time / batch
 - Deployment Targets
 - Database
 - Self-service web portal
 - Operational line-of-business system



Output considerations: What deliverables help you meet your goal?



Complex Visualizations



Reports

Management Portals / Dashboards
Business Performance Overview

	Jan
Cash and Short Term Investments	1,817,058
Receivable, Net of Non-Current	-643,863
Other Current Assets	451,859
Total Current Assets	1,625,054
Total Assets	1,682,966
Accounts Payable	-1,791,146
Other Current Liabilities	6,915
Total Current Liabilities	-1,697,514

New Mobile Applications

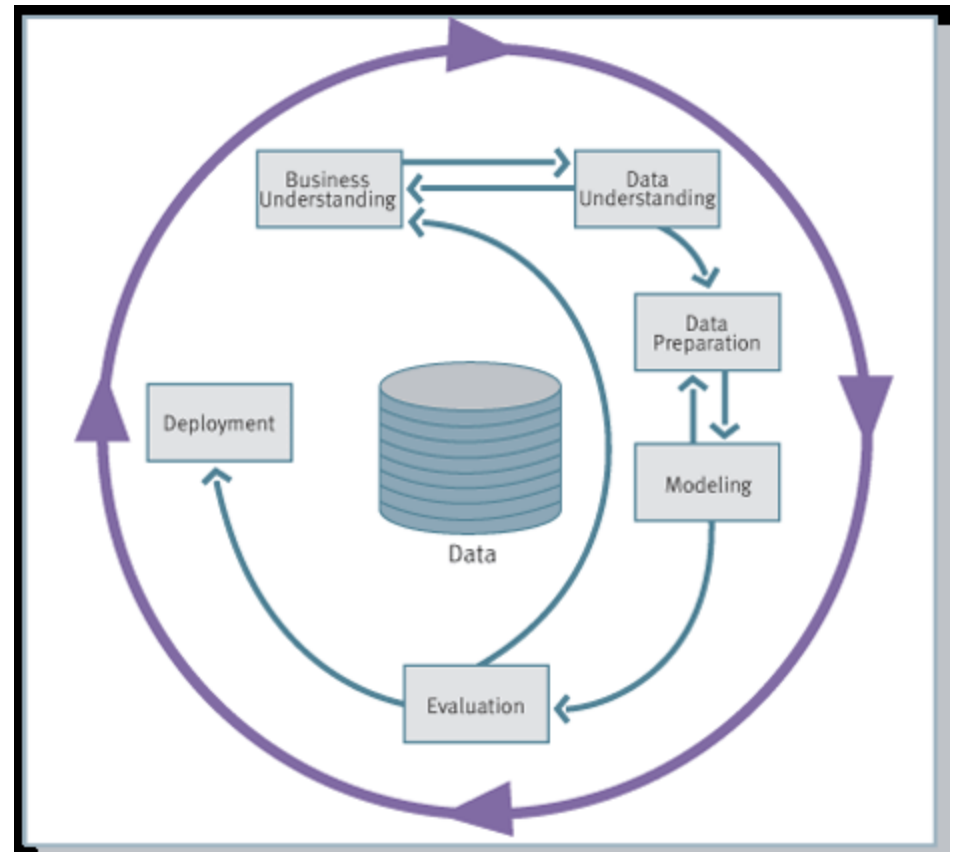


Customizations of
Point of Interaction Systems



The Predictive Analytics Process

- CRISP-DM
 - Cross Industry Standard Process for Data Mining
- Business understanding
- Data understanding
- Data preparation
- Modeling
- Evaluation
- Deployment

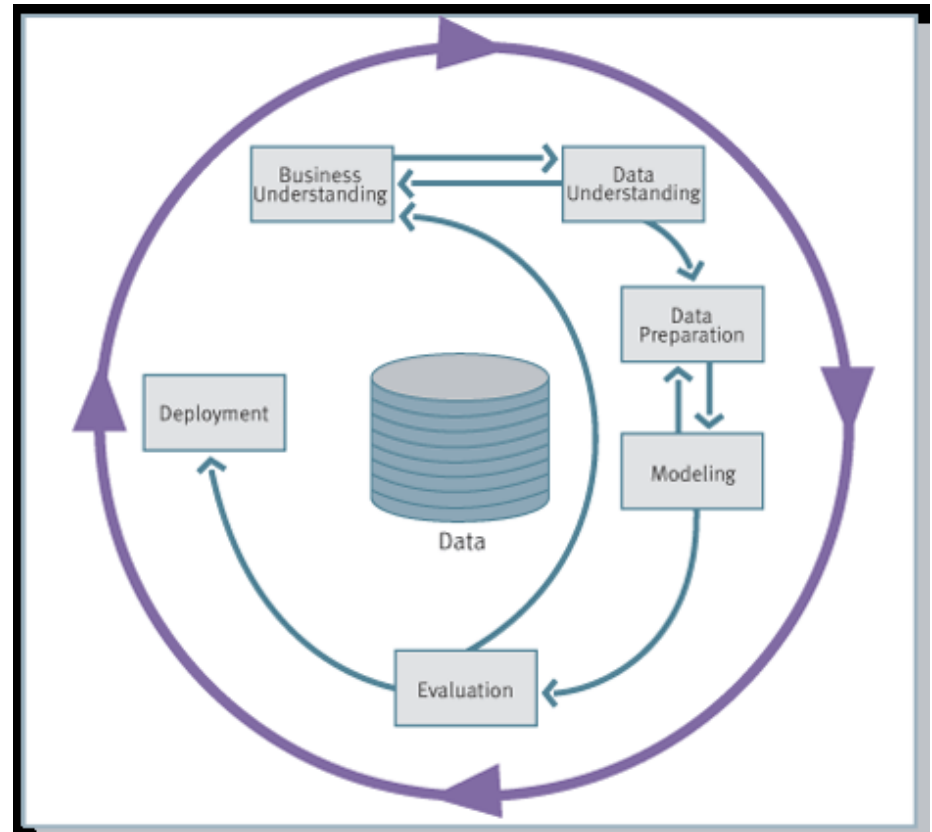




It must be iterative....

- How often will you need to update the solution?
 - Automatically
 - By an analyst
- How do you want this done?
 - Solutions can quickly become outdated
- Time to market of the new solution is critical
 - Solutions can quickly become outdated

Analytics aren't an end point
they are an iterative process





What is Deployment?

- It's about how it gets into the DNA of your organization
- A predictive analytics 'solution' is more than just creating analytic output.
- Think about how the entire process will be deployed in your organization.



Consumer facing points of interaction?

The screenshot shows a 'Deployment Portal' interface. At the top, there are navigation links: Home, Content Repository, Preferences, My Jobs, Model Management, About, and Help. Below this is a search bar and a 'Go!' button. A breadcrumb trail reads: Content Repository > Glenn Van Renschooten > Reports > Customer Status Report.rpt. The main content area displays a file named 'Customer Status Report.rpt' with a 'Modified On: 4/29/09 4:50 PM' timestamp. Below the file information, there are two main sections: 'Dashboards & Reports?' and 'Customer Status Report'. The 'Dashboards & Reports?' section contains several charts and tables, including a bar chart for 'YTD Shipped' and a table for 'YTD Shipped and Revenue'. The 'Customer Status Report' section shows a summary table with columns for 'City, State, Zip', 'Total Sales', and 'Total Cost'. The summary table shows a total sales of \$108,425.00 and a total cost of \$69,825.00.

System Level Deployment

The screenshot shows a customer details page. At the top, there are navigation links: my account, team activities, products, items, introduction. Below this is a user profile for 'Welcome, John Palmer' with a 'log out' link. The page is divided into several sections: 'Customer Details', 'Products', 'Contact History', 'Details current call', and 'Recommendation'. The 'Customer Details' section includes fields for Last name, First name, Address, City, Zipcode, Gender, Age, and Profession. The 'Products' section lists items like 'Teen Visa Card', 'Home Equity Loan', and 'Easi Access Account'. The 'Contact History' section shows a table with columns for Description, Date, and Result. The 'Details current call' section has a 'Type of contact' dropdown and a 'Description' text area. The 'Recommendation' section is highlighted with a red box and contains a table with columns for Interaction, Other, and Action. The 'Action' column has a dropdown menu with options like 'Get more info', 'Add to cart', and 'Remove from cart'. A 'submit' button is located below the recommendation table.

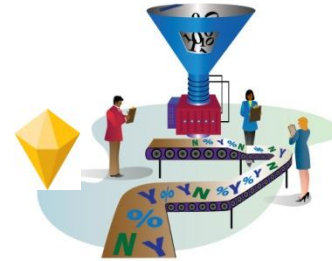




Degrees of deployment



Visualization & Reporting



Batch Scoring



Optimized Decisions



Ad-hoc Discovery



Manual Scoring

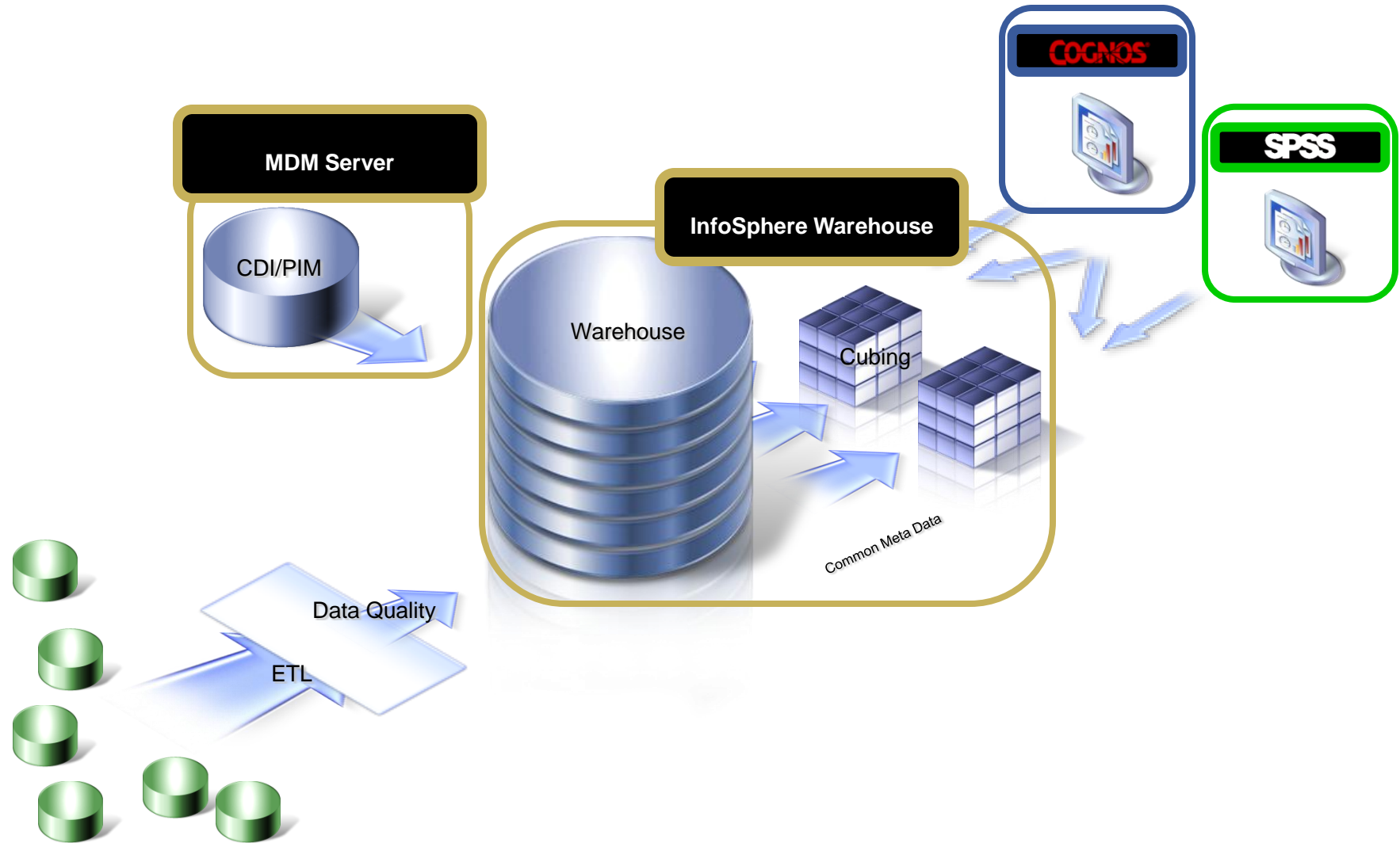


Automating Decisions



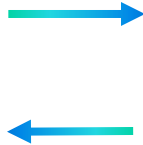


integration to Enterprise Data Infrastructure





Enterprise Transformation in Action



Hi, How can I help you?

Cross-sell?
<context data>



<customer data>



Example
Leveraging Service Calls as Sales Opportunities

Potential Campaign	Valid in this case?	Margin	Response Probability	Expected Value
A	No			
B	Yes	90	54%	49
C	Yes	85	62%	64



Transformation at the point of interaction....

FKCRG030 Cross-sell information screen

Relation nr : 1409797 Mr. HM van Ee

Product Cross-sell entry Sales arguments
 Cont. Travel RPP - 25 pt Protection against trip cancellation
 Insurance Emergency Medical Expenses
 Yes For travellers with more than 4 wk/year

RPP-member : Yes Group contract : K60
 Age : 32 (18-04-1970) Customer since : 04 years 05 Months
 Family topology : Married/children

Products in use :
 Third party liab. | Line of Credit | Gold 5 Polis
 Property Insurance

Products in use in same household :
 Orange Investment Fund | Savings account

Contact information

Cont date	Product / Campaign	Summary
2002-11-18		Log on website
2002-11-15	WVA upsell OLP	Brochure send by marketing
2002-11-15		Log on website

Risk Profiling on Legacy System

SPARBELEG

INVESTEREN OVERZICHTEN

Filtering: 07/2014/05 TELEFONISCHE OVERZICHTING Restwaarde: NLS 0,00

Houdert 1: 90000017 Bergama, V. 10/05/2010 N
 BALEN

Van Rentalbeleg Actual saldo: 0,00 Afschaf saldo: 0,00

Resting: NLS afbreken
 Houd tegenwoordig 90000017 Bergama, V.

Diskeerfonds: 07/2014/05

Spaard Sponderaal afgeven
 Deure

Overeenstemming Product Begindatum Bron Einddatum

0762914165	RenteEekoning	20010326	ADV	0
0762914270	RenteEekoning	20010326	KLANT	0
0762914289	RenteEekoning	20010326	KLANT	0
0762914300	RenteEekoning	20010403	KLANT	0
0762914319	RenteEekoning	20010403	KLANT	0
0762914335	RenteEekoning	20010409	KLANT	0

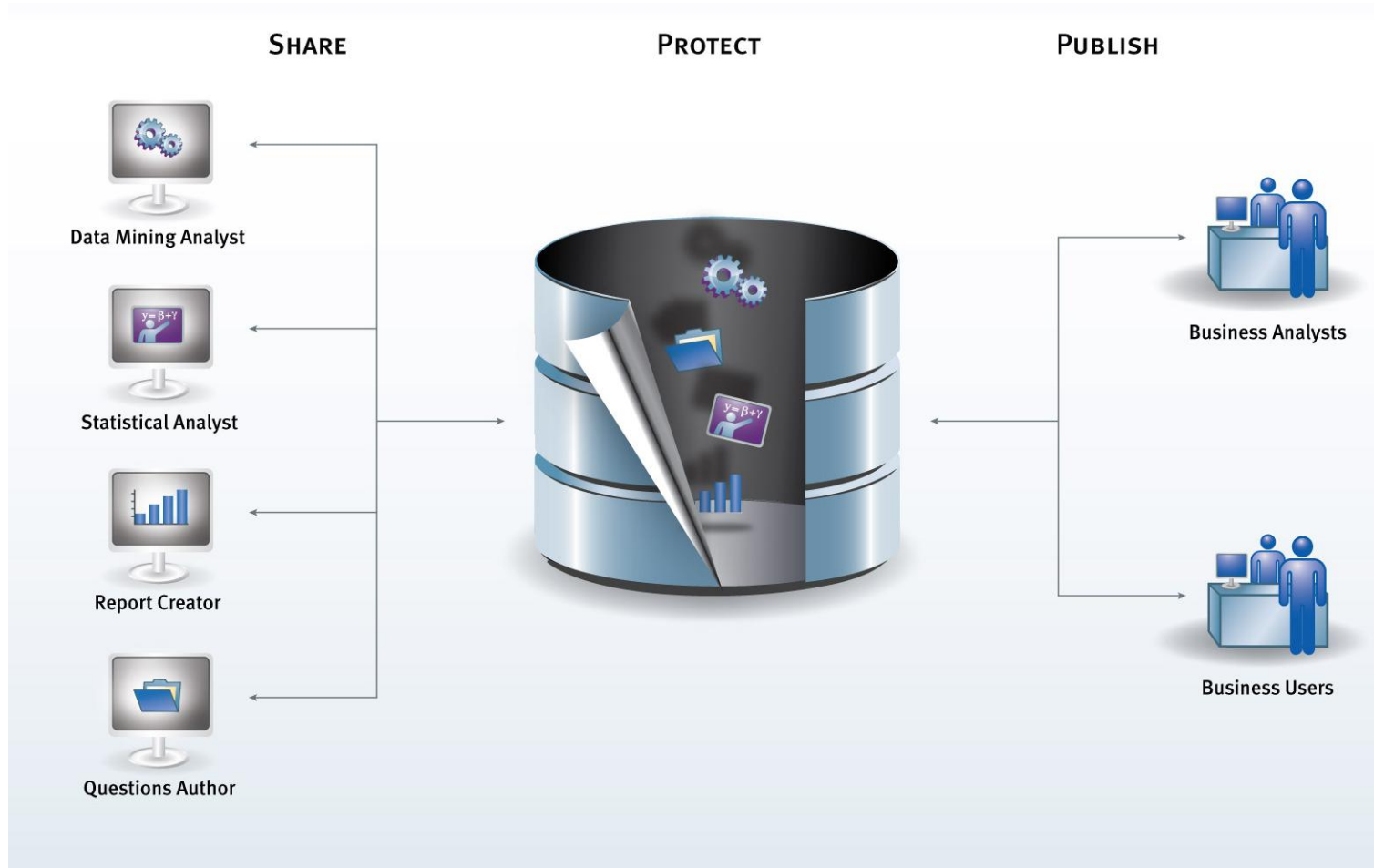
Real Time Cross Sell Recommendation



Mobile Secondary Screening Notification



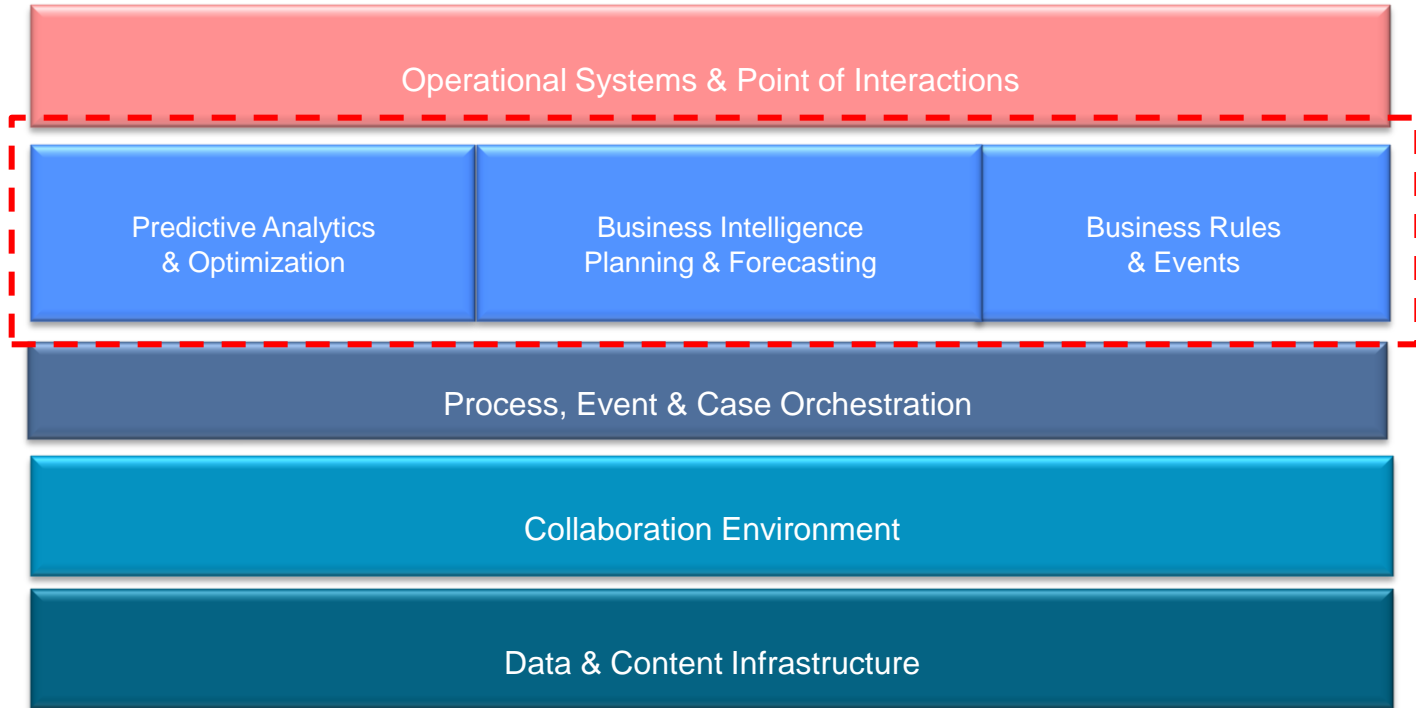
SPSS Fits in the I.T. Governance Model





Predictive Analytics in the IT Stack

Key Technologies
for optimizing
the point of
Interaction



Remember: Predictive Analytics Context



Broad Usage



Focused Usage

Analytical Reporting
Drill

- Top down view
- Drillable reports
- Sort top & bottom
- Review then query
- Market shifts
- Product ranking

Trending
Slice and Dice

- Personal exploration
- Compare & contrast
- Rotate and nest
- Work disconnected
- Sales trend analysis
- Market analysis

Scenario Modeling
What-if

- Model scenarios
- Reorganize, reshape
- Compare scenarios
- Save versions
- Financial analysis
- Profitability analysis

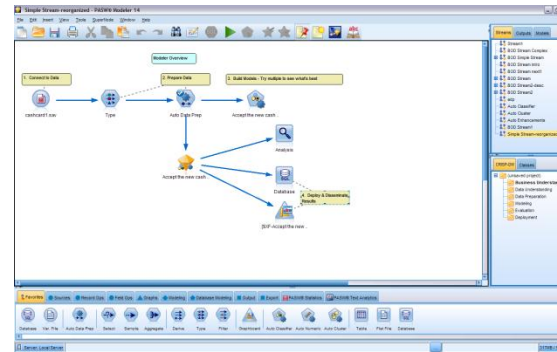
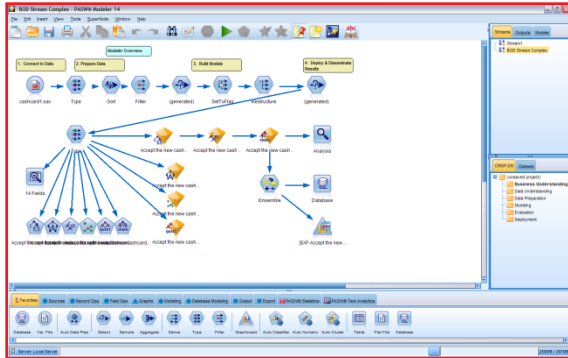
Predictive Analytics
What might be

- Uncover patterns
- Apply algorithms
- Mine data and text
- Predict outcomes
- Fraud prevention
- Churn analysis





SPSS Technology is Iterative and Collaborative



Automated Modeler Streams
Built by Non-Professionals

Sophisticated Modeler Streams
Built by Professionals

IBM® SPSS® Modeler Advantage

Data **Modeling** **Score**

Model type: Auto model Interactive model

Data source: Target:

Records used to

Point and Shoot Modeling
for the Business

IBM® SPSS® Decision Management for Claims

What If...

Simulation Data Source: sdbank claims data | Simulation Date: 2009-11-02 12:41:27 | Claim Area: | Interaction Point: All interaction points

Combine matrix	Model actions			Results		
	Refer	Standard	Fast Track	Action	Count	Percent
Rules actions	Refer	Refer	Standard	Refer	77	9.872%
	Standard	Refer	Standard	Standard	187	23.974%
	Fast Track	Refer	Fast Track	Fast Track	516	66.154%
					780	100%

Names: Run 3

Total Simulation Records: 975

Display: Count | Number of runs retained: 3

Action	Run 1 3%	Run 2 2%	Run 3 95%
Refer	22	25	77
Standard	52	229	187
Fast Track	793	516	516
Total	780	780	780

Outcome Optimization for the Business



Real World Example

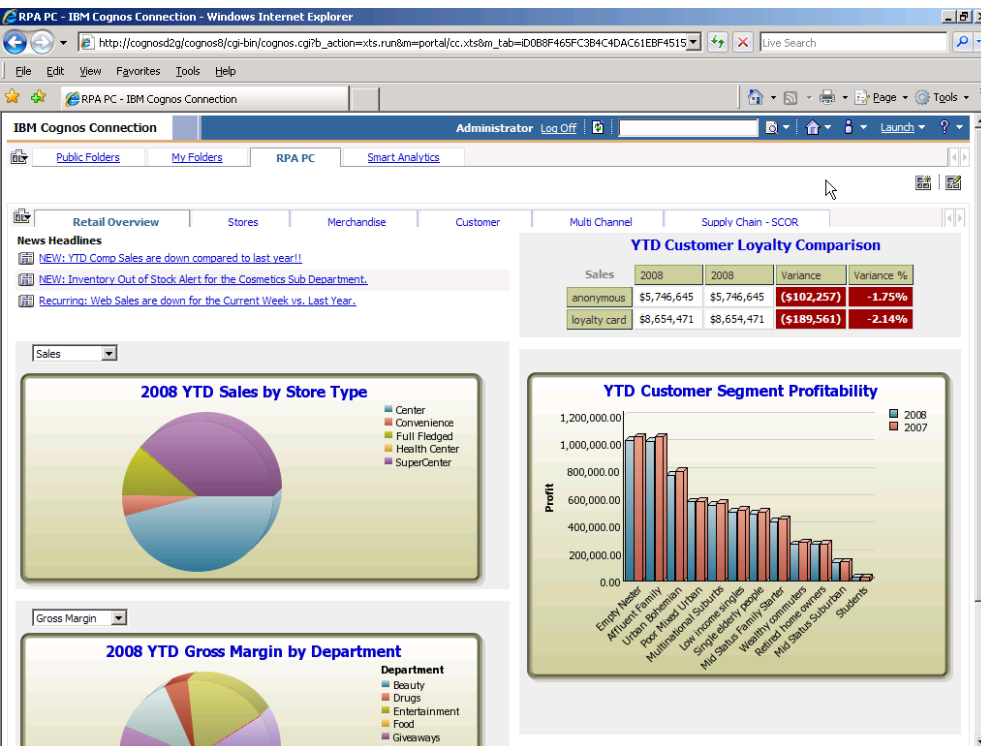


Market Basket Analysis

- Example: a retail dashboard within in a Cognos BI Tool

Key questions: Who is purchasing snacks and what else are they purchasing with those snacks?

- Breakdown of sales into customer segments which have been built into this dashboard using a data mining, cluster model generated by IBM SPSS Modeler





Transformation is Limited Only By Your Creativity

Reduce Maintenance costs for truck fleets?

Identify Store Locations?

Recruit the best employees?

Reduce fraudulent claims?

Choose the right Product Features?

Optimize Campaign Effectiveness?

Reduce Border incursions?

Retain Frequent Flyers?

Catch smugglers?

Reduce Churn?

Prevent crime?

Optimize product mix?





Questions to Ask

- What usage patterns fit best in your organization?
- Who has the data that you need?
- How will your solution evolve?





Thank You!

Your Feedback is Important to Us



IBM Software

Rob Risany

Program Director, Enterprise Solutions

+1-312-651-3408

rrisany@us.ibm.com

