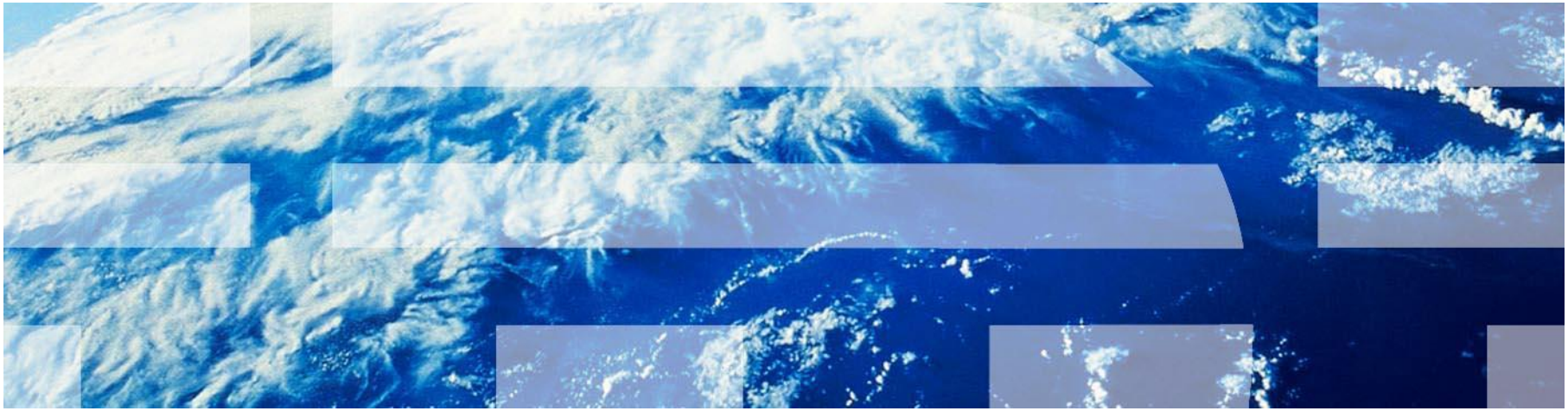
IBM SPSS Data Mining Workshop



© 2013 IBM Corporation

Agenda

Welcome and Introductions
Introduction to Predictive Analytics
Exercise: Navigating IBM SPSS Modeler
Exercise: Predictive in 20 Minutes
Data Mining Methodology and Application
Exercise: Data Mining Techniques
Exercise: Deployment
Wrap-up

© 2013 IBM Corporation

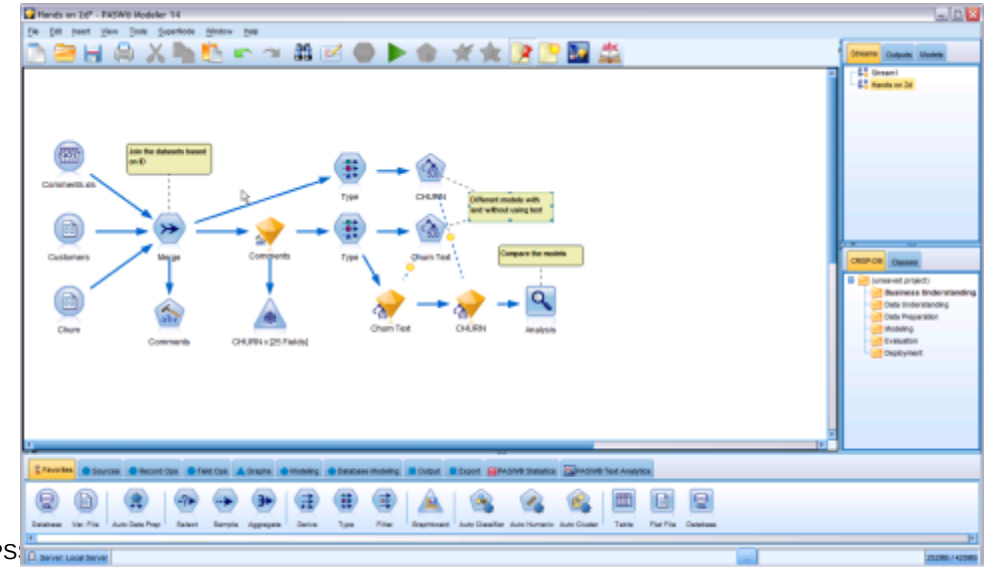
Objectives

- Introduction to predictive analytics and data mining
- Stimulate thinking about how data mining would benefit your organization
- Demonstrate ease of use of powerful technology
- Get experience in “doing” data mining
- See examples of existing customers and their realized ROI/benefits

© 2013 IBM Corporation

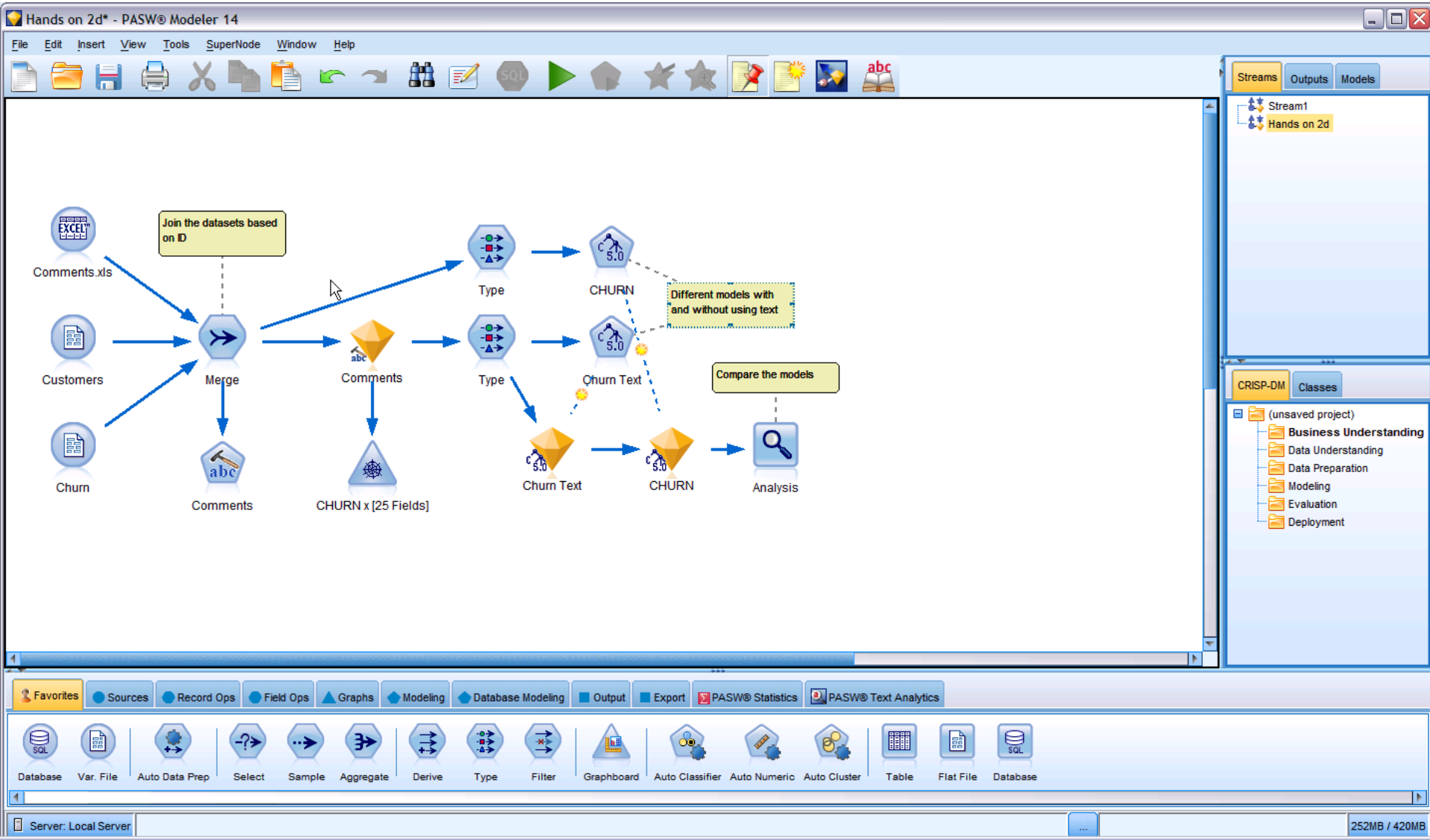
IBM SPSS Modeler

- High performance data mining and text analytics workbench
- Used for the proactive
  - Identification of revenue opportunities
  - Reduction of costs
  - Increase in productivity
  - Forecasting
- Allows analytics to be repeated and integrated within business systems

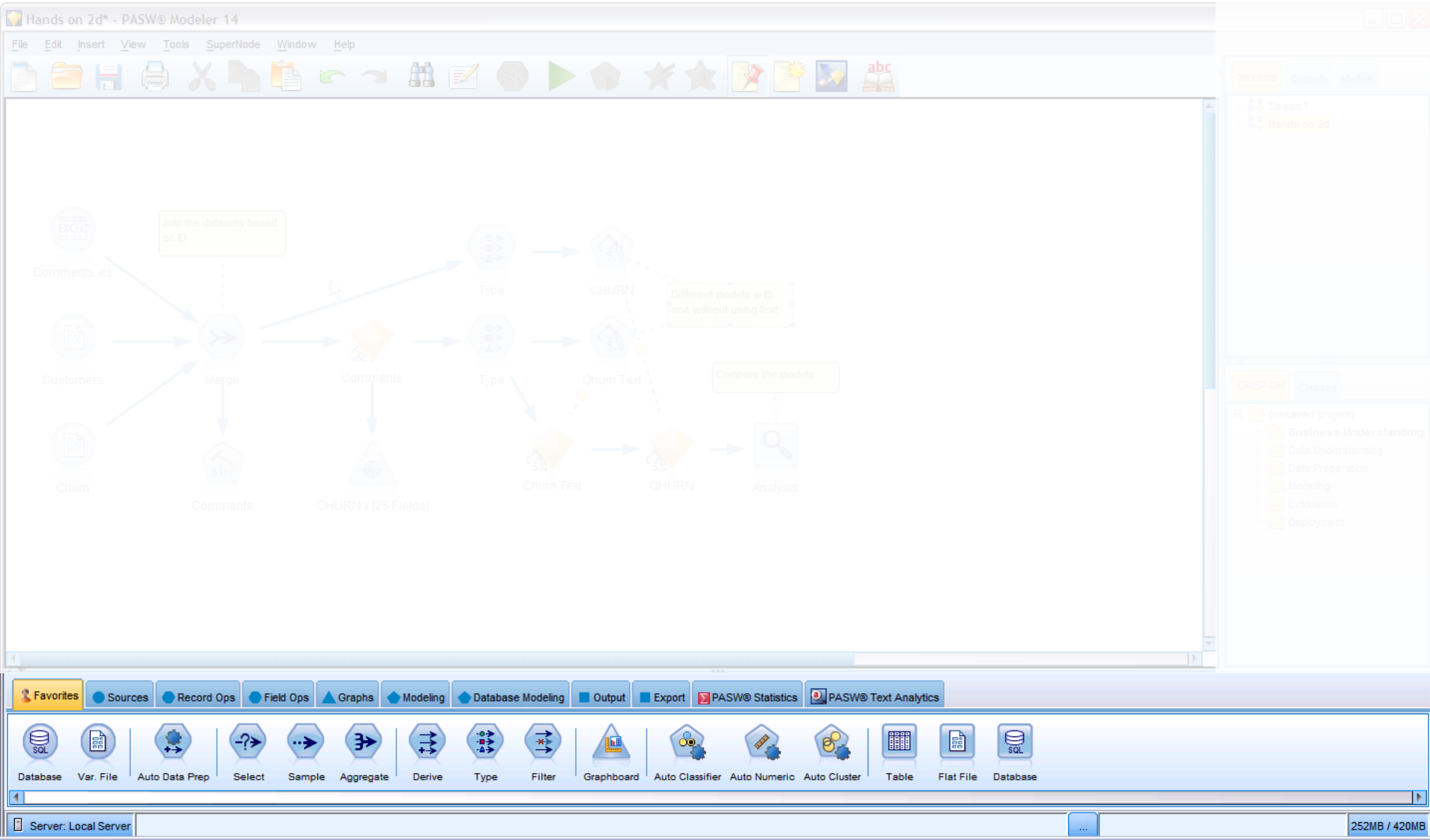


© 2013 IBM Corporation

IBM SPSS Modeler



IBM SPSS Modeler



Exercise: Predictive in 20 Minutes

Goal:

- **Identify who has cancelled their contract**

Approach:

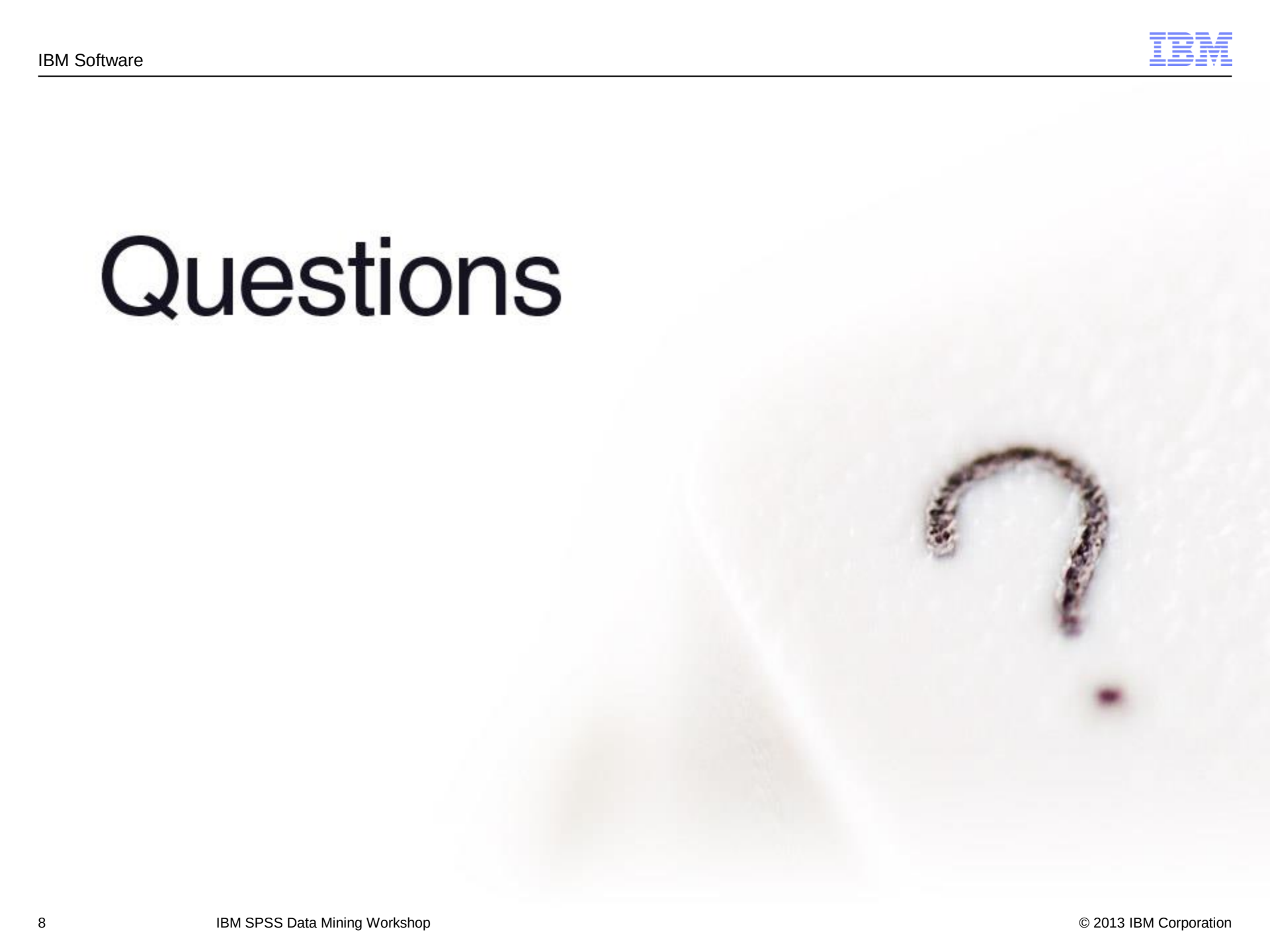
- Use a data extract from a CRM
- Define which fields to use
- Choose the modeling technique
- Automatically generate a model to identify who has cancelled
- Review results

Why?

- To prevent customers cancelling, by proactively identifying those likely to cancel before they do.

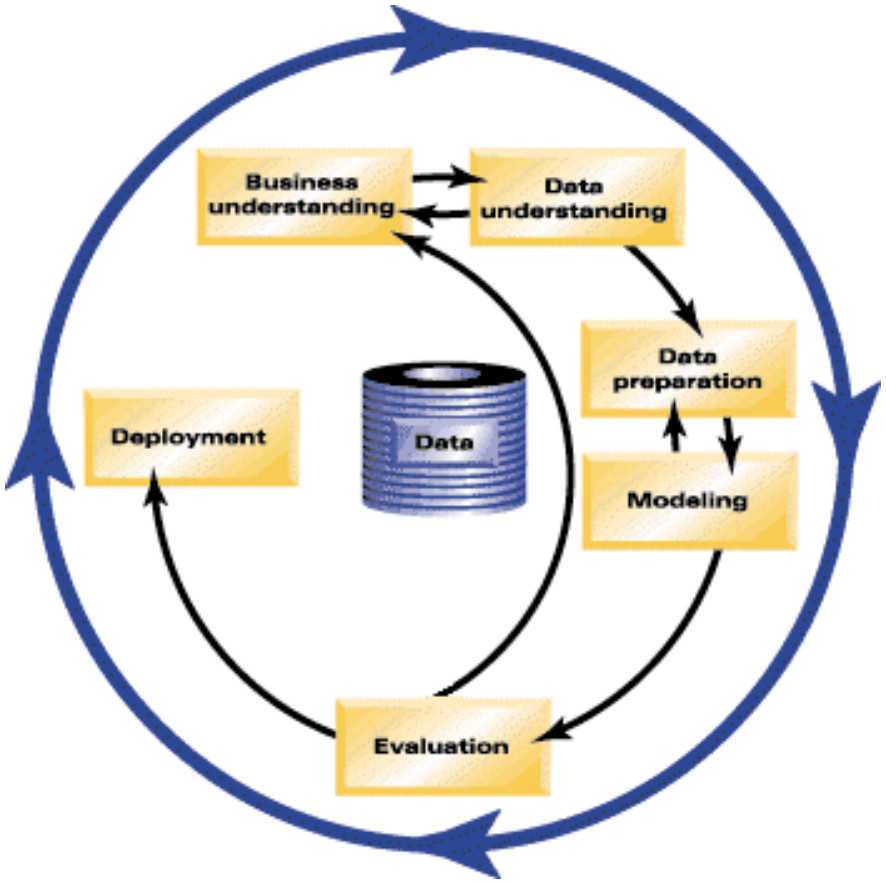
© 2013 IBM Corporation

Questions



© 2013 IBM Corporation

Data mining methodology



- **CR**oss-**I**ndustry **S**tandard **P**rocess **M**odel for **D**ata **M**ining
- Describes Components of Complete Data Mining Project Cycle
- Shows Iterative Nature of Data Mining
- Vendor and Industry Neutral

© 2013 IBM Corporation

Data mining techniques

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

© 2013 IBM Corporation

Data mining techniques

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

© 2013 IBM Corporation

Data mining techniques

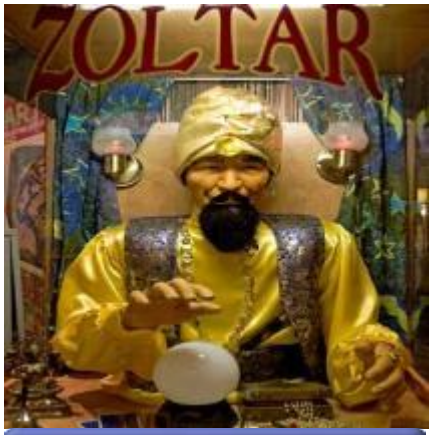
Technique	Usage	Algorithms
Classification (or prediction)	• Used to predict group membership (e.g., will this employee leave?) or a number (e.g., how many widgets will I sell?)	• Auto Classifiers, Decision Trees, Logistic, SVM, Time Series, etc.
Segmentation	• Used to classify data points into groups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

© 2013 IBM Corporation

Exercises

Association

Segmentation

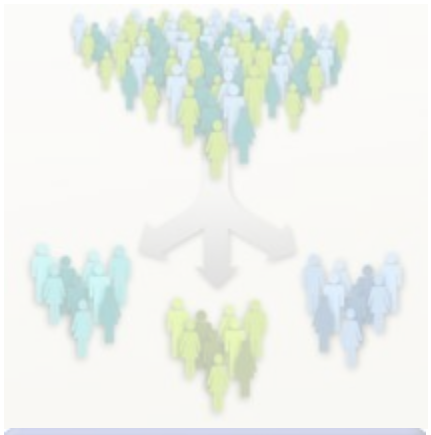


Classification

© 2013 IBM Corporation

Exercises

Association



Segmentation

Classification

© 2013 IBM Corporation

Association model

Goal:

- **Identify what products are being sold together**

Approach:

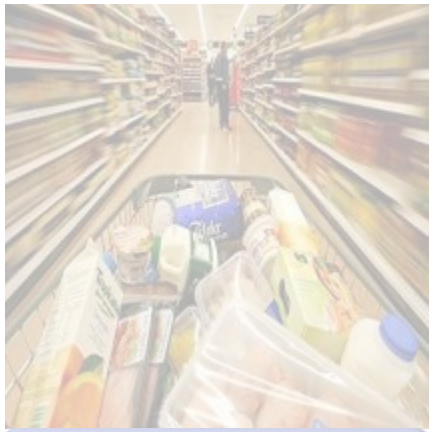
- Use a data extract from a transactional system
- Define which fields to use
- Visualize relationship between products
- Generate association model
- Review results

Why?

- Identify next likely purchase
- Create bundles to increase $ value

© 2013 IBM Corporation

Exercises



Association

Segmentation

Classification

© 2013 IBM Corporation

Segmentation model

Goal:

- **Create segments of customers**

Approach:

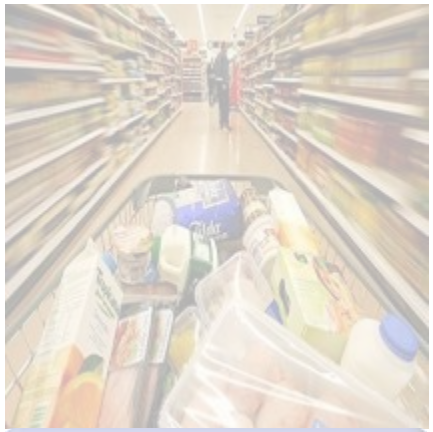
- Customer data from a database or CRM
- Define which fields to use
- Automatically generate a model to group customers

Why?

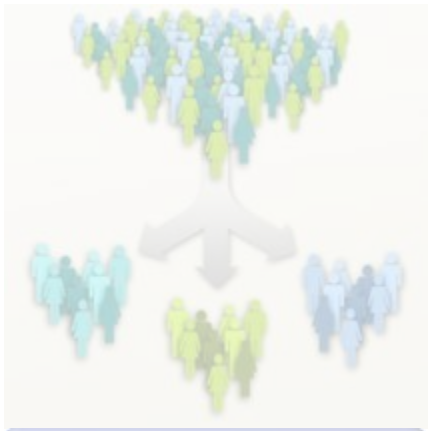
- Better customer understanding (demographics, socio-economic etc)
- Tailored messages for each group/segment
- Personal and more relevant for consumers

© 2013 IBM Corporation

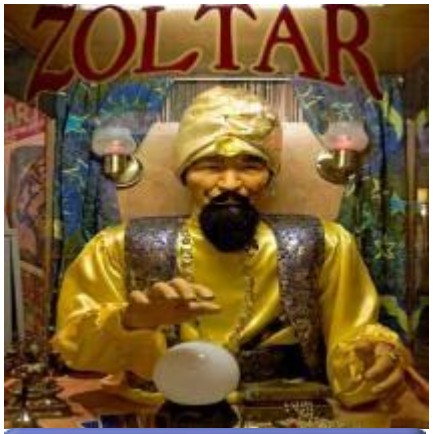
Hands on sessions



Association



Segmentation



Classification

© 2013 IBM Corporation

The importance of text

Because people communicate with **words**, not numbers, it has become critical to be able to **mine text** for its **meaning** and to sort, analyse, and understand it in the same way that data has been tamed. In fact, the two basic types of information complement each other, with data supplying the "what" and text supplying the "**why**".

Source IDC: "Text Analytics: Software's Missing Piece?"

© 2013 IBM Corporation

Classification model

Goal:

- **Identify who is likely to cancel their contract**

Approach:

- Use a data extract from a CRM
- Use open ended comments from call center
- Extract concepts from the text
- Define which fields to use
- Choose the modeling technique
- Automatically generate a model to identify who has cancelled
- Review results

Why?

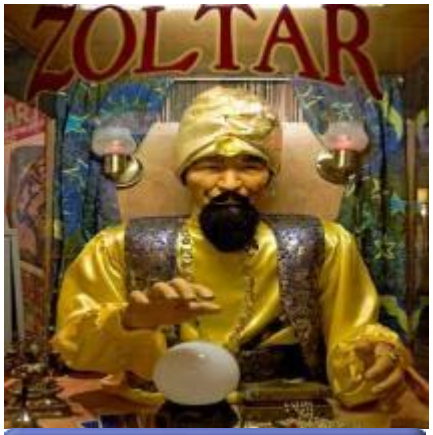
- Identify customers at risk before they churn
- Unstructured data can provide insight into customers actions and improve model accuracy

© 2013 IBM Corporation

Exercises

Association

Segmentation



Classification

© 2013 IBM Corporation

Data mining and text analytics

Data mining

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

Text analytics

- Extract, analyze and create structure for 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

© 2013 IBM Corporation

Deployment

© 2013 IBM Corporation

Deployment

Goal:

- **Deploy a predictive model**

Approach:

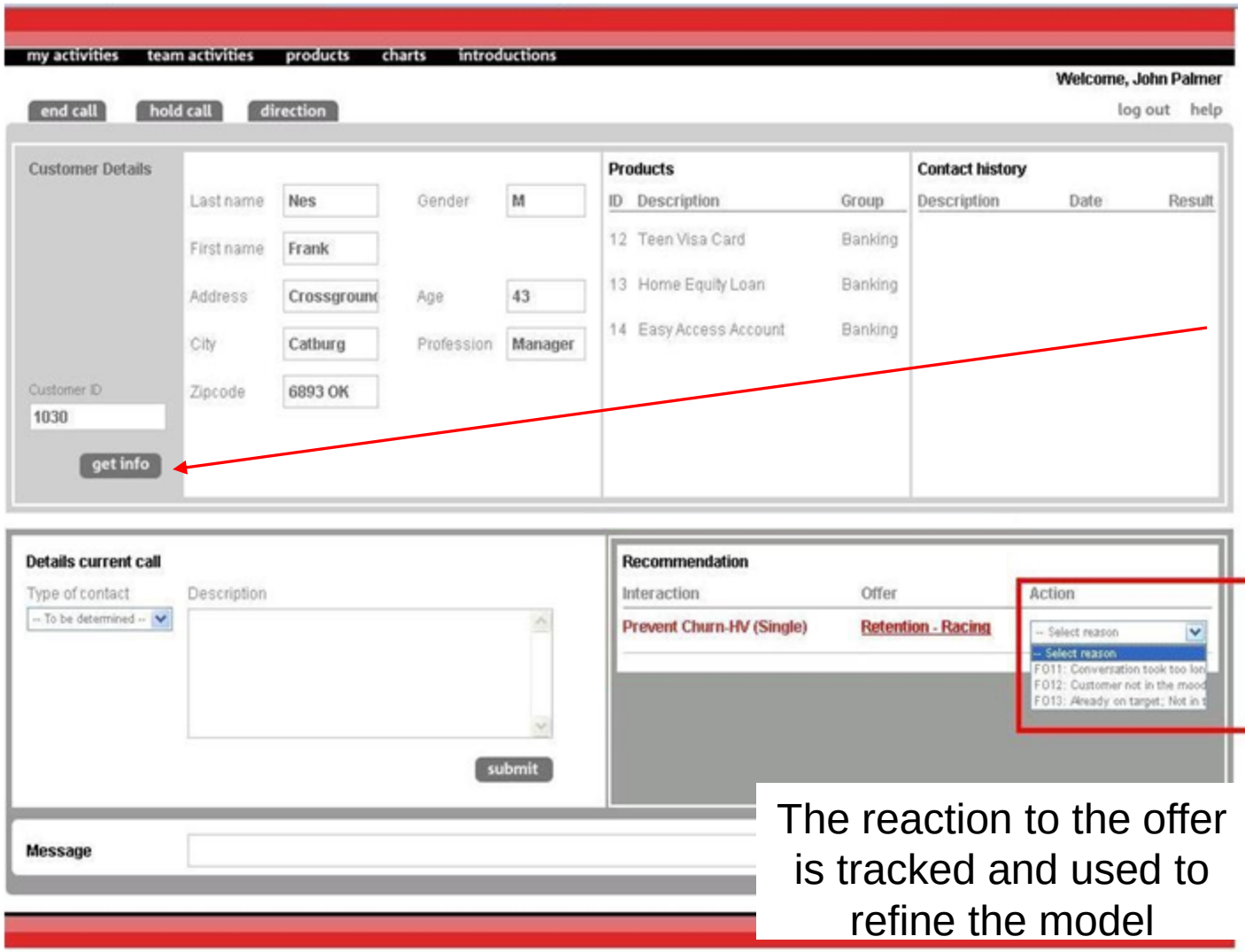
- Use the stream generated in the earlier session
- Pass new data through the stream and ‘score’ the data
- Identify those likely to cancel
- Export an .xls file with 50 most likely to cancel

Why?

- Extend the reach of analytics in an organization
- Allows analytics at the point of impact rather than being reactive

© 2013 IBM Corporation

Deployment – integrating with existing systems



A call center agent submits customer information during an interaction

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

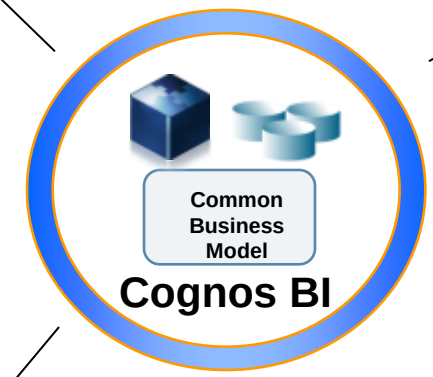
The reaction to the offer is tracked and used to refine the model

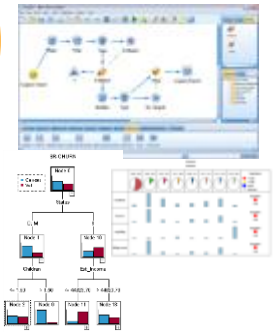
© 2013 IBM Corporation

Deployment – integrating with Cognos BI

1) Leveraging BI, identify problem or situation needing attention

2) SPSS predictive analytics feed results back into the BI layer

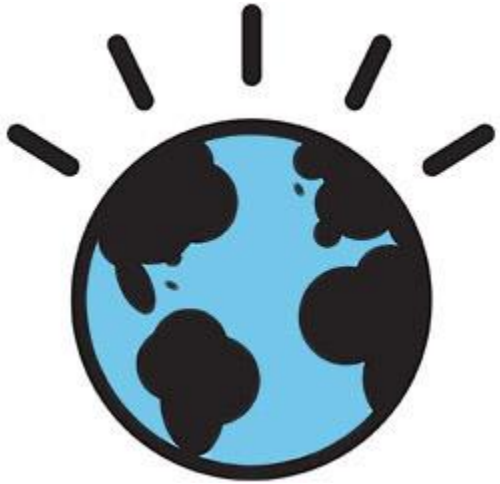




3) Results widely distributed via BI for consumption by business Users

© 2013 IBM Corporation

Smarter planet



Instrumented

Interconnected

Intelligent

© 2013 IBM Corporation

Workshop takeaways

Easy to use, visual interface
- Short timeframe to be productive with actionable results
- Does not require knowledge of programming language

Business results focused
- Cost effective solution that delivers powerful results across organization
- Flexible licensing and deployment options
- Full range of algorithms for your business problems

End-to-end solution
- Data preparation through real time interactions
- Use structured, unstructured and survey data
- Full suite of products, from data collection through deployment

© 2013 IBM Corporation

Workshop takeaways

Flexible architecture

- Leverages the investments already made in technology
- Does not require data in a proprietary format or DB
- Structured and unstructured data
- Open architecture (both inputs and outputs)
- SQL Pushback

© 2013 IBM Corporation

Predictive analytics customer success

- "**94% achieved a positive return on investment** with an average payback period of **10.7 months**."
- "Returns were achieved through reduced costs, increased productivity, increased employee and customer satisfaction, and greater visibility."
- "Flexibility, performance, and price were all key factors in purchase decisions."

Nucleas Research, *An independent provider of Global Research and Advisory Services.*

cablecom

"Reduced churn from 19 to 2%"

FBTO

"35% reduction in mailing cost, 2X response rate, 29% more profit"

BT

"100% increase in campaign effectiveness"

AEGON

"30 Million Euro in new revenue"

© 2013 IBM Corporation

We appreciate your feedback.
Please fill out the survey form in order to improve this educational event.

© 2013 IBM Corporation