



Business Analytics **Forum**

The Premier Conference for Information and Analytics

November 3-5, 2010

Gold Coast Convention and Exhibition Centre Queensland, Australia

Unveiling **IBM SPSS Modeler New and Improved**

Jason Burke
IBM



IBM SPSS Modeler

Why this is Important to know...

IBM SPSS Modeler helps commercial, government and academic organizations discover the predictive intelligence held in their data.

The new features improve performance and flexibility while making IBM SPSS Modeler a more integrated part of the IBM technology stack.

Speaking to you today...

Jason Burke

Manager, Technical Sales
Business Analytics

IDC - Independent Financial Impact Studies



*“The median ROI for the projects that **incorporated predictive technologies** was **145%**, compared with a median ROI of **89%** for those projects that did not.”*

Source: IDC, “Predictive Analytics and ROI: Lessons from IDC’s Financial Impact Study”





2010 focus remains on reducing costs and improving processes, but shift to top line and change initiatives is on the horizon

Business expectations for IT call for greater productivity and continued cost-efficiencies

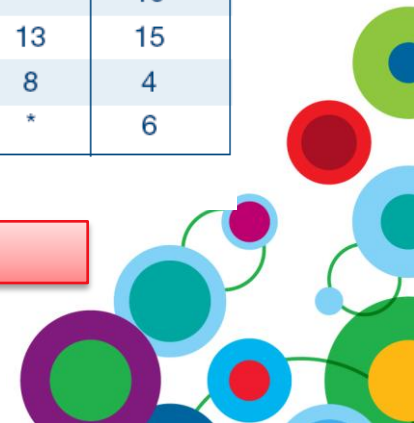
Business expectations

Ranking of business priorities CIOs selected as one of their top 5 priorities in 2010, and projected for 2013

Ranking	2010		2009	2008	2007	2013
Improving business processes	1	↔	1	1	1	2
Reducing enterprise costs	2	↔	2	5	2	8
Increasing the use of information/analytics	3	↑	5	8	7	5
Improving enterprise workforce effectiveness	4	↓	3	6	4	7
Attracting and retaining new customers	5	↓	4	2	3	3
Managing change initiatives	6	↑	8	3	10	1
Creating new products or services (innovation)	7	↓	6	12	*	12
Targeting customers and markets more effectively	8	↓	7	9	*	9
Consolidating business operations	9	↑	11	13	*	16
Expanding current customer relationships	10	↓	9	7	*	10
Supporting regulation, reporting and compliance	11	↑	12	14	13	15
Creating new sources of competitive advantage	12	↑	13	11	8	4
Expanding into new markets and geographies	13	↓	10	4	*	6

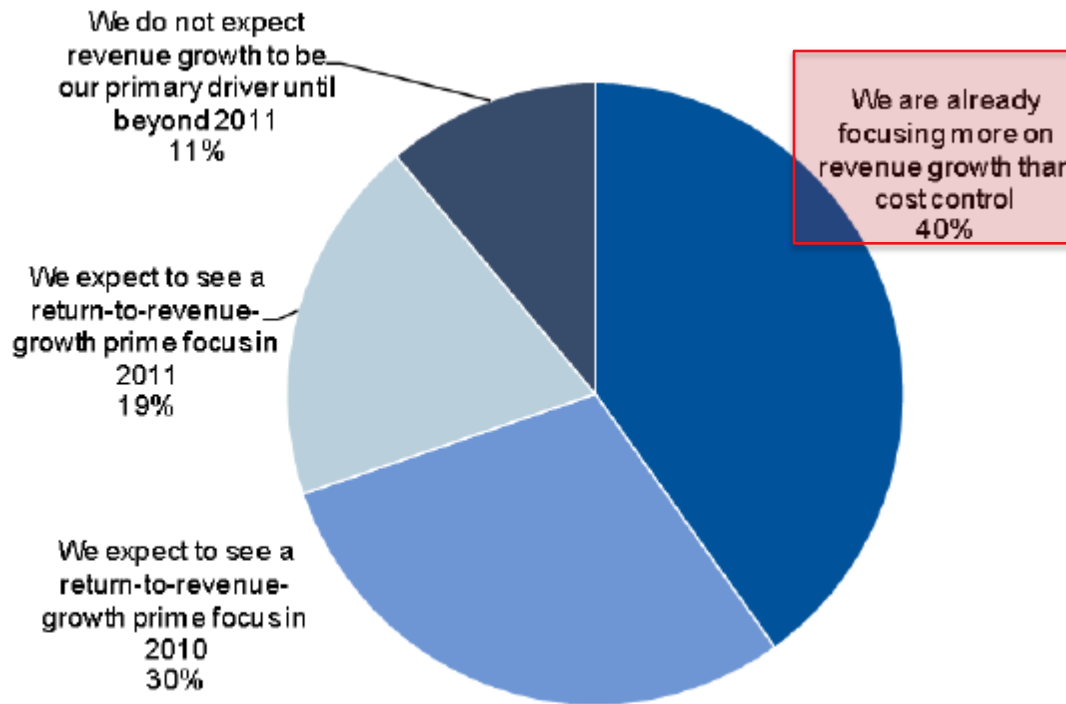
* New question for that year

And increasing the use of information and analytics is seen as key





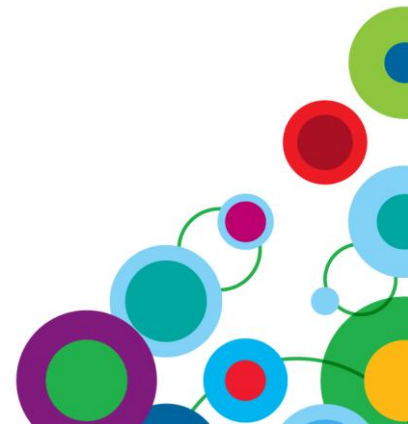
40% of CIO's already suggesting that additional focus on revenue growth initiatives is occurring



Source: Gartner (March 2010)

- And through 2010, this will increase to 70%
 - 30% of respondents still feeling the impact of the global recession

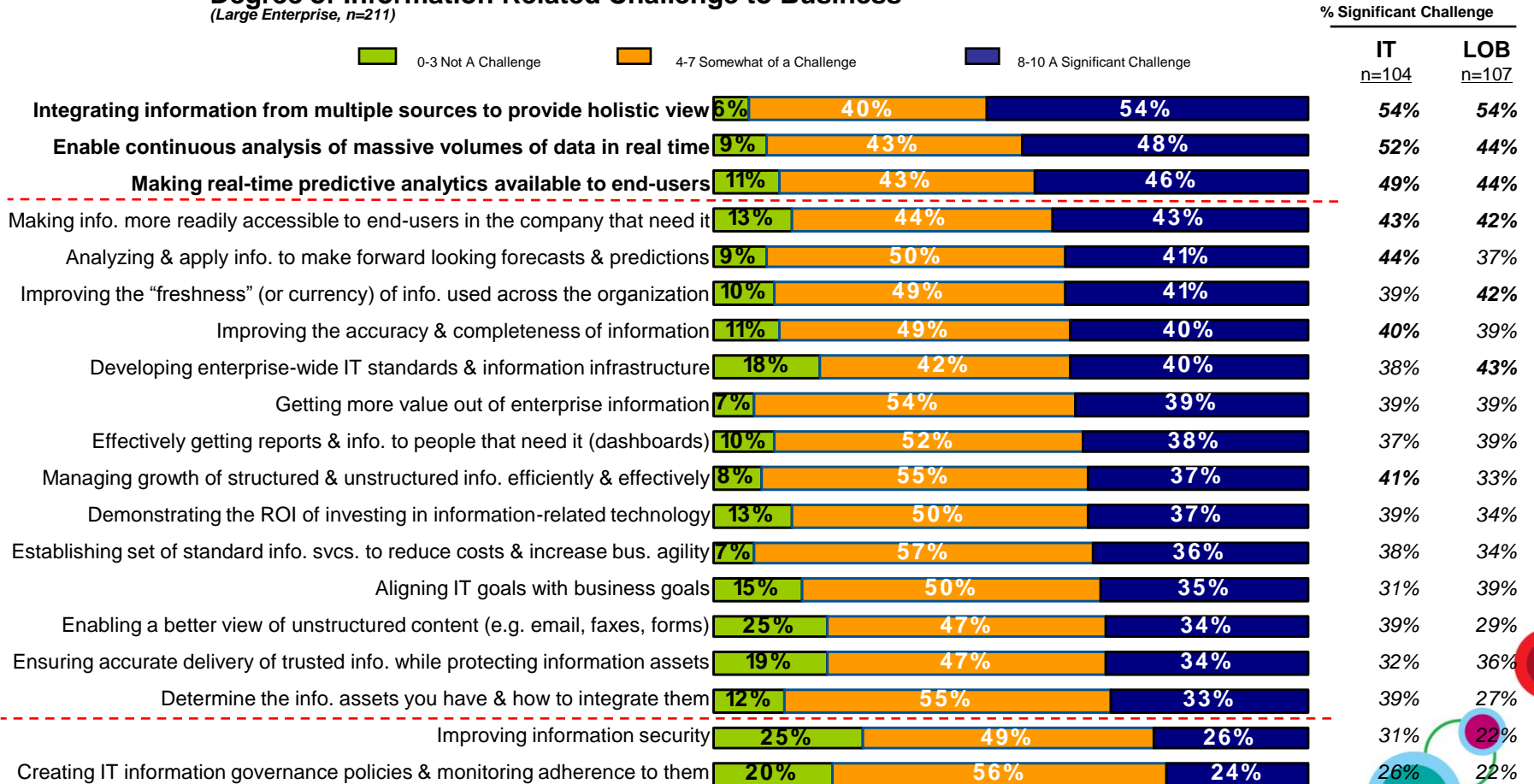
Source: Gartner CEO and Senior Business Executive Survey (March 2010)



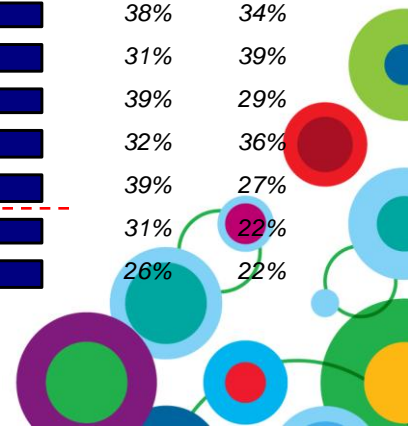


Top information related challenges are access to timely, relevant information, integrating information from multiple sources and making advanced analytics available to all levels of users of the organization

Degree of Information Related Challenge to Business (Large Enterprise, n=211)



Source: New Intelligence Value Proposition Research (Fall 2009)





Modeler Puts “Predictive” into Predictive Analytics

IBM SPSS Modeler is a high-performance data mining and text analytics workbench that quickly delivers positive ROI by creating the predictive intelligence that allows organizations to proactively and repeatedly exploit revenue opportunities, reduce costs and increase operational productivity.

Modeler Professional is a data mining workbench for the analysis of **structured numerical data to model outcomes and make predictions that inform business decisions with predictive intelligence**

Modeler Premium allows organizations to tap into the predictive intelligence held in all forms of data. Modeler Premium goes beyond the analysis of **structured numerical data alone and includes information from unstructured data such as web activity, blog content, customer feedback, e-mails, articles, and more to create the most accurate predictive models possible.**





Introducing IBM SPSS Modeler

- **Two *Editions***
- **IBM SPSS Modeler Professional**
 - Data access and transformation
 - Automatic data preparation
 - Modeling algorithms
 - Automatic model generation
 - Output and visualization
- **IBM SPSS Modeler Premium**
 - All of the above PLUS
 - Interactive text mining workbench
 - Learn more at: *Discover the Valuable Insights Hidden in Text Data (CPA3259)*

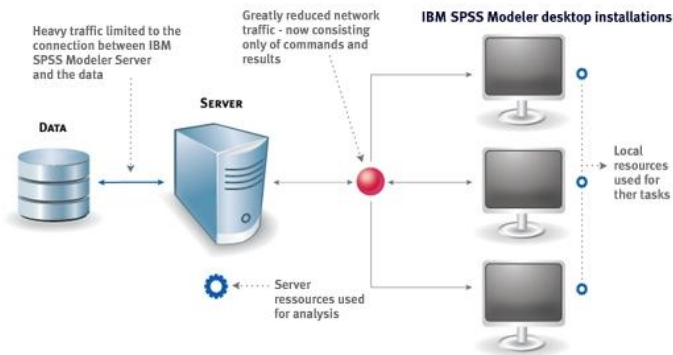
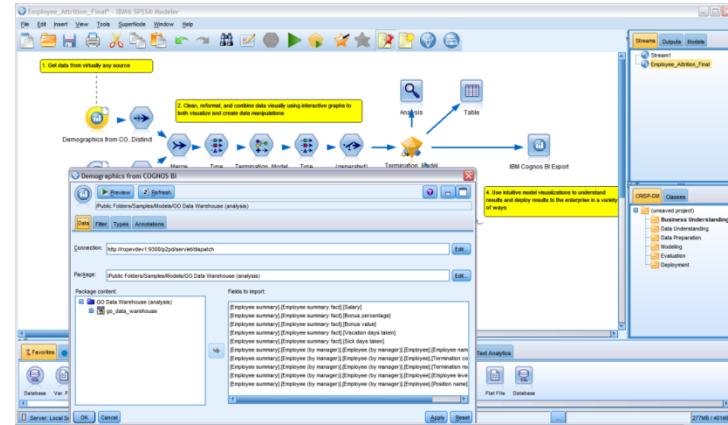




IBM SPSS Modeler Deployment Options

Client (Desktop)

- Access local files
- Connect to operational databases
- ★ Connect to Cognos BI
- Processing performed on local installation



Client/Server

- Data operations/processing on server
- In-database data mining
- SQL pushback
- Modeler Batch
- ★ SuSE Linux Enterprise Server 10 (zLinux)
- ★ Inclusion in Smart Analytics System for Power (AIX)

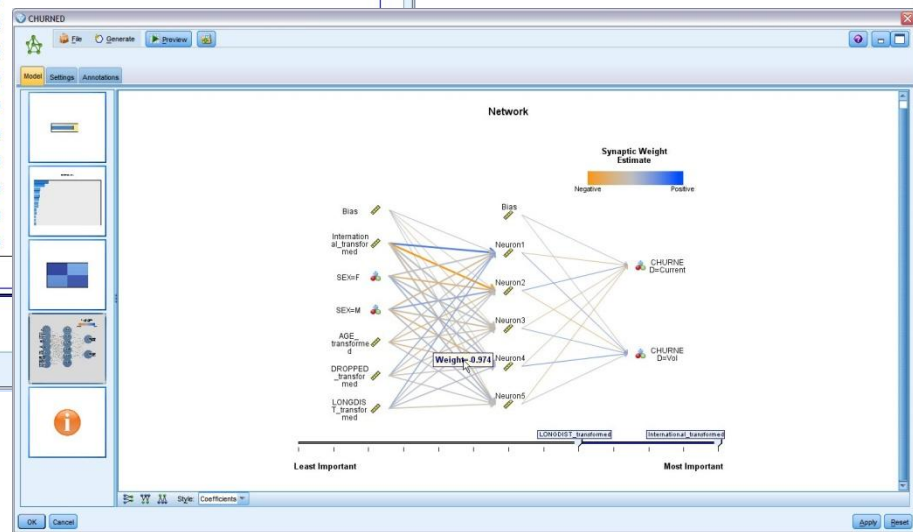
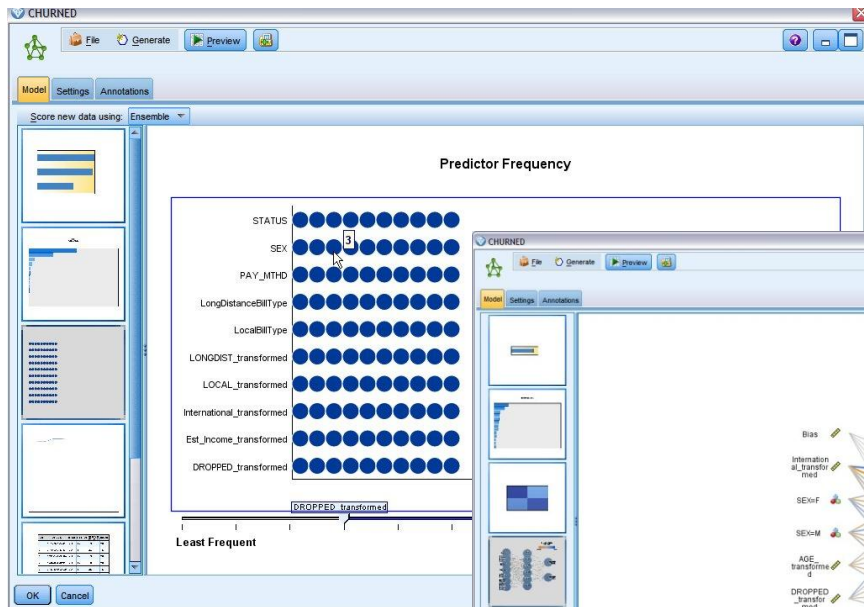
★ New!



What's New in IBM SPSS Modeler?

- More Accessible and Easier to Use
 - Improved visualizations
 - Consistency with IBM SPSS Statistics
 - Easier to prepare models for deployment
- More Powerful
 - New and improved techniques
 - New deployment options
 - Better enterprise performance
- More Integrated with IBM Technologies
 - IBM Cognos 8 Business Intelligence
 - IBM InfoSphere Warehouse





More Accessible and Easier to Use

- Improved visualization
- Better integration with IBM SPSS Statistics
- Easier to prepare models for deployment

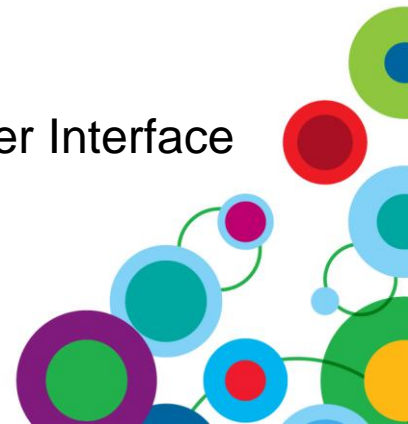
IBM SPSS Modeler Overview and New Feature Highlights

DEMONSTRATION



Feature Highlights of What's New

- Interactive Ensemble Model Visualization
 - Available when specifying an optimization technique
 - Implemented for Linear, Neural Networks and trees (C&RT, Chaid and Quest)
- Interactive Neural Net Model Visualization
- Run-time Interactive Parameter Prompts
- Better Control Over Excel Export
- Tighter Integration Between IBM SPSS Statistics and IBM SPSS Modeler
 - Consistency Between IBM SPSS Statistics and IBM SPSS Modeler Terminology
 - Shorter learning curve for IBM SPSS Statistics users
 - Easier to use both products for analytics together
 - Use of Additional IBM SPSS Statistics Features From Modeler Interface
 - Custom dialogues
 - GLMM modeling algorithm

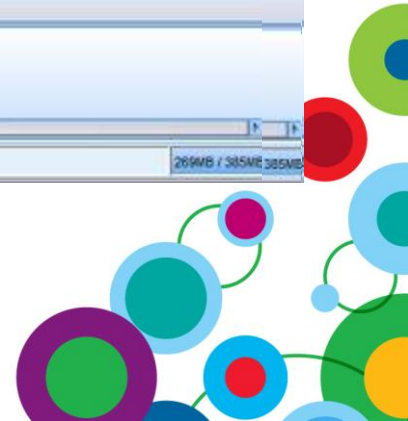




Preparing Models for Deployment

The screenshot shows the IBM SPSS Modeler interface with a workflow diagram. The workflow starts with a data source 'churn.bt', followed by 'Partition', 'type', 'Auto Data Rep', and 'CHURNED'. A red arrow points from a callout box to the 'Open in IBM SPSS Modeler Advantage' button in the toolbar.

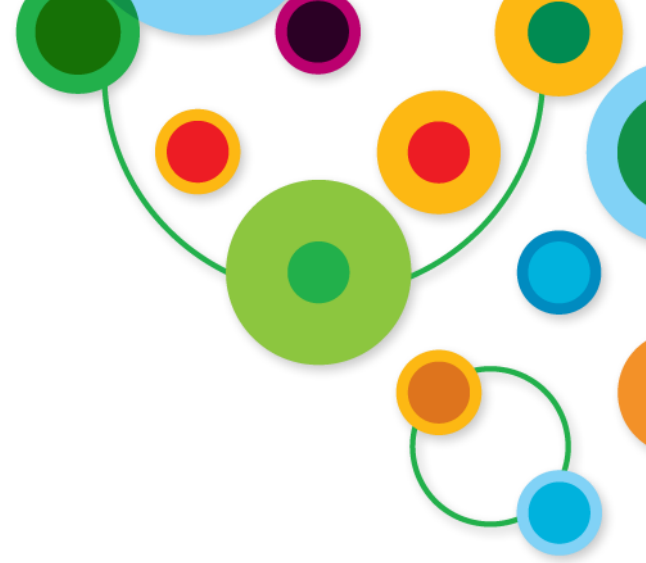
- Open Modeler Advantage directly from Modeler interface
- Better integrations with IBM SPSS Decision Management
- Easier deployment of predictive intelligence





More Powerful


- New and improved techniques
- New deployment options
- Better enterprise performance





Feature Highlights

New and Improved Modeling Techniques

- New Neural Net Algorithm
 - Interactive model visualizations makes it easy to understand the model
 - Includes multilayer perceptron and radial basis functions
 - Leverages ensemble optimization methods and new ensemble viewer
 -  Heritage Neural Net is removed from the palette and a message box cautions when an old one is present
- New Linear Algorithm
 - Includes automatic data preparation within the technique
 - Interactive model visualization makes it easy to understand the model created
 - Leverages ensemble optimization methods and new ensemble viewer
 - Heritage linear algorithm remains in the model selection palette as “Regression”



Feature Highlights

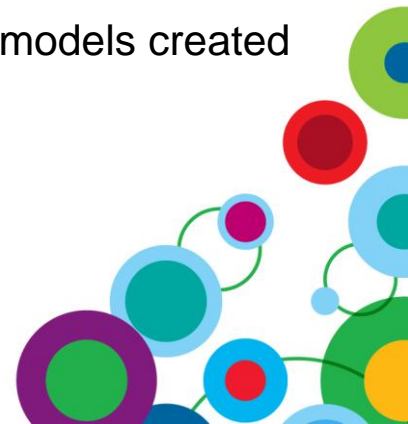
New Methods for Model Optimization: Boosting, Bagging and Large Dataset Processing



- Algorithms Implemented For:
 - Decision Tree Algorithms: C&R Tree, QUEST, CHAID
 - New Neural Net
 - New Linear
- Boosting
 - Enhance model accuracy
- Bagging
 - Enhance model stability
- Large Dataset Processing
 - Requires server
 - Can handle unlimited data volumes
- Standard Model
 - Single model to predict the target using the predictors.
- Ensemble
 - Occurs when boosting, bagging, or very large datasets are requested in Objectives
 - A rule is used to combine the predicted values from base models to compute ensemble score
 - New interactive ensemble viewer helps understand the overall model and the component models created



Also supported within the automatic modeling nodes when any of the relevant algorithms are in use



New Deployment Options for IBM SPSS Modeler Server

SPSS on System z



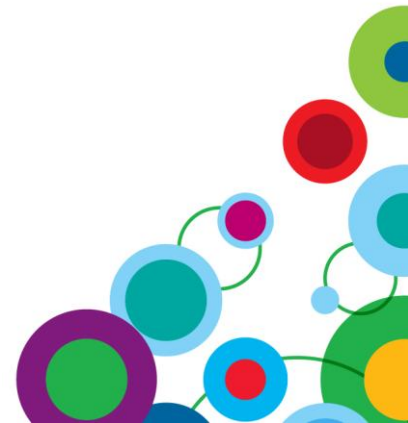
- What is System z?
 - IBM's mainframe computers
 - Known for their reliability - **z stands for zero down time**
 - Customers include the largest financial, insurance and government agencies
- Why SPSS on System z?
 - Analytics are increasingly a mission critical process
 - Single, best-in-class vendor solution for hardware, warehousing, business intelligence and (now) predictive analytics
- What SPSS products are available on Linux for System z?
 - IBM SPSS Modeler 14.1 Server
 - IBM SPSS Statistics 19 Server
 - IBM SPSS Collaboration and Deployment 4.2



New Deployment Options for IBM SPSS Modeler Server IBM Smart Analytics POWER System (AIX)




- What is IBM Smart Analytics System?
 - Analytical Data Warehouse Platform: Pre-Integrated Hardware + Software + Services
 - Total Data Warehouse Appliance delivered "Table Create-Ready"
- What role does IBM SPSS Modeler 14.1 play?
 - SPSS Development providing installation best-practices to Warehouse Development Labs
 - Performance benchmarks will be published
 - Explicit testing performed to ensure optimal performance and "best practices" integration



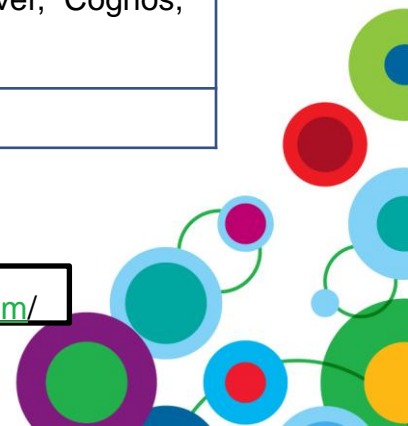


Smart Analytic System Benefits

	
	Smart Analytics System
One Call support	<ul style="list-style-type: none">▪ Included
Coordinated Stack certification (SW, OS, & Firmware)	<ul style="list-style-type: none">▪ Included
Services Implementation Premium Support Health Check	<ul style="list-style-type: none">▪ Included▪ Included▪ 1st year Included
Server, Storage, SW	IBM Systems, Storage, Modeler 14.1 Server, Cognos, & InfoSphere Warehouse Software
Bottom Line	All of the above part of the System

Faster time to value at a lower cost!

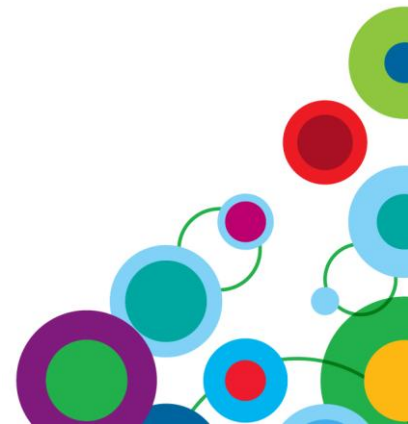
To learn more, please visit: <http://www-01.ibm.com/software/data/infosphere/smart-analytics-system/>





Extending Support for Enterprise Data Sources

- Mining History: Mainframe Data Support
 - IBM Classic Federation Server Support
 - Direct virtualized SQL access to mainframe data sources.
 - Can access data stored in VSAM, IMS, CA-IDMS, CA-Datcom, Software AG Adabas, and DB2® UDB for z/OS® databases (zDB2)
 - Access both structured and unstructured data from those sources
 - Read-only
 - IBM Classic Federation Server does NOT ship with Modeler
 - It can be purchased from IBM
- Mining Alternative Data Sources: XML
 - Read and write
 - Read a single file or a set of files

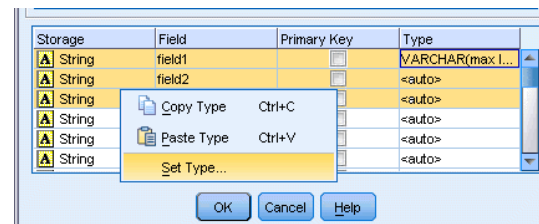
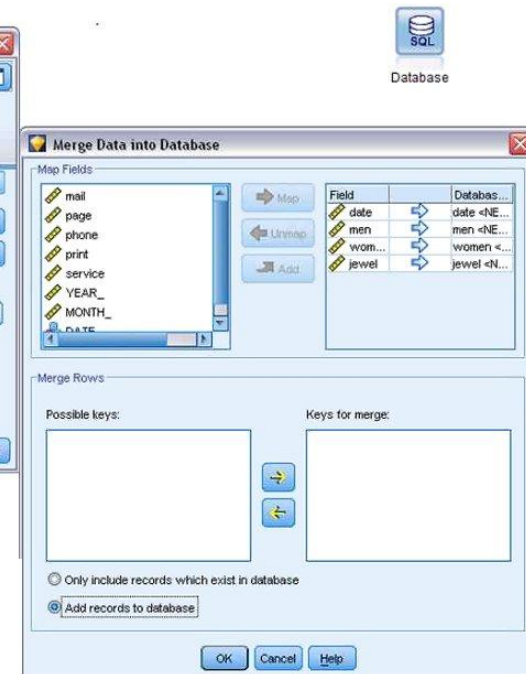




Feature Highlights

Better Enterprise Performance with Operational Databases

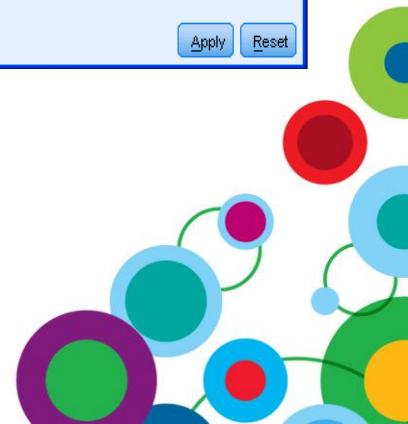
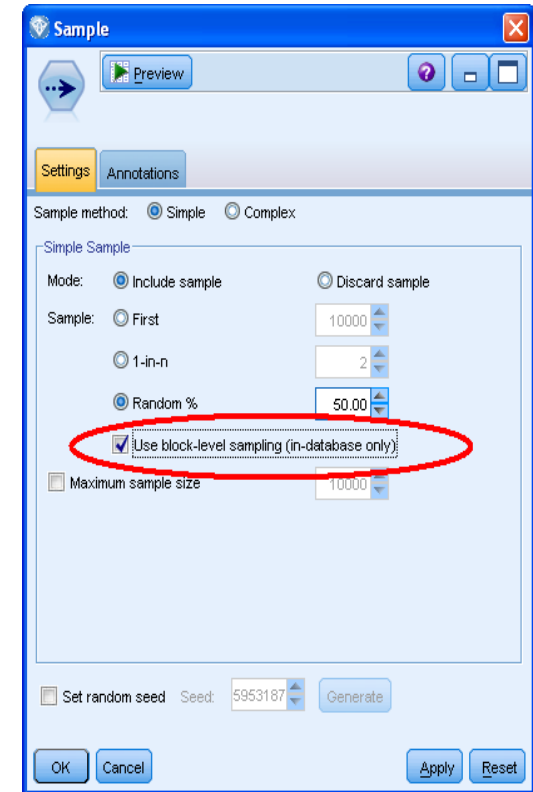
- Ease of use
 - Ability to update or add data to an existing operational table
 - Schema sub-dialog supports multi-select
 - GUID Support
 - Default Database Configuration
- More in-database mining options
 - Microsoft Time Series & Sequence Clustering
 - Oracle Attribute Importance





Feature Highlights – SQL Pushback Improved Enterprise Performance and Data Support

- Sampling Enhancements
 - SQL Pushback for 1 in N sampling and Rand Sampling (DB2)
 - Block level in db sampling (DB2 / Oracle)
- Updated support for nested SQL
 - Unlimited nesting level for supported databases
 - Used to be set to 8 by default for all DBs
 - Now set according to each database limitation
 - Reduced size of SQL Generated when a large expression is used multiple times in subsequent expressions (i.e. derive + filler node) via nesting
- Date / Time Function Pushback
 - (i.e. date_before, date_days_difference, datetime_in_seconds, dt_minute_T)
- Stripchar pushback





More Integrated With IBM Technologies

- IBM Cognos 8 Business Intelligence
- IBM InfoSphere Warehouse

Integrating Predictive Intelligence with Business Intelligence: Yesterday



- Uses Framework Manager
- Maps operational data to packages
- Understand data, databases, and enterprise reporting needs



IT

- Uses Report Studio to create reports
- Accesses reporting data prepared by IT
- Understands enterprise reporting needs



BI Professional

- Uses Modeler to create models
- Accesses operational data
- Understands data from an analysis perspective



Modeler

- Interacts with enterprise data through a dashboard
- Interested in what is happening now and what is happening tomorrow



Business Manager



Executive

Managers and Executives have visibility into metrics

BI Professional creates report(s)

IT sets up data for enterprise reporting

Modeler updates database(s) and informs IT that data is ready

Modeler gets data or data access to create models and score

IT sets up data for enterprise reporting



Integrating Predictive Intelligence with Business Intelligence: ~~Yesterday~~ Today



- Uses Framework Manager
- Maps operational data to packages
- Understand data, databases, and enterprise reporting needs



IT

- Uses Report Studio to create reports
- Accesses reporting data prepared by IT
- Understands enterprise reporting needs



BI Professional

- Uses Modeler to create models
- Accesses operational data
- Understands data from an analysis perspective



Modeler

- Interacts with enterprise data through a dashboard
- Interested in what is happening now and what is happening tomorrow



Business Manager



Executive

Managers and Executives have visibility into metrics

BI Professional creates reports

IT sets up data for enterprise reporting

Modeler updates data and exports package directly from Modeler

Modeler accesses enterprise reporting data from Modeler

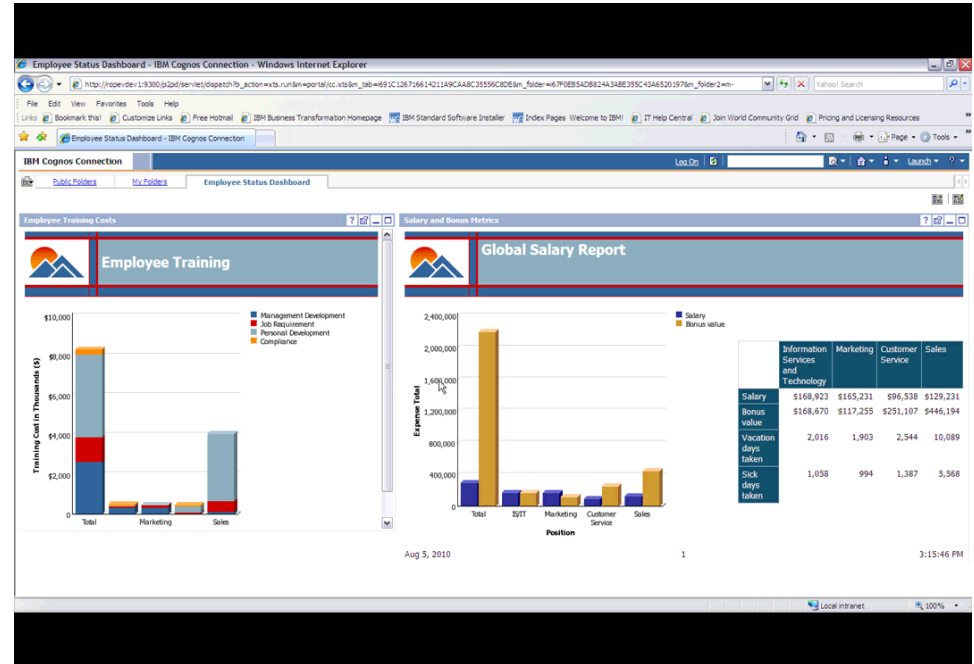
IT sets up data for enterprise reporting





Access Cognos BI Data inside SPSS Modeler

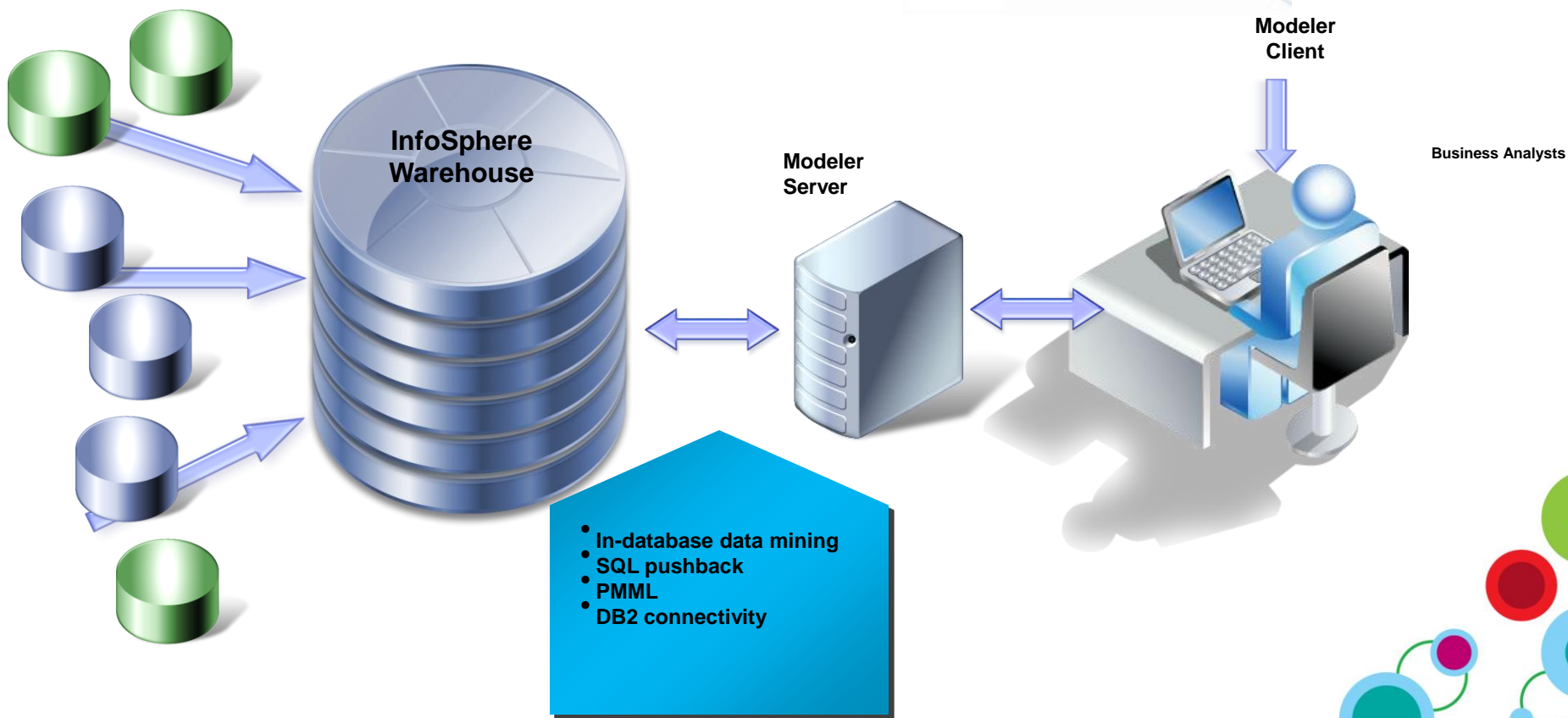
- New Cognos Nodes
 - Source & Export
 - Read / Write Cognos 8 Packages
- Benefits:
 - Leverage data investment in Cognos BI
 - Enable easy access to Cognos BI data inside Modeler
 - Avoid manual extraction & import of data
 - Create and use consistent enterprise data view for reports and modeling



IBM SPSS Modeler is Complementary to InfoSphere Warehouse



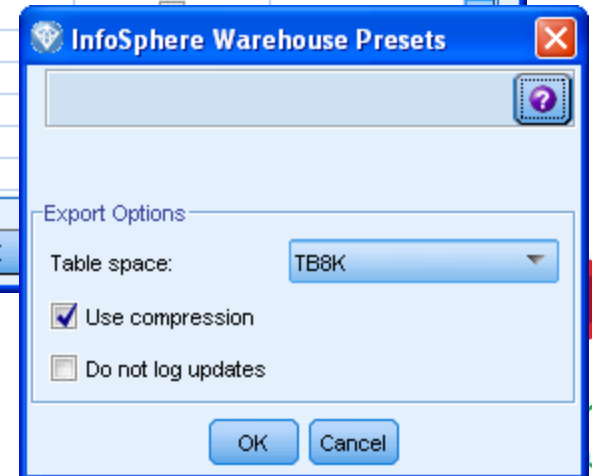
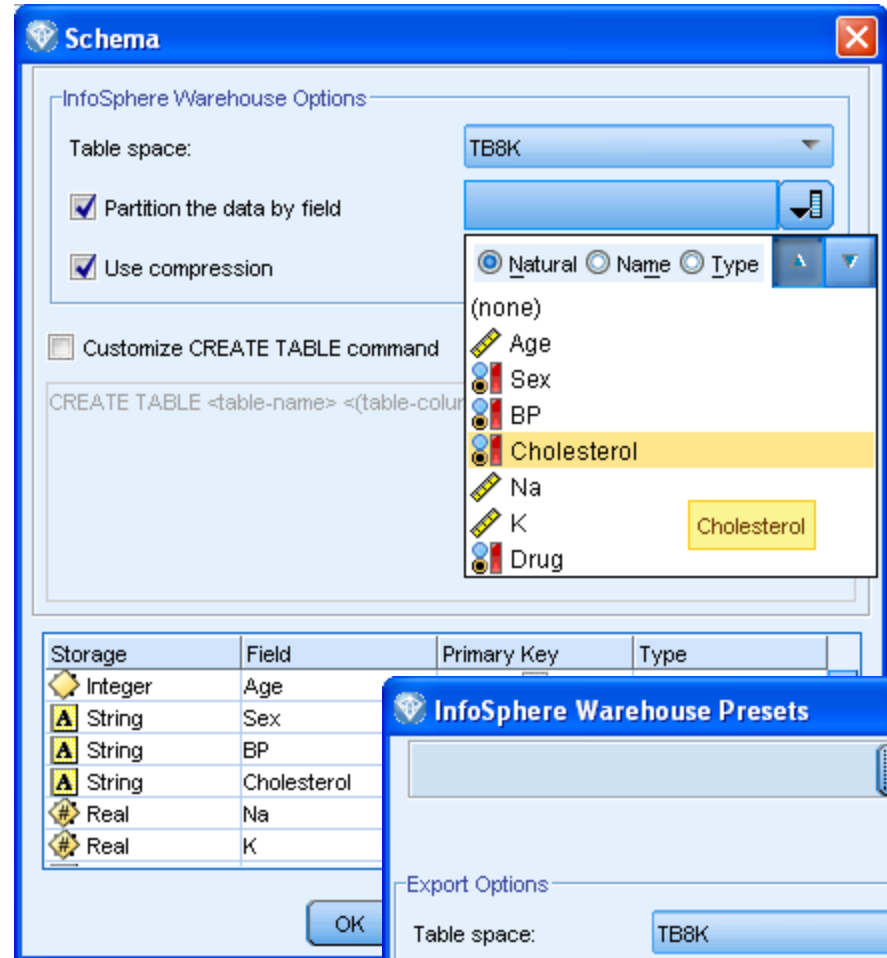
InfoSphere™
software





Enhanced InfoSphere / DB2 Support

- DB2 Specific
 - Compression
 - Not Logged Initially
 - Partitioning
 - Default Tablespace
 - Key field
- Ability to setup some parameters as default presets on Database Configuration



Enhanced InfoSphere / DB2 Support New In Database Algorithms



- New config settings
- New Nodes / Algorithms
 - Logistic Regression
 - Naïve Bayes
 - Time Series
 - Creates unrefined Model
 - RBF
 - Added to Regression

Helper Applications

IBM SPSS Statistics Oracle Microsoft IBM

Enable InfoSphere Warehouse Data Mining Integration

DB2 Connection: nmccarroll@DB2_hurspsperf4_TEST

Warn when about to overwrite an InfoSphere Warehouse Data Mining model

List InfoSphere Warehouse Data Mining models

Enable launch of InfoSphere Warehouse Data Mining Visualization

Path for Visualization executable:

Time Series Visualization Plug-in Directory: C:\Program Files\IBM\SW\Shared\plugins\c...

Enable InfoSphere Warehouse Data Mining Visualization

Memory limit (MB):

Other power options:

Check InfoSphere Warehouse Ver...

Visualizer

TIME

Y1: PASSENGERS_AB

Y2: PASSENGERS_BC

2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010

Summary

- IBM SPSS Modeler is a high-performance data mining and text analytics workbench
- IBM SPSS Modeler allows organizations across industry and regardless of size
 - Quickly realize a positive ROI from analytical investment
 - Uncover predictive intelligence
 - Proactively reduce costs and increase productivity
 - Repeatedly reduce costs and increase productivity
- New features in IBM SPSS Modeler
 - Make predictive analytics more accessible
 - Support the ever-more strategic role of predictive analytics to an organization
 - Increases synergy with IBM technology for a best in class, single vendor solution for hardware, data warehousing, business intelligence and predictive analytics



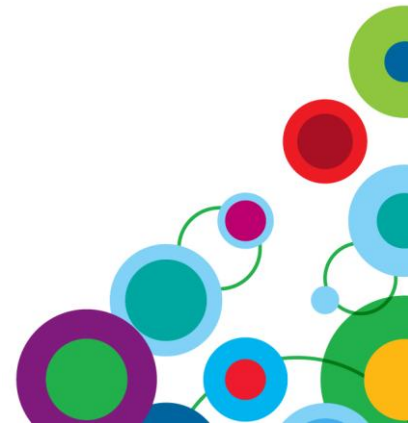
WIN *an Apple[®] iPad*

Please remember to complete your session evaluation online at the Communication Station or point your Smart Phone browser to:

www.spss.com/goldcoast



For a chance to win an Apple[®] iPad



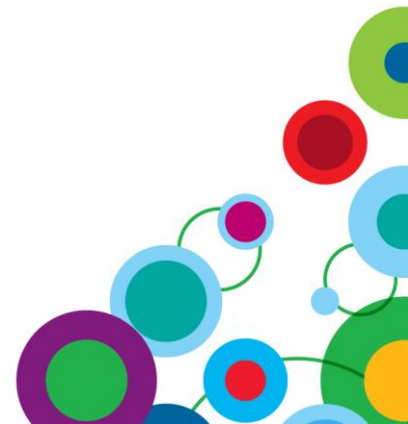
Thank you!

Q&A



Additional Resources

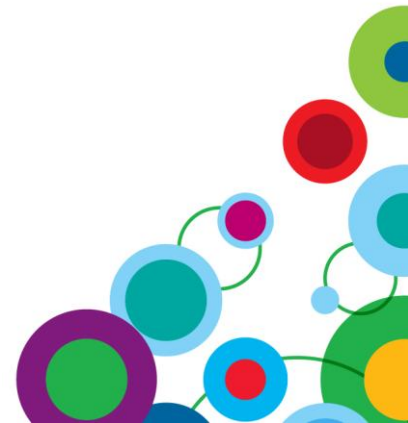
- [IBM SPSS Modeler](http://www.ibm.com) on www.ibm.com
 - <http://www.ibm.com/software/analytics/spss/products/modeler/>
- NEW & Updated Collateral
 - IBM SPSS Modeler Professional Spec Sheet
 - IBM SPSS Modeler Premium Spec Sheet
 - IBM SPSS Modeler and IBM Cognos 8 Business Intelligence Brochure
 - IBM SPSS Modeler and IBM InfoSphere Warehouse Brochure
- Questions about this presentation:
 - Jason Burke (jason.burke@au1.ibm.com)





Information and Analytics Communities

- **On-line communities, User Groups, Technical Forums, Blogs, Social networks, and more**
 - Find a community that interests you at...
 - ibm.com/software/data/community
- **Information Champions**
 - Recognizing individuals who have made the most outstanding contributions to Information Management communities
 - ibm.com/software/data/champion



Disclaimer



© Copyright IBM Corporation 2010. All rights reserved.

U.S. Government Users Restricted Rights - Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

THE INFORMATION CONTAINED IN THIS PRESENTATION IS PROVIDED FOR INFORMATIONAL PURPOSES ONLY. WHILE EFFORTS WERE MADE TO VERIFY THE COMPLETENESS AND ACCURACY OF THE INFORMATION CONTAINED IN THIS PRESENTATION, IT IS PROVIDED “AS IS” WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. IN ADDITION, THIS INFORMATION IS BASED ON IBM’S CURRENT PRODUCT PLANS AND STRATEGY, WHICH ARE SUBJECT TO CHANGE BY IBM WITHOUT NOTICE. IBM SHALL NOT BE RESPONSIBLE FOR ANY DAMAGES ARISING OUT OF THE USE OF, OR OTHERWISE RELATED TO, THIS PRESENTATION OR ANY OTHER DOCUMENTATION. NOTHING CONTAINED IN THIS PRESENTATION IS INTENDED TO, NOR SHALL HAVE THE EFFECT OF, CREATING ANY WARRANTIES OR REPRESENTATIONS FROM IBM (OR ITS SUPPLIERS OR LICENSORS), OR ALTERING THE TERMS AND CONDITIONS OF ANY AGREEMENT OR LICENSE GOVERNING THE USE OF IBM PRODUCTS AND/OR SOFTWARE.

IBM, the IBM logo, ibm.com, IBM Cognos 8 , DB2, IBM InfoSphere, IBM SPSS Modeler, IBM SPSS Statistics, IBM SPSS Decision Management and IBM SPSS Collaboration and Deployment Services are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at “Copyright and trademark information” at www.ibm.com/legal/copytrade.shtml

Other company, product, or service names may be trademarks or service marks of others.

