



Predictive Analytics / Data Mining im Controlling

Sprecher: Dr. Karsten Oehler



Das Buch zum Film.....

Data-Mining im Controlling	
Abweichungsanalysen und Prognoserechnungen durch Data Mining verbessern	
<ul style="list-style-type: none"> ■ Das Controlling verfügt über große Datenbestände, sodass dort Data-Mining-Verfahren zielbringend eingesetzt werden können. ■ Insbesondere für die Abweichungsanalyse, die Früherkennung und die Prognoserechnungen sind durch den Einsatz von Data-Mining-Verfahren Prozessverbesserungen zu erwarten. ■ Bestehende Verfahren des Performance Managements werden nicht obsolet, sondern durch Data-Mining-Verfahren sinnvoll ergänzt. ■ Der Beitrag zeigt potenzielle Möglichkeiten für den Einsatz von Data-Mining-Werkzeugen für Controller. ■ Data-Mining-Verfahren können mittlerweile auch ohne tiefgreifende Spezialkenntnisse, etwa in Statistik, angewendet werden. 	
Inhalt	Seite
<hr/> <ul style="list-style-type: none"> ■ Die Autoren 	
1	

Data-Mining im Controlling	
1 Überblick: IT-Einsatz im Controlling	
1.1 Business-Intelligence-Werkzeuge und Excel	
<p>Business Intelligence, also das interaktive Arbeiten mit Unternehmensdaten zur Analyse, ist kein ganz neues Thema mehr im Controlling. Seit die ersten mehrdimensionalen Anwendungen vor längerer Zeit am Markt erschienen sind, hat Business Intelligence (BI) zwar die Tabellenkalkulation nicht ersetzen können, es hat sich jedoch eine gesunde Synthese entwickelt: Standardisierbare Controlling-Prozesse wie Budgetierung oder Abweichungsanalysen werden meistens mit BI-Werkzeugen, eher ad hoc orientierte Aufgaben und Berechnungen über Excel abgewickelt.</p> <p>Allerdings hat man das Gefühl, dass die Entwicklung in Bezug auf bessere Systemunterstützung etwas stagniert. Natürlich werden die BI-Umgebungen komfortabler und die Ergebnisse schneller bereitgestellt. Auch die Tabellenkalkulation ist schneller und bunter geworden und einige Werkzeuge sind auch intuitiver zu handhaben (manche allerdings auch nicht). Aber der nächste große Entwicklungsschritt bei der Controlling-Unterstützung lässt auf sich warten: Wie kann die Arbeit des Controllers durch IT-Werkzeuge signifikant verbessert, also effektiver (neue Einsichten) oder effizienter (weniger Routine) werden?</p>	<p>Entwicklungstau bei Systemunterstützung im Controlling</p>
1.2 Anwendungen früher und heute	
<p>In diesem Zusammenhang soll ein uralter Ausspruch von Professor Peter Mertens thematisiert werden: Er sagte Ende der Siebzigerjahre sinngemäß, existierende Anwendungen beschränkten sich auf das phantasielose Abprogrammieren manueller Methoden.¹ Das war damals auf die Anbieter klassischer Kostenrechnungssysteme gemünzt, die sich darauf beschränkten, etablierte Verfahren in Programme zu gießen.</p> <p>So ist es schon lange in Anbetracht moderner Dashboards nicht mehr. Simulationen oder auch die mehrdimensionale Analyse sind erst durch die IT-Unterstützung möglich geworden. Methoden wie die Monte-Carlo-Simulation sind allerdings nur mit ausreichender IT-Unterstützung sinnvoll umsetzbar. Aber trotzdem steht die Effizienz im Vordergrund, denn die meisten Verfahren gab es bereits vor der IT-Unterstützung und haben sich nur wenig geändert. Es wäre schade, wenn sich die neuen Möglichkeiten</p>	<p>Einfallsloses Abprogrammieren in den Siebzigerjahren</p> <p>Neue Möglichkeiten der IT-Unterstützung</p>
<hr/> <p>¹ Mertens, (1984), S. 88.</p>	
2	

Erscheint im Controlling Berater, Ende 2011, PDF vorab über karsten.oehler@de.ibm.com erhältlich.



Agenda

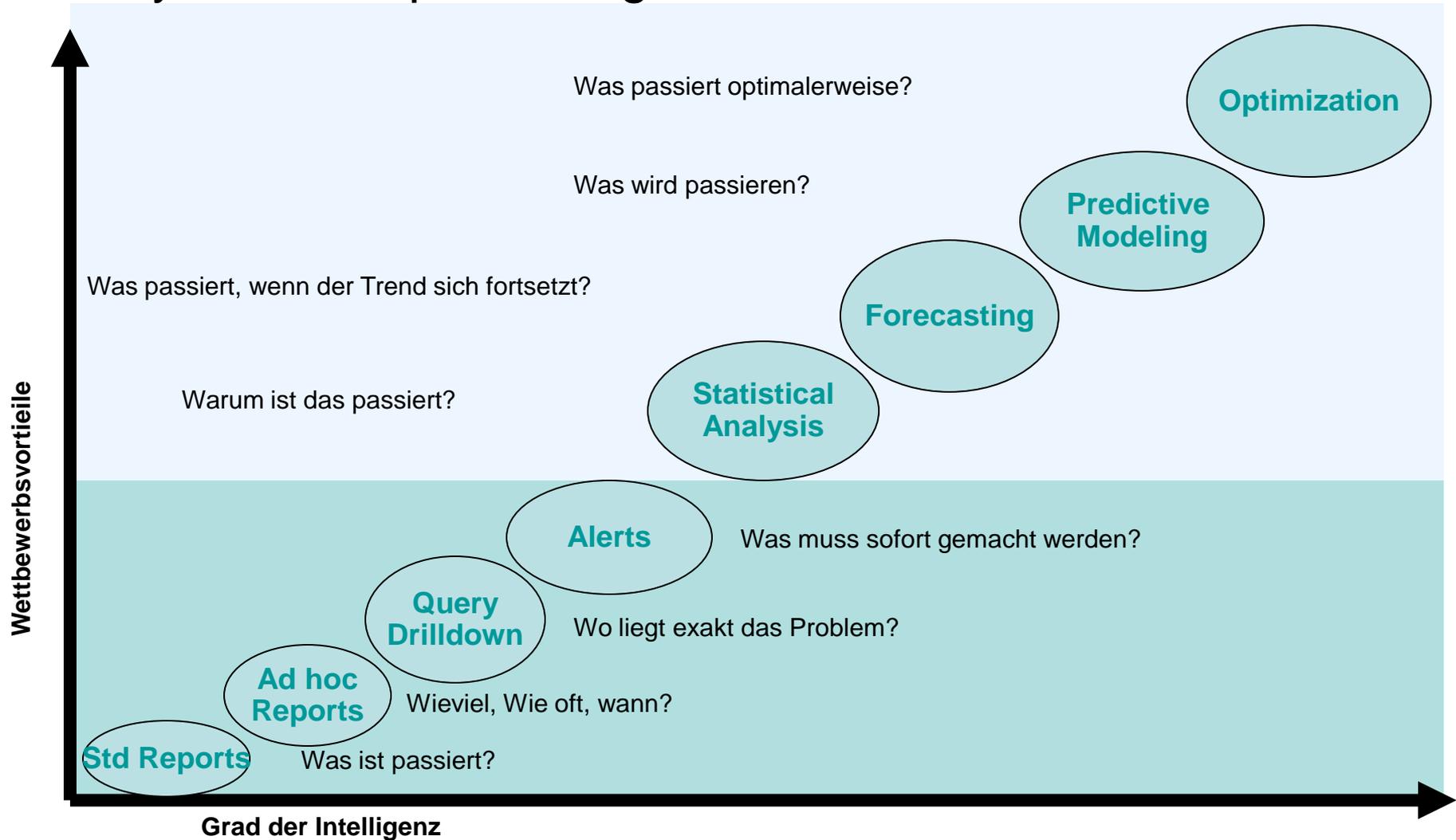
- Motivation
- Predictive Analytics versus OLAP?
- Predictive Analytics im Controlling
- Was ist nun konkret zu tun?



Motivation

- Standard-BI und Performance Management sind weitgehend in mittleren bis größeren Unternehmen etabliert.
- IBMs Stärke ist das Advanced Analytics und die zugehörige Lösungskompetenz.
- Hierzu müssen wir den Value Added für die Fachbereiche herausstellen. Das ist auch Ziel dieses Webinars.

Analytics und Optimierung



Quelle: Davenport, "Competing on Analytics"

Thesen zum BI-Stand im Controlling

- Unternehmen haben in Datenversorgung investiert. Diese Daten stehen dem Controlling zur Verfügung.
- Controlling nutzt seit langem klassisches BI, hängt aber bzgl. intelligenter Nutzung hinter Vertrieb und Marketing weit zurück.
- Klassisches Controlling ist „deterministisch“. Die Methoden sind überaltet.
- Innovative Controlling-Abteilungen interessieren sich stärker für „Beyond BI“.
- Externe Daten sind (endlich) auch im Controlling von höherer Bedeutung



Pains

- Durch ERP / trotz DW – Einsatz: Ertrinken in Daten
- Hohe Forecast-Ungenauigkeit
- Hoher Aufwand bei der Abweichungsanalyse
- Unsicherheit bei Abschätzung von Erfolgs-Einflussfaktoren



Agenda

- Motivation
- Predictive Analytics versus OLAP?
- Predictive Analytics im Controlling
- Was ist nun konkret zu tun?

Was ist Predictive Analytics





Was ist Predictive Analytics





Predictive Analytics
generiert aus *Daten* operative *Aktionen*, indem
verlässliche Schlüsse zur aktuellen Situation
und zukünftigen Ereignissen *erkannt* bzw.
prognostiziert werden.

Neue Methoden braucht das Land



So geht's...

The screenshot displays the IBM SPSS Modeler interface. The main workspace shows a workflow diagram starting with a data source 'tree_credit.sav', followed by a 'Type' node, and then a 'Creditrating' node. A 'Credit rating' node is also shown below, connected to the 'Type' node. A 'Table' node and an 'Analysis' node are also visible. The right-hand pane shows a project tree under 'CRISP-DM' with a 'Geschäftsverständnis' (Business Understanding) folder expanded, containing sub-items: 'Datenverständnis', 'Datenvorbereitung', 'Modellierung', 'Evaluierung', and 'Bereitstellung'. The bottom toolbar includes various modeling tools like 'Datenbank', 'Datei (var.)', 'Autom. Datenvorbereitung', 'Auswählen', 'Stichprobe', 'Aggregieren', 'Ableiten', 'Typ', 'Filter', 'Diagrammtafel', 'Autom. Klassifizierer', 'Autonumerisch', 'Autom. Cluster', 'Tabelle', 'Textdatei', and 'Datenb.'. The status bar at the bottom indicates 'Server: Lokaler Server' and '276 MB / 345 MB'.

So geht's

The screenshot shows the IBM SPSS Modeler interface. The main workspace displays a workflow with three nodes: 'tree_credit.sav', 'Type', and 'Creditrating'. The 'Creditrating' node is highlighted, and its output is visible in the 'Modelle' pane on the right, showing a CHAID model icon and the text 'Credit rating'.

In the foreground, the 'tree_credit.sav' window is open, showing the 'Vorschau' (Preview) tab. The preview window displays a table with 10 data rows and 6 columns: Credit rating, Age, Income level, Number of credit cards, Education, and Car loans. The first row is highlighted.

	Credit rating	Age	Income level	Number of credit cards	Education	Car loans
1	Bad	36.220	Medium	5 or more	College	More than 2
2	Bad	21.990	Medium	5 or more	College	More than 2
3	Bad	29.170	Low	5 or more	High school	More than 2
4	Bad	32.753	Low	5 or more	College	None or 1
5	Bad	36.771	Medium	5 or more	College	More than 2
6	Bad	39.325	Medium	5 or more	College	More than 2
7	Bad	31.699	Medium	5 or more	College	More than 2
8	Bad	34.718	Low	5 or more	High school	More than 2
9	Bad	31.531	Low	5 or more	High school	More than 2
10	Bad	24.780	Medium	5 or more	College	More than 2

So geht's

The screenshot shows the IBM SPSS Modeler interface. The main workspace displays a workflow starting with a file named 'tree_credit.sav', followed by a 'Type' node, and then a 'Creditrating' node. A 'Type' dialog box is open in the foreground, showing a table of field settings.

Feld	Messung	Werte	Fehlend	Überprüfen	Rolle
Credit rating	Nominal	Bad,Good,"No cr...	*	Keine	Ziel
Age	Stetig	[20.0026995233...		Keine	Eingabe
Income level	Ordinal	High,Low,Medium	Aus	Keine	Eingabe
Number of credit c...	Nominal	"Less than 5","5 ...		Keine	Eingabe
Education	Nominal	"High school","Col...		Keine	Eingabe
Car loans	Nominal	"None or 1","Mor...		Keine	Eingabe

At the bottom of the dialog box, there are radio buttons for 'Aktuelle Felder anzeigen' (selected) and 'Nicht verwendete Feldeinstellungen anzeigen', along with 'OK', 'Abbrechen', 'Anwenden', and 'Zurücksetzen' buttons.

So geht's

The screenshot shows the IBM SPSS Modeler interface. In the background, a workflow diagram is visible with nodes: 'tree_credit.sav', 'Type', and 'Creditrating'. The 'Creditrating' node is highlighted, and a dialog box is open over it.

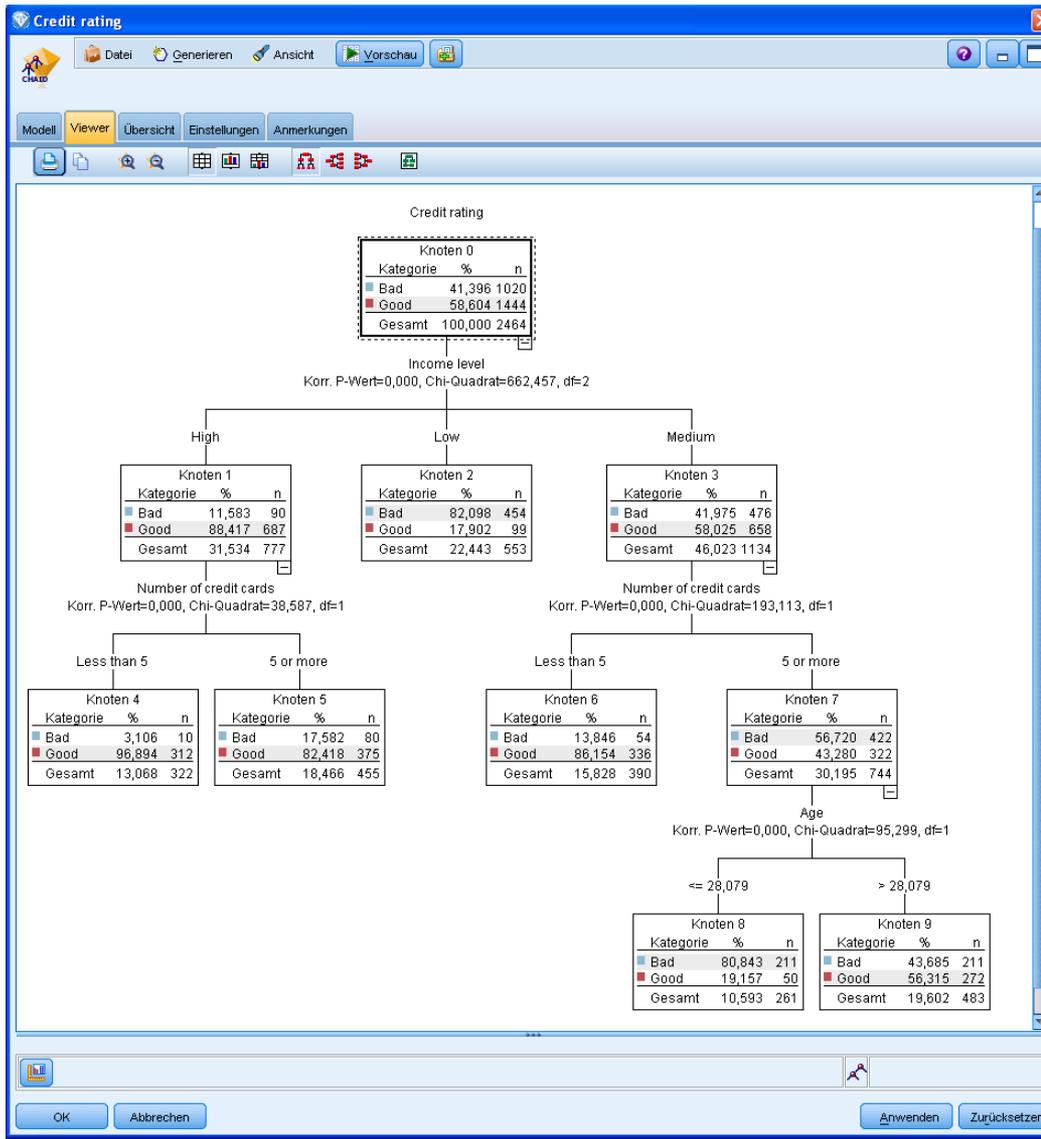
Creditrating Dialog Box:

- Ziel:** Standardmodell
- Erstellungsoptionen:**
 - Wählen Sie ein Element aus:
 - Ziel
 - Grundlagen
 - Stoppregeln
 - Kosten
 - Ensembles
 - Erweitert
 - Wie möchten Sie vorgehen?
 - Neues Modell aufbauen
 - Training des bestehenden Modells fortsetzen
 - Was ist Ihr Hauptziel?
 - Einen einzelnen Baum aufbauen
 - Einzelner Baum
 - Modus: Modell erzeugen Interaktive Sitzung starten
 - Interaktiv erstellte Aufbauregeln verwenden
 - Modellgenauigkeit erweitern (Boosting)
 - Modellstabilität erweitern (Bagging)
 - Modell für sehr umfangreiche Datensets erstellen (erfordert Server)
 - Beschreibung:

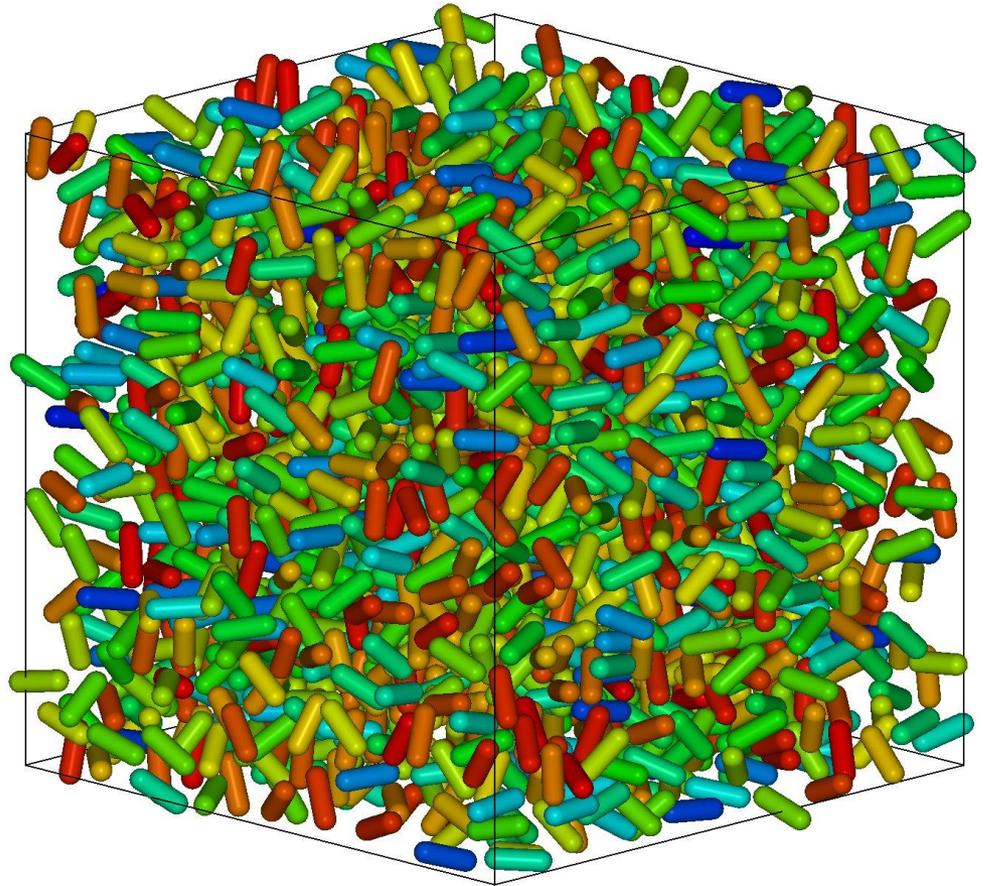
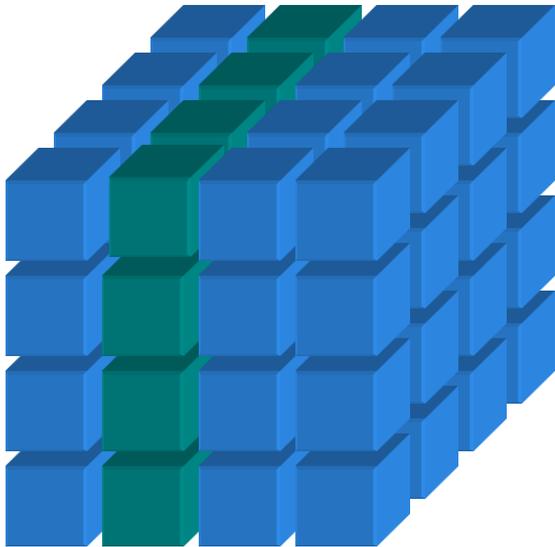
Erstellt ein einzelnes Standardmodell zur Erklärung der Beziehungen zwischen Feldern. Standardmodelle lassen sich einfacher interpretieren und ermöglichen schnelleres Scoring als verstärkte, geboostete, verpackte oder umfangreiche Datenset-Ensembles.

Buttons at the bottom of the dialog: OK, Ausführen, Abbrechen, Anwenden, Zurücksetzen.

So geht's



OLAP versus Predictive Analysis

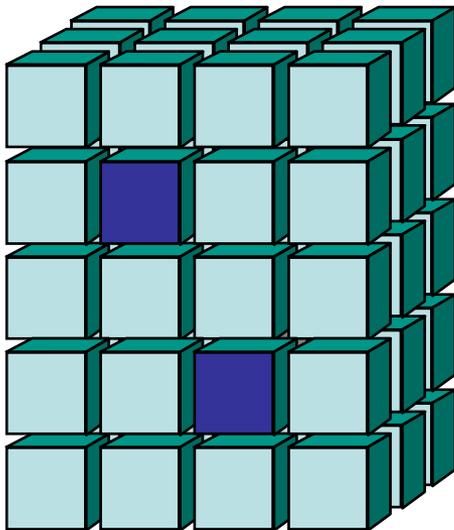


Von OLAP zu PA

Welche Kunden wurden letztes Jahr insolvent?

Welche besonderen Merkmale hatten diese Kunden?

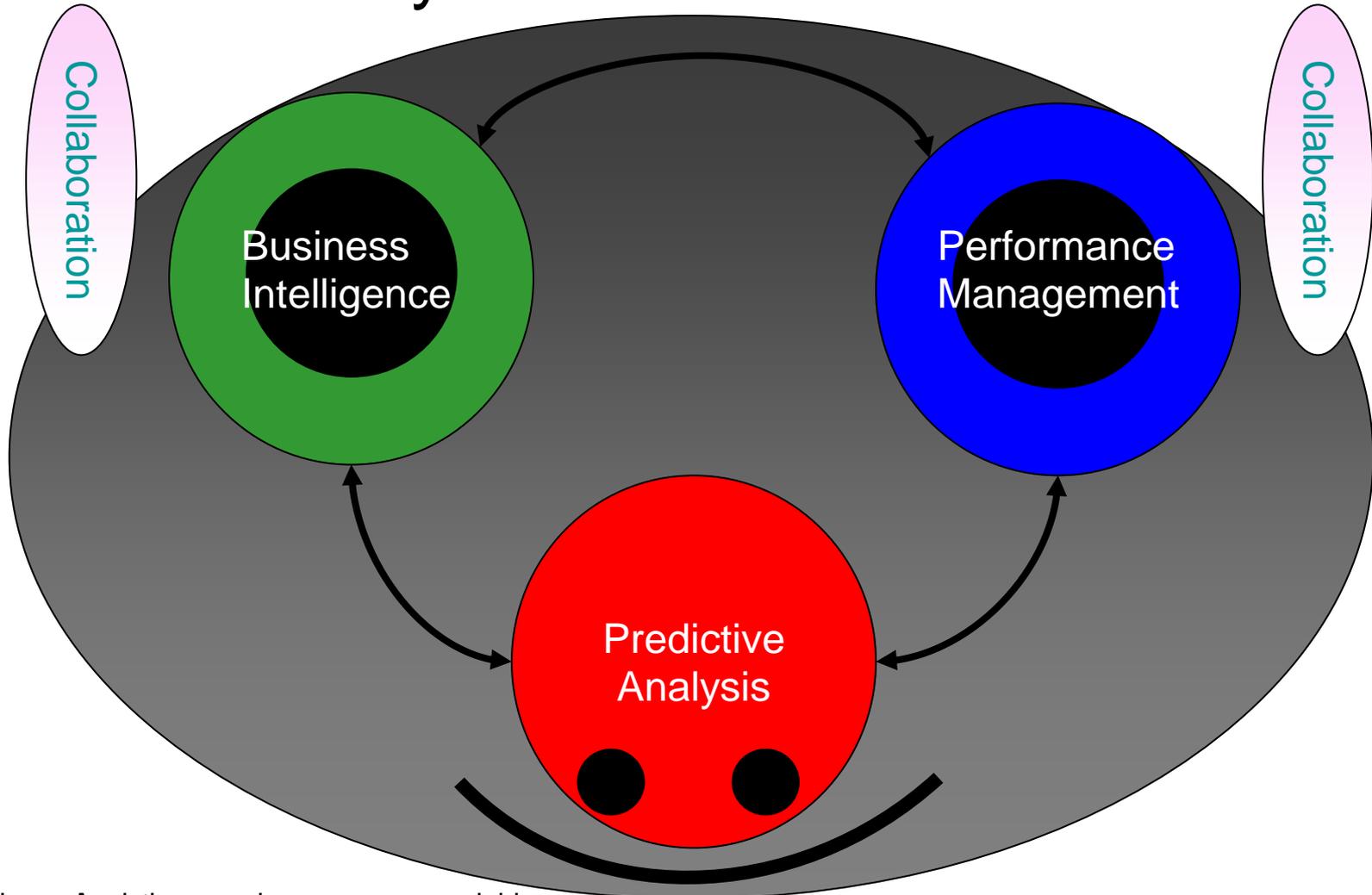
Mit welchen Kunden sollten wir besser kein Geschäft machen?



© Hans Snoek / pixelio.de © Rainer Sturm / pixelio.de

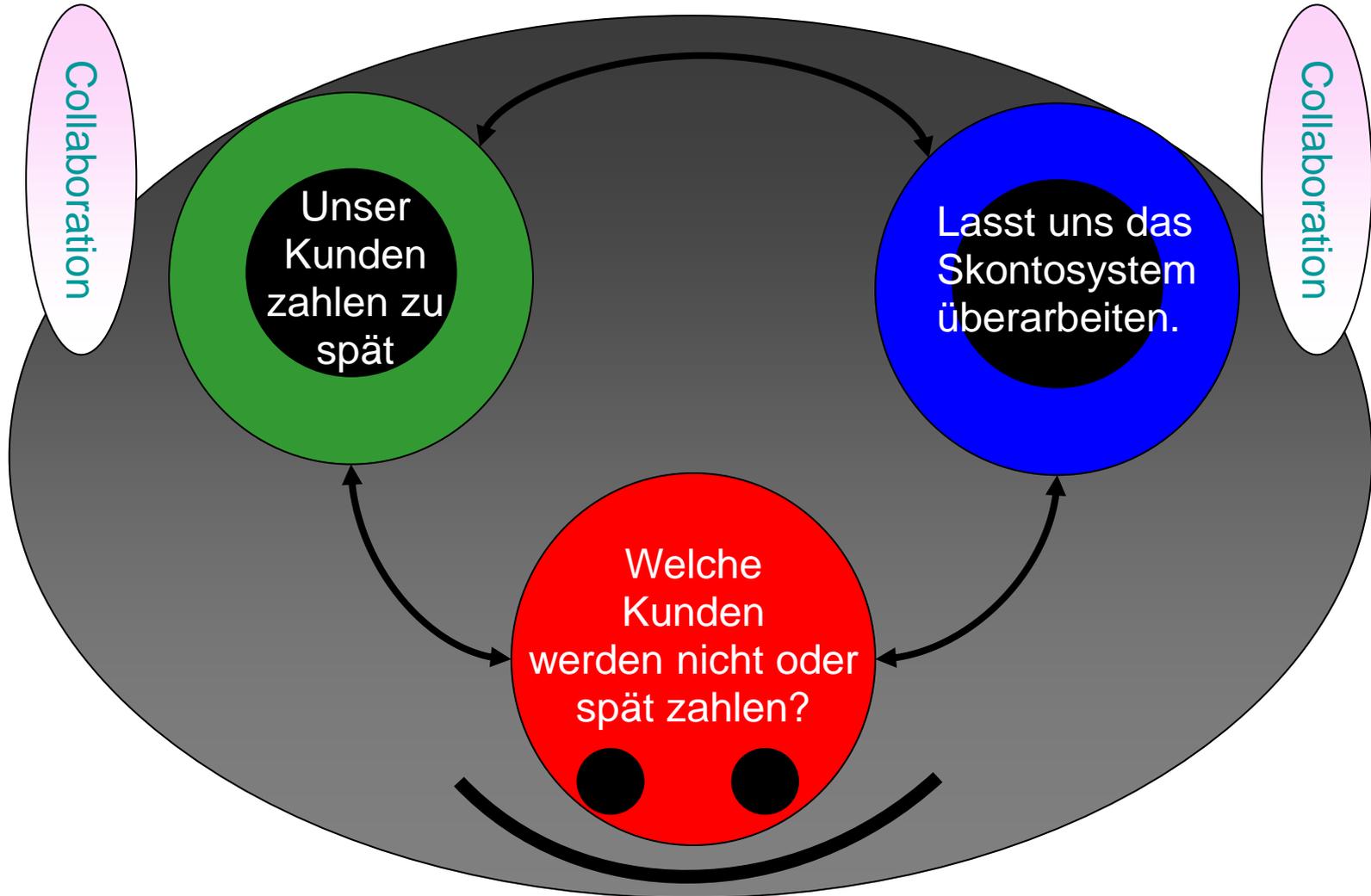
Die Trias im Business Analytics

– das ist **BARney***

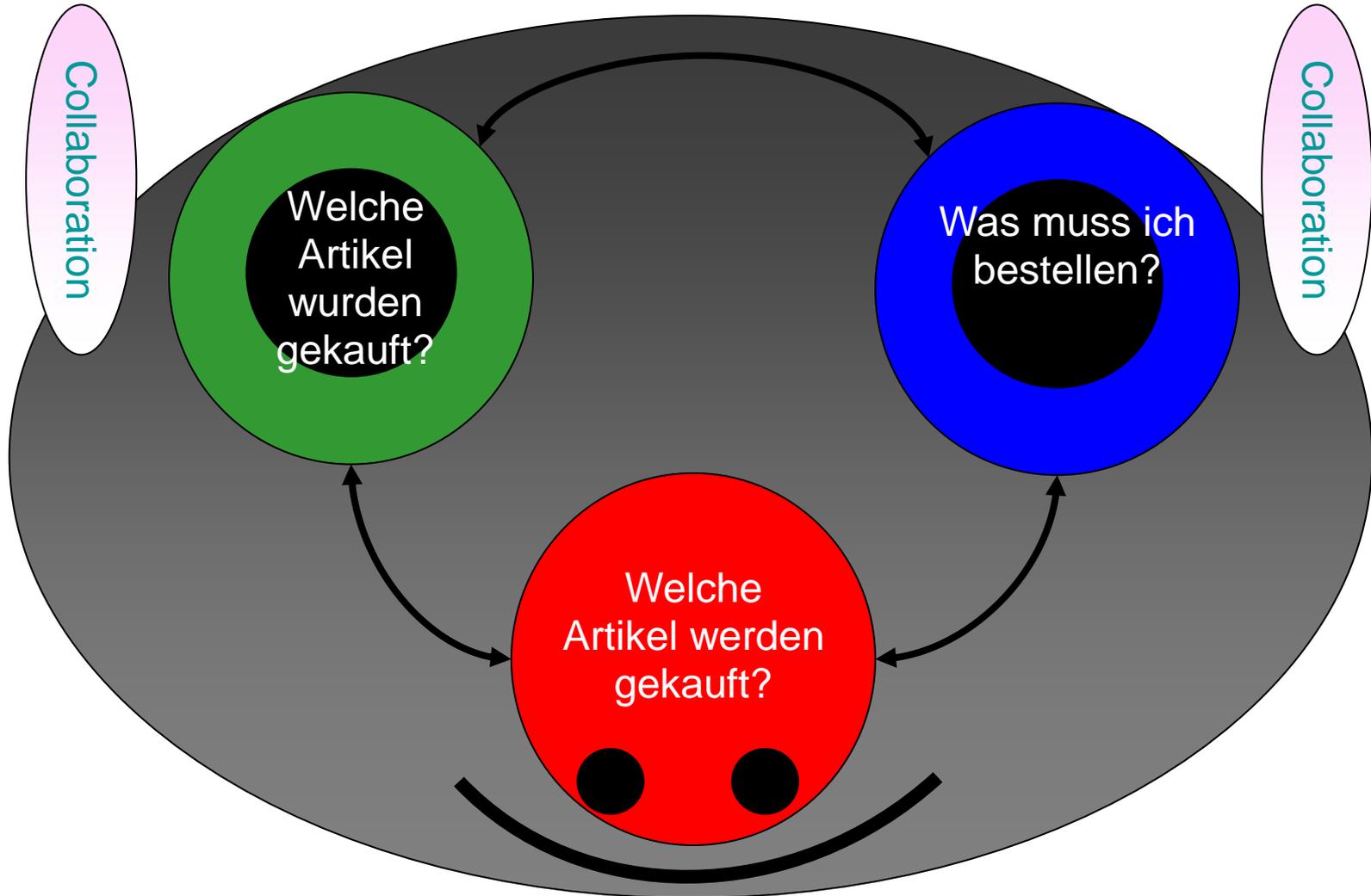


*Business Analytics = real new enormous yield

BARney in Aktion.....



BARney in Aktion.....



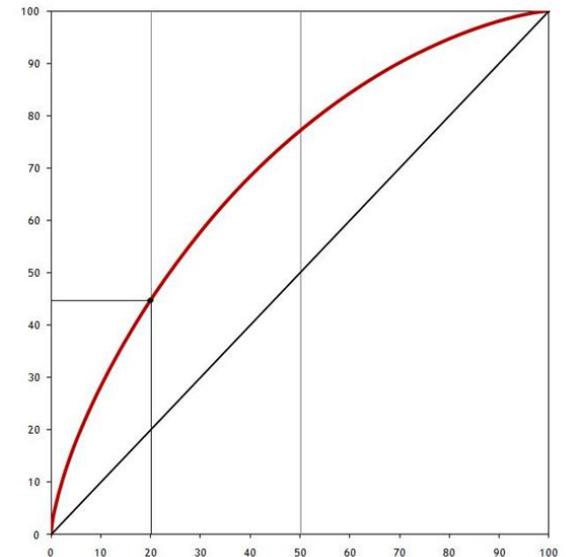
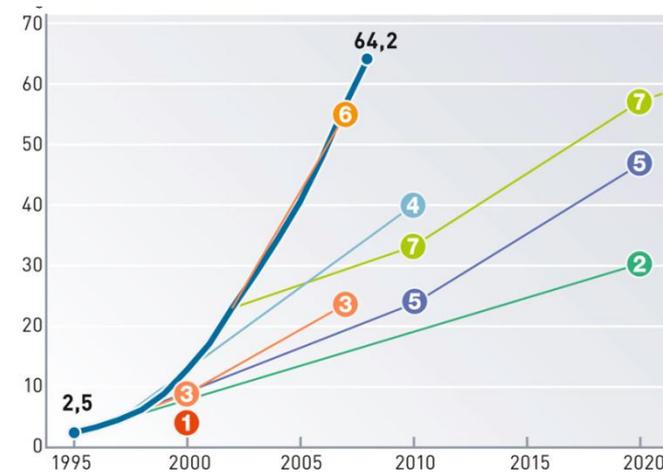


Agenda

- Motivation
- Predictive Analytics versus OLAP?
- Predictive Analytics im Controlling
- Was ist nun konkret zu tun?

PA & Controlling

- Forecasting / Planung
- Szenario Management
- Risiko Management
- Working Capital Management
- Vertriebscontrolling
- Abweichungsanalyse
-



Abweichungsanalyse

- Statt „wo sind Abweichungen aufgetreten?“ eher „wo werden mit einer hohen Wahrscheinlichkeit Abweichungen auftreten?“
- „Predictive Maintenance“
- Ermittlung von kritischen Kunden, Produkten, Kostenstellen (wo die Auswirkungen von Abweichungen am größten sind)



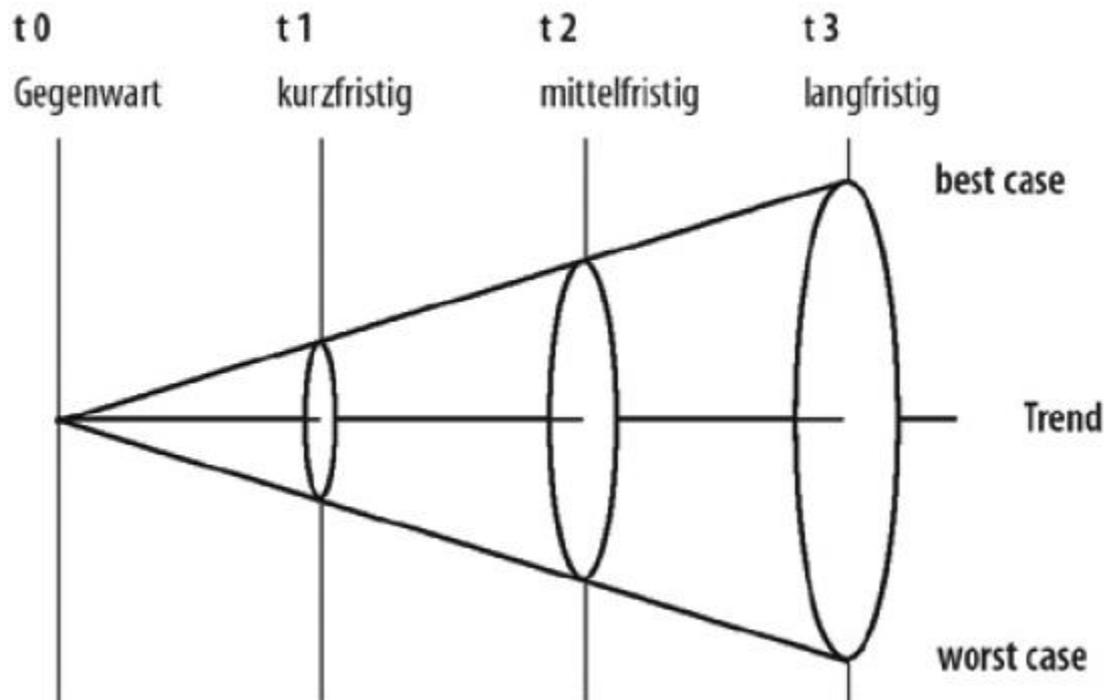
Forecasting

- Statt Fortschreibung
kausale Abhängigkeiten
berücksichtigen (z.B. Kauf
zu Marktentwicklung, Preis-
Absatzkurve)
- Arbeiten mit
Regressionskoeffizienten
- Z.B. Restwertermittlung



Szenario-Analyse

- Wo kommen eigentlich Wahrscheinlichkeiten für Best, Normal, Worst her?
- Einbezug von Risiken aus einer Falldatenbank



Assoziationsanalyse

The screenshot displays the IBM PASW Modeler 13 interface for an association analysis. The main workspace shows a flowchart starting with a data source 'FamilyTronics_Orders..' leading to a 'Type' node. From 'Type', arrows point to '20 fields' nodes, 'aPriori Associations', and 'Carma Associations'. 'aPriori Associations' is linked to a yellow box labeled 'High Affinity Flat Panels and Receivers', and 'Carma Associations' is linked to a yellow box labeled 'Potential Cross Sell candidates'.

The 'aPriori Associations' window is open, showing a table of results sorted by Confidence %:

Consequent	Antecedent	Support %	Confidence %
65 Inch Class 1080p	1400W 10.2 Ch Receiver	28.469	56.979
1400W 10.2 Ch Receiver	65 Inch Class 1080p	34.882	46.502
40 Inch Class 1080p	805W 7.1 Ch Receiver	16.034	43.285
65 Inch Class 1080p	910W 7.1 Ch Receiver	16.374	42.549
65 Inch Class 1080p	55 Inch Class 1080p	17.114	41.488
65 Inch Class 1080p	40 Inch Class 1080p	19.241	40.021
40 Inch Class 1080p	910W 7.1 Ch Receiver	16.374	39.658
65 Inch Class 1080p	22 Inch Class 720p	17.855	38.947
65 Inch Class 1080p	980W 7.1 Ch Receiver	23.522	38.804
65 Inch Class 1080p	32 Inch Class 1080p	23.335	38.771
26 Inch Class 720p	46 Inch Class 1080p	18.235	37.55
1050W 7.1 Ch Receiver	55 Inch Class 1080p	17.114	37.242
1050W 7.1 Ch Receiver	22 Inch Class 720p	17.855	36.482
980W 7.1 Ch Receiver	26 Inch Class 720p	19.881	36.318
805W 7.1 Ch Receiver	40 Inch Class 1080p	19.241	36.071
65 Inch Class 1080p	805W 7.1 Ch Receiver	16.034	36.05

The bottom of the interface shows a navigation bar with tabs for Favorites, Sources, Record Ops, Field Ops, Graphs, Modeling, Output, Export, and PASW Statistics. The 'Modeling' tab is active, showing icons for Automated, Classification, Association, and Segmentation. The 'Association' sub-tab is selected, showing icons for GRI, Apriori, Carma, and Sequence. The status bar at the bottom indicates 'Server: Local Server' and '181MB / 286MB'.



Cognos Viewer - Marketing Analysis - Windows Internet Explorer

http://cognosprod/cognos8/cgi-bin/cognos.cgi?b_action=dashboard&pathinfo=/cm&frag-header=true&path=storeID(%22130)

Cognos Viewer - Marketing Analysis

2010 High Margin Products

Rows: Product SKUs [Product SKUs] | Columns: 2009 [Time] | Context: San Francisco [Channels] | Get and Receive [Campaigns] | Units Sold (Campaigns) [Marketing Measures] | Actual/Forecast [Scenarios]

	2009	Jan-09	Feb-09	Mar-09	Apr-09	May-09	Jun-09	Jul-09	Aug-09	Sep-09	Oct-09	Nov-09	Dec-09
Flat Panel TVs	190	23	28	31	25	25	29	28	0	0	0	0	0
65 Inch Class 1080p	33	4	5	6	4	4	5	6	0	0	0	0	0
60 Inch Class 1080p	35	4	5	5	6	4	5	5	0	0	0	0	0
55 Inch Class 1080p	33	4	5	5	4	5	5	4	0	0	0	0	0
50 Inch Class 1080p	35	4	6	6	4	4	6	5	0	0	0	0	0
46 Inch Class 1080p	27	4	4	4	4	4	4	4	0	0	0	0	0

Marketing Revenue & Expenses

Rows: Income Statement [Income Statement] | Columns: 2009 [Time] | Context: Actual/Forecast [Scenarios] | San Francisco [Channels] | \$ [- Percent of Sales] | 1400W 10.2 Ch Receiver [Product SKUs]

	2009	Jan-09	Feb-09	Mar-09	Apr-09	May-09	Jun-09	Jul-09	Aug-09	Sep-09	Oct-09	Nov-09	Dec-09
Total Units Sold	1,772	168	169	160	113	156	173	139	119	148	176	115	136
Units Sold (Campaigns)	656	70	81	62	25	58	85	41	31	50	88	17	48
Quantity	1,116	98	88	98	88	98	88	98	88	98	88	98	88
Marketing Expenses	\$1,782.20	\$134.75	\$174.40	\$188.65	\$155.87	\$154.21	\$144.83	\$111.50	\$167.10	\$108.05	\$143.45	\$150.10	\$149.30
Total Revenue	12,723.45	1,208.73	1,211.20	1,147.64	807.68	1,119.56	1,239.91	1,001.35	853.82	1,064.84	1,266.75	827.13	974.83
Operating Margin	8,630.20	868.71	831.34	764.31	512.58	774.43	884.74	718.80	539.87	754.86	865.57	496.23	618.76
Operating Margin %	67.83%	71.87%	68.64%	66.60%	63.46%	69.17%	71.36%	71.78%	63.23%	70.89%	68.33%	59.99%	63.47%
Corporate Net Income	8,630.20	868.71	831.34	764.31	512.58	774.43	884.74	718.80	539.87	754.86	865.57	496.23	618.76
Unit Price	11,401	7,179	7,179	7,179	7,179	7,179	7,179	7,179	7,179	7,179	7,179	7,179	7,179

Hidden

Aus der Praxis

Hintergrund

- #2 Autoteile-Händler in den USA
- 3.300 Filialen in 40 Staaten

Ziele

- Lagerhaltung mit richtigem Mix aus SKUs in jeder Filiale
- Umsatzsteigerung im “Do it for me”-Bereich
- Umsatzsteigerung im Bereich der wenig gängigen Teile
- Reduktion der Handlingkosten durch die Bereitstellung des richtigen Artikel-Mixes an der richtigen Position der Supply Chain

Lösung

- Modelle werden zur Vorhersage der Artikelnachfrage auf Filialebene generiert
- Die Lösung betrachtet mehr als 500.000 Artikel
- Alle analytischen Komponenten werden zentral verwaltet und automatisch alle 120 Tage neu kalkuliert
- Direkte Integration in ERP Systeme

Ergebnisse

- Reduktion des non-working capitals (wenig und schlecht gehende Artikel SKUs über 13 Monate) um \$54.7Mio
- Umsatzsteigerung in den Filialumsätzen um \$109M pro Jahr
- Vorhersagemodelle für Nachfrage waren mindestens zu 70+% genau
- Signifikante Kosteneinsparungen durch Automatisierung

Return on Investment



94 % der SPSS Kunden erreichen einen positiven ROI in weniger als 11 Monaten



ROI: 1192%
Payback: 2 months
Ertrag \$1.875.000
Kosten \$145.731

3 Ideen

- Die Beschäftigung mit Unsicherheit erfordert andere Methoden
- PA ist ein Werkzeug nicht nur für Spezialisten (Statistiker)
- PA ist eine natürliche Ergänzung zu BI und PM





Agenda

- Motivation
- Predictive Analytics versus OLAP?
- Predictive Analytics im Controlling
- Was ist nun konkret zu tun?



Der ideale Ansprechpartner

- Zielgruppe sind Innovatoren in größeren Controlling- / Finance-Abteilungen.
- Fachlich: Leiter Controlling, CFO
- Idealerweise Kunden, die schon Erfahrung mit BI/PM haben (idealerweise natürlich Cognos), der Sprung von TK zu Predictive Analytics dürfte zu groß sein.



Mögliche Einstiegsfragen

- Kennen Sie Ihre wesentlichen Erfolgs-Einflussfaktoren?
- Haben Sie häufig hohe Forecast-Abweichung?
- Nutzen Sie im Unternehmen bereits Mining-Technologien? Wenn ja, in welchen Bereichen und was sind Ihre Erfahrungen?
- Nutzen Sie externe Einflussfaktoren beim Forecasting? Wenn ja, wie ermitteln Sie die Einflussstärke?
- Wie viele Abweichungen analysieren Sie monatlich? Wie priorisieren Sie?
- Nutzen Sie Szenario-Analyse? Wie fundieren Sie Wahrscheinlichkeiten und Bandbreiten?

Typische Einwände

- Das ist nur was für Experten.
 - Die Werkzeuge sind deutlich benutzerfreundlicher geworden.
 - Fachabteilungen, nicht ITler nutzen diese Werkzeuge.
- Der Aufwand rechtfertigt sich nicht.
 - ROIs einsetzender Unternehmen sprechen eine andere Sprache.
 - Aufwand liegt in der Datenbereinigung, wobei ein Großteil des Weges durch DW/BI bereits gegangen wurde.
- Das ist alles noch Zukunftsmusik.
 - Predictive Analytics existiert seit vielen Jahren.
 - Werkzeuge sind lange aus dem Experimentierstadium heraus



Dr. Karsten Oehler
Leiter Fachvertriebsberatung



Mobile 0151 147 16 306
Email Karsten.oehler@de.ibm.com

Thank you !

