

Pentaho Community Edition

Uma Introdução a Business Intelligence e ao Pentaho BI Open Source

Rafaela Raganham

rafaela @ ambientelivre.com.br

Marcio Junior Vieira

marcio @ ambientelivre.com.br



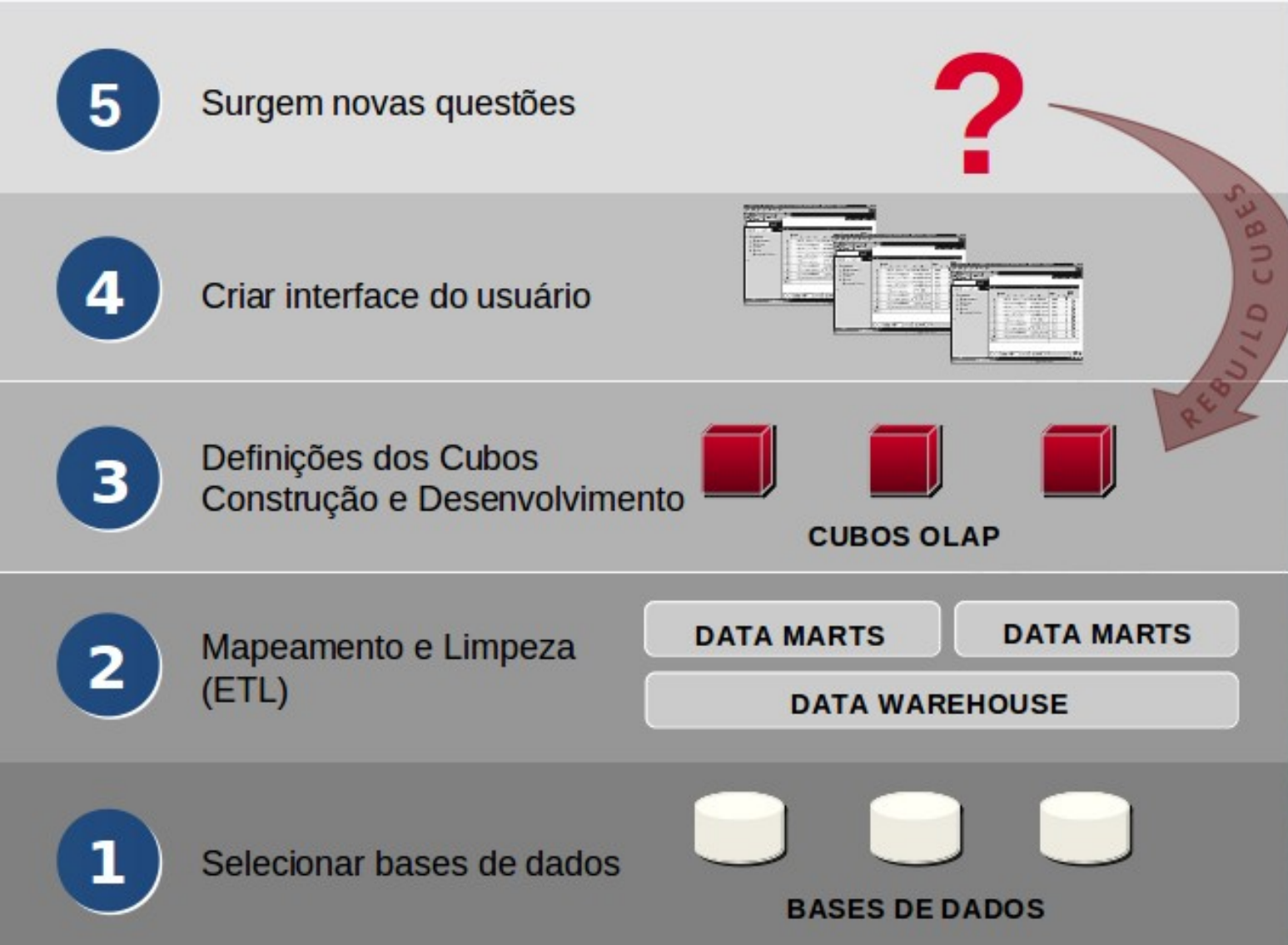
Sobre o Palestrante Marcio Junior Vieira



- A 10 anos trabalha nas áreas de análise, consultoria e desenvolvimento de sistemas.
- Formato em Tecnologia em Informática, Pós-Graduado em Software Livre - UFPR.
- Fundador e Consultor de soluções em Software Livre da Ambiente Livre Tecnologia.
- Palestrante em diversos eventos de Software Livre. (FISL, CONISLI, SOLISC, LATINOWARE, SFD, etc.)
- Estuda o Pentaho desde 2006 Trabalha desde 2009.

- Conjunto de **conceitos, métodos e recursos** tecnológicos que habilitam a obtenção e distribuição de informações geradas a partir de dados. Operacionais e históricos visando gerar resultados para tomada de decisão;
- O termo Business Intelligence (BI), pode ser traduzido como **Inteligência de negócios**, refere-se ao processo de coleta, organização, análise, compartilhamento e monitoramento de informações que oferecem suporte a gestão de negócios.

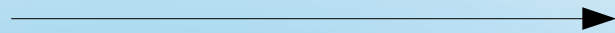
O que é Business Intelligence?



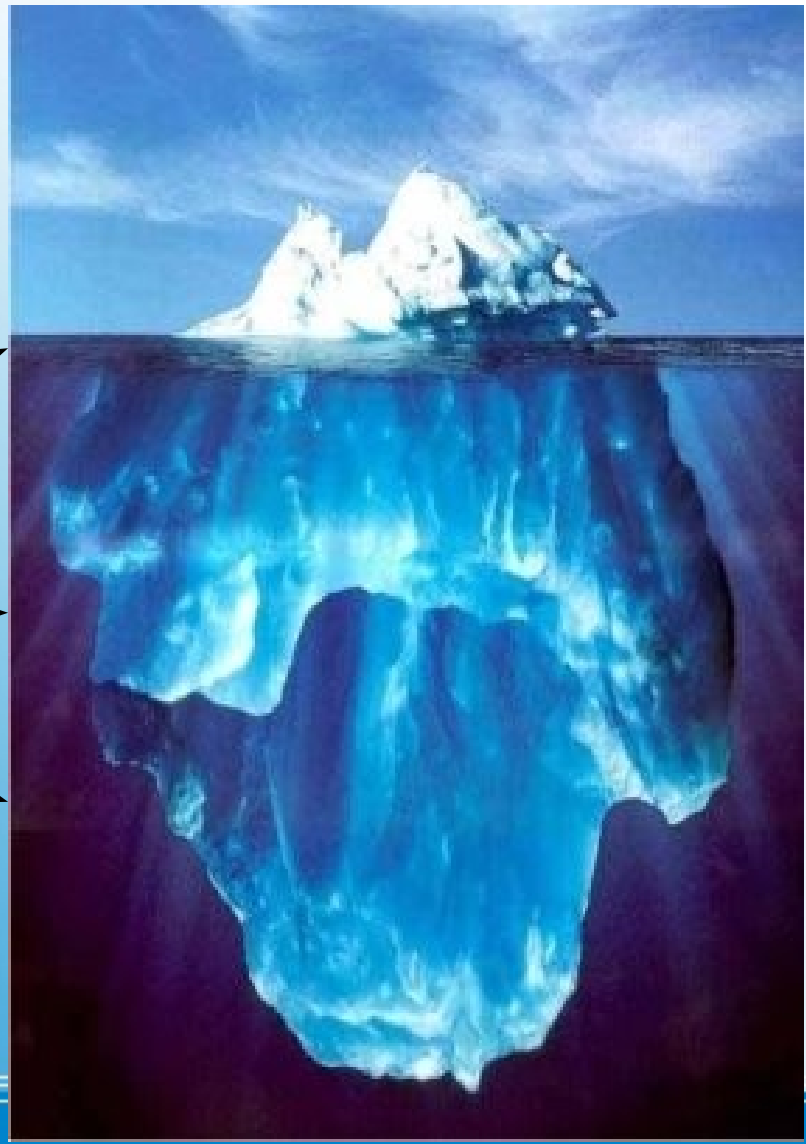
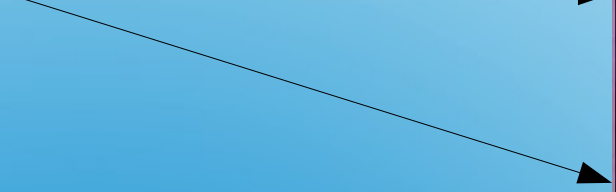
Modelo Tradicional do BI



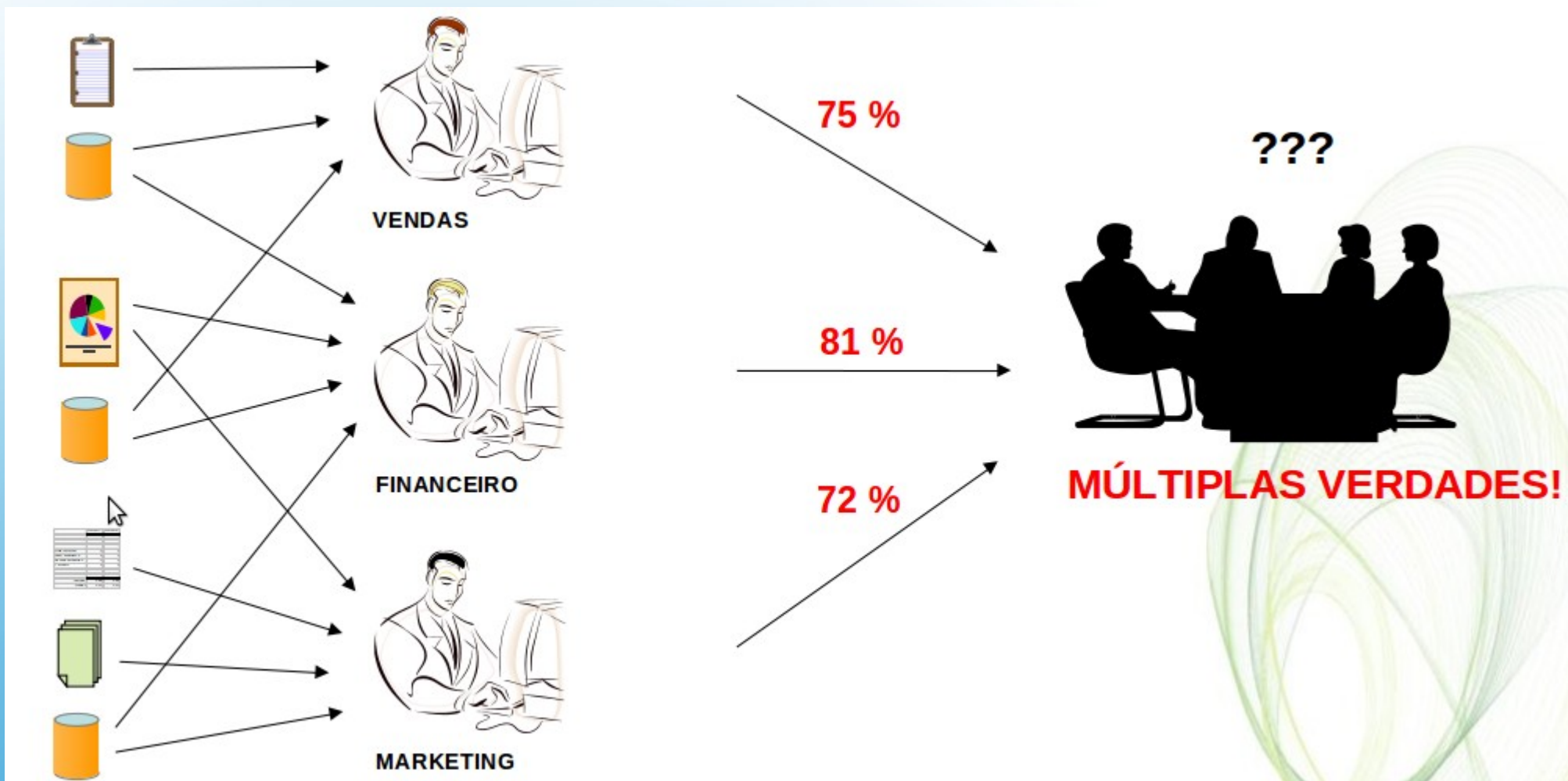
BI ?



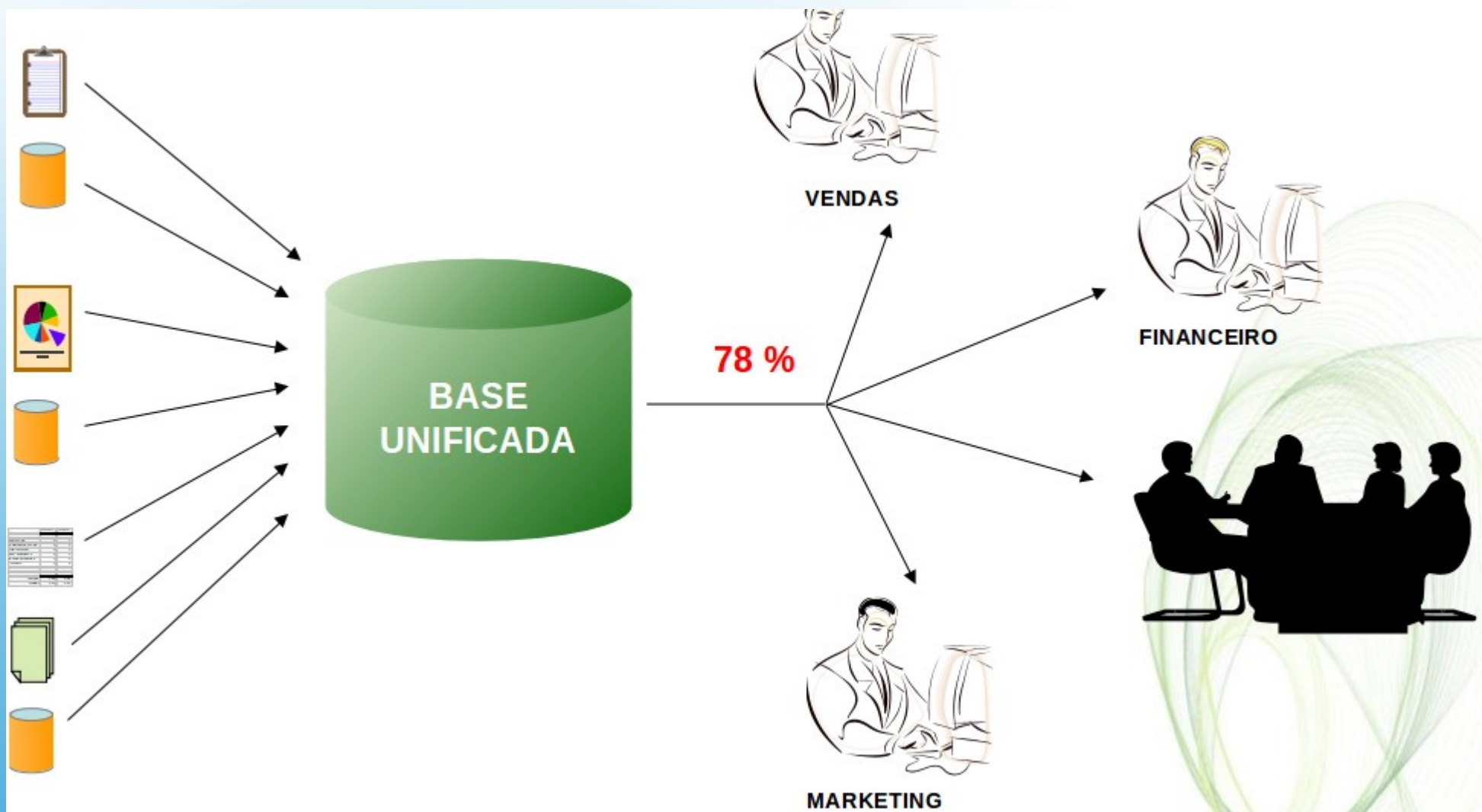
Isso é BI



BI



Desafios



Centralização de resultados

- Em 2011 o faturamento mundial do setor deve atingir 7 bilhões de dólares, o segmento deve crescer 9%.



- BI e realidade para pequenas e médias
- Um dos cinco maiores orçamentos de TI

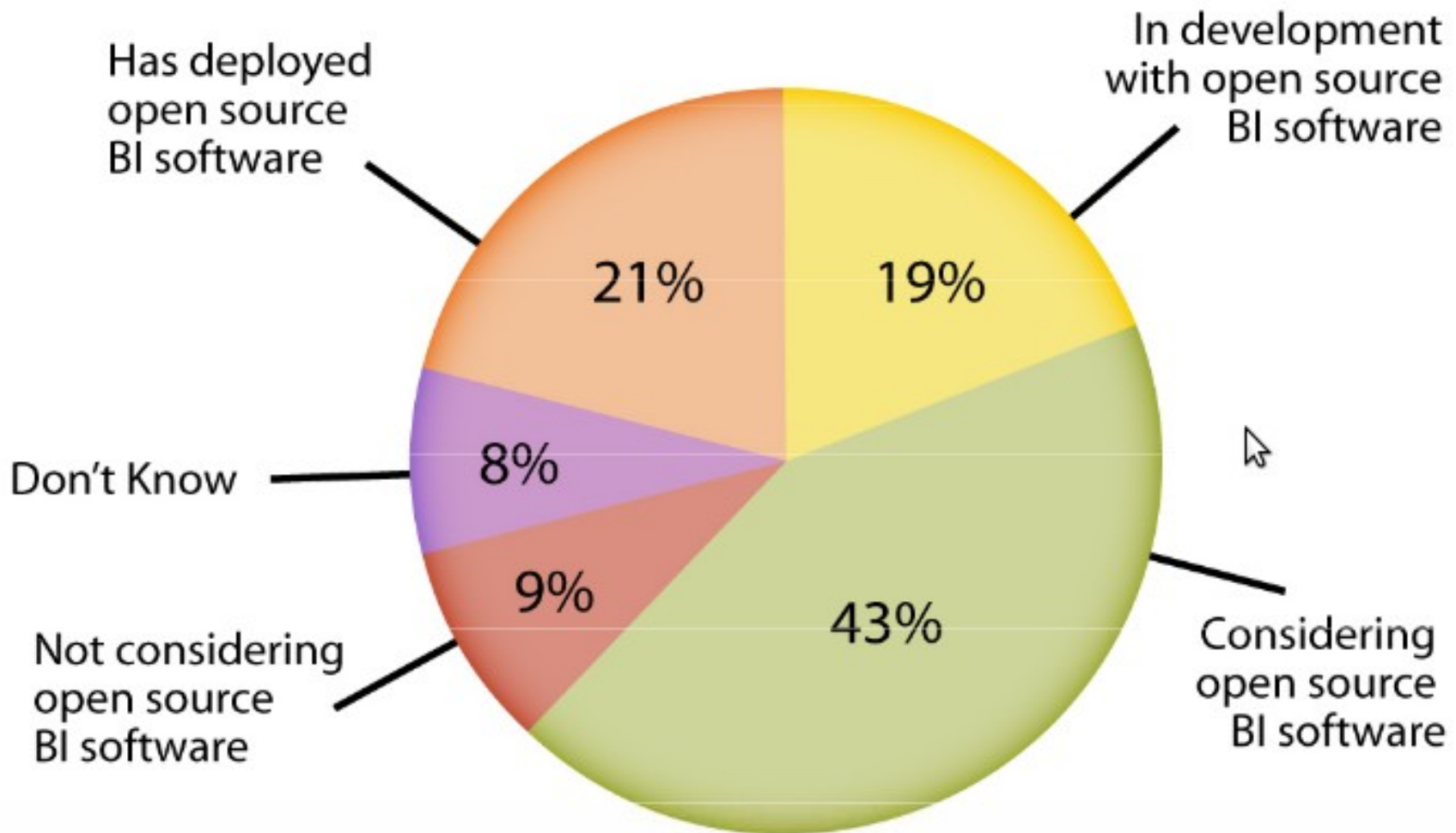


Mercado de BI



- 1,5 mil CIOs em todo mundo mostram que orçamentos continuam o mesmo, mas mudam os projetos. Valores crescem na América latina
- As prioridades 2010/2011:
 - 1 - BI.
 - 2 - ERP.
 - 3 - Servidores de Armazenamento.
 - 4 – Modernização de aplicativos.
 - 5 – Tecnologias de colaboração.

Tendências do Mercado



Mercado de BI – Open Source

- Crescimento na adoção de Open Source BI, 3 vezes superior a plataforma tradicional em 2012.
- Em 2013 85% de todos os software comerciais vão incluir elementos de código aberto em geral.

Mercado BI open Source

- Desempenha funções de análise e desenvolvimento de sistemas desde 2008, atualmente trabalha na Ambiente Livre Tecnologia onde presta serviços de desenvolvimento e customizações de sistemas open source.
- Formada em Sistemas de Informação, pela Faculdades Integradas do Brasil (Unibrasil).

Rafaela Raganham



Pentaho Corporation

- A Pentaho é uma empresa de origem norte-americana, sediada na Florida, EUA.
- Agrupou projetos open source maduros e contratou alguns dos principais desenvolvedores.

- **Matt Casters**

Chief Architect, Pentaho Data Integration
Kettle Project Founder



- **Julian Hyde**

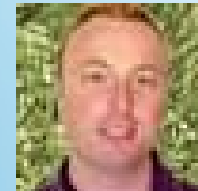
Chief Architect, Pentaho Analysis (OLAP)
Mondrian Project Founder



- **Thomas Morgner** - Chief Architect,
Pentaho Reporting



- **Mark Hall** - Sr. Pentaho Data Mining
Consultant / Engineer



fonte: <http://www.pentaho.com/team/>

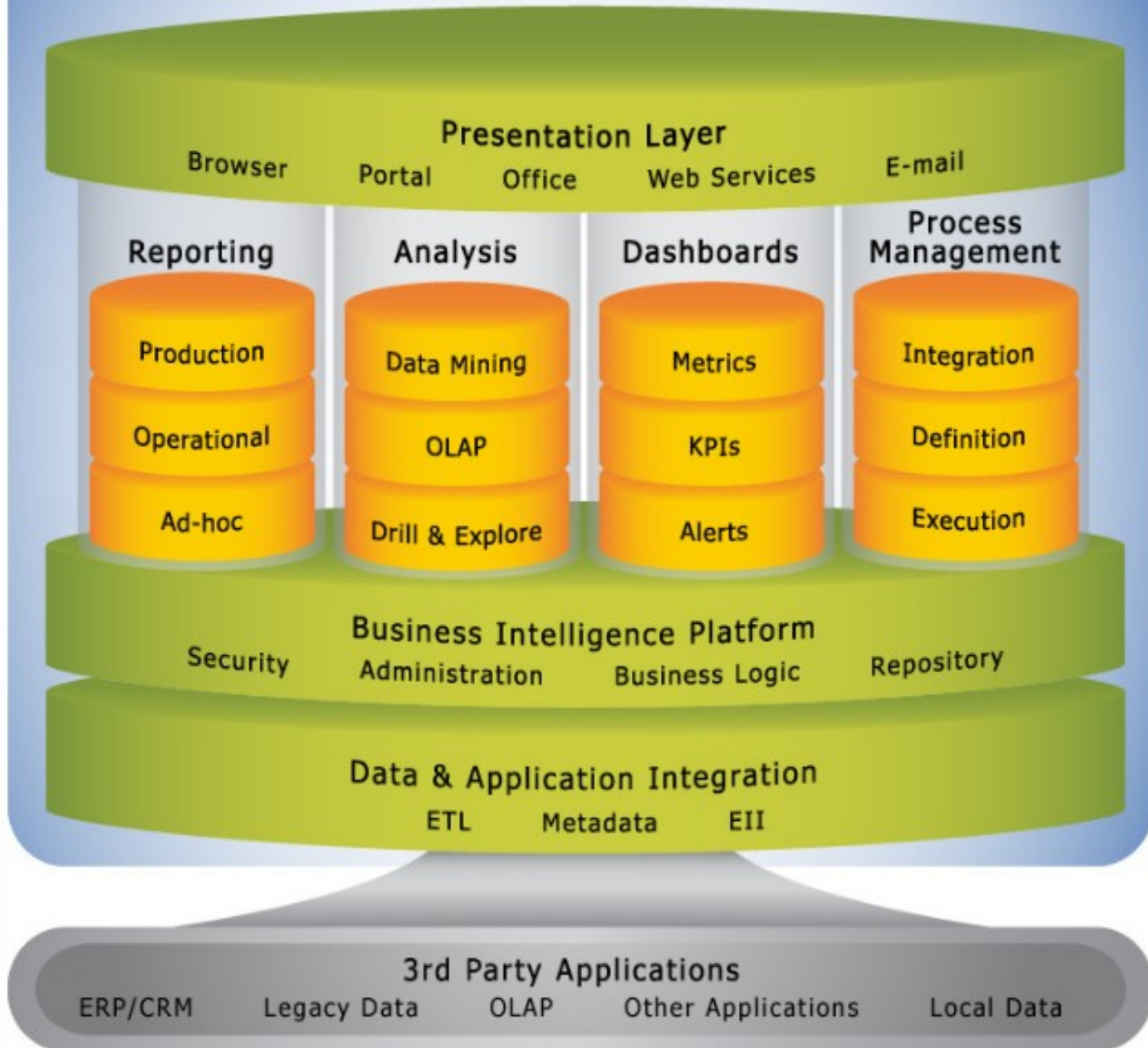
Pentaho Team



- + de 8 Milhões de Downloads (média de 1 a cada 30 segundos)
- 9.000 Produtos desenvolvidos sobre a plataforma Pentaho.
- Utilizado em mais de 185 países.
- Comunidade no Brasil estimada de mais de 1123 usuários cadastrados (lista yahoo).

Números

Pentaho Open BI Suite



Vantagens Pentaho BI

- **Usabilidade** – Componentes prontos para serem executados.
- **Habilidades** – Utilização de padrões para facilitar integrações (XML, W3C, J2EE, JSP, SQL, Javascript...).
- **Orientação a ferramentas** – Sistema completamente Integrado, concentrado em processos e soluções.

Vantagens Pentaho BI

- **Customizações/Extensão** – Toda Infra-estrutura, componentes, mecanismos e interfaces podem ser modificados (disponibilidade de código-fonte) e componentes re-utilizáveis.
- **Foco em Relatórios e Análise** – concentrada em processos e workflow e direcionada a regras de negócios.
- **Auditoria** – Robustos processos de Rastreamento.

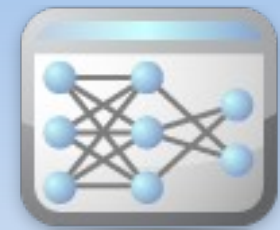
Vantagens Pentaho BI

- Custo zero de licença de software (CE).
- Portal WEB.
- Consolidação de Informações da empresa.
- Extrair e integrar dados de múltiplas fontes.
- Faz uso da experiência.
- Procurar relações de causa e efeito.
- Transformar os registros obtidos em informação útil para o conhecimento empresarial.

Funcionalidades



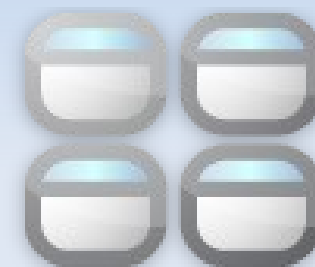
Reporting



Data Mining



Analysis



Dashboards



Data
Integration



Pentaho Data Integration

Data
Integration

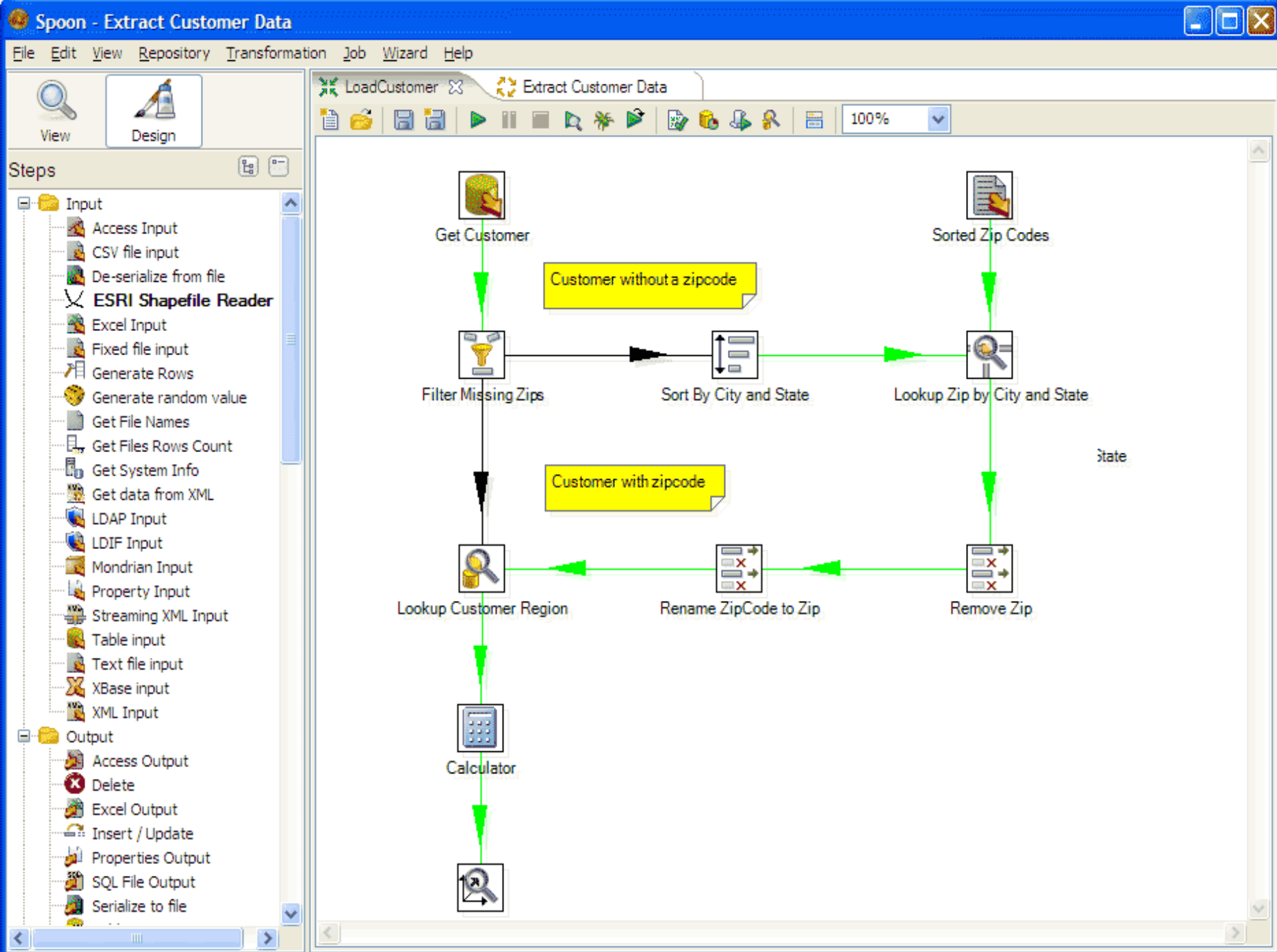
- Rica biblioteca de transformação com mais de 100 objetos de mapeamento.
- Ampla fonte de dados, incluindo pacotes de aplicações, mais de 30 plataformas de código aberto e proprietário do banco de dados, documentos do Excel, etc
- Suporte avançado de armazenagem de dados para Dimensões.

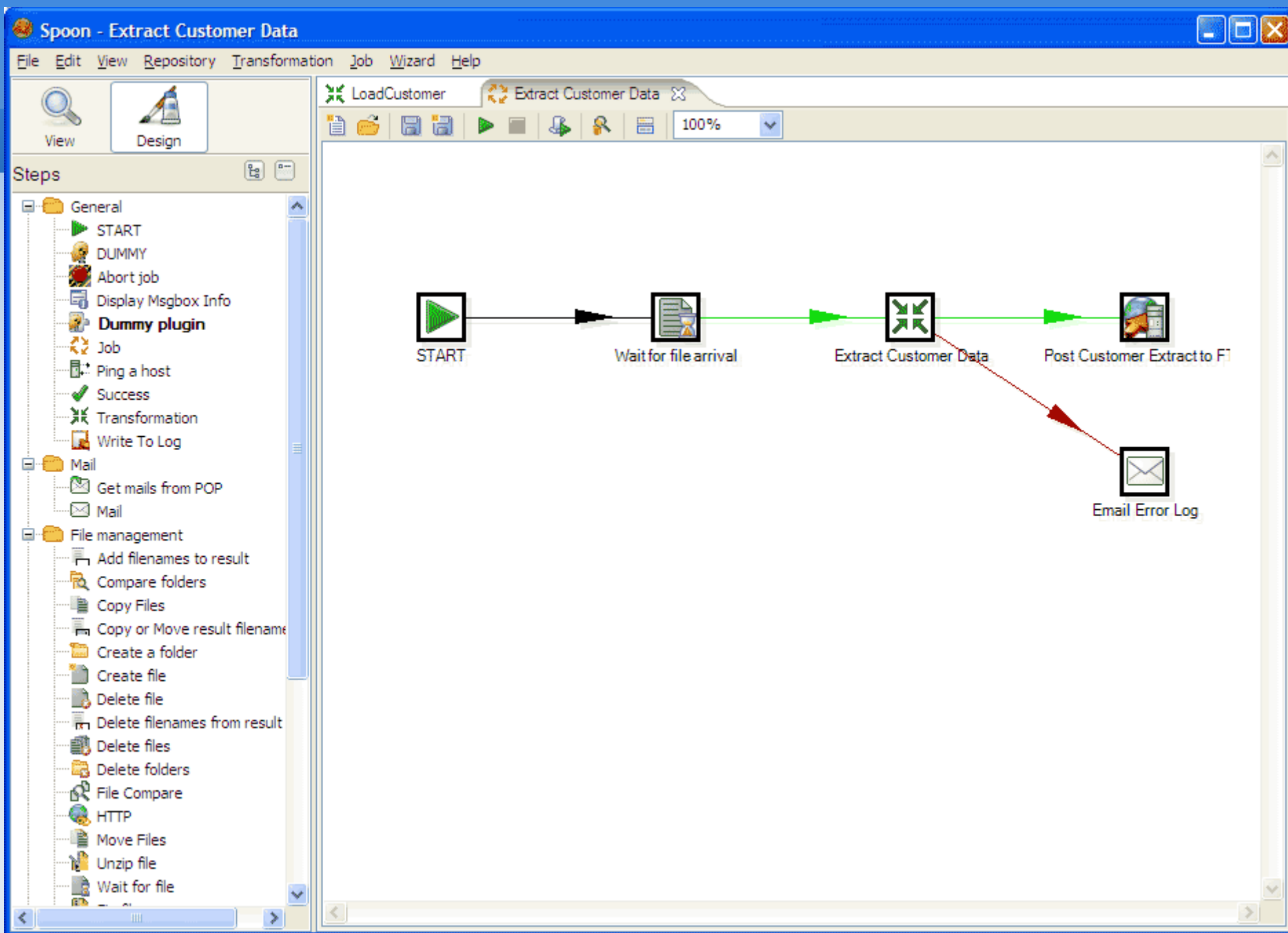


Casos de Uso - Data Integration

Data
Integration

- Armazém de dados.
- Migração de dados entre aplicações
- Importação de dados em DB, de arquivos de texto, planilhas Excel, sistemas relacionais, etc.
- Exportações de dados para outras bases de dados ou arquivos de texto
- Limpeza de dados através da aplicação de condições complexas transformações de dados
- Informação de enriquecimento através da integração de dados de várias fontes







Pentaho Analysis

Analysis

- Explorar livremente informações comerciais através de perfuração e tabulação de dados.
- Velocidade de resposta as consultas analíticas complexas.
- Visualiza a informação multi-dimensional (OLAP), escolhendo de métricas e atributos específicos para analisar

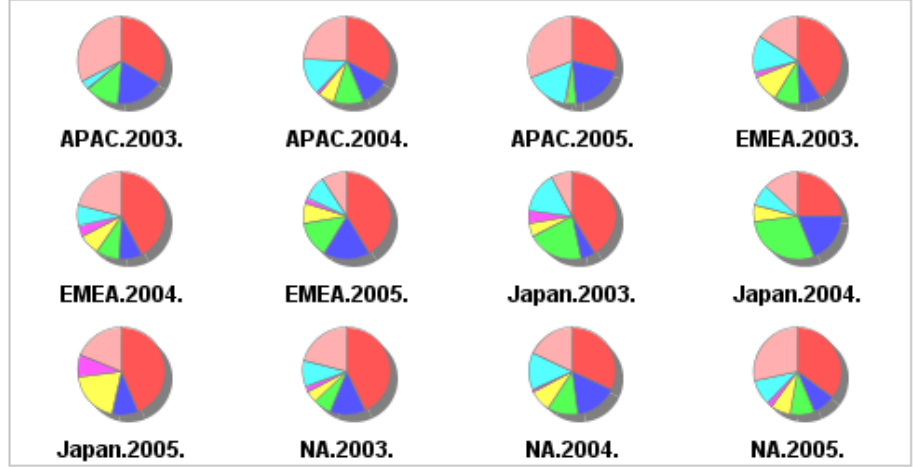


Pentaho Analysis

Analysis

- Totalmente Web, drag-and-drop e criação de relatórios.
- Classificação e filtragem avançadas
- Totais personalizados e definidos pelo usuário
cálculos
- Gráfico visualizações

- BI Developer Examples
 - Analysis
 - Business Rules
 - CDA
 - CDF
 - ChartBeans Examples
 - Data Integration with Ke
 - Data Source
 - Printing
 - Reporting
 - SVG
 - Web Service



Slicer: Measure=Sales

- Classic Cars. Motorcycles. Planes. Ships. Trains. Trucks and Buses. Vintage Cars.

Product	Markets											
	+ APAC			+ EMEA			+ Japan			+ NA		
	Time	Time	Time	Time	Time	Time	Time	Time	Time	Time	Time	
	+ 2003	+ 2004	+ 2005	+ 2003	+ 2004	+ 2005	+ 2003	+ 2004	+ 2005	+ 2003	+ 2004	+ 2005
+ Classic Cars	115.011	199.372	97.574	691.273	1.015.790	384.538	120.696	42.071	18.835	587.428	581.043	237.791
+ Motorcycles	60.789	63.159	65.870	141.836	204.042	161.260	16.485	31.959	4.176	178.109	291.421	55.020
+ Planes	42.663	67.681	11.082	154.519	209.128	128.008	60.556	49.177		90.016	202.942	60.985
+ Ships		35.323	3.070	172.428	186.992	67.845	14.156	10.453	8.407	58.238	142.904	48.856
+ Trains	1.681	8.226		29.538	90.973	17.995	13.279		3.524	28.304	25.551	15.398
+ Trucks and Buses	11.298	80.634	53.735	228.699	185.421	86.859	44.498	13.349		135.936	252.572	61.281
+ Vintage Cars	111.639	147.212	105.688	263.695	504.062	83.324	22.888	21.470	7.979	281.727	324.815	191.727

Slicer: [Measure=Sales]



Cubes

SteelWheelsSales

Dimensions

▼ Markets

- All Markets
- Territory
- Country
- State Province
- City

▼ Customers

- All Customers
- Customer

▶ Product

▶ Time

▼ Order Status

- All Order Status

Measures

▼ Measures

- Quantity
- Sales



Columns

Territory

Rows

Years

Sales

Line

Filter

Years	MeasuresLevel	Line	APAC	EMEA	Japan	NA
2003	Sales	Classic Cars	115,011	691,273	120,696	587,428
		Motorcycles	60,789	141,836	16,485	178,109
		Planes	42,663	154,519	60,556	90,016
		Ships		172,428	14,156	58,238
		Trains	1,681	29,538	13,279	28,304
		Trucks and Buses	11,298	228,699	44,498	135,936
		Vintage Cars	111,639	263,695	22,888	281,727
2004	Sales	Classic Cars	199,372	1,015,790	42,071	581,043
		Motorcycles	63,159	204,042	31,959	291,421
		Planes	67,681	209,128	49,177	202,942
		Ships	35,323	186,992	10,453	142,904
		Trains	8,226	90,973		25,551



Pentaho Reporting Designer

Reporting

- Consulta totalmente Web.
- **Amplo suporte de fonte de dados, incluindo relacionais, OLAP, XML, Objetos Java.**
- Opções de saída Popularer (PDF, HTML, Microsoft Excel, Rich Text Format, ou texto simples)
- Implantação flexível de relatórios incorporados ao BI.
- Design de relatórios gerador rapidamente com o assistente de relatório
- Publique diretamente para o servidor de BI.



Pentaho Report Designer

Reporting

- Adicionar visualizações rica de dados com mais de 15 tipos de gráficos personalizados
- Internacionalização, apoio multi-lingua(arquivo único)
- Embed HTML e JavaScript para controles dinâmicos e interativos relatórios on-line.
- Filtrar relatórios usando o modo built-in de visualização interativa.

Gerencial



Reporting

TopN Least Available

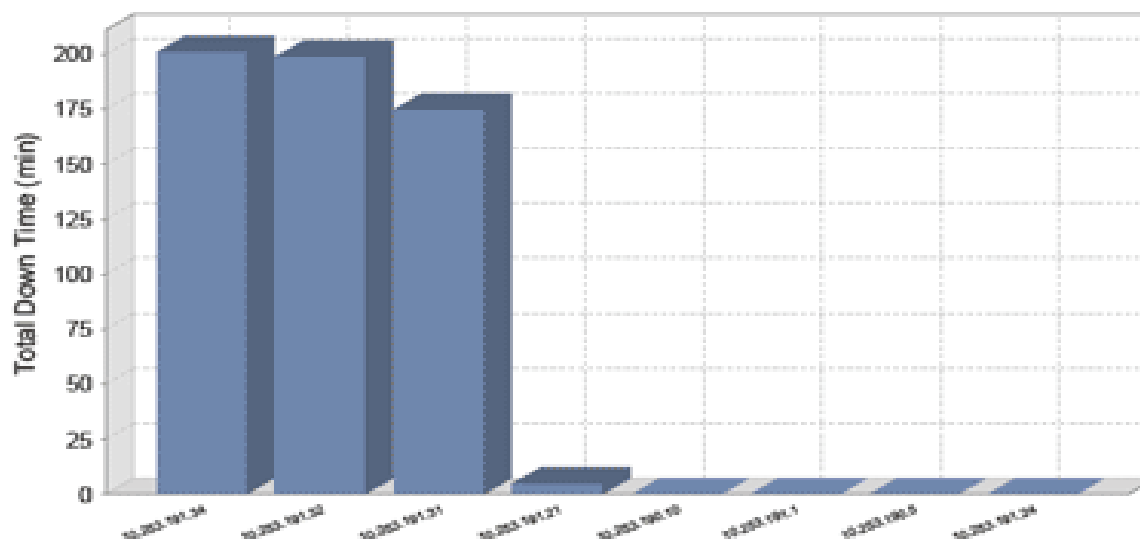
TopN least available managed assets

Summary

Report Period: Feb 01, 2006 12:00 AM to Feb 10, 2006 5:46 PM

Number of Elements Included: 10
 Availability Target: 99.000%
 Types of Outages Contributing to Downtime: Unplanned
 Business Day Policy: Full Day including Weekends

Availability Graph



Report Details

Name	IP Address	Type	Total Downtime	No. Of Outages	Availability %	
10.253.191.34	10.253.191.34	Windows Host	03h 20m 50s	11	98.568%	Red
10.253.191.32	10.253.191.32	Windows Host	03h 18m 39s	12	98.584%	Red
10.253.191.31	10.253.191.31	Windows Host	02h 55m 05s	15	98.752%	Red
10.253.191.21	10.253.191.21	Windows Host	00h 05m 00s	1	99.964%	Yellow
10.253.190.10	10.253.190.10	RS-3000	00h 00m 00s	0	100.000%	Green
10.253.191.1	10.253.191.1	Cisco/WSX5302	00h 00m 00s	0	100.000%	Green
10.253.190.5	10.253.190.5	RS-3000	00h 00m 00s	0	100.000%	Green
10.253.191.24	10.253.191.24	Windows Host	00h 00m 00s	0	100.000%	Green

Operacional



Reporting

Steel Wheels

Steel Wheels
500 International Speedway, Daytona, FL 32114
(123) 455-7890 <http://www.callsteel.com>
Run Date: 2/2/06 1:24 PM

TO: Reims Collectables
59 rue de l'Abbaye, null
Reims, null 51100 France

INVOICE

Attn: Paul Henriot
Sales Rep: 1337
Terms: Net 30 days

Invoice #: 10121
Account Number: 353
Date: May 07, 2003

SKU	Product Description	Price/Unit	Qty Ordered	Total Price
S50_4713	2002 Yamaha YZR M1	\$74.85	44	\$3,293.40
S24_2360	1982 Ducati 900 Monster	\$76.88	32	\$2,460.16
S32_4485	1974 Ducati 350 M93 Desmo	\$86.74	25	\$2,168.50
S12_2823	2002 Suzuki XREO	\$165.68	50	\$8,284.00
S10_1678	1969 Harley Davidson Ultimate Chopper	\$81.35	34	\$2,765.90
				<u>\$18,971.96</u>

Send Payment and Remittance Slip to:

Steel Wheels
500 International Speedway
Daytona, FL 32114

Thank you for your business!

Account Number: 353

Reims Collectables

Page 1 / 5

REMITTANCE

Invoice #: 10121
Account Number: 353

Reims Collectables
59 rue de l'Abbaye, null
Reims, null 51100 France

AMOUNT ENCLOSED: _____

Produção



Reporting

Steel Wheels

Steel Wheels, Inc.
Human Resources Actual vs Forecast
Period ending June 30, 2005

Region: Central

Department	Position	Actual	Budget	Variance
Executive Management				
	SVP Partnerships	\$387,415	\$392,100	\$24,685
	SVP WW Operations	\$476,000	\$725,887	\$249,887
	SVP Strategic Development	\$383,242	\$403,405	\$20,163
	CEO	\$549,625	\$522,250	-\$27,375
	Total	\$1,776,282	\$2,043,642	\$267,360

Department	Position	Actual	Budget	Variance
Finance				
	Controller	\$570,373	\$577,070	\$6,697
	Payroll	\$367,415	\$432,100	\$64,685
	Administrative Assistant	\$827,861	\$760,990	-\$66,871
	IS	\$570,759	\$577,346	\$6,587
	CFO	\$770,272	\$719,855	-\$50,417
	Total	\$3,106,680	\$3,067,361	-\$39,319

Department	Position	Actual	Budget	Variance
Human Resource				
	Sexual Harassment	\$530,473	\$538,570	\$8,097
	EOE	\$530,207	\$538,380	\$8,173
	HR Generalists	\$856,190	\$771,225	-\$84,965
	HR Training	\$397,473	\$443,570	\$46,097
	Administration	\$549,625	\$552,250	\$2,625
	SVP HR	\$574,895	\$570,300	-\$4,595
	Total	\$3,438,863	\$3,414,295	-\$24,568

Department	Position	Actual	Budget	Variance
Marketing & Communication				
	Graphics	\$782,375	\$728,500	-\$53,875
	Writer	\$405,985	\$459,850	\$53,865
	Analyst Relations	\$383,375	\$443,500	\$60,125
	Press Relations	\$497,296	\$524,872	\$27,576
	CMO	\$827,861	\$760,990	-\$66,871
	Product Marketing Mgr	\$693,531	\$665,040	-\$28,491



Inventory List [X]

Line: Classic Cars | Motorcycles | Planes | Ships | Trains | Tr






Show Bar Section: TRUE

Autoupdate on selection Update

Steel Wheels

Steel Wheels, Inc - Buying Department
Inventory List
June 30, 2005

Classic Cars, Planes, Ships, Trains

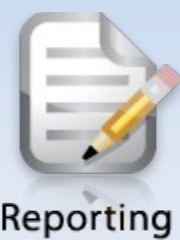
LINE: Classic Cars							
Vendor	BAR CODE	SKU	Name	Scale	On Hand	Cost	MSRP
Autoart Studio Design		512_3099	1968 Ford Mustang	1:12	87 units	\$ 95	\$ 195
Description: float, doors and trunk all open to reveal highly detailed interior features. Steering wheel actually turns the front wheels. Color dark green.							
8,107							
Camozel DieCast Legends		524_3528	1966 Shelby Cobra 427 S/C	1:24	8,287 units	\$ 29	\$ 30
Description: This diecast model of the 1966 Shelby Cobra 427 S/C includes many authentic details and operating parts. The 1:24 scale model of this iconic lightweight sports car from the 1960s comes in silver and it's own display case.							
8,107							
		524_2840	1958 Chevy Corvette Limited Edition	1:24	2,542 units	\$ 35	\$ 35
Description: The operating parts of this 1958 Chevy Corvette Limited Edition are particularly delicate due to their precise scale and require special care and attention. Features: rotating wheels, working steering, opening doors and trunk. Color dark green.							
2,542							
		5700_2104	1982 Camaro Z28	1:58	5,324 units	\$ 47	\$ 91
Description: Features include opening and closing doors. Color: White. Measures approximately 9 1/2" long.							
6,034							
Classic Metal Creations		530_3949	1952 Alpine Renault 1300	1:30	7,305 units	\$ 99	\$ 214
Description: Turnable front wheels; steering function; detailed interior; detailed engine; opening hood; opening trunk; opening doors; and detailed chassis.							
7,305							

Structure Data

- Page Header
- Report Header
- Groups
 - Group Header
 - Details Body
 - Details Header
 - Details
 - vertical-line
 - horizontal-line
 - simple-barcode: PRODUCTCODE
 - number-field: QUANTITYINSTOCK
 - rectangle
 - rectangle
 - message-field: Description: \$(PR
 - number-field: MSRP
 - number-field: BUYPRICE
 - number-field: QUANTITYINSTOCK
 - string-field: PRODUCTSCALE
 - string-field: PRODUCTNAME
 - string-field: PRODUCTCODE
 - string-field: PRODUCTVENDOR

Style Attributes

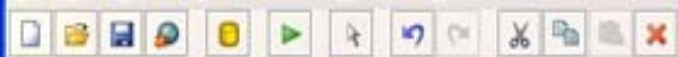
Name	Value	Formula
common		
type	number-field	
field	MSRP	
value		
name		
format	\$ #,###	
if-null		
query-meta...		
data-format	<input type="checkbox"/>	
style-format	<input type="checkbox"/>	
wizard		
aggregation-gr...		
aggregation-type		
labels-detail-h...		
show-changes		
html	<input type="checkbox"/>	
class		



Reporting



Ambia



Edit Chart

Bar Line Chart

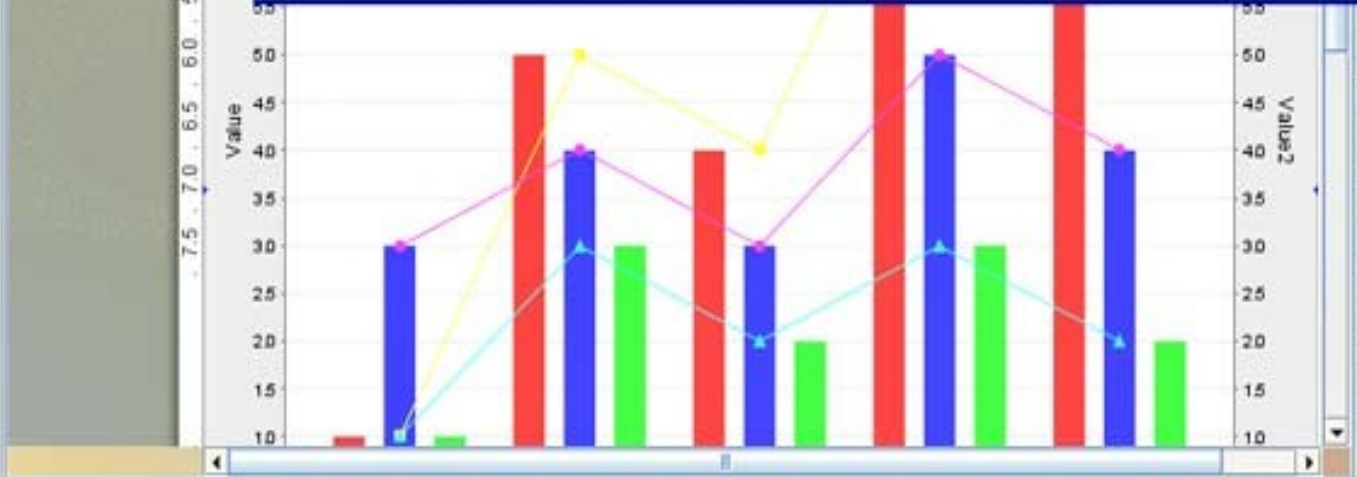
Name	Value
Required	
name	
bar-data-source	Data Collector 0
line-data-source	Data Collector 1
no-data-message	CHART.USER_NO_DATA_
Title	
chart-title	
chart-title-field	
title-font	SansSerif-BOLD-14
Bar Options	
stacked	False
stacked-percent	False
horizontal	False
series-color	

Primary DataSource **Secondary DataSource**

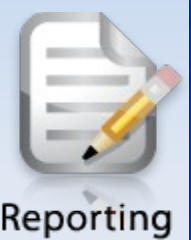
CategorySet Data Collector

Name	Value
Common	
name	
category-column	YEAR_ID
value-columns	[SALES]
Series	
series-by-value	
series-by-field	[PRODUCTLINE]
Group	
group-by	
reset-group	
Advanced	
dependencyLevel	0

OK Cancel



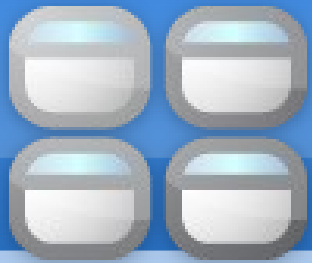
Field	Value	Formula
type	legacy-chart	
field		
value		
name		
if-null		
query-meta...		
data-format	<input type="checkbox"/>	
style-format	<input type="checkbox"/>	
encoding		
encoding-quality		
encoding-type		
wizard		
labels-detail-h...		
show-changes	<input type="checkbox"/>	
html		
class		



Reporting



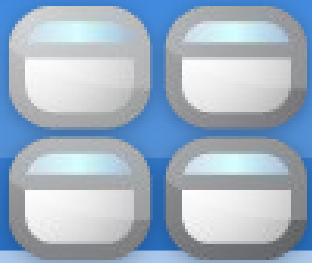
Amb



Pentaho Dashboards

Dashboards

- oferece métricas em uma atrativa interface intuitiva e visual, proporcionam informações crítica de que necessitam para entender e melhorar o desempenho organizacional.
- Self-Service Designer dashboard que permite aos usuários criar facilmente painéis de negócio personalizados.



Pentaho Dashboards

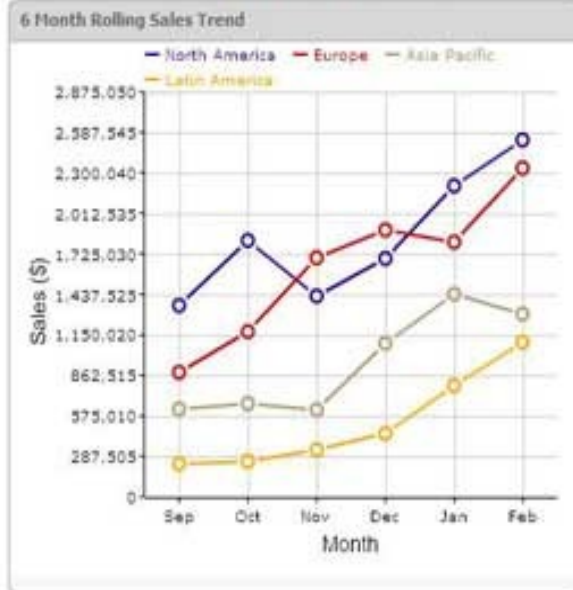
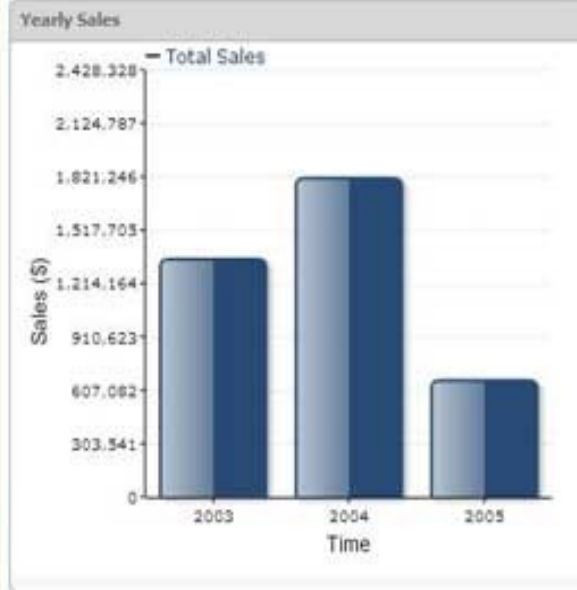
Dashboards

- Integração com o Pentaho Reporting and Analysis Pentaho.
- Integrado para monitorar continuamente a alertar para os problemas e notificar os usuários a tomar medidas.



Executive Sales Da...

Executive Sales Dashboard



Supplier Name	Location	Sales (\$)
Vintax Art Galleries	London	\$353,315.35
Classic Metal Creations	New York	\$343,896.11
Second Gear Diecast	London	\$338,931.95
Exotic Designs	Beijing	\$328,610.02
Gearbox Collectibles	Los Angeles	\$303,849.73
Motor City Art Classics	Detroit	\$301,302.50
Min Lin Diecast	New York	\$293,025.40
Red Start Diecast	Chicago	\$283,500.57
Highway 66 Mini Classics	Phoenix	\$281,965.58
Vinty Diecast Productions	Liverpool	\$279,000.44
Autobart Studio Design	Boston	\$277,587.22
Carousel DieCast Legends	Seattle	\$268,694.55
Studio M Art Models	New York	\$207,721.99
Speedway Sports	Nashville	\$190,234.34
Autocast	Boston	\$180,234.24
Distinct Designs	Orlando	\$135,309.02
Vroom Vehicle	Dallas	\$108,898.44
Average Supplier Sales		\$85,984.43

Product Sales Analysis



Measures	Time	Markets											
		APAC				EMEA				NA			
		Classic Cars	Motorcycles	Planes	Ships	Classic Cars	Motorcycles	Planes	Ships	Classic Cars	Motorcycles	Planes	Ships
Sales	2003	115,011	60,789	42,663		691,273	141,836	154,519	172,428	587,428	178,109	90,016	58,238
	2004	199,372	63,159	67,681	35,323	1,015,790	204,042	209,128	186,992	581,043	291,421	202,942	142,904
	QTR1	16,995	28,888	31,951	6,139	255,628	44,732		49,413	76,199		25,808	32,255
	QTR2	66,000			2,829	190,241	12,776	26,533	6,840	27,954	117,579	73,806	46,811
	QTR3	51,784				202,903	67,419	90,042	32,810	207,325	60,531	19,072	41,517
	QTR4	64,593	34,271	35,730	26,355	367,018	79,114	92,553	97,929	269,564	113,310	84,255	22,321
	2005	97,574	65,870	11,082	3,070	384,538	161,260	128,008	67,845	237,791	55,020	60,985	48,856

Slicer:



Dashboards



Pentaho User Console - Dashboard - Mozilla Firefox

http://localhost:8080/pentaho/Home

File View Tools Help

pentaho

Dashboard

Product Line Performance

Yearly Sales

Sales (\$)

Product Lines

Product Line	Year 1	Year 2	Year 3
Ships	~1,400,000	~1,200,000	~800,000
Planes	~900,000	~700,000	~400,000
Motorcycles	~600,000	~500,000	~200,000
Trucks and Buses	~700,000	~500,000	~200,000
Vintage Cars	~600,000	~500,000	~200,000
Classic Cars	~400,000	~300,000	~100,000

Product Line Mix

Product Line	Value	Percentage
Vintage Cars	790,268.69	21%
Classic Cars	1,408,261.42	37%
Trucks and Buses	353,943	9%
Motorcycles	240,997.15	6%
Trains	60,253.56	2%
Ships	524,540.10	14%
Planes	440,788.4	12%

Sales Analysis

Measures	Time	Markets				Product	
		APAC		EMEA		Product	
		Classic Cars	Motorcycles	Planes	Ships	Classic Cars	Motorcycles
Sales	2003	115,011	60,789	42,663		691,273	141,6
	2004	199,372	63,159	67,681	35,323	1,015,790	204,0
	QTR1	16,995	28,888	31,951	6,139	255,628	44,7
	QTR2	66,000			2,829	190,241	12,7
	QTR3	51,784				202,903	67,4
	QTR4	64,593	34,271	35,730	26,355	367,018	79,1
	2005	97,574	65,870	11,082	3,070	384,538	161,2

Slicer:



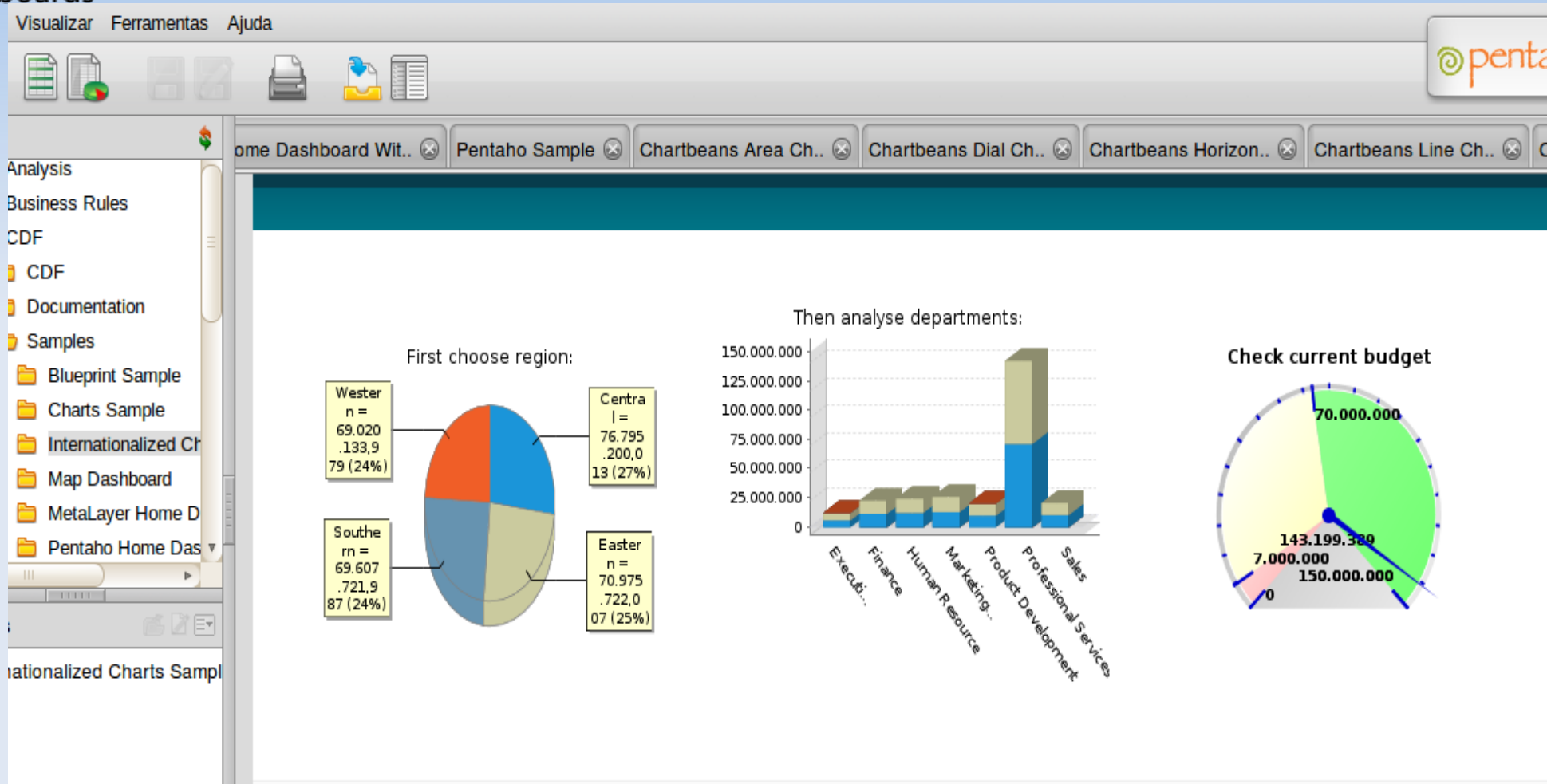
Dashboards





Indicadores

Dashboards



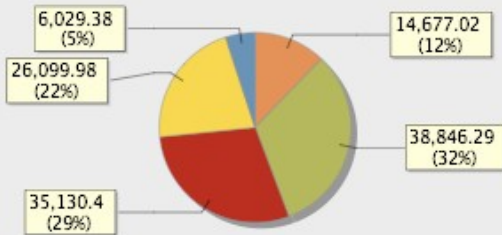
Pentaho Google Maps Dashboard

Select Sales Thresholds

View: **West Coast** | East Coast

📍 < 60000 < 📍 < 130000 < 📍

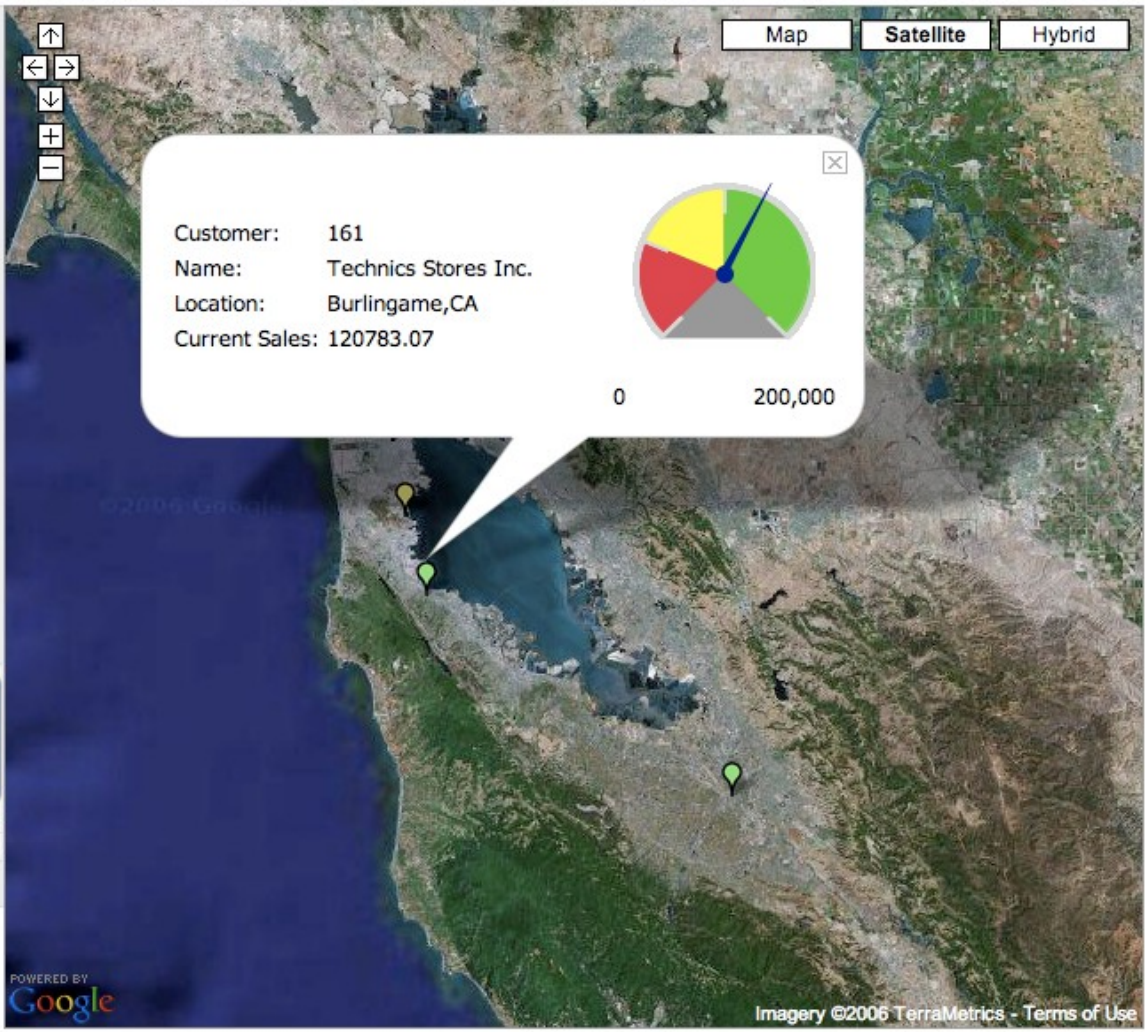
Customer Product Mix



- Classic Cars
- Motorcycles
- Planes
- Trucks and Buses
- Vintage Cars

Sales History

Date	Amount
05 Jan 2005	\$13,530
02 Nov 2004	\$2,916
28 Oct 2003	\$62,305
24 Jul 2003	\$42,032
Total	\$120,783

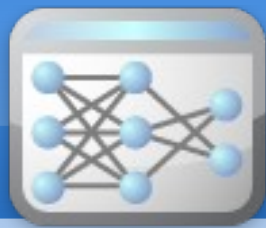


Dashboards

Done

- Pentaho User Console
- <http://localhost:8080>

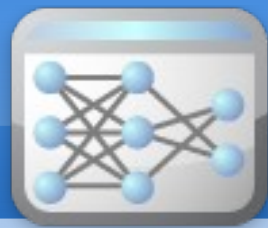
Demonstração



Pentaho Data Mining

Data Mining

- É um nível de excelência de BI.
- Fornece informações sobre padrões e relacionamentos ocultos nos seus dados.
- Permite-lhe explorar correlações para melhorar o desempenho organizacional
- Fornece indicadores de desempenho futuro.
- Permite a incorporação de recomendações em suas aplicações
- Permite-lhe tirar o máximo partido de um conjunto de algoritmos de mineração de dados



P

Data Mining

Weka Explorer

Preprocess **Classify** Cluster Associate Select attributes Visualize

Classifier: Choose **Logistic -R 1.0E-8 -M -1**

Test options:

- Use training set
- Supplied test set (Set...)
- Cross-validation Folds:
- Percentage split %:

 More options...

(Nom) class:

Start Stop

Result list (right-click for options)

Classifier output:

```

num_dependents
own_telephone
foreign_worker

Time taken to build:
=== Stratified cross
=== Summary ===

Correctly Classified
Incorrectly Classified
Kappa statistic
  
```

Weka Classifier Visualize: ThresholdCurve. (Class value good)

X: False Positive Rate (Num) Y: True Positive Rate (Num)

Colour: Threshold (Num) Select Instance

Reset Clear Open Save Jitter

Plot (Area under ROC = 0.7849)

Class colour:

Weka classification boundary visualizer

Dataset: iris

Class Attribute: class (Nom)

Class color: Iris-setosa Iris-versicolor Iris-virginica

Classifier: Choose **IBk -K 5 -W 0 -A "weka.core.neighboursearch.LinearNNS**

Visualization Attributes:

- X: sepalength (Num)
- Y: sepalwidth (Num)

Add / remove data points:

- Add points Iris-setosa
- Remove points Remove all

 Open a new window

Sampling control:

- 2 Base for sampling (r)
- 1 Num. locations per pixel
- 3 Kernel bandwidth (k)

Plotting:

- Plot training data

89.1751 %

1000

Precision	Recall	F-Measure	ROC Area	Class
0.798	0.864	0.83	0.785	good
0.607	0.49	0.542	0.785	bad
0.741	0.752	0.744	0.785	

Log



Data Mining

Weka Explorer

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize

Open file... | Open URL... | Open DB... | Generate... | Undo | Edit... | Save...

Filter: Choose **None** Apply

Current relation: Relation: german_credit | Instances: 1000 | Attributes: 21 | Sum of weights: 1000

Selected attribute: Name: credit_history | Missing: 0 (0%) | Distinct: 5 | Type: Nominal | Unique: 0 (0%)

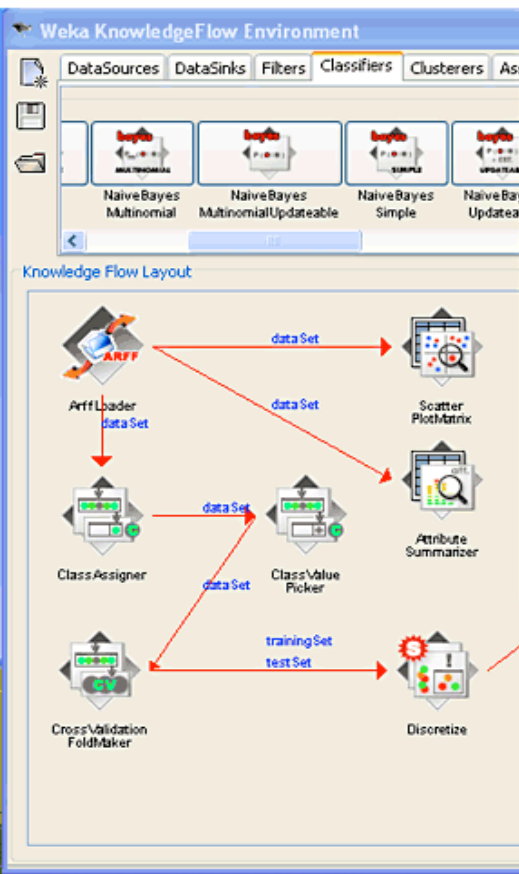
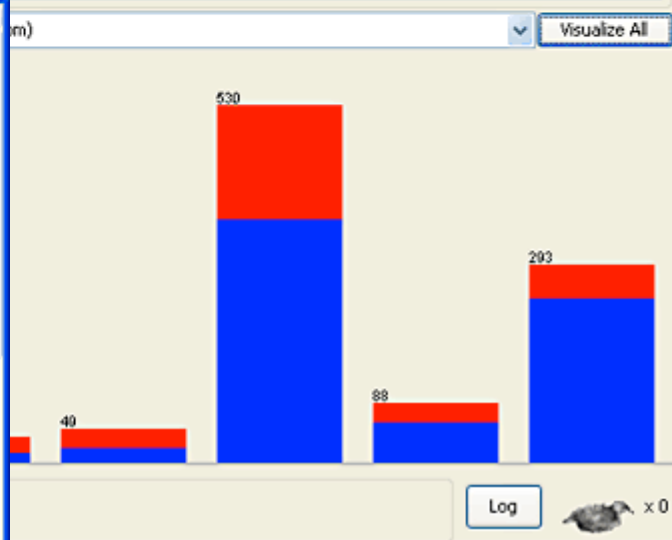
No.	Label	Count	Weight
1	no credits/all paid	40	40.0
2	all paid	49	49.0
3	existing paid	530	530.0
4	delayed previously	88	88.0
5	critical/other existing credit	293	293.0

Attributes: All | None | Invert | Pattern

No.	Name
1	<input type="checkbox"/> checking_status
2	<input type="checkbox"/> duration
3	<input checked="" type="checkbox"/> credit_history
4	<input type="checkbox"/> purpose
5	<input type="checkbox"/> credit_amount
6	<input type="checkbox"/> savings_status
7	<input type="checkbox"/> employment

Visualize

Visualize All



Liferay - Pentaho

http://localhost:8080/user/joebloggs/3?p_p_id=RegionFilterPortlet_WAR_pentaho&p_p_action=1& Google

LIFERAY
Enterprise. Open Source. For Life.

Welcome Joe Bloggs!

Home Plugins Pentaho Add Page

Filters

Select filters to apply to other controls on this page

Select a Region
 Central Eastern Southern Western

Select one or more Departments
 Choose...
 Executive Management
 Finance
 Human Resource
 Marketing & Communication

Select one or more Position Titles
 Choose...
 CEO
 CMO
 IS
 EOE

Year

Update...

Actual Headcount - % Variance from Budget

CEO SVP Partnerships SVP Strategic Development SVP WW Operations

Headcount Costs

Headcount Data

Position	Actual	Budget	Variance
CEO	\$549,625	\$522,250	-\$27,375
Total	\$549,625	\$522,250	-\$27,375

Integração com Portal CMS

Versões Pentaho

- **Community Edition**

Contém todos os principais recursos sem certificação. comunidade de usuários ativa.

- **Enterprise Edition**

Oferece algumas funcionalidades adicionais que facilitam diversos processos e documentação especializada, assim como contato direto com suporte da Pentaho.

http://www.pentaho.com/products/enterprise/enterprise_comparison.php

Passo a passo para aprender Pentaho BI

- Conceitual.
- Pentaho Business Intelligence Server
- Pentaho Data Integration – Kettle
- Bases OLAP no MySQL e Ferramentas de Apoio
- Pentaho Analysis – Mondrian
- Pentaho Report
- Pentaho MetaData Editor.
- Pentaho Report Studio.
- Pentaho Data Mining - Weka.

Prêmios do Pentaho Project



Referências e links Interessantes!

- Projeto Pentaho <http://www.pentaho.com/>
- CE <http://community.pentaho.com/>
- www.ambientelivre.com.br (Suporte) Tutoriais Livres.
- Lista de Discussões:
pentahobr@yahoogrupos.com.br

Nossos clientes Pentaho

- Atualmente atendemos á empresas que usam o Pentaho como solução de Negócios nas seguintes áreas:
- Transporte, farmácias , Comercial
- Recursos Humanos
- Tecnologia
- Conhecimento Industrial
- Saúde ocupacional
- Gestão Empresarial (ERP)
- Serviços Financeiros (Banco)
- Inteligência de Negócio - Web

Usando o Pentaho



SOLUÇÕES



Obrigado

- Contatos:
marcio @ ambientelivre.com.br
rafaela @ ambientelivre.com.br

- <http://twitter.com/ambientelivre>
@ambientelivre
@RafaelaRaganham



- Blog no Software Livre
<http://softwarelivre.org/marciojv>