

BigData

VANTAGENS

- Interpretar grandes quantidades de dados
- Analisar dados não estruturados
- Interpretar tendências de eventos

DESVANTAGENS

- Privacidade dos dados
- Segurança dos dados
- Discriminação dos dados



A complex, abstract pattern of blue lines and glowing nodes resembling a circuit board or a network of data flow. The lines are various shades of blue, from light cyan to dark navy, and the glowing nodes are small, bright blue circles. The pattern is dense and organic, filling the left side of the slide.

05

ETL

Samuel Reine Alves Ribeiro

E - EXTRACT

ETL é o nome de um processo de gerenciamento de dados, dividido em três partes: extração, transformação e carregamento.

T - TRANSFORM

Esse processo é muito comumente utilizado para a organização de [Data Warehouses](#), principalmente quando o fluxo de dados a ser gerenciado é muito grande. O processo de ETL ajuda não somente na organização de dados, mas também na otimização do funcionamento de um sistema.

L - LOAD

Data Warehouse



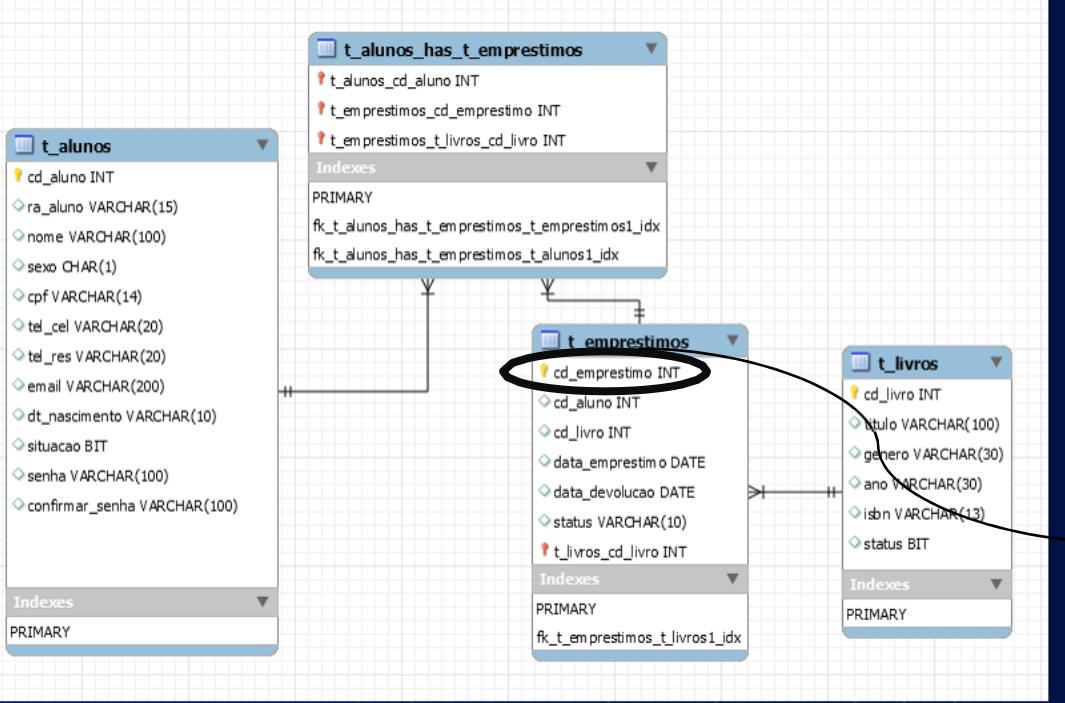
Empresa

Banco de Dados

Os bancos de dados armazenam de maneira eficiente dados relacionados a uma empresa ou sistema: histórico de vendas de um colaborador, dados de um cliente (e-mail, telefone, endereço, CPF, etc), dados de fornecedores.

Os bancos de dados também possuem um sistema para acesso às suas informações armazenadas: SGDB – Sistema de Gerenciamento do Banco de dados.

Data Warehouse



Os sistemas gerenciais dos banco de dados geralmente trabalham com uma tabela de dados. Cada item da tabela de dados pode conter uma planilha, ou uma lista, ou uma própria tabela, em relação a si mesmo.

	id	dt	id_produto	id_vendedor	id_cliente	qt
1	1	2014-06-01 13:31:39.383	3	10	1	6
2	2	2008-07-18 13:31:39.383	7	8	1	9
3	3	2010-07-06 13:31:39.383	5	10	7	14
4	4	2015-05-10 13:31:39.383	7	2	10	12
5	5	2016-11-09 13:31:39.383	3	1	2	9
6	6	2010-10-27 13:31:39.383	4	2	7	13
7	7	2016-03-02 13:31:39.383	6	2	3	4
8	8	2012-10-07 13:31:39.383	5	4	10	5
9	9	2007-12-06 13:31:39.383	2	11	7	10
10	10	2015-03-05 13:31:39.383	3	7	9	13
11	11	2009-09-21 13:31:39.383	7	7	4	6
12	12	2008-02-06 13:31:39.383	3	7	3	11

Data Warehouse

Com o crescimento das empresas, as tabelas de dados começam a se tornar cada vez mais complexas e as listas armazenadas nos bancos de dados não são mais suficientes: precisamos de relatórios a partir dessas listas. Precisamos analisar os dados armazenados no nosso banco de dados. Isso requer sempre PODER DE PROCESSAMENTO.

id	dt	id_produto	id_vendedor	id_cliente	qt
1	2014-06-01 13:31:39.383	10	1	6	
2	2008-07-18 13:31:39.383	7	8	1	9
3	2010-07-06 13:31:39.383	9	10	7	14
4	2015-05-10 13:31:39.383	7	2	10	12
5	2016-11-09 13:31:39.383	3	1	2	9
6	2010-10-27 13:31:39.383	4	2	7	13
7	2016-03-02 13:31:39.383	6	2	3	4
8	2012-10-07 13:31:39.383	5	4	10	5
9	2007-12-06 13:31:39.383	2	11	7	10
10	2015-03-05 13:31:39.383	3	7	9	13
11	2009-09-21 13:31:39.383	7	7	4	6
12	2008-02-06 13:31:39.383	3	7	3	11

Tabela de vendas

Relatório para o RH						
Mes_Reférencia	Colaborador	Cargo	Total_Vendas	Salario	Comissões	Salario_do_mes
Jan/2017	Vitor Festeiro	DBA	R\$ 1.862.400,00	R\$ 30.000,00	R\$ 27.936,00	R\$ 57.936,00
Jan/2017	Marcelino Pinga	CEO	R\$ 914.900,00	R\$ 40.000,00	R\$ 13.723,50	R\$ 53.723,50
Jan/2017	Saturnino Besta	Coordenador de IT	R\$ 1.049.500,00	R\$ 30.000,00	R\$ 15.742,50	R\$ 45.742,50
Jan/2017	Rolando Rocha	Analista de IT	R\$ 1.707.400,00	R\$ 20.000,00	R\$ 25.611,00	R\$ 45.611,00
Jan/2017	Salvador das Dores Fortes	Gerente de IT	R\$ 600.500,00	R\$ 30.000,00	R\$ 9.007,50	R\$ 39.007,50
Jan/2017	Email Suarez	Gerente de Vendas	R\$ 999.500,00	R\$ 20.000,00	R\$ 14.992,50	R\$ 34.992,50
Jan/2017	Vespertino Cedo	Gerente Administrativo	R\$ 1.239.900,00	R\$ 15.000,00	R\$ 18.598,50	R\$ 33.598,50

Relatório para o Comercial

Mes_Reférencia	Produto	Meta	Vendas	% meta atingida
Jan/2017	CPU Nasa 10Ghz	R\$ 1.000.000,00	R\$ 1.034.000,00	103,40 %
Jan/2017	HD SSD 1TB	R\$ 1.000.000,00	R\$ 1.012.000,00	101,20 %
Jan/2017	Monitor 32 Pol Full HD	R\$ 500.000,00	R\$ 537.000,00	107,40 %
Jan/2017	Mouse Gammer 25 botões	R\$ 100.000,00	R\$ 98.800,00	98,80 %
Jan/2017	Notebook Gamer Pro	R\$ 12.000.000,00	R\$ 13.025.000,00	108,54 %
Jan/2017	RAM DDR4 4GB Powerturbo	R\$ 300.000,00	R\$ 206.000,00	68,67 %
Jan/2017	Teclado Gammer 350 teclas	R\$ 100.000,00	R\$ 175.500,00	175,50 %

Data Warehouse

Se a empresa cresce muito, o processamento do banco de dados pode começar a ser muito requerido e, com o volume cada vez maior de dados a ser armazenado, gerenciado, processado e analisado, o sistema da empresa pode acabar por ficar sobrecarregado e lento.

Como o banco de dados fica dentro do servidor da empresa, melhorar o poder de processamento do servidor poderia resolver o problema da sobrecarga do sistema. Porém, essa solução pode ser muito onerosa.



Empresa

Sobrecarregado



Banco de Dados

Data Warehouse & ETL

Qual a solução para o problema de armazenamento, gerenciamento e análise de dados com o crescimento de uma empresa?

Uma das soluções mais comuns é chamada de CRESCIMENTO VERTICAL COM DATA WAREHOUSE E ETL.

Consiste em manter o banco de dados originais, com as listas e tabelas anteriores, e ao mesmo tempo criar um novo banco de dados, chamado de Data WareHouse no qual serão carregados os dados tratados (informações) provenientes do banco de dados original.

Banco de produção - Relacional

The diagram illustrates the ETL process. It shows three tables from the original database: 'Vendas' (Sales), 'Cooperadores' (Partners), and 'Produtos' (Products). These tables are connected by arrows pointing towards a large arrow labeled 'ETL' on its side, which points to a new database on the right.

Vendas																																																																														
<table border="1"><thead><tr><th>id</th><th>d</th><th>id_produto</th><th>id_vendedor</th><th>id_cliente</th><th>qt</th></tr></thead><tbody><tr><td>1</td><td>2016-08-01 03:51:29,363</td><td>3</td><td>10</td><td>1</td><td>1</td></tr><tr><td>2</td><td>2016-07-18 03:51:29,363</td><td>7</td><td>8</td><td>1</td><td>3</td></tr><tr><td>3</td><td>2016-07-16 13:23:28,363</td><td>7</td><td>8</td><td>2</td><td>14</td></tr><tr><td>4</td><td>2016-07-18 13:23:29,363</td><td>7</td><td>9</td><td>10</td><td>12</td></tr><tr><td>5</td><td>2016-07-29 13:23:29,363</td><td>3</td><td>1</td><td>10</td><td>9</td></tr><tr><td>6</td><td>2016-07-27 13:23:29,363</td><td>4</td><td>2</td><td>7</td><td>13</td></tr><tr><td>7</td><td>2016-07-02 13:23:29,363</td><td>6</td><td>2</td><td>3</td><td>4</td></tr><tr><td>8</td><td>2012-04-07 13:23:29,363</td><td>5</td><td>4</td><td>13</td><td>5</td></tr><tr><td>9</td><td>2007-10-26 13:23:29,363</td><td>2</td><td>11</td><td>7</td><td>10</td></tr><tr><td>10</td><td>2015-05-15 13:23:29,363</td><td>3</td><td>7</td><td>9</td><td>13</td></tr><tr><td>11</td><td>2009-09-21 13:23:29,363</td><td>7</td><td>7</td><td>4</td><td>6</td></tr><tr><td>12</td><td>2009-09-09 13:23:29,363</td><td>3</td><td>7</td><td>3</td><td>11</td></tr></tbody></table>	id	d	id_produto	id_vendedor	id_cliente	qt	1	2016-08-01 03:51:29,363	3	10	1	1	2	2016-07-18 03:51:29,363	7	8	1	3	3	2016-07-16 13:23:28,363	7	8	2	14	4	2016-07-18 13:23:29,363	7	9	10	12	5	2016-07-29 13:23:29,363	3	1	10	9	6	2016-07-27 13:23:29,363	4	2	7	13	7	2016-07-02 13:23:29,363	6	2	3	4	8	2012-04-07 13:23:29,363	5	4	13	5	9	2007-10-26 13:23:29,363	2	11	7	10	10	2015-05-15 13:23:29,363	3	7	9	13	11	2009-09-21 13:23:29,363	7	7	4	6	12	2009-09-09 13:23:29,363	3	7	3	11
id	d	id_produto	id_vendedor	id_cliente	qt																																																																									
1	2016-08-01 03:51:29,363	3	10	1	1																																																																									
2	2016-07-18 03:51:29,363	7	8	1	3																																																																									
3	2016-07-16 13:23:28,363	7	8	2	14																																																																									
4	2016-07-18 13:23:29,363	7	9	10	12																																																																									
5	2016-07-29 13:23:29,363	3	1	10	9																																																																									
6	2016-07-27 13:23:29,363	4	2	7	13																																																																									
7	2016-07-02 13:23:29,363	6	2	3	4																																																																									
8	2012-04-07 13:23:29,363	5	4	13	5																																																																									
9	2007-10-26 13:23:29,363	2	11	7	10																																																																									
10	2015-05-15 13:23:29,363	3	7	9	13																																																																									
11	2009-09-21 13:23:29,363	7	7	4	6																																																																									
12	2009-09-09 13:23:29,363	3	7	3	11																																																																									
Cooperadores	<table border="1"><thead><tr><th>id</th><th>nm</th><th>ds_cargo</th><th>v_efetivo</th><th>pc_comissoes</th></tr></thead><tbody><tr><td>1</td><td>Marcelino Pingo</td><td>CEO</td><td>40000,00</td><td>1,00</td></tr><tr><td>2</td><td>Everaldo Sauer</td><td>Gerente de Vendas</td><td>30000,00</td><td>1,50</td></tr><tr><td>3</td><td>Feliciano Triste</td><td>Assistente de Vendas</td><td>7000,00</td><td>1,50</td></tr><tr><td>4</td><td>Edsona Damascos</td><td>Assistente de Vendas</td><td>8000,00</td><td>1,50</td></tr><tr><td>5</td><td>Averbauskae Nunes</td><td>Assistente de Vendas</td><td>8000,00</td><td>1,50</td></tr><tr><td>6</td><td>Heitorinho Cebó</td><td>Chefe Administrativo</td><td>15000,00</td><td>1,00</td></tr><tr><td>7</td><td>Adriano Almeida</td><td>Chefe</td><td>7000,00</td><td>1,00</td></tr><tr><td>8</td><td>Flávio da Costa</td><td>Assistente Administrativo</td><td>7000,00</td><td>1,00</td></tr><tr><td>9</td><td>João Coelho</td><td>Assistente Técnico</td><td>10000,00</td><td>1,00</td></tr><tr><td>10</td><td>Salvador das Flores Farias</td><td>Gerente de IT</td><td>30000,00</td><td>1,00</td></tr><tr><td>11</td><td>Vitor Fazenda</td><td>DBA</td><td>30000,00</td><td>1,00</td></tr><tr><td>12</td><td>Saturnino Beata</td><td>Coordenador de IT</td><td>30000,00</td><td>1,00</td></tr></tbody></table>	id	nm	ds_cargo	v_efetivo	pc_comissoes	1	Marcelino Pingo	CEO	40000,00	1,00	2	Everaldo Sauer	Gerente de Vendas	30000,00	1,50	3	Feliciano Triste	Assistente de Vendas	7000,00	1,50	4	Edsona Damascos	Assistente de Vendas	8000,00	1,50	5	Averbauskae Nunes	Assistente de Vendas	8000,00	1,50	6	Heitorinho Cebó	Chefe Administrativo	15000,00	1,00	7	Adriano Almeida	Chefe	7000,00	1,00	8	Flávio da Costa	Assistente Administrativo	7000,00	1,00	9	João Coelho	Assistente Técnico	10000,00	1,00	10	Salvador das Flores Farias	Gerente de IT	30000,00	1,00	11	Vitor Fazenda	DBA	30000,00	1,00	12	Saturnino Beata	Coordenador de IT	30000,00	1,00												
id	nm	ds_cargo	v_efetivo	pc_comissoes																																																																										
1	Marcelino Pingo	CEO	40000,00	1,00																																																																										
2	Everaldo Sauer	Gerente de Vendas	30000,00	1,50																																																																										
3	Feliciano Triste	Assistente de Vendas	7000,00	1,50																																																																										
4	Edsona Damascos	Assistente de Vendas	8000,00	1,50																																																																										
5	Averbauskae Nunes	Assistente de Vendas	8000,00	1,50																																																																										
6	Heitorinho Cebó	Chefe Administrativo	15000,00	1,00																																																																										
7	Adriano Almeida	Chefe	7000,00	1,00																																																																										
8	Flávio da Costa	Assistente Administrativo	7000,00	1,00																																																																										
9	João Coelho	Assistente Técnico	10000,00	1,00																																																																										
10	Salvador das Flores Farias	Gerente de IT	30000,00	1,00																																																																										
11	Vitor Fazenda	DBA	30000,00	1,00																																																																										
12	Saturnino Beata	Coordenador de IT	30000,00	1,00																																																																										
Produtos	<table border="1"><thead><tr><th>id</th><th>nm</th><th>vl</th><th>id_Marca</th><th>meta</th><th>quatu</th></tr></thead><tbody><tr><td>1</td><td>Mouse Gammer 25 botões</td><td>200,00</td><td>1</td><td>100000,00</td><td>100,00</td></tr><tr><td>2</td><td>Teclado Gammer 350 teclas</td><td>300,00</td><td>1</td><td>200000,00</td><td>195,00</td></tr><tr><td>3</td><td>Monitor 32 Pol Full HD</td><td>1000,00</td><td>1</td><td>500000,00</td><td>650,00</td></tr><tr><td>4</td><td>RAM DDR4 4GB PowerTurbo</td><td>600,00</td><td>1</td><td>100000,00</td><td>625,00</td></tr><tr><td>5</td><td>CPU Nasa 100x</td><td>2000,00</td><td>1</td><td>100000,00</td><td>1000,00</td></tr><tr><td>6</td><td>HD SSD 1TB</td><td>2000,00</td><td>1</td><td>100000,00</td><td>1000,00</td></tr><tr><td>7</td><td>Notebook Gamer Pro</td><td>2900,00</td><td>1</td><td>120000,00</td><td>1025,00</td></tr></tbody></table>	id	nm	vl	id_Marca	meta	quatu	1	Mouse Gammer 25 botões	200,00	1	100000,00	100,00	2	Teclado Gammer 350 teclas	300,00	1	200000,00	195,00	3	Monitor 32 Pol Full HD	1000,00	1	500000,00	650,00	4	RAM DDR4 4GB PowerTurbo	600,00	1	100000,00	625,00	5	CPU Nasa 100x	2000,00	1	100000,00	1000,00	6	HD SSD 1TB	2000,00	1	100000,00	1000,00	7	Notebook Gamer Pro	2900,00	1	120000,00	1025,00																													
id	nm	vl	id_Marca	meta	quatu																																																																									
1	Mouse Gammer 25 botões	200,00	1	100000,00	100,00																																																																									
2	Teclado Gammer 350 teclas	300,00	1	200000,00	195,00																																																																									
3	Monitor 32 Pol Full HD	1000,00	1	500000,00	650,00																																																																									
4	RAM DDR4 4GB PowerTurbo	600,00	1	100000,00	625,00																																																																									
5	CPU Nasa 100x	2000,00	1	100000,00	1000,00																																																																									
6	HD SSD 1TB	2000,00	1	100000,00	1000,00																																																																									
7	Notebook Gamer Pro	2900,00	1	120000,00	1025,00																																																																									

Banco de dados original

ETL

Novo Banco de dados,
com informações
tratadas – Data
WareHouse

Banco datawarehouse

folha_pagamentos

Mes_Reférencia	Colaborador	Cargo	Total_Vendas	Salario	Comissões	Salario_liquido
Jan/2017	Wor Fettino	DBA	R\$ 1.932.400,00	R\$ 30.000,00	R\$ 27.936,00	R\$ 57.936,00
Jan/2017	Marcelino Pingo	CEO	R\$ 914.900,00	R\$ 40.000,00	R\$ 13.723,50	R\$ 53.723,50
Jan/2017	Saturnino Beata	Coordenador de IT	R\$ 1.649.500,00	R\$ 15.742,50	R\$ 45.124,50	R\$ 50.877,50
Jan/2017	Ricardo Rocha	Analista de IT	R\$ 1.797.000,00	R\$ 20.000,00	R\$ 25.200,00	R\$ 45.200,00
Jan/2017	Salvador das Flores Farias	Gerente de IT	R\$ 400.000,00	R\$ 60.000,00	R\$ 8.075,50	R\$ 59.024,50
Jan/2017	Everaldo Sauer	Gerente de Vendas	R\$ 999.500,00	R\$ 25.000,00	R\$ 14.982,50	R\$ 34.982,50
Jan/2017	Heitorinho Cebó	Gerente Administrativo	R\$ 1.239.000,00	R\$ 15.000,00	R\$ 18.598,50	R\$ 33.598,50

acompanhamento_metas

Mes_Reférencia	Produto	Meta	Vendas	% meta atingida
Jan/2017	CPU Nasa 100x	R\$ 1.000.000,00	R\$ 1.034.000,00	103,40 %
Jan/2017	HD SSD 1TB	R\$ 1.000.000,00	R\$ 1.012.000,00	101,20 %
Jan/2017	Monitor 32 Pol Full HD	R\$ 500.000,00	R\$ 537.000,00	107,40 %
Jan/2017	Mouse Gammer 25 botões	R\$ 100.000,00	R\$ 98.000,00	98,80 %
Jan/2017	Notebook Gamer Pro	R\$ 12.000.000,00	R\$ 13.025.000,00	109,54 %
Jan/2017	RAM DDR4 4GB PowerTurbo	R\$ 300.000,00	R\$ 206.000,00	68,67 %
Jan/2017	Tecleido Gammer 350 teclas	R\$ 100.000,00	R\$ 175.500,00	175,50 %

Data Warehouse & ETL

O ETL então vai extrair os dados do banco de produção de dados original (Extract), transformar nos relatórios que eu preciso, analisa-los (Transform) e carregar para o novo banco de dados, chamado Data WareHouse as informações já prontas para consulta (Load).

Banco de produção - Relacional					
Vendas					
1	id	id_produto	id_vendedor	id_cliente	dt
2	2014-06-01 13:31:29.363	3	10	1	2014-06-01 13:31:29.363
3	2016-07-06 13:31:29.363	7	8	1	2016-07-06 13:31:29.363
4	2016-09-18 13:31:29.363	5	10	7	2016-09-18 13:31:29.363
5	2016-11-29 13:31:29.363	3	2	13	2016-11-29 13:31:29.363
6	2016-12-19 13:31:29.363	4	2	7	2016-12-19 13:31:29.363
7	2017-01-02 13:31:29.363	6	2	3	2017-01-02 13:31:29.363
8	2017-01-27 13:31:29.363	5	11	13	2017-01-27 13:31:29.363
9	2017-01-29 13:31:29.363	2	11	7	2017-01-29 13:31:29.363
10	2017-01-30 13:31:29.363	3	7	8	2017-01-30 13:31:29.363
11	2017-01-31 13:31:29.363	7	7	4	2017-01-31 13:31:29.363
12	2018-02-09 13:31:29.363	3	7	3	2018-02-09 13:31:29.363

Cooperadores					
1	id	nm	ds_cargo	vl_salaro	pc_comissao
2	NULL	Administrador Pega	CEO	40000.00	1.00
3	0	Enival Sucess	Gerente de Vendas	20000.00	1.00
4	1	Feliciano Triste	Assistente de Vendas	7000.00	1.50
5	2	Emilene Gomes	Assistente de Vendas	8000.00	1.00
6	3	Averaunes das Neves	Assistente de Vendas	9000.00	1.00
7	4	Vesperino Cedo	Gerente Administrativo	15000.00	1.00
8	5	Frederico da Silva	Cash	7000.00	1.50
9	6	Henrique da Luz	Assistente Administrativo	7000.00	1.50
10	7	José Coucha	Assistente Técnico	10000.00	1.00
11	8	Sabedoria dos Corais Fofos	Gerente de IT	30000.00	1.00
12	9	Voor Fattans	DBA	20000.00	1.00
13	10	Saturiano Beta	Coordenador de IT	30000.00	1.00

Produtos					
1	nm	vl	ic_ativo	meta	data
2	Mouse Gamer 25-botões	200,00	1	1000000,00	130/90
3	Teclado Gamer 250 teclas	300,00	1	2000000,00	195/90
4	Monitor 32 Pol Full HD	1000,00	1	5000000,00	650/50
5	RAM DDR4 4GB PowerTurbo	500,00	1	3000000,00	325/90
6	CPU Nossa 10GHz	2000,00	1	10000000,00	1300/00
7	HD SSD 1TB	2000,00	1	10000000,00	1300/00
8	Notebook Gamer Pro	25000,00	1	120000000,00	16250/00

Banco de dados original



Novo Banco de dados, com informações tratadas – Data WareHouse

Banco datawarehouse

folha_pagamentos

Mes_Reférencia	Colaborador	Cargo	Total_Vendas	Salario	Comissao	Salario_Ato_Vend
Jan/2017	Vitor Festoso	DBA	R\$ 1.962.400,00	R\$ 30.000,00	R\$ 27.336,00	R\$ 57.936,00
Jan/2017	Marcelino Pingo	CEO	R\$ 914.900,00	R\$ 40.000,00	R\$ 13.723,50	R\$ 53.723,50
Jan/2017	Sáurimo Beta	Coordenador de IT	R\$ 1.045.500,00	R\$ 30.000,00	R\$ 15.742,50	R\$ 45.742,50
Jan/2017	Roland Rocha	Analista de IT	R\$ 1.707.400,00	R\$ 20.000,00	R\$ 25.611,00	R\$ 45.611,00
Jan/2017	Selvador das Dores Fofas	Gerente de IT	R\$ 600.000,00	R\$ 10.000,00	R\$ 15.007,50	R\$ 39.007,50
Jan/2017	Enival Sucess	Gerente de Vendas	R\$ 959.500,00	R\$ 20.000,00	R\$ 14.952,50	R\$ 34.952,50
Jan/2017	Vesperino Cedo	Gerente Administrativo	R\$ 1.239.900,00	R\$ 15.000,00	R\$ 18.358,50	R\$ 33.358,50

acompanhamento_metas

Mes_Reférencia	Produto	Meta	Vendas	% meta atingida
Jan/2017	CPU Nossa 10GHz	R\$ 1.000.000,00	R\$ 1.034.000,00	103,40 %
Jan/2017	HD SSD 1TB	R\$ 1.000.000,00	R\$ 1.012.000,00	101,20 %
Jan/2017	Monitor 32 Pol Full HD	R\$ 500.000,00	R\$ 537.000,00	107,40 %
Jan/2017	Mouse Gamer 25-botões	R\$ 100.000,00	R\$ 98.000,00	98,00 %
Jan/2017	Notebook Garner Pro	R\$ 12.000.000,00	R\$ 13.025.000,00	108,54 %
Jan/2017	RAM DDR4 4GB PowerTurbo	R\$ 300.000,00	R\$ 206.000,00	68,67 %
Jan/2017	Teclado Garner 350 teclas	R\$ 100.000,00	R\$ 175.500,00	175,50 %

Data Warehouse

Vantagens

- Alta Performance.
- Fácil de se consultar e analisar os dados, já que no DW as informações já são tratadas.
- Reaproveitamento de processamento: uma vez transformados, os dados já ficam prontos para consulta e análise.

Desvantagem

- Alto custo gerencial para tratar e transformar os dados. Pode sair caro para pequenas empresas.

A complex, abstract pattern of blue lines and glowing nodes resembling a circuit board or neural network. The lines are various shades of blue, and small glowing circles are placed at the intersections and endpoints of the lines.

05

Data Mining

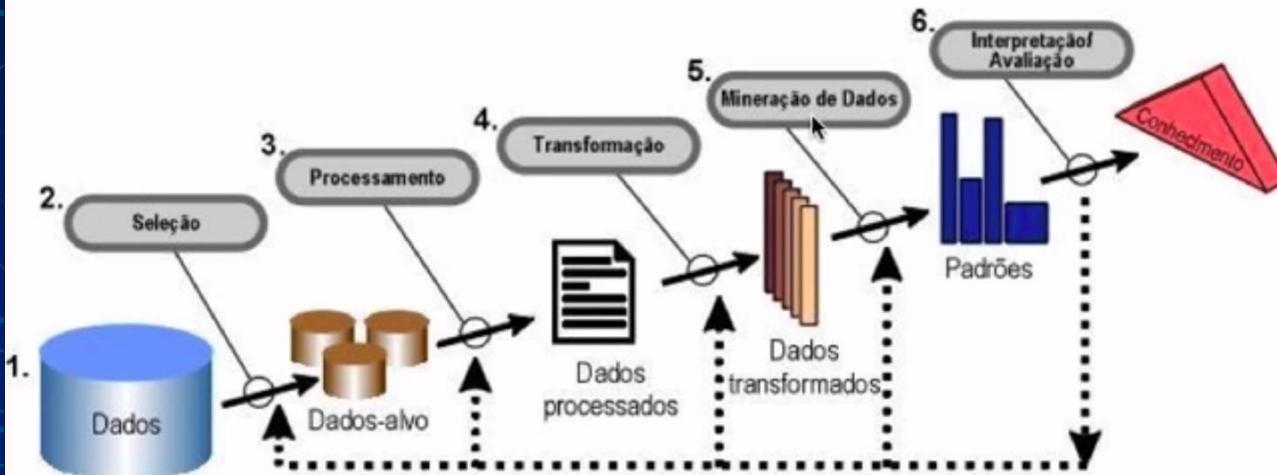
Breno Palma Miele Aniceto

O que é data mining?

Data mining é formada por um conjunto de ferramentas e técnicas que através do uso de algoritmos de aprendizagem ou classificação baseados em redes neurais e estatística. Estes são capazes de explorar um conjunto de dados, extraíndo ou ajudando a evidenciar padrões nestes dados e auxiliando na descoberta de conhecimento.



Etapas do Processo



Para que serve?

Buscamos com a MD (Mineração de Dados) conhecer melhor os clientes, seus padrões de consumo e motivações. Data Mining resgata em organizações grandes o papel do dono atendendo no balcão e conhecendo sua clientela.



Veja algumas aplicações possíveis do data mining em diversos segmentos:

- Identificação do perfil dos possíveis compradores de determinado produto, para captação de clientes;
- Decisão sobre alocação de produtos nas prateleiras de acordo com o perfil de consumo dos clientes;
- Identificação de atividades suspeitas e prevenção a fraudes;
- Análise de competências dos candidatos a uma vaga;
- Previsão de volume de pacientes em determinada categoria.

Algumas ferramentas:

- Redes neurais
- Indução de regras
- Árvores de decisão
- Visualização



A complex, abstract pattern of blue lines and glowing nodes, resembling a circuit board or a network of data flow, serves as the background for the slide. It features various line thicknesses and glowing circular endpoints.

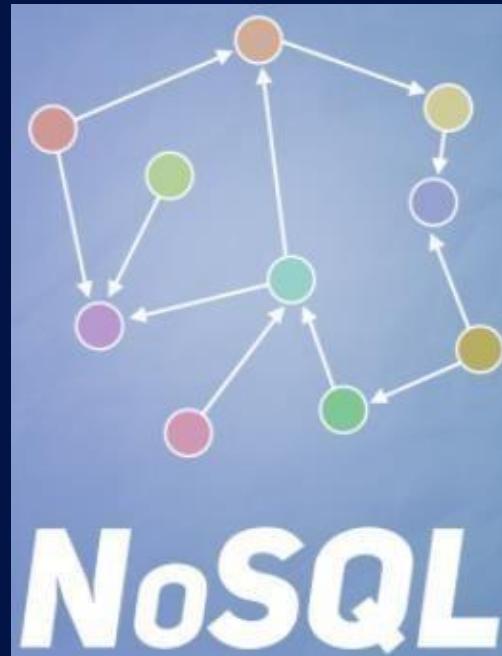
06

NoSQL

Carlos Livius da Silva

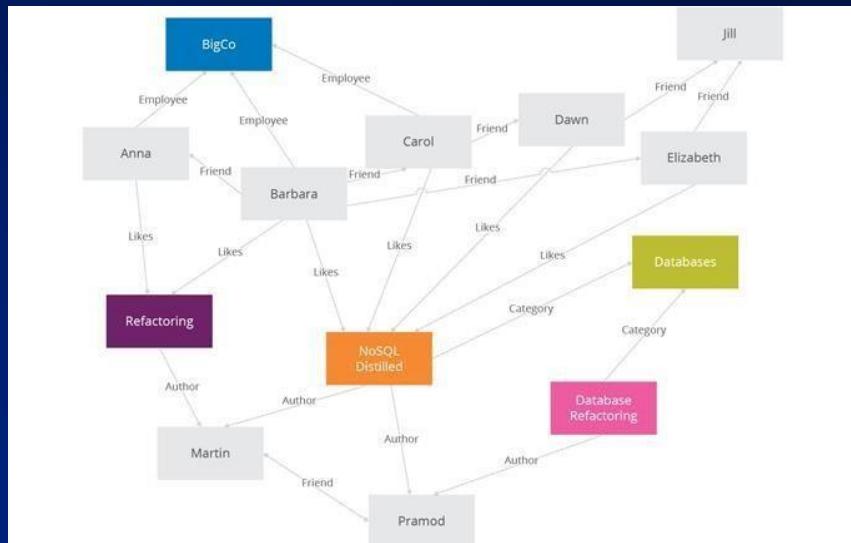
O que é?

- NoSQL - Banco de Dados Não Relacional
- Para que serve?
- Como surgiu?



Quais seus tipos e funcionamentos?

Bancos
de dados gráficos:

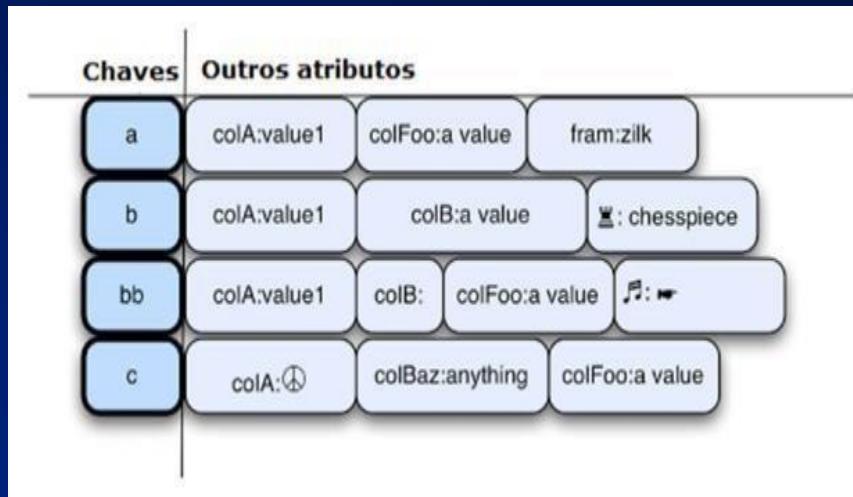


Banco de dados
Documentos:

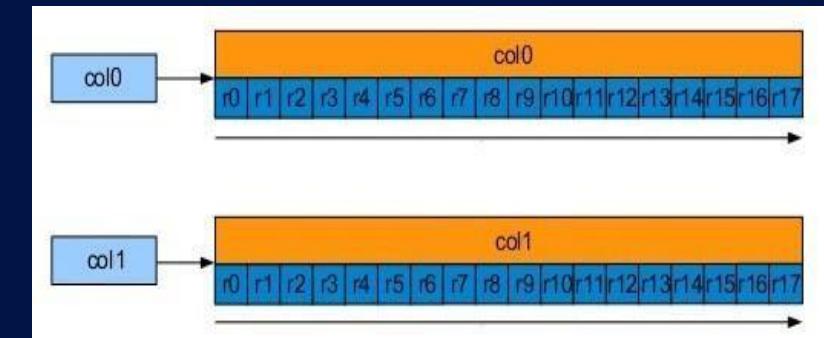
```
{  
  "id": 55,  
  "Pais": "Brasil",  
  "Regiao": "América do Sul",  
  "Populacao": 201032714,  
  "PrincipaisCidades": [  
    {  
      "NomeCidade": "São Paulo",  
      "Populacao": 1182876,  
    },  
    {  
      "NomeCidade": "Rio de Janeiro",  
      "Populacao": 6323037,  
    }  
  ]  
}
```

Quais seus tipos e funcionamentos?

Bancos de dados
Chave-Valor:



Banco de dados
Coluna:



SQL vs NoSQL

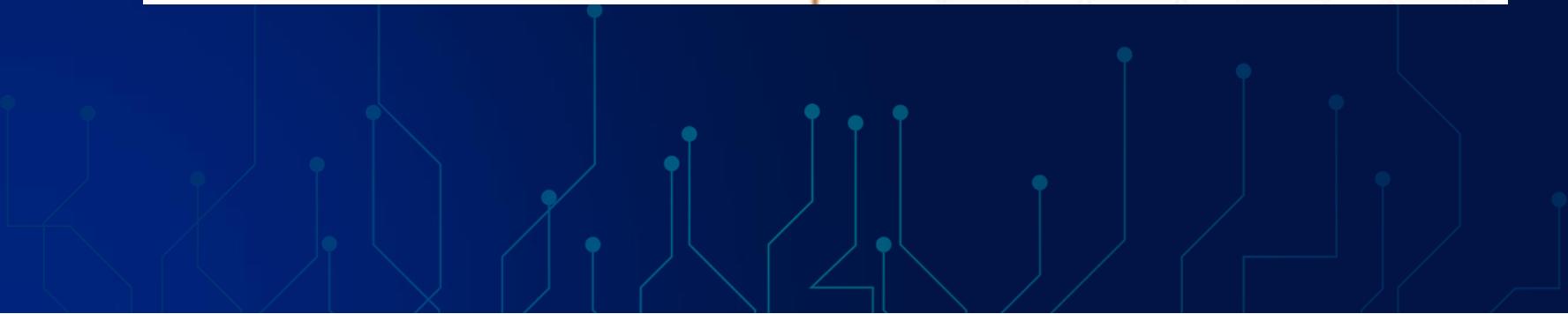
SQL

Table: Users

	id	name	createdAt	updatedAt
1.	1	Ashton	2018-08-12 1...	2018-08-12 1...
2.	2	Jakeem	2018-08-12 1...	2018-08-12 1...
3.	3	Uma	2018-08-12 1...	2018-08-12 1...

NoSQL →

Users Collection



Vantagens

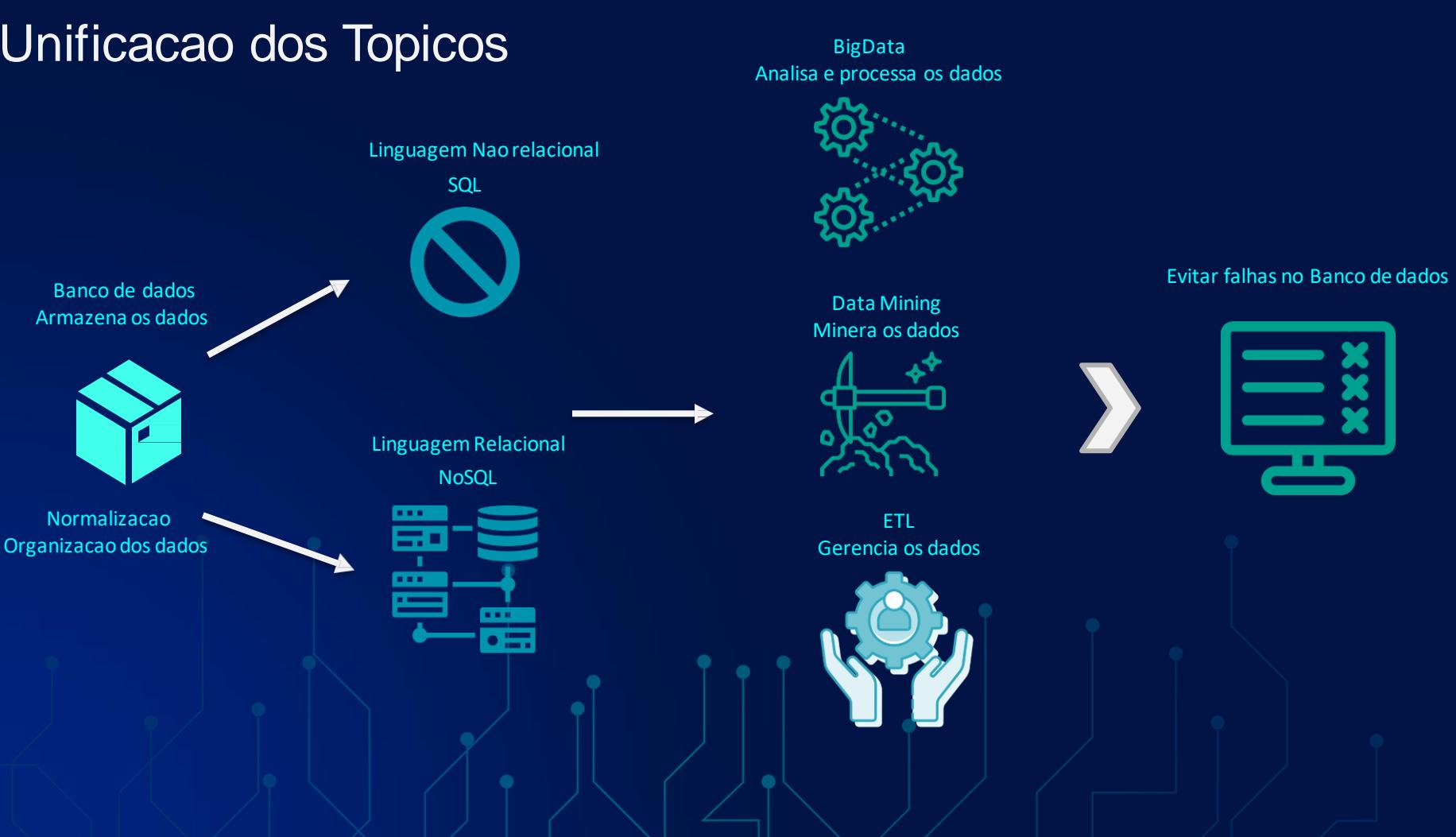
- Escalabilidade
- Altamente funcional
- Alta performance
- Flexibilidade

Desvantagens

- Recente
- Pouco reconhecido
- Suporte técnico
- Difícil instalação



Unificacao dos Topicos



Referências

- ALVES, Gustavo Furtado de Oliveira. **O que é um banco de dados?**. DICAS DE PROGRAMAÇÃO. Disponível em: <https://dicasdeprogramacao.com.br/o-que-e-um-banco-de-dados/>. Acesso em 14 set 2020
- FIA. **Big Data: o que é, como aplicar, a importância e exemplos.** Disponível em: <<https://fia.com.br/blog/big-data/amp/>>. Acesso em: 13 set. 2020
- ORACLE. **O Que é Big Data?.** Disponível em: <<https://www.oracle.com/br/big-data/what-is-big-data.html>>. Acesso em: 07 set. 2020
- NASCIMENTO, R. **Afinal, o que é Big Data?.** Disponível em: <<http://marketingpordados.com/analise-de-dados/o-que-e-big-data-%f0%9f%a4%96/>>. Acesso: 10 set. 2020
- PATEL, N. **Big Data: O Que é, Ferramentas e Como Aplicar na Sua Empresa.** Disponível em: <<https://neilpatel.com/br/blog/big-data-o-que-e/>>. Acesso em: 07 set 2020
- Josué. "O que é DATA WAREHOUSE? Você precisa de um?". Youtube. 26 de set. de 2017. <<https://www.youtube.com/watch?v=Q81zwSmajo0>>. Acessado em 11 de set. de 2020
- MACORATTI, José Carlos. MACORATTI.NET. **SQL – Consultas T-SQL básicas.** Disponível em: <http://www.macoratti.net/14/05/sql_tsql3.htm>. Acesso: em 12 set 2020
- DEVMEDIA. **Guia completo de SQL.** Disponível em <<https://www.devmedia.com.br/guia/guia-completo-de-sql/38314>>. Acesso: em 11 set 2020

Referências

- MICROSOFT. **CREATE PROCEDURE (Transact - SQL)**. Disponível em <<https://docs.microsoft.com/pt-br/sql/t-sql/statements/create-procedure-transact-sql?view=sql-server-ver15>>. Acesso em: 11 set 2020
- Cetax. "DATA MINING: O QUE É, CONCEITO E DEFINIÇÃO". Blog Cetax. abr. 2016. Disponível em: <<https://www.cetax.com.br/blog/data-mining/>> . Acesso em 10 de set. de 2020
- AMAZON . **O que é NoSQL?**. Disponível em: <<https://aws.amazon.com/pt/nosql/>>. Acesso em: 15 set 2020
- GEEKHUNTER, Penélope Victória. **Banco de dados NoSQL: um manual prático e didático**. Disponível em: <<https://blog.geekhunter.com.br/banco-de-dados-nosql-um-manual-pratico-e-didatico/#O que são bancos de dados NoSQL>> Acesso em: 14 set 2020
- CIÊNCIAS E DADOS, Davi Matos. **Top 6 NoSQL Databases**. Disponível em: <<http://www.cienciaedados.com/top-6-nosql-databases>> Acesso em: 15 set 2020
- MONGODB, Lauren Schaefer. **What is NoSQL**. Disponível em: <<https://www.mongodb.com/nosql-explained>> Acesso em: 14 set 2020
- MARQUES FERNANDES, Henrique Marques Fernandes. **Banco de Dados Relacional(SQL) e Não Relacional(NoSQL) – O que são, para que servem e qual a diferença?**. Disponível em: <<https://marquesfernandes.com/tecnologia/banco-de-dados-relacional-sql-e-nao-relacional-nosql-o-que-sao-para-que-servem-e-qual-a-diferenca/>> Acesso em: 15 set 2020
- Sehn Körting. Thales. "How data mining works". Youtube. set. 2014. Disponível em: <<https://www.youtube.com/watch?v=W44q6qszdqY&t=125s>>