



Business Intelligence

DIMENSÕES DE ALTERAÇÃO LENTA

Business Intelligence - BI

Conteúdo:

- 1) Business Intelligence
- 2) Data Warehouse / Data Mart / Data Lake
- 3) Modelagem Dimensional
- 4) ETLs
- 5) Dimensões alteradas lentamente
- 6) Ferramentas OLAP

Business Intelligence - BI

Conteúdo:

5) Slowly Changing Dimensions:

- Introdução
- Conceitos
- Regras
- Características
- OLTP

1 - Introdução

DIMENSÕES QUE MUDAM LENTAMENTE

6 - Dimensões de alteração Lenta

Exemplos

- um Cliente que está cadastrado há 5 anos e que agora está casado, após cinco anos de cadastramento, só agora faremos uma mudança em seu cadastro, informando o seu casamento;
- Registros das vendas, por vendedores. O vendedor é transferido de um escritório regional para o outro. Como gravar essa alteração na nossa dimensão de vendas?

6 - Dimensões de alteração Lenta

Introdução

- Quando essas mudanças ocorrem, você pode optar por seguir uma dessas duas grandes opções:
 - Gravar o histórico de alterações.
 - Substituir os valores conforme necessário.
- Dimensões de Alteração Lenta Ou SCD (Slowly Changing Dimensions)
 - Ralph Kimball inicialmente considerava três estratégias para o tratamento da SCD;
 - Ao longo dos anos aprofundou-se as definições iniciais e incluiu novos tipos de tratamento para as SCD;

2 – Tipos de tratamento para as SCD

CONHECENDO CADA TIPO

6 - Dimensões de alteração Lenta

Tipos de tratamento para as SCD

- SCD Tipo 1: Substituir;
- SCD Tipo 2: Adicionar uma linha;
- SCD Tipo 3: Adicionar coluna;
- SCD Tipo 4: Tabela da história separada;
- SCD Tipo 6: Hybrid.

6 - Dimensões de alteração Lenta

Tipos de tratamento para as SCD

- Em alguns casos, será conveniente combinar várias técnicas;
- É importante notar que, embora existam diferentes maneiras de executar cada técnica é essencial dispor de chaves substitutas em tabelas de dimensão para aplicar a aplicar estas técnicas;
- Ao aplicar diferentes técnicas, há a necessidade de modificar a estrutura da tabela de dimensão;

6 - Dimensões de alteração Lenta

SCD Tipo 1: Substituir

- É o mais básico e simples de implementar, apesar de não salvar as mudanças históricas, não requer qualquer modelagem especial e não precisa de adicionar novos registros para a tabela;
- Substitui os dados antigos com os novos dados e, portanto, não controla os dados históricos em tudo;
- Este é o mais adequado para corrigir certos tipos de erros de dados, tais como a ortografia de um nome (Supondo que você não vai precisar saber como ele costumava ser incorreto no passado).

6 - Dimensões de alteração Lenta

SCD Tipo 1: Substituir

- O Cliente mora na cidade de Curitiba.

ID	Nome do Cliente	Cidade
1	Wilquison Silva	Curitiba

- Depois de algum tempo, o cliente muda-se para a cidade de Bagé.
- No Tipo 1, simplesmente há a substituição da informação.

ID	Nome do Cliente	Cidade
1	Wilquison Silva	Bagé

6 - Dimensões de alteração Lenta

SCD Tipo 2: Adicionar uma linha

- Esta estratégia adiciona algumas COLUNAS à tabela, para armazenar o histórico de alterações;
- Cria vários registros nas tabelas com chaves separadas;
- Preserva o histórico do registro, pois um novo registro é inserido cada vez que uma alteração é feita;

6 - Dimensões de alteração Lenta

SCD Tipo 2: Adicionar uma linha

- As colunas normalmente adicionadas são:

Coluna	Descrição
Data de inicio	Data em que o registro entrou em vigor. Por padrão usado frequentemente uma data muito precoce, por exemplo, "1000/01/01".
Data de inatividade	Data em que o registro atual não estava mais em vigor. Por padrão, a data é usada frequentemente muito futurista, como "01/01/9999".
Versão	Número seqüencial que aumenta a cada nova mudança. Por padrão, geralmente começa em "1".
Versão atual	Especifica se o campo atual é o atual. Este valor pode ser, se é verdade "true" ou "1" e se ela é falsa ". 0" "flase" ou

6 - Dimensões de alteração Lenta

SCD Tipo 2: Adicionar uma linha

- O Cliente mora na cidade de Curitiba.

ID	Nome do Cliente	Cidade
1	Wilquison Silva	Curitiba

- Depois de algum tempo, o cliente muda-se para a cidade de Bagé.
- No Tipo 2:

ID	Nome do Cliente	Cidade	Inicio	Final	versao
1	Wilquison Silva	Curitiba	01/01/2000	30/12/2020	1
2	Wilquison Silva	Bagé	31/12/2020	31/12/9999	2

6 - Dimensões de alteração Lenta

SCD Tipo 3: Adicionar Coluna

- Esta estratégia requer a adição de uma coluna para cada coluna cujos valores deseja-se acompanhar as mudanças;
- Armazena as alterações em uma tabela adicional feitas na tabela;

6 - Dimensões de alteração Lenta

SCD Tipo 3: Adicionar Coluna

- Exemplo: O Cliente morava em Curitiba e mudou-se para Bagé:

ID	Nome do Cliente	Cidade
1	Wilquison Silva	Curitiba

- Aplicando o Tipo 3

ID	Nome do Cliente	Cidade	Cidade_Anterior
1	Wilquison Silva	Bagé	Curitiba

6 - Dimensões de alteração Lenta

SCD Tipo 4: Tabela de histórico

- Neste método, uma tabela histórica separada é usada para rastrear todas as alterações históricas de atributo da dimensão para cada uma das dimensões;
- A tabela de dimensão 'principal' mantém apenas os dados atuais;
- É, habitualmente, conhecido apenas como o uso de “tabelas de histórico”, onde se mantém uma tabela os dados atuais, e uma tabela adicional é utilizado para manter um registro de algumas ou todas as alterações;

6 - Dimensões de alteração Lenta

SCD Tipo 4: Tabela de histórico

- Exemplo: O Cliente morava em Curitiba e mudou-se para Bagé:

ID	Nome do Cliente	Cidade
1	Wilquison Silva	Curitiba

ID	Nome do Cliente	Cidade
1	Wilquison Silva	Bagé

- Tabela Cliente_histórico:

ID	Nome do Cliente	Cidade	Data_Inicio	Data_Final
1	Wilquison Silva	Curitiba	01/01/2000	31/05/2020
1	Wilquison Silva	Bagé	01/06/2020	31/12/9999

6 - Dimensões de alteração Lenta

SCD Tipo 6: HÍBRIDO

- Combina as abordagens dos tipos 1, 2 e 3 ($1 + 2 + 3 = 6$);
- A abordagem é usar um SCD tipo 2, mais a adição de um par adicional de colunas de data para indicar o período de validade de uma linha específica e um sinalizador para indicar se o registro é o registro atual;
- Não é frequentemente utilizado porque tem o potencial de complicar o acesso do usuário final, mas tem algumas vantagens sobre as outras abordagens, especialmente quando técnicas são empregadas para mitigar a complexidade.

6 - Dimensões de alteração Lenta

SCD Tipo 6: HÍBRIDO

- Valor atual: valor atual do atributo. Todos os registros de histórico para determinado item de atributo têm o mesmo valor atual;
- Valor histórico: valor histórico do atributo. Todos os registros de histórico para determinado item de atributo podem ter valores diferentes;
- Data de início: data de início da 'data efetiva' do histórico do atributo;
- Data final: data de término da 'data efetiva' do histórico do atributo;
- Status atual: para manter informações sobre o registro mais recente.

6 - Dimensões de alteração Lenta

SCD Tipo 6: HÍBRIDO

- Exemplo:

ID	Nome do Cliente	Cidade atual	Cidade anterior	Data Inicio	Data Final	Status Atual
1	Wilquison Silva	Curitiba		01/01/2000	31/05/2010	N
1	Wilquison Silva	Videira	Curitiba	01/06/2010	15/10/2020	N
1	Wilquison Silva	Bagé	Videira	16/10/2020	31/12/9999	Y

Obrigado!