

BANCO DE DADOS

**Introdução a Views e conceitos
básicos sobre Data Warehouse e
Data Lake**

VIEW

- **Caminho alternativo** para visualização de dados de uma ou mais tabelas.
- **VIEWS** são muito utilizadas para informações calculadas ou derivadas como idade e preço total, ou então, para evitar a exposição de dados restritos (como salário, dados de auditoria entre outros).
- As **VIEWS** às vezes são chamadas de tabelas virtuais.

VIEW

- No contexto tão latente de Big Data e Análise de Dados, as **VIEWS** muitas vezes representam soluções extremamente valiosas.
- É muito comum que empresas tenham sistemas diferentes que compartilham os mesmos dados, nesse caso o acesso a dados por **VIEWS** é muito utilizado.
- As **VIEWS** agilizam processos de consulta e muitas vezes resguardam regras de negócios!

VIEW

- É possível operar comandos de definição de dados como INSERT, DELETE e UPDATE nas VIEWS, e essas operações “podem” ser replicadas para as tabelas originais.
- Geralmente VIEWS criadas a partir de tabelas únicas recebem operações e já replicam para as tabelas originais. Produtos de bancos de dados diferentes podem executar de forma diferente essas operações.
- O uso de VIEWS deve ser bem orquestrado dentro de um projeto de banco de dados.

VIEW

- Assim como são executados os comandos SELECTs em tabelas, podem também ser executados nas VIEWS.
- É possível executar consultas que combinem dados oriundos de tabelas e de views.

CRIAR UMA VIEW

- Uma view pode ser uma representação exata de uma única tabela.
- Uma view pode ser originada de um filtro de registros e colunas de uma única tabela.
- Também é possível que uma view seja constituída a partir de dados oriundos de diversas tabelas (por JOIN ou UNION) com filtros de registros ou colunas, de acordo com as necessidades do projeto.

CRIAR UMA VIEW

```
CREATE VIEW <nome da view>  
(<nome das colunas - opcional>) AS  
SELECT <nome das colunas>  
FROM <tabelas>  
WHERE <condição>
```

CRIAR UMA VIEW

```
CREATE VIEW vProduto AS  
SELECT idproduto, nome FROM produto  
WHERE valor<=10
```

	idproduto	nome
1	1	Cenoura
2	2	Iogurte Natural
3	3	Iogurte de Morango
4	5	Vassoura Azul
5	6	Batata Ondulada
6	8	Choco Bolotas

CRIAR UMA VIEW

```
CREATE VIEW vProduto2 (codigo, produto, total) AS  
SELECT idproduto, nome, qtde_estoque*valor  
FROM produto
```

	codigo	produto	total
1	1	Cenoura	807.00
2	2	Iogurte Natural	235.62
3	3	Iogurte de Morango	529.24
4	4	Queijo Prato	275.88
5	5	Vassoura Azul	706.82
6	6	Batata Ondulada	489.02
7	7	Limpa Tela	226.50
8	8	Choco Bolotas	3800.10

ELIMINAR UMA VIEW

DROP VIEW <nome da view>

DROP VIEW vProduto;

DROP VIEW vProduto2;

DATA WAREHOUSE e DATA LAKE

- Ao tratar de Banco de Dados é muito comum que temas como Data Warehouse e Data Lake sejam lembrados.
- Contexto atual de muitos projetos que envolvam BI (Business Intelligence) e Big Data favorecem criações de Data Warehouses e Data Lakes.
- Bancos de Dados relacionais (como vimos) podem ser parte da estrutura de Data Warehouses e Data Lakes.

DATA WAREHOUSE

- É uma estrutura de Banco de Dados para análises de dados e Business Intelligence (BI).
- Tem uma arquitetura constituída para atender processamento e armazenamento de grandes quantidades de dados.
- Os dados em um Data Warehouse geralmente são derivados de uma ampla variedade de fontes como bancos de dados de sistemas variados, planilhas, arquivos de texto, entre outros.

DATA WAREHOUSE

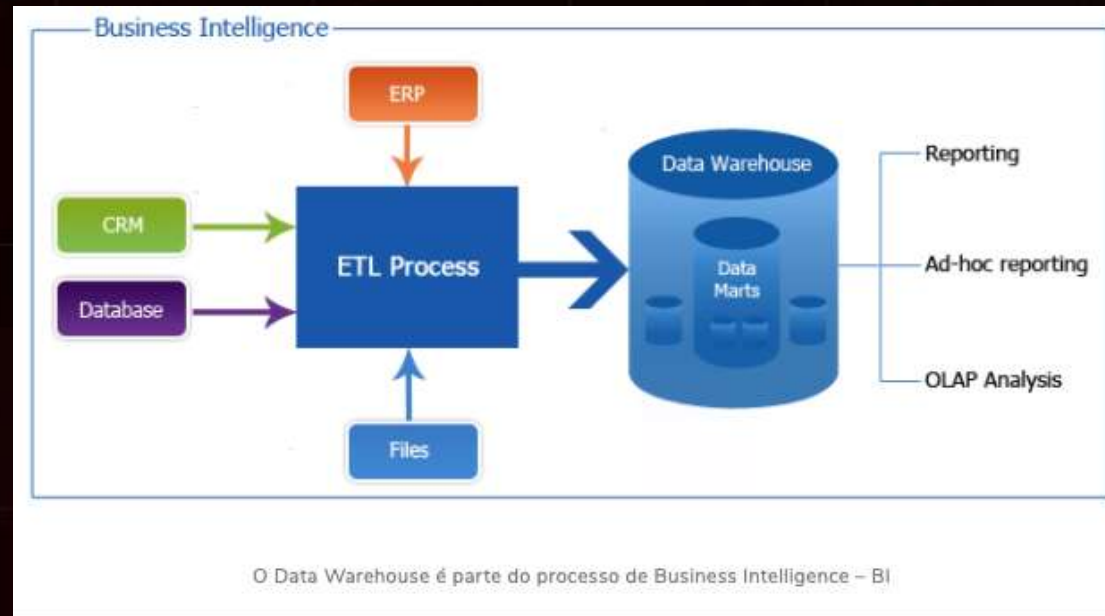
- O registro de dados constituídos com o tempo é um registro histórico muito valorizado em Data Warehouses.
- Os dados devem ser bem estruturados antes de ingressarem no Data Warehouse (seguem uma lógica).
- Os Data Warehouses, por medida de integridade, nunca permitem que dados sejam alterados. São bases estritamente para inserção de dados e posteriores consultas.
- Uso da Modelagem Dimensional (não usa dados normalizados), sendo os principais tipos as modelagens Star Schema (Modelo Estrela) e Snowflake Schema (Floco de Neve).

DATA WAREHOUSE

- **Principais características (por William Inmon, pai do Data Warehouse):**
 - **Orientados para o assunto.** Eles podem analisar dados sobre um determinado assunto ou área funcional (como vendas).
 - **Integrados.** Os data warehouses criam consistência entre diferentes tipos de dados de fontes distintas.
 - **Não volátil.** Quando os dados estão em um data warehouse, eles são estáveis e não mudam.
 - **Variáveis de acordo com o tempo.** Um data warehouse permite analisar as mudanças ao longo do tempo.

DATA WAREHOUSE

- Composição:
 - Banco de Dados Relacional
 - Uma solução de ETL (Extração, Transformação e Carga de Dados)
 - Recursos de Análise e Mineração de Dados
 - Recursos de Visualização de Dados



DATA LAKE

- É uma estrutura de Banco de Dados do tipo repositório que armazena um grande conjunto de dados em seu formato nativo.
- Os dados em um Data Lake normalmente não permitem uma visão refinada dos dados.
- Um Data Lake recebe os dados sem qualquer tipo de tratamento específico, em estado natural, podendo posteriormente ser utilizados para uma variedade de finalidades.

DATA LAKE

- A arquitetura do Data Lake é simples, porque ela não é definida anteriormente ao recebimento dos dados.
- Como são coletados de várias fontes e armazenados em seu estado original, são apenas replicações de conjuntos de dados estruturados, semiestruturados ou não estruturados.
- Uma grande vantagem de um Data Lake é a velocidade de acumulação de dados, porque não exige qualquer tipo de limpeza ou tratamento anterior à coleta dos dados.

DATA WAREHOUSE X DATA LAKE

- DATA WAREHOUSE

- Dados limpos, estruturados e organizados (isso é caro e pode representar perda de tempo)
- Grandes volumes
- Utilizados por analistas que já sabem o que querem como resultado
- Hardware geralmente caro

- DATA LAKE

- Dados brutos
- Grandes Volumes
- Usados por analistas de dados que esperam encontrar informações novas
- Hardware mais acessível (custo)

BANCO DE DADOS

**Introdução a Views e conceitos
básicos sobre Data Warehouse e
Data Lake**