O Snowflake é uma tecnologia que foi projetada exclusivamente para data warehouse na nuvem que permite a análise de dados em alta performance, larga escala, escalabilidade e simplicidade operacional. Assim conectando várias empresas. Casos de uso Snowflake:

1. Data Warehousing Moderno

Implementação de data warehouses modernos na nuvem para substituir ou complementar os data warehouses on-premises tradicionais (Grandes servidores dentro de empresas, ocupando espaço que poderia estar na nuvem). Tendo alta escalabilidade, elasticidade e redução de custos com infraestrutura e manutenção.

2. Análise de Dados em Tempo Real

Análise de grandes volumes de dados em tempo real para obter visão de dados e comparações imediatas. Tendo decisões mais rápidas, suportando operações dos negócios.

3. Big Data

Análise de grandes volumes de dados com escalabilidade elástica. Assim com grande processamento eficiente dos dados com custos controlados e performance.

4. Compliance e Segurança de Dados

Implementação de medidas de segurança e de acordo com regulamentações (LGPD). Criptografia, controle de acesso granular, auditoria e conformidade com regulamentações.

Exemplo de uso

- E-commerce: Análise de comportamento de clientes.

- Finanças: Análise de risco e detecção de fraudes financeiras.

- Saúde: Análise de dados de pacientes, pesquisa das clínicas, gestão de operações hospitalares (caso da MV aqui em recife).

- Varejo: Gestão de inventário e análise de vendas.

O Star Schema é um banco de dados também utilizado em data warehouses para organizar dados de maneira que facilite a análise dos dados. Tendo em vista à sua simplicidade e eficiência em consultas. casos de Star Schema:

1. Análise de Vendas

Estruturar dados das vendas para análise detalhada por produto, região, tempo e cliente. Contém dados de transações de Produto, Cliente, Tempo, Região. Permite relatórios detalhados e análises de tendências de vendas, desempenho de produtos, e comportamento de clientes.

2. Marketing e Análise de Campanhas

Analisar a eficácia das campanhas de marketing através de diferentes canais e demografias. Contém dados sobre interações de campanhas (impressões, cliques, conversões). Campanha, Canal, Tempo, Demografia. Facilita a análise de ROI (retorno sobre investimento) de campanhas, identificação de canais mais eficazes e segmentação de mercado.

3. Análise Financeira

Organizar dados financeiros para análise de desempenho e fluxos de caixa. Contendo dados financeiros das contas. Simplifica a criação de relatórios financeiros, análise financeira por período.

7. Educação e Análise de Desempenho Acadêmico

Analisar o desempenho de alunos e professores.

Tabela de Fato: Contendo dados de desempenho (notas, frequência, avaliações). Tabelas Dimensional: Aluno, Tempo, Curso, Professor, assim facilita a análise de desempenho geral.

Vantagens do Star Schema

- Simplicidade: Estrutura fácil de entender e de usar para consultas.

- Performance: Melhor performance de consultas devido à organização otimizada dos dados.

- Escalabilidade: Facilmente adaptável para grandes volumes de dados.

**Cubos OLAP**

Cubos OLAP (Online Analytical Processing) são estruturas de dados multidimensionais usadas para facilitar a consulta e a análise rápidas de grandes volumes de dados, típicas em aplicações de Business Intelligence (BI). A arquitetura e a implementação de cubos OLAP envolvem várias etapas e componentes. Abaixo está uma descrição detalhada da arquitetura e do processo de implementação:

- Arquitetura de Cubos OLAP

Dimensões: Representam os eixos de análise, como tempo, produtos e regiões. Cada dimensão é composta por hierarquias que permitem diferentes níveis de detalhe. Estruturas multidimensionais que contêm medidas associadas a diversas dimensões. Cada célula no cubo representa uma medida específica em um ponto de interseção das dimensões.

- Tipos de Cubos OLAP

ROLAP (Relational OLAP): Baseado em bancos de dados relacionais. As consultas são executadas diretamente nas tabelas relacionais usando SQL.

MOLAP (Multidimensional OLAP): Utiliza um banco de dados multidimensional para armazenar dados. Oferecendo melhor performance para consultas complexas.

HOLAP (Hybrid OLAP): Combina as abordagens ROLAP e MOLAP, armazenando dados agregados em um banco multidimensional e dados detalhados em um banco relacional.

- Implementação dos Cubos OLAP

Modelagem de Dados:

Determine quais métricas ou indicadores de performance serão analisados. Identificação das Dimensões: Determine os eixos de análise Definição das Hierarquias: Estruturado em cada dimensão em diferentes níveis de granularidade (ano, trimestre, mês, dia).

- Design do Esquema de Dados (modelos pedidos na atividade)

Star Schema: Tabelas fato centralizadas conectadas a tabelas dimensão.

Snowflake Schema: Extensão do Star Schema onde as tabelas dimensão são normalizadas.

- Construção do Cubo OLAP

Utilização de ferramentas de OLAP para definir o cubo, suas medidas e dimensões. Popular o cubo com dados pré-calculados e agregados, o que pode envolver a criação de índices e pré-computação de resultados.

- Benefícios dos Cubos OLAP

Performance: Consultas rápidas devido ao pré-cálculo de agregações.

Flexibilidade: Facilidade de explorar dados em diferentes perspectivas e granularidades.

Interatividade: Suporte a análises dinâmicas e interativas.

Complexidade Reduzida: Simplificação de consultas complexas através de tela intuitiva.