

A Importância da Modelagem de Dados para Data House

Introdução

Neste capítulo, iremos abordar a importância da modelagem de dados no contexto de Data House (DW). Conforme discutido no capítulo anterior, todo projeto de implementação de um DW está frequentemente associado a um projeto maior de Business Intelligence (BI). A primeira etapa envolve definir os objetivos e determinar quais áreas participarão do projeto. Com essas definições claras, inicia-se a fase de modelagem do DW, começando pela modelagem de negócio.

Exemplo: Projeto de BI para uma Rede de Supermercados

Contexto:

Uma rede de supermercados deseja melhorar suas operações, entender melhor o comportamento do cliente e aumentar as vendas através de insights mais profundos. Para isso, decide implementar um Data Warehouse como parte de uma iniciativa maior de Business Intelligence.

Passos do Projeto:

1. **Definição dos Objetivos:**
 - **Objetivo 1:** Aumentar as vendas através de campanhas de marketing mais eficazes.
 - **Objetivo 2:** Melhorar a gestão de estoque para reduzir perdas e garantir que produtos populares estejam sempre disponíveis.
 - **Objetivo 3:** Entender o comportamento do cliente para personalizar ofertas e melhorar a experiência de compra.
2. **Determinação das Áreas Participantes:**
 - **Marketing:** Para campanhas de marketing direcionadas e análise de eficácia.
 - **Vendas:** Para análise de vendas por produto, loja, e região.
 - **Gestão de Estoque:** Para otimização de estoque e redução de perdas.
 - **Atendimento ao Cliente:** Para entender feedbacks e melhorar o serviço.
3. **Modelagem do Data Warehouse (DW):**
 - **Modelagem de Negócio:**
 - **Fontes de Dados:** Sistemas de ponto de venda (POS), sistemas de gestão de estoque, CRM (Customer Relationship Management), entre outros.
 - **Dimensões:** Produto, Tempo, Localização, Cliente.
 - **Fatos:** Vendas, Estoque, Transações de Clientes.

Modelagem de Negócio

Na modelagem de negócio para um DW, o objetivo é compreender o negócio da organização ou de um departamento específico. Eles devem verificar como o negócio é visto pela empresa ou pelos funcionários daquele departamento, identificando os dados utilizados no processo de tomada de decisão. Esta é a primeira etapa crucial: a modelagem de negócio.

É comum que um projeto de DW comece com uma ou duas áreas de negócio. Este método pode ser trabalhoso, mas uma vez implementado, outras áreas podem ser integradas ao DW existente. No entanto, essa abordagem pode ser problemática, pois cada área de negócio possui seus próprios requisitos e necessidades, resultando, muitas vezes, em um sistema complexo e difícil de gerenciar.

Alternativa: Data Marts

Para evitar a complexidade de integrar múltiplas áreas de negócio de uma só vez, uma alternativa recomendada é criar data marts, pequenos segmentos do DW, para cada departamento. Posteriormente, eles podem consolidar esses data marts em um DW maior para toda a organização. Este método é mais trabalhoso e envolve um projeto maior, mas pode ser mais eficaz a longo prazo.

Modelagem de Dados

A modelagem de dados se preocupa em compreender os dados, criando uma visão que permite entender suas características e relacionamentos. Depois que o modelo de negócio está definido, eles podem explorar os relacionamentos entre os dados, identificando as entidades, seus principais atributos e os tipos de relacionamentos implementados. Com isso, é possível avançar para a fase de modelagem dimensional, seguida pela modelagem física e, finalmente, pela implementação do Data House.

Benefícios do Modelo de Dados

O modelo de dados utilizado em um DW é crucial para tornar as informações compreensíveis e navegáveis para o usuário. Além disso, determina a flexibilidade do DW, permitindo modificações para atender novos requisitos que surgem à medida que os usuários utilizam o sistema.

Existem diferentes abordagens sobre a melhor forma de criar um DW. Alguns autores sugerem construir o DW primeiro e apresentá-lo à área de negócio para validação e ajustes. Outros, incluindo aqueles que preferem uma abordagem centrada no usuário, acreditam que é ideal iniciar com a participação ativa do usuário de negócio, pois eles conhecem melhor suas necessidades. Apesar das divergências, é consenso que os requisitos vão mudar e o DW precisa ser flexível o suficiente para acomodar essas mudanças.

Desempenho e Eficiência

O modelo de dados também influencia o desempenho de acesso a grandes volumes de dados, essencial para o processamento analítico nos sistemas de apoio à decisão. Uma

boa modelagem de dados permite identificar e evitar possíveis gargalos, garantindo um ambiente analítico eficiente.

Características de um Bom Modelo de Dados

Um bom modelo de dados para um DW deve possuir as seguintes características:

1. Reflete a visão que o usuário tem do negócio e suas medidas.
2. É simples e facilmente compreensível.
3. É flexível para a incorporação de novos elementos.
4. Gera esquemas de banco de dados que permitem bom desempenho em consultas com grande volume de dados.