# **UNIDADE 1 - INTRODUÇÃO À BI**

Um grande problema é a falta de conhecimento em relação à grande quantidade de dados e informações disponíveis nas organizações. Isso é motivado pela dificuldade de integração de dados de diversas fontes e pelo desafio de extrair conhecimento útil desses dados. Além disso, as organizações enfrentam pressões do mercado, dos consumidores e da evolução tecnológica, tornando necessário tomar decisões ágeis com informações corretas, no momento certo e facilmente acessíveis.

A solução proposta é o Business Intelligence (BI), que envolve metodologias, processos, arquiteturas e tecnologias para transformar dados brutos em informações úteis para o processo de tomada de decisão. O BI ajuda a reduzir custos, aumentar a eficiência, prover insights e integrar dados de maneira on demand.

Além disso, é importante mencionar a importância do Data Warehousing (DW) como parte do BI, destacando suas características, como ser orientado por assuntos, não volátil e variável com o tempo e integrado. O DW é fundamental para armazenar e integrar dados de diferentes fontes e facilitar a análise.

Também é importante abordar a importância do Data Analytics, que envolve a exploração dos dados do DW para obter insights e suportar a tomada de decisão. Ele menciona ferramentas e tecnologias utilizadas no campo do BI, como o uso de Machine Learning, catalogação de analytics, preparação de dados, visualização de dados e muito mais.

As organizações estão enfrentando desafios devido à grande quantidade de dados, fontes heterogêneas, qualidade ruim dos dados e a necessidade de consultas complexas que não são adequadas para bancos de dados operacionais (OLTP).

Algumas das principais ferramentas e serviços de BI: Amazon QuickSight, Looker, Qlik, Tableau, PowerBI, Snowflake, Google Big Query, Amazon Redshift e Microsoft Azure Synapse Analytics.

# **UNIDADE 2 - MODELAGEM DE DW**

Aqui abordamos a diferença entre a modelagem tradicional e a dimensional, com foco na modelagem multidimensional utilizada em Data Warehouses (DW). Aqui está um resumo das principais informações:

Modelagem Tradicional vs. Dimensional: A modelagem tradicional visa refletir relacionamentos entre dados, fornecendo uma visão global e normalizada dos dados. No entanto, é ineficiente para tomada de decisões devido à sua complexidade. Por outro lado, a modelagem dimensional é otimizada para análise de dados, simplificando o esquema do banco de dados e incluindo tabelas de dimensões e fatos.

Modelagem Multidimensional: O modelo multidimensional é projetado para análise de dados e é reconhecido por ferramentas OLAP/BI. Ele consiste em tabelas de dimensões (com propriedades descritivas) e tabelas de fatos (com medidas agregáveis). O objetivo é envolver fatos com contexto das dimensões para facilitar consultas e análises.

Exemplo de um Cubo: Um cubo multidimensional contém dimensões como Produto, Localização, Tempo e medidas como Vendas. Isso permite análises multidimensionais.

Esquema Estrela: É uma estrutura de Data Warehouse que inclui tabelas de dimensões e uma tabela de fatos, onde os IDs das dimensões são FKs na tabela de fatos.

Esquema Snowflake: É uma variação do esquema estrela em que as tabelas de dimensões são normalizadas para economizar espaço.

Esquema Constelação: Permite que várias tabelas de fatos compartilhem as mesmas dimensões.

Tabela de Fato: Contém eventos de um processo de negócios, medidas (normalmente numéricas) e é normalmente normalizada.

Medidas Aditivas, Semi-aditivas e Não Aditivas: Medidas aditivas podem ser somadas ao longo de todas as dimensões, enquanto as semi-aditivas e não aditivas têm restrições na agregação.

Tipos de Tabela de Fato: As tabelas de fatos podem ser de três tipos: transação (captura eventos detalhados), instantâneo periódico (registra resumos em intervalos regulares) e instantâneo acumulado (registra múltiplas datas representando estágios do fato).

# **UNIDADE 3 - MODELAGEM MULTIDIMENSIONAL - DIMENSÕES**

Esses são os principais conceitos abordados no texto relacionados à modelagem multidimensional em um ambiente de data warehousing:

Dimensões: As dimensões são perspectivas usadas para analisar os dados em um data warehouse. Por exemplo, uma dimensão pode ser um cubo tridimensional com dimensões como Produto, Tempo e Cliente.

Tabela de Dimensão: Essas tabelas correspondem aos pontos de entrada das consultas e armazenam as descrições do negócio, rótulos e agrupamentos dos relatórios. Elas são fundamentais para tornar o data warehouse compreensível.

Projeto de um DW - Hierarquias: Hierarquias são formadas quando uma dimensão é quebrada em sub-dimensões de forma hierárquica, permitindo a análise em diferentes níveis de detalhe.

Hierarquia versus Não Hierarquia: A decisão de usar hierarquias ou não afeta a estrutura do data warehouse, a normalização dos dados e o processamento de consultas. Hierarquias podem ser mais semânticas e facilitar análises específicas.

Atualizações de Dados nas Dimensões: Quando os dados nas dimensões mudam, é necessário lidar com essas mudanças usando estratégias como SCD1 (sobrescrever), SCD2 (adicionar uma nova linha), SCD3 (adicionar uma coluna) ou SCD4 (adicionar uma tabela).

Projeto Multidimensional: Isso envolve a modelagem de dimensões e fatos, considerando a granularidade, dimensões degeneradas, minidimensões, dimensões com papéis, entre outros.

Bridge Tables (Tabelas Pontes): São usadas para lidar com relacionamentos muitos-para-muitos entre dimensões ou entre fato e dimensão. Elas incluem informações de peso ou fator que são importantes em alguns cálculos.

Dimensões M para N: Quando uma dimensão possui um relacionamento multivalorado com a tabela de fatos, podem ser usados fatores de peso para representar a contribuição de cada valor em um grupo.

# **UNIDADE 5.2 - KPI E DASHBOARDS**

O conteúdo trata de Business Intelligence (BI) e Visualização de dados com foco em Key Performance Indicators (KPIs) e Dashboards. Aqui estão os principais pontos do material:

KPIs (Key Performance Indicators) são indicadores de desempenho essenciais para avaliar o sucesso de um negócio. Eles são definidos de acordo com as estratégias da empresa e são apresentados em Dashboards de forma gráfica, interativa e com vários elementos de interface, como relatórios, gráficos, mapas e semáforos.

A criação de Dashboards eficazes requer simplicidade (usando a estratégia KISS - Keep It Simple Stupid) e um profundo entendimento do negócio, curiosidade e a capacidade de identificar problemas.

A visualização de dados é uma técnica antiga, mas sua popularidade aumentou significativamente nos últimos anos, especialmente com o surgimento de ferramentas de BI modernas.

Diferentes tipos de gráficos podem ser usados para representar dados, e a escolha depende do objetivo. Por exemplo, gráficos de barras são adequados para comparar valores numéricos em categorias discretas.

Os KPIs podem rastrear várias métricas relacionadas à eficiência, qualidade, governança, comportamento, economia, desempenho do projeto e pessoal, utilização de recursos e sustentabilidade.

Os KPIs podem ser classificados de acordo com diferentes critérios, como dimensão temporal, input/output, questões abordadas e medida (qualitativo ou quantitativo).

Exemplos de KPIs incluem taxa de conversão, taxa de retenção de clientes, custo por aquisição de cliente, retorno sobre o investimento (ROI), entre outros.

O processo de definição de KPIs envolve montar uma equipe, categorizar métricas potenciais, fazer um brainstorming, priorizar métricas, definir precisamente cada métrica escolhida e montar o conjunto final de métricas.

Boas práticas ao criar Dashboards incluem manter a clareza, coesão, hierarquia e posicionamento adequados, além de seguir convenções e entender o objetivo do público-alvo.

Também foram apresentados exemplos de Dashboards em várias áreas, como marketing, vendas, recursos humanos, logística e finanças.

O Balanced Scorecard (BSC) é mencionado como uma metodologia amplamente usada na gestão de negócios para avaliar o desempenho com base em quatro perspectivas: financeira, cliente, processos internos e aprendizado.