

FACULDADE DE TECNOLOGIA DE SÃO PAULO
FATEC MAUÁ
DESENVOLVIMENTO DE SOFTWARE MULTIPLATAFORMA

NILTON DIONISIO GUERRA

DATA WAREHOUSE

MAUÁ
2023

NILTON DIONISIO GUERRA

O QUE É DATA WARHOUSE:

Subtítulo do Trabalho

Trabalho apresentado para a Disciplina de banco de dados não relacional, pelo Curso de Desenvolvimento de software multiplataforma da Faculdade de tecnologia de São Paulo, ministrada pelo Prof. Anderson Silva Vanin.

SUMÁRIO

INTRODUÇÃO	3
1. QUE É DATA WAREHOUSE.....	4
2. ETL	4
2.1 Extrair	4
2.2 Transformar	5
2.3 Carregar	6
3. RELAÇÃO A BI (BUSINESS INTELLIGENCE).....	6
3.1 Ferramentas De Consulta E Emissão De Relatórios	6
3.2 Dashboards (Painéis Digitais)	7
3.3 Ferramentas Olap.....	7
3.4 Ferramentas De Data Mining	7
3.5 Dss E Eis	8
REFERÊNCIAS.....	10

INTRODUÇÃO

Neste trabalho iremos abordar as etapas no ETL, sendo elas a extração, transformação e carregamento; Além de também explicar de maneira sucinta o BI (Business Intelligence), tendo como foco as ferramentas de consulta e emissão de relatórios, dashboards (Painéis Digitais), Ferramentas OLAP, ferramentas de DataMining e DSS e EIS, além de que também iremos abordar o que é DataMarts.

1. QUE É DATA WAREHOUSE

Um data warehouse é um tipo de sistema de gerenciamento de dados projetado para ativar e fornecer suporte às atividades de business intelligence (BI), especialmente a análise avançada. Os data warehouses destinam-se exclusivamente a realizar consultas e análises avançadas e geralmente contêm grandes quantidades de dados históricos. Os dados em um data warehouse geralmente são derivados de uma ampla variedade de fontes, como arquivos de log de aplicativos e aplicativos de transações.

Um data warehouse centraliza e consolida grandes quantidades de dados de várias fontes. Seus recursos analíticos permitem que as organizações obtenham informações de negócios úteis de seus dados para melhorar a tomada de decisões. Com o tempo, cria-se um registro histórico que pode ser inestimável para cientistas de dados e analistas de negócios. Devido a esses recursos, um data warehouse pode ser considerado como a “única fonte confiável” de uma organização.

2. ETL

ETL é um acrônimo para Extração, Transformação e Carga (Extract, Transform, Load), que é um processo essencial no contexto de um Data Warehouse. O ETL é o conjunto de atividades que envolvem a obtenção, preparação e carregamento dos dados no Data Warehouse.

2.1 Extrair

Extrair é uma das etapas do processo ETL (Extração, Transformação e Carga), que consiste na obtenção de dados de diferentes fontes, como bancos de dados, sistemas externos, arquivos, APIs, entre outros, e trazê-los para um local centralizado onde possam ser processados e analisados.

A extração de dados envolve a identificação das fontes de dados relevantes, a definição de critérios de seleção e a obtenção dos dados a partir dessas fontes. Isso pode incluir a extração de dados brutos, como registros de transações, informações de clientes, registros de vendas, entre outros, em diferentes formatos, como texto, planilhas, bancos de dados ou arquivos estruturados.

A extração também pode envolver a aplicação de transformações iniciais nos dados, como a limpeza de dados, a validação de formatos e a conversão de tipos de dados, para prepará-los para a próxima etapa do processo ETL, que é a transformação.

A etapa de extração é fundamental para o sucesso de um projeto de integração de dados, pois é responsável por obter os dados necessários das fontes corretas e garantir a integridade e qualidade desses dados, para que possam ser utilizados em análises e relatórios no contexto do Business Intelligence.

2.2 Transformar

A transformação é uma etapa chave no processo ETL (Extração, Transformação e Carga) de um projeto de integração de dados. É a fase em que os dados extraídos de diversas fontes são manipulados, convertidos e enriquecidos para atender às necessidades de análise e armazenamento em um destino, como um Data Warehouse ou um Data Mart.

A transformação de dados pode incluir várias atividades, como limpeza de dados, validação, enriquecimento, agregação, padronização, normalização, cálculos, filtragem, entre outros. Essas transformações são aplicadas aos dados brutos extraídos, a fim de prepará-los para uso em análises e relatórios.

As transformações podem ser feitas usando técnicas e ferramentas específicas de ETL, como mapeamento de colunas, expressões SQL, scripts, ou até mesmo algoritmos complexos de processamento de dados. O objetivo é garantir que os dados sejam formatados, estruturados e enriquecidos de acordo com os requisitos de negócio e as regras de validação definidas.

A transformação é uma etapa crítica no processo ETL, pois permite a limpeza, enriquecimento e preparação dos dados para serem utilizados em análises e relatórios, garantindo a qualidade, integridade e consistência dos dados que serão armazenados e utilizados nas atividades de tomada de decisões em uma organização.

2.3 Carregar

O carregamento no processo ETL (Extração, Transformação e Carga) refere-se à etapa em que os dados extraídos de diversas fontes são transformados e carregados em um destino, como um Data Warehouse ou um Data Mart. Nessa etapa, os dados são limpos, validados, enriquecidos e convertidos em um formato adequado para o armazenamento e análise.

O carregamento geralmente envolve a aplicação de regras de negócio, transformações e integrações de dados para garantir a qualidade e a integridade dos dados. Essa etapa também pode incluir a combinação e o enriquecimento dos dados provenientes de diferentes fontes, a fim de obter uma visão consolidada e integrada dos dados.

O carregamento no processo ETL é uma etapa crítica para garantir a confiabilidade e a qualidade dos dados que serão usados nas análises e relatórios do Business Intelligence. É fundamental para preparar os dados de forma adequada, garantindo que estejam prontos para serem utilizados nas atividades de análise e tomada de decisões dentro de uma organização.

3. RELAÇÃO A BI (BUSINESS INTELLIGENCE)

O Business Intelligence (BI) é um conjunto de processos e tecnologias que transformam dados em informações úteis para a tomada de decisões. Envolve coleta, análise e visualização de dados para insights e otimização de negócios. É usado em várias áreas de uma organização e pode oferecer vantagens competitivas.

3.1 Ferramentas De Consulta E Emissão De Relatórios

As ferramentas de consulta e emissão de relatórios são componentes do Business Intelligence (BI) que permitem aos usuários acessar, analisar e visualizar dados de forma eficiente. Incluem dashboards, ferramentas de consulta ad-hoc,

relatórios interativos, ferramentas de geração de relatórios, ferramentas de análise de dados e integração com outras ferramentas de BI. A escolha das ferramentas adequadas depende dos requisitos e objetivos de análise de dados da organização.

3.2 Dashboards (Painéis Digitais)

Dashboards, ou painéis digitais, são componentes do Business Intelligence (BI) que fornecem uma visualização rápida e intuitiva de informações-chave em tempo real. São compostos por gráficos, tabelas e indicadores-chave de desempenho (KPIs) que permitem aos usuários monitorar o desempenho do negócio, identificar tendências e padrões importantes de forma visual e interativa. Os dashboards são uma ferramenta valiosa para tomar decisões informadas e acompanhar o progresso em relação aos objetivos de negócio de maneira eficiente.

3.3 Ferramentas Olap

Ferramentas OLAP (Online Analytical Processing) são componentes do Business Intelligence (BI) que permitem aos usuários realizar análises multidimensionais em dados. Essas ferramentas permitem a exploração e análise de dados em várias dimensões, como tempo, produto, localização, entre outros, para identificar tendências, padrões e insights ocultos. As ferramentas OLAP oferecem recursos avançados, como a criação de cubos de dados, consultas multidimensionais, análise de fatos e dimensões, e recursos de drill-down e drill-up para aprofundar ou resumir os dados em diferentes níveis de granularidade. Essas ferramentas são amplamente utilizadas em atividades de análise de dados e suportam a tomada de decisões informadas em uma ampla variedade de setores e indústrias.

3.4 Ferramentas De Data Mining

Ferramentas de Data Mining são componentes do Business Intelligence (BI) que utilizam técnicas avançadas de análise de dados para descobrir padrões, tendências e relações ocultas em grandes volumes de dados. Essas ferramentas permitem a identificação de informações valiosas e insights significativos para apoiar

a tomada de decisões informadas. As técnicas de Data Mining incluem algoritmos de aprendizado de máquina, análise estatística, mineração de texto, análise de redes sociais e outros métodos avançados de análise de dados. As ferramentas de Data Mining são usadas em várias indústrias e setores, como marketing, finanças, saúde, varejo e telecomunicações, para identificar oportunidades de negócios, melhorar a eficiência operacional e obter uma vantagem competitiva.

3.5 Dss E Eis

DSS (Sistemas de Suporte à Decisão) e EIS (Sistemas de Informação Executiva) são componentes do Business Intelligence (BI) que fornecem informações estratégicas e insights para apoiar a tomada de decisões em uma organização.

DSS são sistemas que auxiliam os tomadores de decisão a analisar dados, realizar modelagem de cenários, simulações e previsões, além de facilitar a seleção de alternativas e a escolha da melhor estratégia de ação.

EIS são sistemas que fornecem informações relevantes em tempo real para os executivos de alto nível de uma organização, permitindo que tomem decisões estratégicas informadas com base em dados atualizados sobre o desempenho da empresa e o ambiente de negócios.

Esses sistemas ajudam os tomadores de decisão a obter insights valiosos, tomar decisões informadas e direcionar ações estratégicas com base em informações precisas e atualizadas, contribuindo para a eficiência operacional, a competitividade e o sucesso nos negócios.

4. O QUE SÃO DATAMARTS

Datamarts são componentes de um sistema de Business Intelligence (BI) que são projetados para atender a necessidades específicas de análise de uma área ou departamento específico de uma organização. São repositórios de dados que são

projetados e otimizados para atender às necessidades de análise de um grupo de usuários ou de uma função de negócio específica.

Os datamarts são geralmente derivados de um Data Warehouse, que é um repositório central de dados integrados de várias fontes. No entanto, diferentemente do Data Warehouse, os datamarts são projetados para serem menores, mais rápidos e mais específicos para atender às necessidades de análise de uma área de negócio específica.

Os datamarts podem ser organizados por áreas funcionais, como vendas, marketing, finanças, recursos humanos, entre outros, e podem conter dados históricos, atuais e futuros relevantes para a área específica. Esses datamarts podem ser projetados para oferecer análises de autoatendimento para os usuários finais, permitindo que eles acessem, analisem e relatem dados relevantes para suas necessidades específicas de análise de forma eficiente.

Os datamarts são projetados para fornecer análises rápidas e eficientes, com dados organizados de forma a permitir uma fácil exploração e análise pelos usuários finais. Eles são uma parte importante de muitas arquiteturas de Business Intelligence, permitindo uma análise mais detalhada e específica de dados para áreas de negócio específicas, contribuindo para a tomada de decisões informadas e estratégicas.

REFERÊNCIAS

Oracle. O que é um Data Warehouse?.Oracle,2022,Disponível em: <https://www.oracle.com/br/database/what-is-a-data-warehouse/#:~:text=Um%20data%20warehouse%20%C3%A9%20um,%2C%20especialmente%20a%20an%C3%A1lise%20avan%C3%A7ada> acessado em:12/04/2023

ChatGPT, O que é DataWarehouse?.OpenAI,2023, Disponível em: <https://chat.openai.com/chat> acessado em:13/04/2023.

Alves,Pedro, O que é ETL – Extract Transform Load?.cetax, disponível em : <https://cetax.com.br/etl-extract-transform-load/> Acessado em:14/04/2023.

Silva, Silvio, O que é business intelligence? Seu guia sobre o BI e por que ele é importante,Sales force, Disponível em: <https://www.tableau.com/pt-br/learn/articles/business-intelligence>, acessado em:14/04/2023

AWS, O que é um data mart?, Amazon, Disponível em:<https://aws.amazon.com/pt/what-is/data-mart/> acessado em:12/04/2023