**AllMart**

**Projeto da Disciplina**

**de BI**

**Prof. Anderson Nascimento  
Versão: 03-04-2024**

**Componentes do Projeto:**

Ivo Lavacek (202203473107)

João Vitor Araújo (202201107065)

Mateus Norcia (202201038381)

Pedro Lustosa (202203798936)

Rodrigo Silveira (202202064831)

|  |
| --- |
| Histórico de Versões |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Data** | **Versão** | **Descrição** | **Autor** | **Aprovado por** |
| 08/04/2024 | 1.0 | Versão Inicial do Projeto | Grupo | Grupo |
| 19/06/2024 | 2.0 | Versão Final do Projeto | Grupo | Grupo |

Sumário

[1. Introdução 4](#_Toc169731957)

[2. Estudo de Caso 5](#_Toc169731958)

[4. Proposta do Projeto de BI 7](#_Toc169731959)

[5. Elaboração do Data Warehouse 8](#_Toc169731960)

[ **Definição do DW** 8](#_Toc169731961)

[ **Arquitetura** 8](#_Toc169731962)

[ **Abordagem de Construção** 8](#_Toc169731963)

[ **Arquitetura Física** 8](#_Toc169731964)

[6. Modelo Multidimensional 9](#_Toc169731965)

[ **Pontos Cardeais do BI** 9](#_Toc169731966)

[ **Matriz Dimensão Indicador** 9](#_Toc169731967)

[ **Modelagem do Projeto** 9](#_Toc169731968)

[7. Projeto de ETL 10](#_Toc169731969)

[12](#_Toc169731970)

[12](#_Toc169731971)

[12](#_Toc169731972)

[8. Dashboard 12](#_Toc169731973)

[<Lembre-se que esta fase deverá estar de acordo com o documento de requisitos elaborado previamente> 12](#_Toc169731974)

[Descrição da Elaboração 12](#_Toc169731975)

[<Criar um texto descrevendo como foi elaborado o dashboard, citando a ferramenta utilizada> 12](#_Toc169731976)

[Telas do Dashboard 13](#_Toc169731977)

[9. Conclusão 15](#_Toc169731978)

[10. Anexos 16](#_Toc169731979)

[11. Arquivos 17](#_Toc169731980)

[12. Conclusão(@) 17](#_Toc169731981)

# **Introdução**

Este projeto foi desenvolvido no âmbito da disciplina de Business Intelligence do curso de Ciência de Dados e Inteligência Artificial do IBMEC. O objetivo é aplicar na prática os conceitos e técnicas aprendidos ao longo do curso. A disciplina de BI proporciona aos alunos uma compreensão profunda das ferramentas e metodologias necessárias para transformar dados brutos em informações valiosas, facilitando a tomada de decisões estratégicas nas organizações.

Este projeto é uma oportunidade de integrar teoria e prática, demonstrando a capacidade de utilizar processos de ETL (Extração, Transformação e Carga), construção de Data Warehouses e desenvolvimento de dashboards interativos.

# **Estudo de Caso**

A empresa AllMart optou pela contratação de nossa equipe para a elaboração de um processo de BI que ajude a empresa a conhecer melhor o seu negócio, entendendo o passado para que possam tomar melhores decisões futuras, obtendo vantagens competitivas no mercado de atuação.

O projeto tratará os dados de compras como fato ativo a ser estudado, levando em consideração as dimensões de data, item, setor e comprador. A proposta será cobrir todas as informações a respeito da junção das dimensões com o fato e consequentemente abstrair informações que possam servir como base para uma estratégia de negócio futuro, visando melhorar o desempenho da empresa.

O objetivo deste estudo é analisar o comportamento de compra dos clientes ao longo de 2023, identificar tendências e padrões, avaliar o desempenho dos setores e entender a contribuição financeira de cada comprador e setor.

Ao final do projeto, a empresa terá acesso a um Data Warehouse e a um processo de ETL automatizado que servirá de base para um dashboard, que também será desenvolvido nesse trabalho. Dessa forma, os gestores terão acesso à milhares de dados relevantes de forma ágil e visual, auxiliando suas tomadas de decisão.

1. **Descrição do Modelo Transacional**

A seguir é apresentado o Diagrama Entidade Relacionamento (DER) dos dados que serão analisados nesse projeto. Apesar de ser um modelo simplificado para um caso real, será suficiente para a aplicação de tudo que foi proposto na disciplina de BI.

* **Modelo DER**

Diagrama

Descrição gerada automaticamente

# **Proposta do Projeto de BI**

Esta seção apresenta o projeto de BI proposto para o projeto.

A blue and white logo

Description automatically generated

Entrada de dados  
elaborados no Mockaroo

Tratativa de dados

Para este projeto, foi proposta a integração e consolidação da base de dados do sistema transacional e dados elaborados no site de geração de dados Mockaroo.

Para isso, foi feita uma base de dados inicial (st\_allmart) usando o SGBD PostgreSQL. Através de um processo de ETL, feito com a ferramenta Pentaho Data Integration (PDI), foi feito o carregamento da área de preparação para o Data Warehouse (dw\_allmart), realizando também a higienização, limpeza e transformação dos dados.

Os dados serão carregados no Data Warehouse configurado como modelo estrela, implementado também no SGBD PostgreSQL, não sendo necessário o uso de Data Marts. A atualização dos dados deverá ser orquestrada de forma regular, apenas sendo necessária a execução de um arquivo .kjb desenvolvido com o PDI, assumindo que a empresa insere os dados das compras corretamente na sua base de dados primária.

A parte da visualização de dados será implementada no Microsoft Power BI, onde serão criados gráficos e indicadores em um dashboard analítico. Sua atualização também deverá ocorrer de forma regular e pode ser automatizada.

# **Elaboração do Data Warehouse**

O Data Warehouse (DW) será a fonte integradora de informações da empresa, a tecnologia será utilizada com o intuito de servir de base para a camada de visualização que será responsável por fornecer dados para a tomada de decisão na organização.

## **Definição do DW**

Nesta seção são apresentados os detalhes de implementação do DW quanto à sua arquitetura, abordagem de construção e disposição física.

### **Arquitetura**

A arquitetura do DW será Global e Centralizada, considerando que a empresa pode conter mais de uma filial, mas a gerência da área de vendas está interessada em todos os dados da empresa.

### **Abordagem de Construção**

Como a arquitetura será Global e Centralizada, não serão construídos Data Marts, dessa forma o processo de construção levará em conta apenas a construção do DW.

### **Arquitetura Física**

Todo o projeto será criado para a arquitetura em Nuvem, ou seja, o DW ficará armazenado em um servidor online, considerando praticidade e custo-benefício.

# **Modelo Multidimensional**

Esta seção apresenta o resultado da modelagem do DW, implementado utilizando o modelo estrela do estudo de caso AllMart.

## **Pontos Cardeais do BI**

Esta seção apresenta a aplicação da técnica dos pontos cardeais para a descoberta das dimensões do projeto.

* Quem: Comprador
* O que: Item
* Onde: Setor
* Quando: Data da compra

## **Matriz Dimensão Indicador**

Técnica não aplicada, pois estamos trabalhando com apenas 1 tabela FATO.

## **Modelagem do Projeto**

Esta seção apresenta o resultado da modelagem de dados dimensional do projeto.

A construção do modelo foi elaborada utilizando o software SQL Power Architect.

A diagram of a sports game

Description automatically generated

# **Projeto de ETL**

Esta seção apresenta e descreve o projeto de ETL realizado na elaboração do projeto de Business Intellingence.

O processo de ETL (Extração, Transformação e Carga) começa com a organização dos dados na Stage Area, nesse projeto é a base de dados ‘st\_allmart’. Essa primeira etapa pode consistir em agregar dados de diversas tabelas em uma base de dados, porém, nesse projeto, os dados foram todos simulados no site moockaroo.com e inclusos num único arquivo (‘script\_insere\_dados\_st\_allmart.sql’). Após a criação e configuração correta da Stage Area, basta rodar esse arquivo e ela estará completa.

Para concluir o restante do processo de ETL, basta rodar os arquivos feitos com o Pentaho Data Integration (PDI). Foram feitos 3 arquivos para popular as dimensões dim\_comprador, dim\_item e dim\_setor. A dim\_data, que é uma constante para qualquer processo de ETL, foi integrada diretamente no Data Warehouse (Database ‘dw\_allmart’ criada diretamente no PostgreSQL) através do arquivo ‘script\_dim\_tempo.sql’ que foi fornecido pelo professor.

E por último, devemos popular a tabela principal do projeto, que é a fato\_compra. Para fazer isso, bastar rodar o arquivo ETL04 feito com o PDI.

* ETL 01 – Carga da Dimensão Setor

Neste ETL é realizado toda da transformação dos dados, carregamento e o versio-namento da dimensão dim\_setor.

A blue arrow pointing to a line

Description automatically generatedA blue arrow pointing to a line

Description automatically generated

* ETL 02 – Carga da Dimensão Item

Neste ETL é realizado toda da transformação dos dados, carregamento e versio- namento da dimensão dim\_item.

A blue arrow pointing to a line

Description automatically generatedA diagram of a project

Description automatically generated

* ETL 03 – Carga da Dimensão Comprador

Neste ETL é realizado toda da transformação dos dados, carregamento e versio- namento da dimensão dim\_comprador.

A computer screen shot of a computer code

Description automatically generatedA diagram of a diagram

Description automatically generated with medium confidence

* ETL 04 – Carga da Tabela Fato

Neste ETL é realizado toda da transformação dos dados e carregamento da tabela fato.

A line of text with green and blue text

Description automatically generated with medium confidenceA close-up of a computer

Description automatically generated

* JOB – Pipeline de Atualização do DW

Neste JOB são realizadas as atualizações dos dados das dimensões e da fato. Ele pode ser agendado para rodar todos os dias. Para esse projeto, não é possível executá-lo pois é exclusivo para a verdão paga do PDI.

A diagram of a cell phone

Description automatically generatedA green line with white squares and green arrows

Description automatically generated with medium confidence

# 

# Dashboard

## Descrição da Elaboração

### O intuito do Dashboard é passar mostrar um insight geral sobre o AllMart relacionado a vendas , com a possibilidade de filtros temporais

O projeto foi construído usando a ferramenta Microsoft Power BI.

## Telas do Dashboard

Interface gráfica do usuário, Aplicativo

Descrição gerada automaticamente

Painel – Dashboard Principal

O Painel principal do Dashboard possui dados mais gerais sobre os principais setores e itens da empresa.

# Conclusão

Este projeto foi uma grande oportunidade de estabelecer todo o processo do BI

convencional. Com a elaboração do modelo multidimensional e a respectiva criação do Data Warehouse, o cliente poderá ter informações atualizadas sobre

o seu negócio de maneira rápida e intuitiva, o que permitirá que o mesmo possa tomar decisões rápidas e assertivas.

# Arquivos

Toda a documentação aberta do projeto está disponível no endereço: https://github.com/mncbl/ProjetoBi2024.1

Lista de Arquivos

|  |  |
| --- | --- |
| Modelo dimensional | Arquivo de criação do *Star Schema* com as devidas relações entre a tabela fato e as tabelas dimensão. |
| Database.rar | Base de dados dos gastos associados a operação da loja, extraído do sistema transacional. |
| Pentaho ETL.rar | Arquivo de realização do ETL do projeto, feito no Pentaho Data Integration, conforme explicado no tópico 7.1. |
|  | Arquivo criado para visualização dos dados disponíveis no Data Warehouse criado. |
| AP2.pbix |  |
|  |  |