

# Análise Exploratória de Dados - Walmart

## Sumário

1. Glossário
2. Objetivo do Estudo
3. Apresentação da Empresa e Problema de Pesquisa
4. Apresentação dos Metadados
5. Análise Exploratória de Dados
6. Conclusão
7. Referencias

### 1. Glossário

- **CPI (Customer Price Index):** Índice de preços ao consumidor, utilizado para medir a inflação acumulada.
- **Holiday Flag:** Indica se a semana analisada contém um feriado.
- **Weekly Sales:** Vendas semanais da loja em dólares americanos.

### 2. Objetivo do Estudo

O estudo visa analisar fatores que afetam o desempenho das lojas do Walmart, identificando padrões em vendas, sazonalidade e impactos de variáveis socioeconômicas como inflação, desemprego, preço de combustível e outros fatores.

### 3. Apresentação da Empresa e Problema de Pesquisa

#### Nome da Empresa

Walmart

#### Missão, Visão e Valores

O Walmart tem como missão "ajudar as pessoas a economizarem dinheiro para que possam viver melhor". Seus valores incluem integridade, respeito ao indivíduo e compromisso com os clientes.

## **Segmento de Atuação**

O Walmart atua no setor varejista, sendo uma das maiores redes do mundo.

## **Número de Colaboradores**

A empresa emprega mais de 2 milhões de funcionários globalmente.

## **Iniciativas na Área de Data Science**

O Walmart utiliza análise de dados para prever demandas, otimizar estoques e personalizar ofertas.

## **Problema do Estudo**

- **O que falta?** Identificação de padrões detalhados que influenciam as vendas.
- **O que incomoda?** A oscilação das vendas devido a fatores externos.
- **O que pode ser melhorado?** Estratégias para aumentar previsibilidade nas vendas.
- **Qual é o gap?** Como sazonalidade, inflação e desemprego impactam as vendas.
- **Há um padrão?** Sim, com base em datas, inflação e eventos sazonais.
- **Há uma afirmação que pode ser contestada?** A suposição de que feriados sempre aumentam as vendas.

## **4. Apresentação dos Metadados**

- **Origem dos dados:** Kaggle (dados públicos) .
- **Sensibilidade:** Não há dados sensíveis.
- **Validade:** Dados de 2010 a 2012 .
- **Proprietário:** Publicado por Mikhail no Kaggle .
- **Restrições de uso:** Dados públicos, sujeitos a direitos de terceiros.
- **Atributos:**
  - STORE: Número da loja.
  - DATE: Data da semana analisada.
  - WEEKLY\_SALES: Total de vendas semanais.

- HOLIDAY\_FLAG: Indica feriado.
- TEMPERATURE: Temperatura média.
- FUEL\_PRICE: Preço do combustível.
- CPI: Índice de inflação.
- UNEMPLOYMENT: Taxa de desemprego.

## 5. Análise Exploratória de Dados

A análise inicial feita em R mostrou que os dados são bem estruturados, sem valores nulos.

### Estatísticas Descritivas

- **Número de Linhas:** 6435
- **Número de Colunas:** 8
- **Tipos de Dados:**
  - Numérico: STORE, WEEKLY\_SALES, TEMPERATURE, FUEL\_PRICE, CPI, UNEMPLOYMENT.
  - Texto: DATE.
  - Binário: HOLIDAY\_FLAG .

### Medidas de Posição e Dispersão

- **Média de Vendas Semanais:** Aproximadamente \$1.046.000.
- **Desvio Padrão:** Aproximadamente \$550.000.
- **Maior Venda Semanal:** Aproximadamente \$3.000.000.

### Distribuição e Frequência

- As vendas seguem uma distribuição assimétrica, com algumas lojas apresentando valores bem acima da média.
- Períodos de alta se concentram em feriados e mudanças sazonais.

### Correlação

- **Vendas e Inflação (CPI):** Tem uma relação negativa e moderada
- **Vendas e Desemprego:** Tem uma relação relativamente negativa e fraca
- **Vendas e Preço de Combustível:** Pouca correlação observada. Tem pouca relação observada

## **Valores Perdidos ou Incorretos**

- Não foram identificados valores nulos.
- O formato da data precisou ser ajustado no R.

## **Anomalias e Outliers**

- Algumas lojas apresentam picos de venda em feriados, enquanto outras não sofrem grande impacto.
- Necessidade de investigação mais profunda sobre o impacto da sazonalidade.

## **Conclusão**

A análise inicial revelou que fatores como inflação e sazonalidade impactam significativamente as vendas do Walmart. Os resultados obtidos podem ser utilizados para otimizar estoques e preços, além de contribuir para previsões futuras.

## **Referências**

- Kaggle Dataset: Walmart Sales .