

## **1. Questão de Negócio**

Walmart, Inc., é uma multinacional estadunidense de lojas de departamento. A companhia foi fundada por Sam Walton em 1962, incorporada em 31 de outubro de 1969 e feita capital aberto na New York Stock Exchange, em 1972. No ano de 2021, obteve um lucro de \$13.51 Bilhões. Sendo uma das principais lojas de varejo do mundo, os dados contemplam as vendas semanais de 45 lojas espalhadas pelos Estados Unidos.

Eu tenho como objetivo, fazer um levantamento do faturamento das lojas nos EUA, e determinar qual loja é a melhor opção para que seu tamanho seja estendido. Para a realização desse projeto utilizaremos a metodologia CRISP-DM, como é um projeto single estarei responsável por todas as tarefas.

## **2. Entendimento de Negócio**

O Walmart é uma das principais lojas de varejo do mundo, os dados contemplam as vendas semanais de 45 lojas espalhadas pelo Estados Unidos, durante um período de 3 anos (05 de fevereiro de 2010 até 01 de novembro de 2012). E o mesmo realiza vários eventos promocionais de descontos ao longo do ano, principalmente semanas que precedem os quatro maiores feriados, que são o Super Bowl, o Dia do Trabalho, o Dia de Ação de Graças e o Natal. Essas semanas têm um peso maior, pois acontece uma maior demanda por alguns produtos e também os eventos promocionais com descontos.

## **3. Coleta de Dados**

### **3.1 Formato dos Dados**

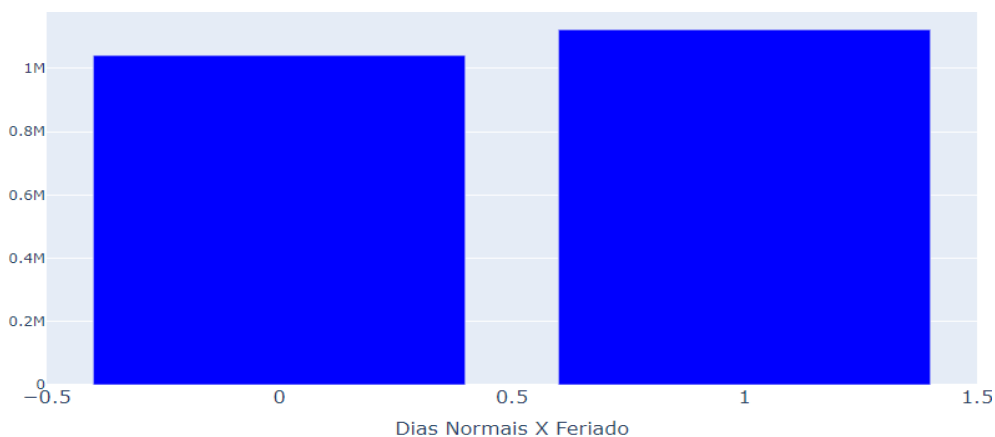
Os dados apresentados pela empresa são bem limpos e de fácil identificação, porém alguns deles devem ser formatados, de forma que facilite a análise, foram eles:

- Formatação da data: As datas estavam em formato de texto comum, então foi transformada em formato correto de data.

### 3.2 Informações Adicionais

Outras informações (Insights) que podem ser extraídos desses dados são:

- **Loja com a maior venda acumulada:** A loja de número 20 teve o maior número de vendas acumuladas (Soma de todas as semanas), outro dado relevante é que ela ultrapassou a média de faturamento geral em todas as semanas do período (143 semanas).
- **O faturamento médio semanal de cada loja:** As 5 lojas mais bem ranqueadas pelo faturamento mensal foram respectivamente: 20, 04, 14, 13 e 02.
- **Correlação do número de vendas com alguma variável:** No geral nenhuma das variáveis apresenta grande correlação com as vendas, porém pode se ressaltar, que se a taxa de desemprego sofrer um aumento significativo, ela pode sim ter um impacto nas vendas. Já que se mais pessoas ficam desempregadas, consequentemente elas não terão poder de compra, gerando assim uma correlação negativa.
- **Influência no período dos feriados:** Os grandes feriados têm um impacto real nas vendas, com é visto no gráfico abaixo. As semanas que apresentam algum feriado, tendem a ter a média de vendas elevada em relação a semanas comuns.



## 4. Limpeza de Dados

Nos dados apresentados não apresentavam valores nulos e faltantes, porém algumas variáveis que seriam irrelevantes na análise foram retiradas, foram eles:

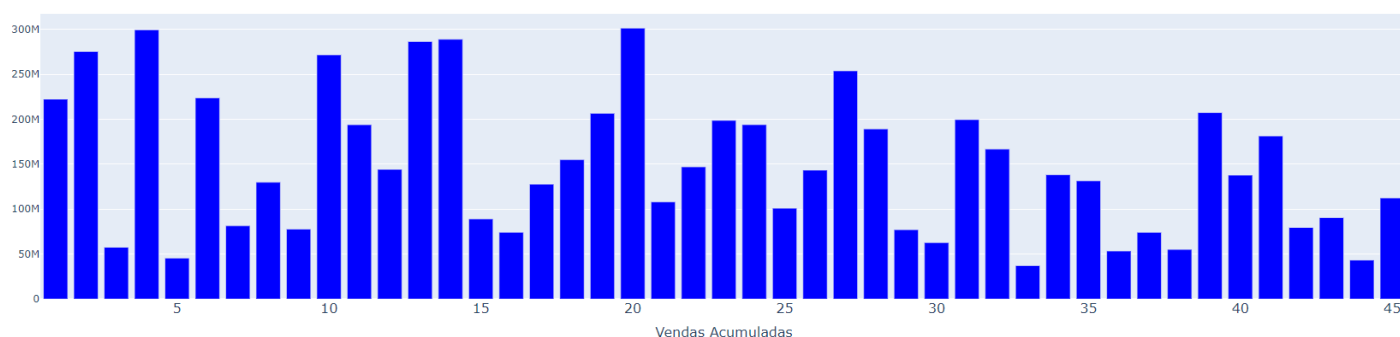
- Temperatura
- Preço do Combustível

## 5. Exploração dos Dados

### 5.1 Início da Análise

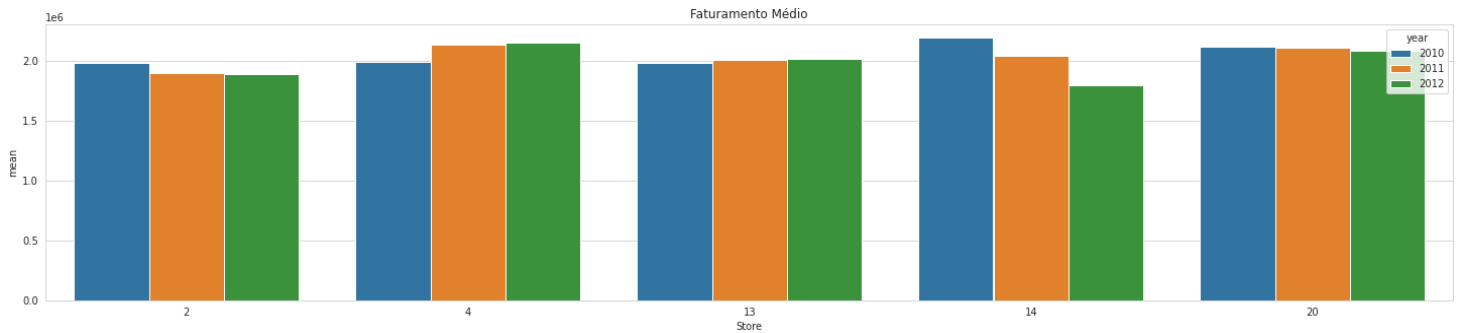
A tabela apresentada pela empresa tem um total de 45 lojas, e as vendas de cada uma delas em 143 semanas, além de algumas métricas econômicas que serão relevantes na análise. As primeiras análises a serem consideradas foram:

#### 1. A loja com maior faturamento acumulado no período



Pode se observar que a loja número 20 e número 04 estão praticamente empatadas no faturamento acumulado, porém esse gráfico traz o faturamento geral do período, ou seja, ele não leva em consideração qual ano teve maior faturamento. Isso nos leva para a próxima métrica.

## 2. Faturamento Médio em torno dos anos



Aqui pode ser visto que a loja número 04 está em uma curva ascendente no faturamento, já a loja 20 é o contrário. E analisando a tabela abaixo pode-se perceber que o índice de desemprego (Unemployment) da loja 04 é abaixo em relação a número 20, logo essa loja pode alcançar um número maior de possíveis clientes que estão empregados, aumentando ainda mais o seu faturamento.

	Store	Weekly_Sales	CPI	Unemployment
0	4	2.09e+06	128.68	5.96
1	20	2.11e+06	209.04	7.37

## Conclusão

Após a análise eu cheguei à conclusão de que a loja 04 será a escolhida para a expansão, pois a mesma já apresenta um faturamento acima da média, e considerando a média da inflação que está abaixo da sua concorrente direta, com isso, ela tem grandes chances do faturamento continuar subindo em torno dos próximos anos.

**NOME:** Samuel Nunes Ferreira

**COLAB:**

<https://colab.research.google.com/drive/1Xh3tvJbCcqBBXNMrBvtv4sAM4i87IB6b?usp=sharing>