UNIVERSIDADE PRESBITERIANA MACKENZIE

FACULDADE DE COMPUTAÇÃO E INFORMÁTICA

GRADUAÇÃO EM CIÊNCIA DE DADOS

Ana Carolina Barros - RA: 10723927

Gabriel De Túlio - RA: 10727793

Laura Rezende - RA: 10722252

Stefan Knorst - RA: 10730822

Thaylor Alcantara - RA: 10723915

**Análise Exploratória - Dados de Ações do Walmart e impactos da pandemia**

SÃO PAULO

2025

**Sumário   
  
Lista de Figuras…………………………………………...……………………………………...…3**

[**Capítulo 1 . Contexto e Objetivo do Estudo 4**](#_r991vlv94gr)

[1.1 . Objetivos Específicos 4](#_sxelbnmmasi4)

[1.2. Apresentação 5](#_cj1z2ofc30bj)

[1.3. Metodologia da Análise Exploratória 6](#_adt4bzd8rgy)

[1.4. Pipeline\* 7](#_noeocwafejeg)

[**Capítulo 2 . Referências e Aquisição dos Dados 9**](#_anysfugvs6ow)

[2.1. Fonte dos Dados 9](#_5y9576g7uovh)

[2.2. Código referência 10](#_ts1ao5vme5ce)

[**Capítulo 3. Dados 10**](#_wcum2mpqasnn)

[3.1 Origem 10](#_aplcgmih43n4)

[3.2. Fonte dos dados 11](#_shu1w4hu8d43)

[3.3. Análise exploratória em Python 11](#_feu4l9ygybyl)

[3.4. Gráfico comparativo dois períodos 12](#_8kuqmc8r409w)

[3.5. Gráfico Março de 2017 a Maio de 2019 12](#_3mjrbz35o2mx)

[3.6. Gráfico Março de 2020 a Maio de 2023 13](#_ktshp97esb89)

[3.7. Análise Quantitativa 14](#_jiagy5nszoix)

[3.8. Visualização e Escalas 14](#_h34nayyqcsy9)

[3.9. Limitações e Observações 15](#_tt9ngld8qzd3)

[3.10. Conclusão 15](#_93aqfd6bgk4b)

[**Capítulo 4 . Descrição do Dataset 15**](#_sl7rlarh4fdm)

[4.1. Análise exploratória do Data Set 16](#_crc68cgybv2w)

[**Capítulo 5. Github\* 20**](#_larmi24kjzcr)

[5.1. Página do Projeto   
Estruturação de etapas em formato “Kaban” de projeto para melhor visualização e distribuição de tarefas e colaboração dentro do grupo. 20](#_e6x7fxgpm4pi)

[5.2. Readme 20](#_bspegglfj4r)

[5.3. Repositório Link 21](#_pybkxw8k4pjv)

[5.4. Google Colab\* 22](#_pap0cjhl778d)

[**Capítulo 7. Cronograma 22**](#_4wunpp1sjrk1)

[7.1. Etapa 1 22](#_j7x4434qd8km)

[7.2. Etapa 2 23](#_abzozyxpbw2j)

[7.3. Etapa 3 23](#_5zsixo1wk0z4)

[7.4. Etapa 4 24](#_jctrgh1jm954)

[**Referências 25**](#_m6d0qxjk499l)

[**Glossário 25**](#_e01934i8p0hs)

**Lista de Figuras**

[Figura 1: Fluxograma 9](#_p1vhb7ljrkzd)

[Figura 2 : Código base - Dataset 10](#_dtvf6925lsfa)

[Figura 3: Código de análise em Python A 11](#_xu3oyufd194k)

[Figura 4: Código de análise em Python B 12](#_nfabhos2vvfk)

[Figura: 5 Gráfico comparativo dois períodos 12](#_2cxleamzyaiw)

[Figura 6: Gráfico Março de 2017 a Maio de 2019 13](#_svaz5yvx2bps)

[Figura: 7: Gráfico Março de 2020 a Maio de 2023 14](#_xg4eiekxzme1)

[Figura 8: Código de análise em R, número de datas únicas, linhas, colunas e sumário dos dados. 16](#_5hhz0gbz632g)

[Figura 9: Código de análise em R. Número de linhas com volume de ações negociadas igual a 0, resumo das datas em que dividendos foram distribuídos e resumo das datas em que houve desdobramento de ações. 17](#_358xkl2awkz1)

[Figura 10: Código de análise em R. Resumo dos dados nos períodos pré e durante a pandemia. 18](#_yqqyqbn47s)

[Figura 11: Boxplot dos valores de fechamento das ações nos dois períodos analisados. 19](#_ru3qrpjjj5ow)

[Figura 12: Página projeto 20](#_1bb0kyjn9q7s)

[Figura 13: Página Readme 20](#_45dmxmxy9xgn)

[Figura 14: Página Repositório 21](#_h968qtpqp3ok)

[Figura 15: Página Repositório Arquivos 21](#_e7s9852g4lul)

[Figura 16: Página Google Colab 22](#_9w6x2f7j1vpo)

[Figura 17: Cronograma etapa 1 23](#_oa54tluw3aou)

[Figura 18: Cronograma etapa 2 23](#_wmh4xjihuokv)

[Figura 19: Cronograma etapa 3 24](#_9q6ws7jc1tcc)

[Figura 20: Cronograma etapa 4 24](#_g508xewzltq7)

# Capítulo 1 . Contexto e Objetivo do Estudo

## 1.1 . Objetivos Específicos

O presente estudo tem como objetivo geral entender os indicadores de volume e negociação das ações\* do Walmart e quais os impactos na volatilidade\* de preços ao longo dos anos. A partir dessa análise, busca-se identificar padrões, tendências e possíveis correlações entre o período da pandemia e as oscilações no mercado acionário da empresa.

Tendo como objetivo principal analisar quais as implicações da pandemia do COVID-19 no comportamento das ações\* do Walmart. A pandemia ocorrida em 2020 representou um evento de grande impacto econômico global, buscamos compreender como as ações\* da empresa reagiram durante esse período para obter insights sobre a mudança do mercado financeiro e qual o comportamento dos investidores diante de cenários de crise.

Para alcançar o objetivo principal, serão desenvolvidas as seguintes análises:

1. Análise da Evolução do Preço das Ações\*:

1.1 Examinar a trajetória dos preços de abertura, fechamento, máxima e mínima ao longo dos anos.

1. 2 Identificar tendências de crescimento ou queda antes, durante e após a pandemia.

1. Identificação da Volatilidade\* do Mercado:

2.1 Calcular a variação percentual diária das ações\* para determinar a volatilidade.

2.2 Comparar a volatilidade\* antes e depois de março de 2020.

3. Impacto da Pandemia no Volume de Negociação:

3.1 Analisar se houve aumento ou redução no volume de transações durante a pandemia.

3.2 Correlacionar volume negociado com oscilações no preço das ações\*.

4. Identificação de Picos e Crises no Período da Pandemia:

4.1 Destacar os dias de maior valorização e depreciação das ações\*.

4.2 Relacionar esses picos com eventos externos relevantes (anúncios governamentais, lockdown\*, pacotes de estímulo econômico).

## **1.2. Apresentação**

O Walmart é uma multinacional americana fundada em 1962 por Sam Walton, consolidando-se como uma das maiores redes varejistas globais. A empresa atua em diversos segmentos, oferecendo produtos que vão desde alimentos e vestuário até eletrônicos e artigos para o lar. Com a missão de "ajudar as pessoas a economizar dinheiro para que possam viver melhor", o Walmart valoriza a liderança servidora, respeito pelo indivíduo e integridade.

Ao longo dos anos, a empresa expandiu sua presença internacionalmente e investiu fortemente no comércio eletrônico para competir com players como a Amazon. Além disso, o Walmart diversificou suas operações, oferecendo serviços financeiros, programas de fidelidade e opções de entrega rápida para atender às novas demandas dos consumidores. Atualmente, emprega aproximadamente 2,2 milhões de funcionários em todo o mundo. No Brasil, o Walmart operava sob a marca Grupo BIG, mas vendeu sua participação majoritária em 2019. Seus principais concorrentes no país incluem Carrefour, GPA (Grupo Pão de Açúcar), Via Varejo e Magazine Luiza.

O Walmart detém uma participação de mercado de 8,59% no varejo dos Estados Unidos. No segmento de comércio eletrônico, possui 6,4% de participação no mercado norte-americano. Já no setor de alimentos, o Walmart lidera com 23,6% de participação de mercado, em 2023, consolidou-se como o principal varejista de alimentos no país.

A pandemia de COVID-19, iniciada em março de 2020, representou um dos maiores desafios da história do varejo global. O Walmart, sendo um dos principais fornecedores de produtos essenciais, teve um aumento expressivo na demanda por itens básicos, como alimentos, produtos de limpeza e artigos de higiene pessoal. A necessidade de adaptação às novas dinâmicas de consumo impulsionou o crescimento do e-commerce\* da empresa, que registrou um aumento significativo no volume de vendas online. Por outro lado, a pandemia também trouxe desafios operacionais, como problemas na cadeia de suprimentos, aumento nos custos logísticos e mudanças no comportamento do consumidor.

Diante desse contexto, analisar os dados históricos das ações\* do Walmart permite entender quais fatores foram determinantes para as variações no preço durante a pandemia. Ao estudar esse impacto, é possível obter insights sobre a resposta da empresa às crises e as tendências do mercado varejista em momentos de incerteza econômica.

## **1.3. Metodologia da Análise Exploratória**

A análise exploratória de dados (EDA) será conduzida em várias etapas para garantir um entendimento abrangente dos impactos da pandemia da COVID-19 nas ações\* do Walmart. Inicialmente, o dataset será obtido da plataforma Kaggle\* e carregado no Google Colab\*, ambiente de análise, utilizando a biblioteca Pandas para manipulação dos dados. Em seguida, será realizada uma etapa de pré-processamento, na qual as colunas do dataset serão inspecionadas para identificar se há valores ausentes, inconsistências e a necessidade de conversões. Com os dados devidamente tratados, será feita a análise temporal da evolução do preço das ações\*, por meio de gráficos de séries temporais que permitam visualizar a trajetória dos preços de abertura, fechamento, máxima e mínima ao longo dos anos, com destaque para o início da pandemia em março de 2020. Além disso, será calculada a volatilidade\* dos preços, utilizando a variação percentual diária das ações\*, e comparando os períodos pré e pós-pandemia para avaliar possíveis alterações no comportamento do mercado. Em seguida, será analisado o volume de negociação das ações\* para identificar possíveis correlações entre a liquidez do ativo e a volatilidade\* dos preços. Para aprofundar a investigação, serão identificados os dias com as maiores altas e quedas nas ações\*, permitindo relacioná-los a eventos econômicos ou decisões estratégicas da empresa. Por fim, os resultados obtidos serão interpretados sob o contexto econômico e financeiro, comparando os padrões encontrados antes, durante e após a pandemia, observando os períodos de estabilidade, de forma a compreender como o mercado reagiu a esse evento e quais foram os impactos para os investidores.

## **1.4. Pipeline**\*

1. Coleta de Dados:

- Extração dos dados históricos das ações\* do Walmart via API\* do Yahoo Finance\*, utilizando a biblioteca `yfinance` em Python.

- Dataset também disponível no Kaggle\*, permitindo acesso alternativo e comparação de fontes.

- Período analisado: 1972 a 2025, com foco em eventos entre 2019 e 2021, marcados pela pandemia da COVID-19.

- Ambiente de desenvolvimento: Google Colab\*, com suporte ao versionamento via GitHub\*.

2. Pré-Processamento:

- Verificação de integridade: conferência de dados ausentes, duplicados e inconsistentes.

Exemplo: 53 registros com volume igual a 0 entre 1972-1979, tratados como não problemáticos.

- Conversão e padronização de formatos:

• Date transformada em tipo datetime.

• Padronização de separadores decimais e formatação monetária.

- Criação de colunas auxiliares:

• Daily Return: variação percentual diária ((Close - Open)/Open).

• Volatility: desvio padrão\* móvel.

- Remoção de colunas não necessárias para a análise central (se houver).

3. Análise Exploratória de Dados (EDA):

a) Análise Temporal dos Preços:

- Visualização da evolução dos preços de abertura, fechamento, máxima e mínima.

- Comparação de tendências antes, durante e após a pandemia.

b) Volatilidade\* de Mercado:

- Cálculo da variação percentual diária.

- Análise de volatilidade\* agregada por mês/ano.

- Comparativo entre períodos pré-março/2020 e pós-março/2020.

c) Volume de Negociação:

- Estudo do comportamento do volume de transações.

- Identificação de picos de negociação e correlação com mudanças nos preços.

d) Identificação de Picos e Crises:

- Análise dos maiores ganhos e perdas diárias.

- Relacionamento com eventos externos: Lockdown\*, pacotes econômicos, comunicados oficiais.

e) Estatísticas Descritivas:

- Média, mediana, desvio padrão\*, valores máximos/mínimos.

- Identificação de outliers\* e possíveis explicações.

4. Geração de Insights:

- Interpretação dos padrões identificados.

- Construção de hipóteses baseadas nos dados.

5. Visualização e Storytelling\*:

- Construção de gráficos e paineis: séries temporais, histogramas, boxplots\* e heatmaps\*.

- Elaboração de storytelling\* com narrativa baseada nos dados.

- Entrega final em formato de apresentação e documentação do projeto.

6. Ferramentas e Linguagens Utilizadas:

- Python (pandas, matplotlib, seaborn, yfinance)

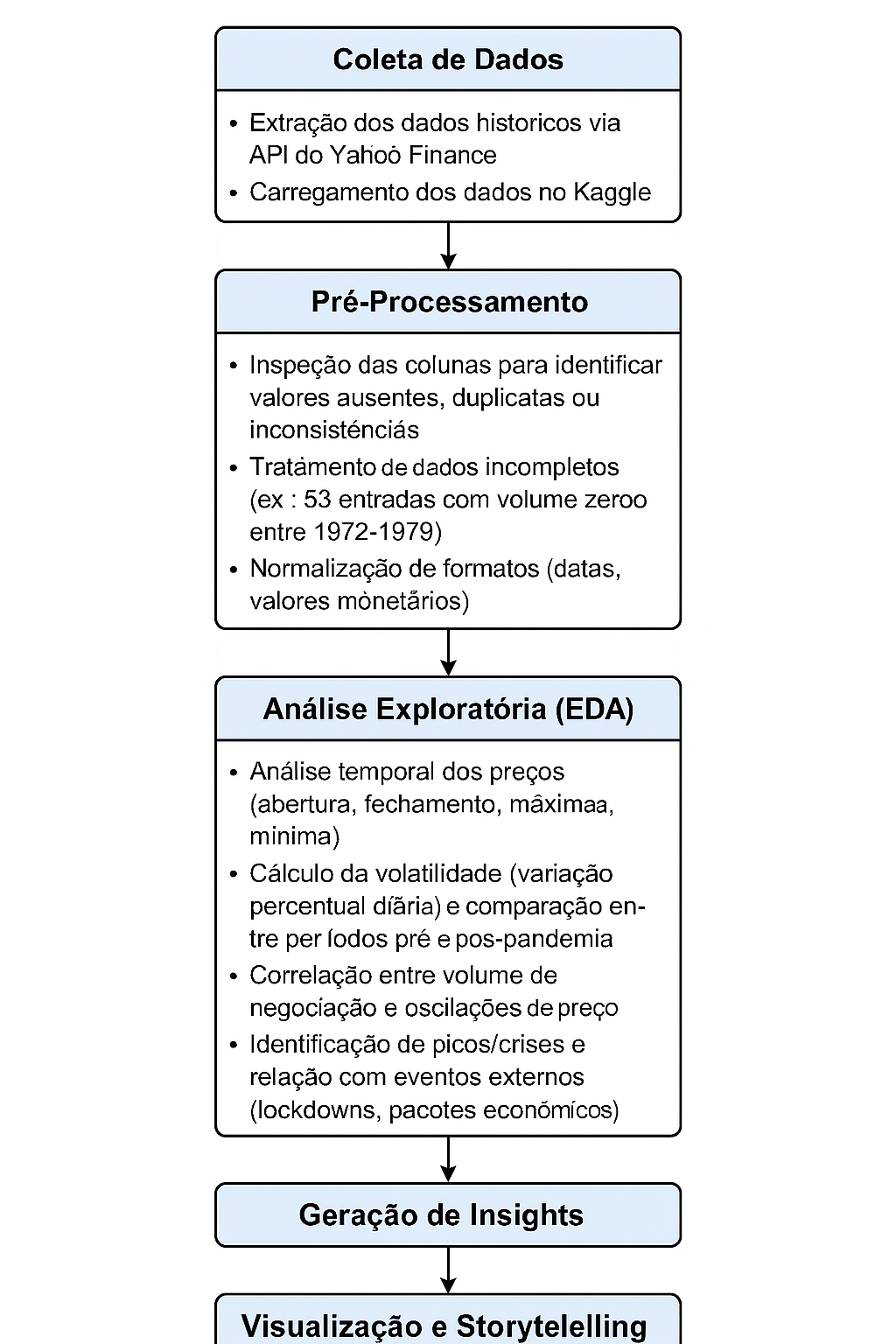
- Google Colab\* (análises e experimentos)

- GitHub\* (repositório do projeto)

- Kaggle\* (repositório complementar)

- Excel/R (complementar, se necessário para estatísticas descritivas).  
  
[Link para fluxograma](https://drive.google.com/file/d/1shMWVa4TiZ2hrMXfftbMl_UvXBSiX36Y/view)

### Figura 1: Fluxograma



Fonte: Print de tela de arquivo gerado via Canva.

# Capítulo 2 . Referências e Aquisição dos Dados

## **2.1. Fonte dos Dados**

O conjunto de dados foi disponibilizado no [Kaggle](https://www.kaggle.com/datasets/meharshanali/walmart-stocks-data-2025?resource=download)\* por Mehar Shan Ali, um criador de conteúdo de dados interessado em análise financeira e aprendizado de máquina. O autor frequentemente compartilha datasets relevantes para a comunidade de ciência de dados, dessa forma proporcionando recursos para estudos, modelagem e testes de algoritmos de previsão de mercado.

Restrições de Uso: Os dados estão em domínio público, podendo ser utilizados livremente por qualquer um;

Licença: [CC0: Public Domain](https://creativecommons.org/publicdomain/zero/1.0/);

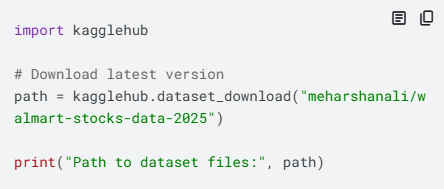
Aquisição dos dados: podem ser extraídos pelo próprio Kaggle\* através do KaggleHub\* ou até mesmo com a linguagem Python por uma IDLE, como VS Code, por exemplo;

Período de Coleta: Os preços das ações\* entre 1972 e 2025 foram coletados utilizando uma API\* do Yahoo Finance;

## **2.2. Código referência**

Código base fornecido pela Kaggle\* para extração dos dados, para formulação de código em Phyton análise.

### Figura 2 : Código base - Dataset



Fonte: Print de tela site Kaggle.

# Capítulo 3. Dados

## 3.1 Origem

Os dados sobre as ações\* da Walmart foram obtidos diretamente do Yahoo Finance\* por meio da API\* Python yfinance. Essa API\* permite a extração automatizada de informações financeiras, garantindo acesso a dados his\*tóricos precisos sobre o desempenho da empresa no mercado de ações.

O conjunto de dados inclui preços diários das ações\*, bem como indicadores financeiros relevantes, permitindo análises detalhadas sobre a variação do ativo ao longo do tempo. Essa abordagem assegura a confiabilidade das informações, uma vez que o Yahoo Finance\* é uma das principais fontes de dados do mercado financeiro global.

## **3.**2**. Fonte dos dados**

Pesquisas realizadas para melhor contextualização dos temas ligados a companhia escolhida para o trabalho Walmart. Extração de informações e banco de dados públicos selecionado para análise e busca por informações de como estruturar melhor a análise:

[Yahoo Finance](https://finance.yahoo.com)\*

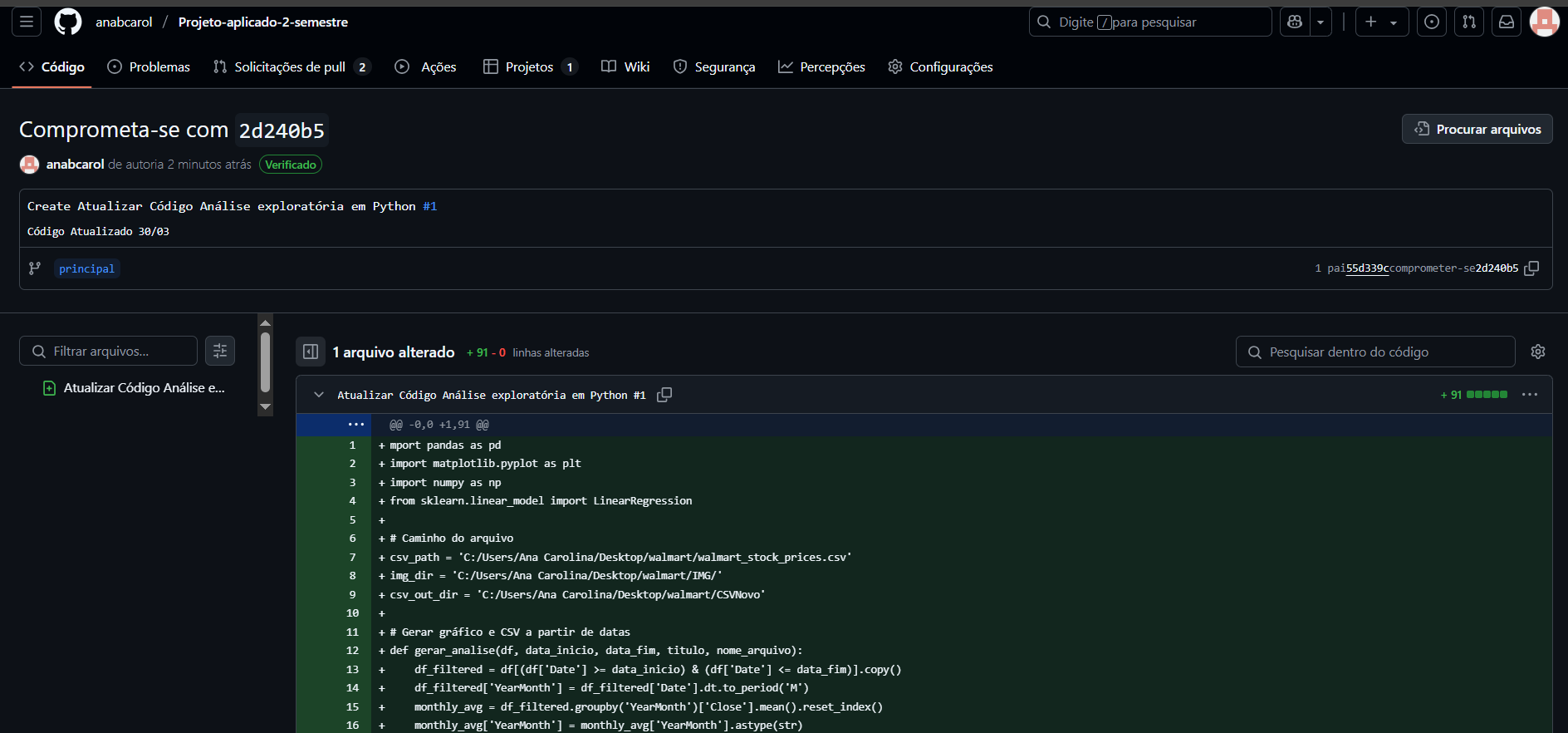
Método de coleta: Python API\* yfinance

Diretamente do [Kaggle](https://www.kaggle.com/datasets/meharshanali/walmart-stocks-data-2025?resource=download)\*

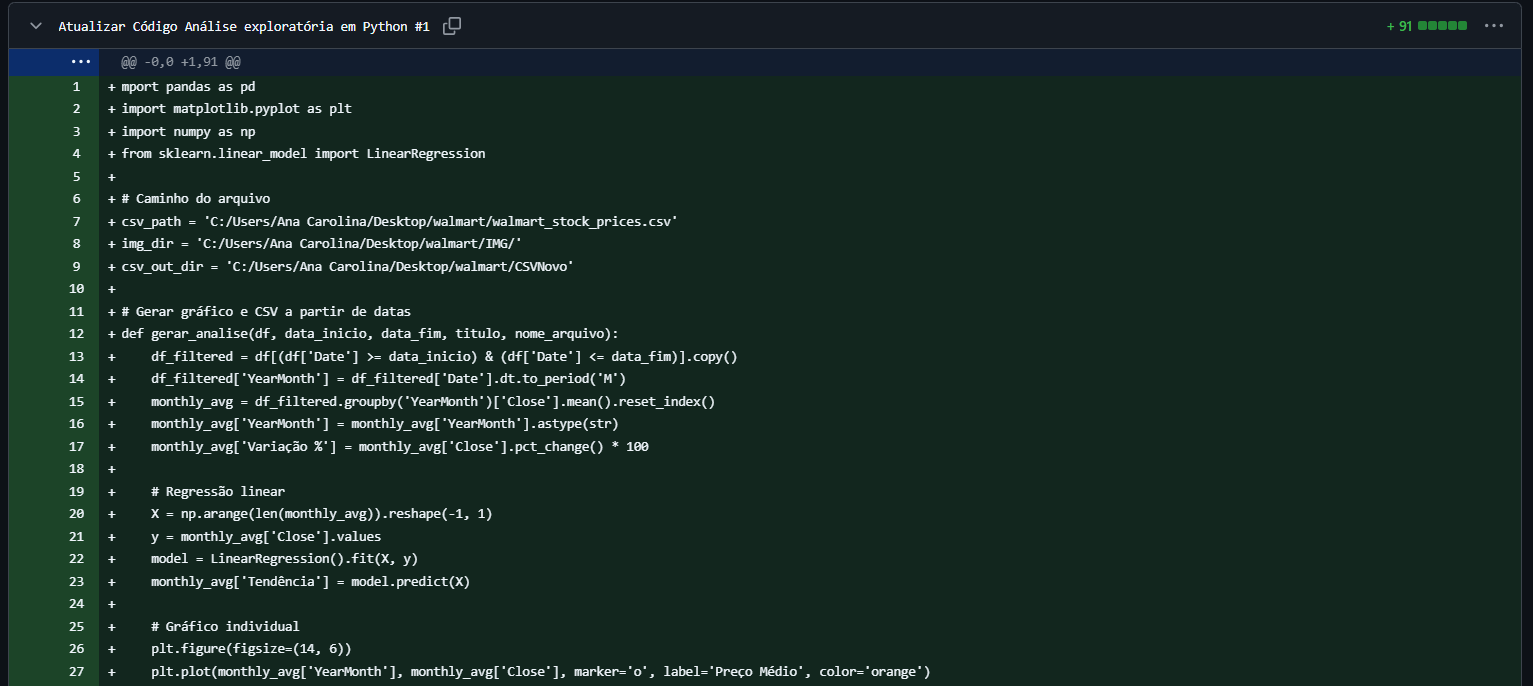
## **3.**3**. Análise exploratória em Python**

Código Criado em Python para análise de dados a fim de comparar a variação e preços de ações\* do período de: Março de 2017 a Maio de 2019 (pré pandemia) e Março de 2020 a Maio de 2023 (pandemia).  
  
Código em Python inserido na plataforma Github\*:

### Figura 3: Código de análise em Python A

  
Fonte: Print de tela do repositório na plataforma Github.

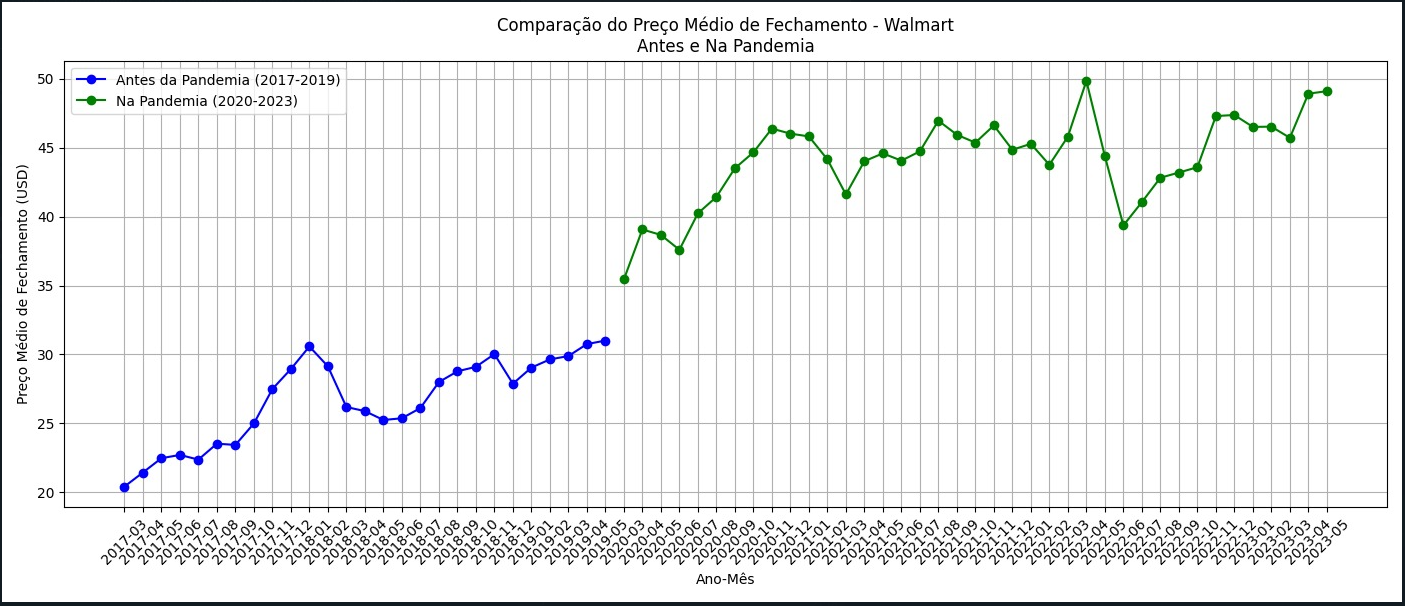
### Figura 4: Código de análise em Python B

  
Fonte: Print de tela do repositório na plataforma Github.

## 3.4. Gráfico comparativo dois períodos

Aqui conseguimos ter uma boa leitura dos dois cenários. Pré pandemia com estabilidade e variações mais sutis com períodos de estabilidade maior.   
No período Pandêmico os preços das ações\* apresentaram variações acentuadas entre altas e baixas nos valores em períodos muito curtos de tempo.

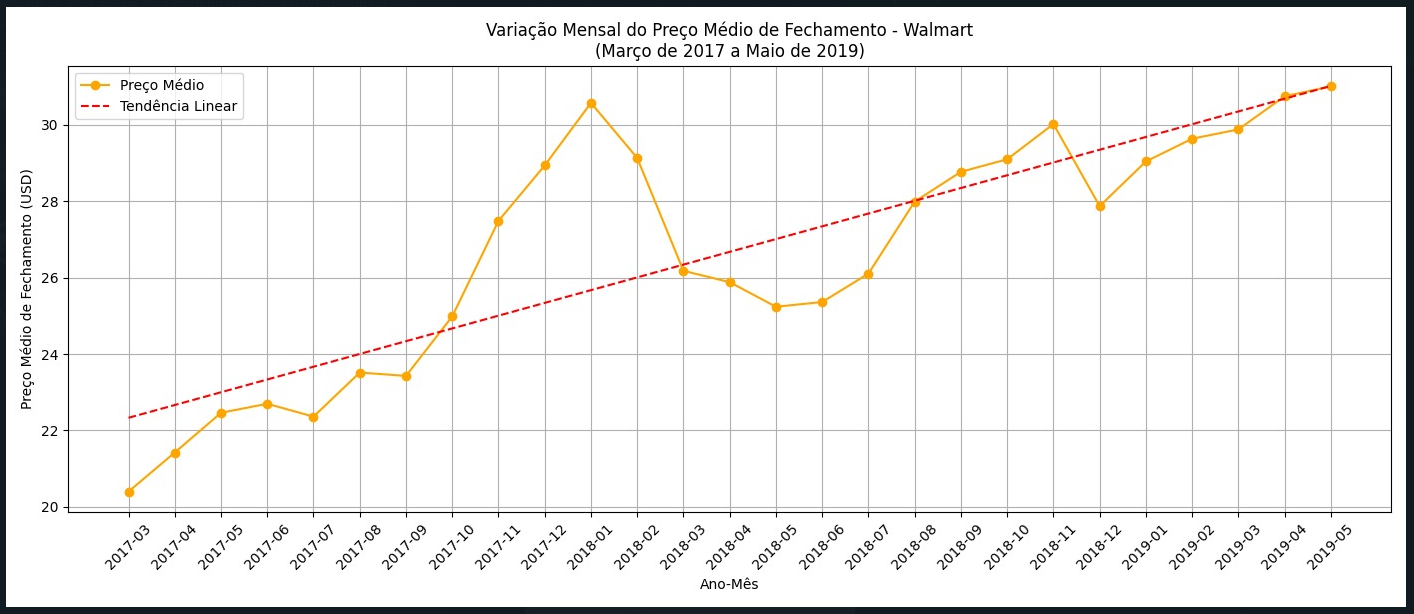
### Figura: 5 Gráfico comparativo dois períodos

  
Fonte: Print de tela do repositório na plataforma Github, gerado via PyCharm.

## 3.5. Gráfico Março de 2017 a Maio de 2019

Estabilidade: Os preços das ações\* do Walmart apresentaram variações moderadas (entre USD 20 e USD 35), refletindo um cenário de crescimento gradual alinhado ao desempenho histórico da empresa.   
Fatores Influentes: Expansão do e-commerce\*, estratégias de preços competitivos e estabilidade macroeconômica nos EUA (ex.: taxa de desemprego abaixo de 4% em 2019).

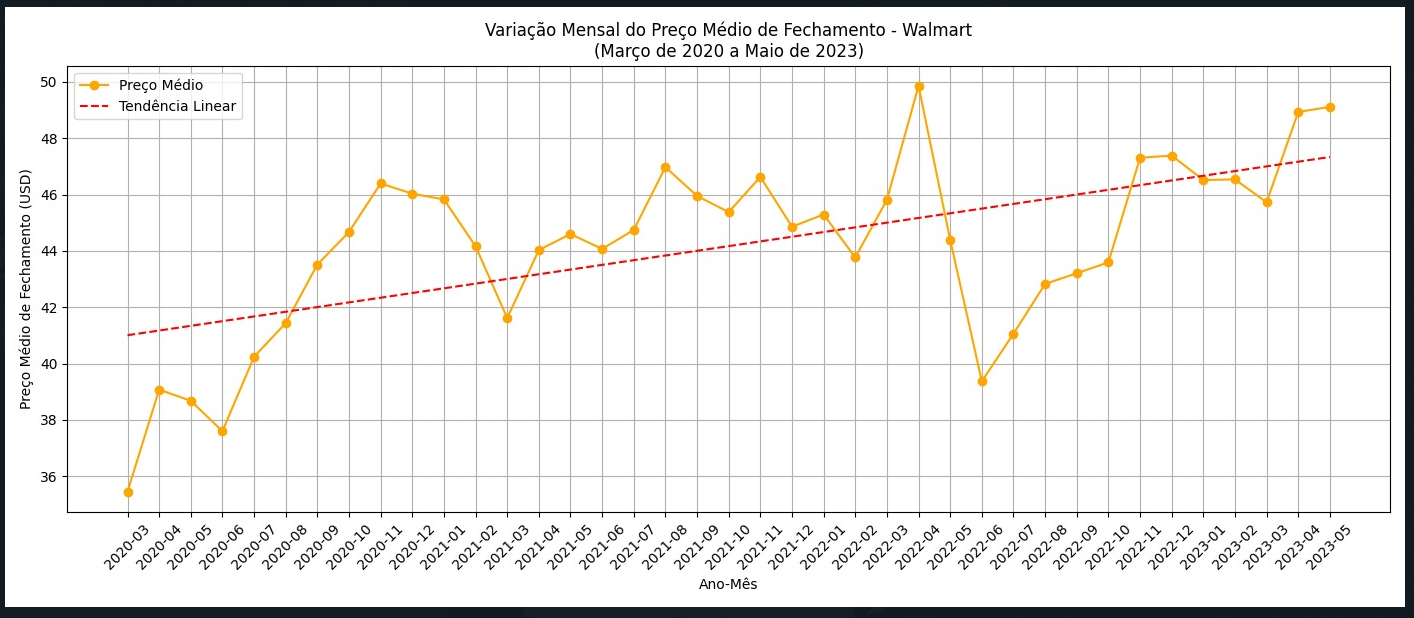
### Figura 6: Gráfico Março de 2017 a Maio de 2019

  
Fonte: Print de tela do repositório na plataforma Github, gerado via PyCharm.

## 3.6. Gráfico Março de 2020 a Maio de 2023

Volatilidade\* Acentuada: Oscilações bruscas (USD 25 a USD 50), com picos de alta (ex.: USD 50 em 2021) e quedas abruptas (ex.: USD 25 em 2020).   
Eventos Relevantes: Q1 2020: Queda inicial devido ao pânico do mercado no início da pandemia.   
Q3 2020–2021: Recuperação impulsionada pelo aumento da demanda por produtos essenciais e investimentos em logística.   
2022–2023:Pressões inflacionárias e custos operacionais elevados impactam a rentabilidade.

### Figura: 7: Gráfico Março de 2020 a Maio de 2023

  
Fonte: Print de tela do repositório na plataforma Github, gerado via PyCharm.

## 3.7. Análise Quantitativa

Volatilidade\*:  
Pré-Pandemia: desvio padrão\* de ~USD 5, indicando baixa volatilidade.   
Pandemia: desvio padrão\* de ~USD 12, refletindo incertezas econômicas e comportamentos especulativos.

Correlações:   
Lockdown\* (2020): Queda de 30% no preço das ações\* no primeiro semestre de 2020.   
Pacotes de Estímulo (2021): Alta de 60% após injeções de capital pelo governo americano.

## 3.8. Visualização e Escalas

Eixo Y: Valores em USD, sem ajuste para inflação\*. Nota-se que a escala inicia em USD 20, o que amplifica visualmente as variações.

Destaques:  
Máximo Histórico (USD 50): Alcançado em 2021, associado ao crescimento recorde do e-commerce\* (aumento de 79% nas vendas online).   
Mínimo Pandêmico (USD 25): Março de 2020, coincidindo com o início das restrições globais.

## 3.9. Limitações e Observações

Projeções vs. Dados Reais: Os dados até 2023 são históricos, mas o dataset inclui projeções até 2025.   
Fatores Externos: Não foram ajustados efeitos da inflação\* (9,1% nos EUA em 2022) ou flutuações cambiais.

## 3.10. Conclusão

Risco-Retorno\*: O período pandêmico evidenciou maior risco (volatilidade), mas também oportunidades de retorno elevado (ex.: alta de 100% entre 2020 e 2021).   
Resiliência do Walmart: A empresa mostrou capacidade de adaptação a crises, com recuperação mais rápida que concorrentes do varejo.

# **Capítulo 4 . Descrição do Dataset**

Este conjunto de dados contém dados históricos de preços e volume de ações\* do Walmart, obtidos diretamente do Yahoo Finance\* usando a API\* de Python *yfinance*. Os dados abrangem preços diários de ações\* e incluem vários indicadores financeiros importantes de 1972 a 2025.

O dataset possui 8 colunas e 13233 linhas de dados. Cada linha corresponde a uma data diferente e cada coluna corresponde a um dado diferente:

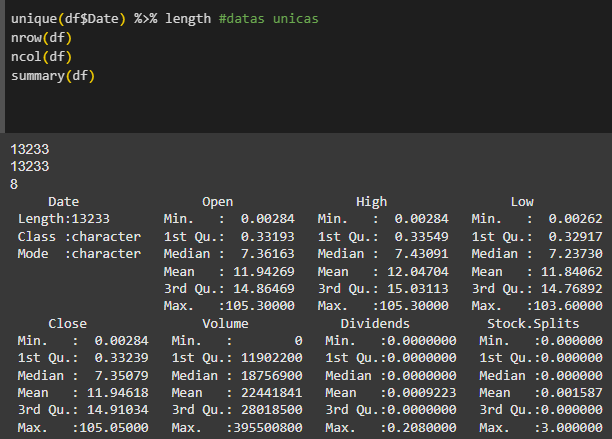
* Date (Data, coluna 1) - Cada entrada corresponde a um dia de negociação.
* Open (Abertura, coluna 2) - Preço registrado na abertura do mercado.
* High (Alto, coluna 3) - Maior preço registrado no dia correspondente.
* Low (Baixo, coluna 4) - Menor preço registrado no dia correspondente.
* Close (Fechamento, coluna 5) - Preço registrado no fechamento do mercado.
* Volume (Volume, coluna 6) - Número total de ações\* negociadas no dia.
* Dividends (Dividendos\*, coluna 7) - Pagamento feito aos acionistas.
* Stock Splits (Desdobramento de ações\*, coluna 8) - Registra casos de desdobramentos de ações\*, aumentando o número de ações\* disponíveis e diminuindo os preços das ações\*.

## 4.1. Análise exploratória do Data Set

Aqui o objetivo é o de investigar todos os dados que o dataset investiga de forma concreta como: número de colunas e linhas, itens descritos, valores máximos e mínimos, presença ou não de outliers\*. Escolhemos usar a linguagem R para o processo.

Em uma primeira vista, podemos ver que todas as datas no dataset são únicas, ocorrendo no período entre 25 de outubro de 1972 e 21 de fevereiro de 2025. Todos os valores de abertura, fechamento, valores altos e baixos são positivos e maiores que 0,00262; indicando não haver nenhum valor aberrante.

### Figura 8: Código de análise em R, número de datas únicas, linhas, colunas e sumário dos dados.

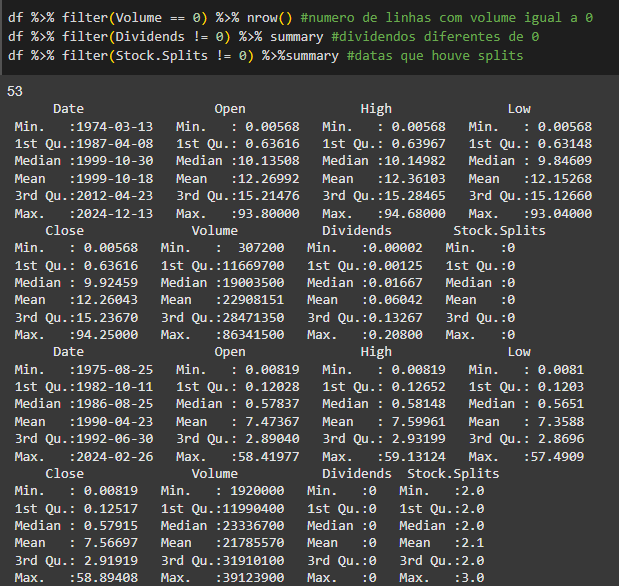
  
Fonte: Print de tela do Google Colab.

Existem 53 entradas em que o volume de ações\* negociadas no dia são iguais a 0, todos entre os anos 1972 e 1979. Como os valores registrados de abertura, fechamento, alto e baixo são todos iguais, acreditamos não haver nada de errado nessas entradas.

Os valores dos dividendos\* possuem apenas 202 entradas diferentes de 0, espalhadas entre os anos 1974 e 2024. Eles possuem um valor mínimo de 0,00002, máximo de 0,20800 e uma média de 0,06042. Isso indica que o pagamento de dividendos\* não é comum para essa ação.

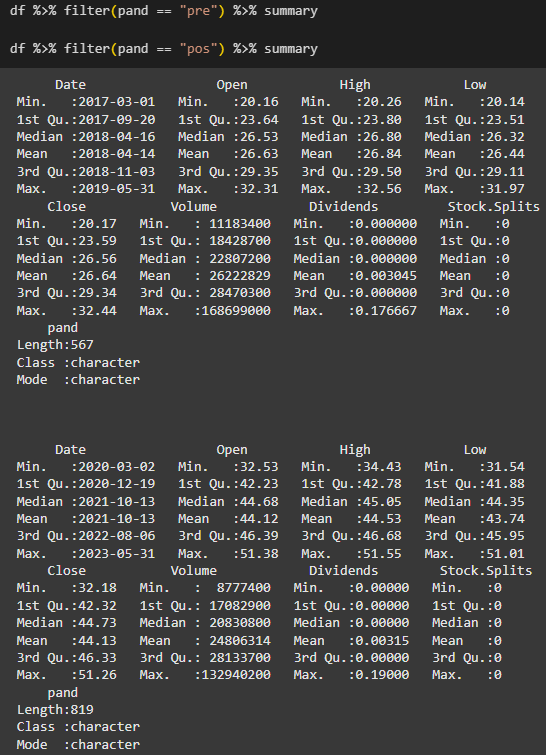
Por último, houveram 10 desdobramentos das ações\* no tempo estudado. 9 deles foram desdobramentos 2 para 1, em que uma ação se torna duas ações\* (todos entre 1975 e 1999). O último desdobramento foi de 3 para 1, em que uma ação se torna 3, que ocorreu em 2024.

### Figura 9: Código de análise em R. Número de linhas com volume de ações negociadas igual a 0, resumo das datas em que dividendos foram distribuídos e resumo das datas em que houve desdobramento de ações.

  
Fonte: Print de tela do Google Colab.

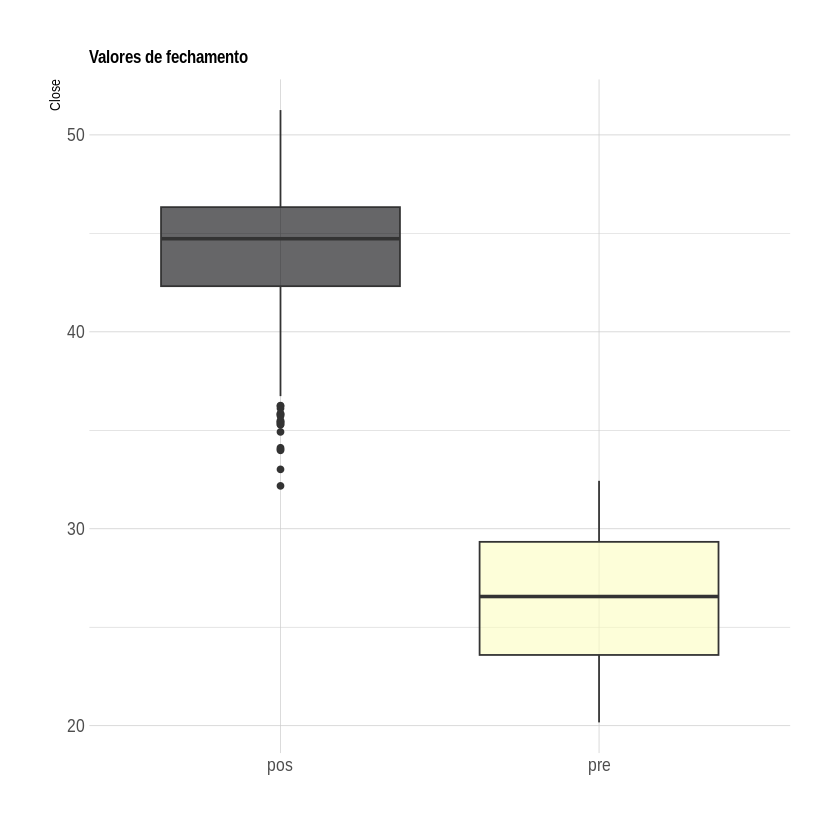
Após uma análise primária, podemos olhar especificamente para os dois períodos a serem estudados, Março de 2017 a Maio de 2019 e Março de 2020 a Maio de 2023. Para facilitar o estudo desses períodos, podemos incluir uma nova coluna nos nossos dados, denominada “pandemia”, que indica se a data está no período pré-pandemia, pós-pandemia ou nenhum dos dois. Nos dois períodos de interesse, vemos que existem 567 valores durante o período pré-pandemia e 819 valores durante a pandemia. Além disso, não houve desdobramento das ações\* no período, portanto qualquer variação abrupta dos valores não pode ser atribuída a isso.

### Figura 10: Código de análise em R. Resumo dos dados nos períodos pré e durante a pandemia.

  
Fonte: Print de tela do Google Colab.

Finalmente, podemos fazer um boxplot\* dos dados, onde podemos observar se existe algum valor aberrante. Assim, para os valores pré-pandemia, observamos que os valores estão homogeneamente distribuídos, enquanto que para os valores pós-pandemia, existem alguns valores que poderiam ser considerados outliers\*. Observando a tabela de dados, vemos que todos esses valores estão agrupados no começo da pandemia, indicando uma demora na valorização das ações\*, possivelmente devido à pandemia.

### Figura 11: Boxplot dos valores de fechamento das ações nos dois períodos analisados.

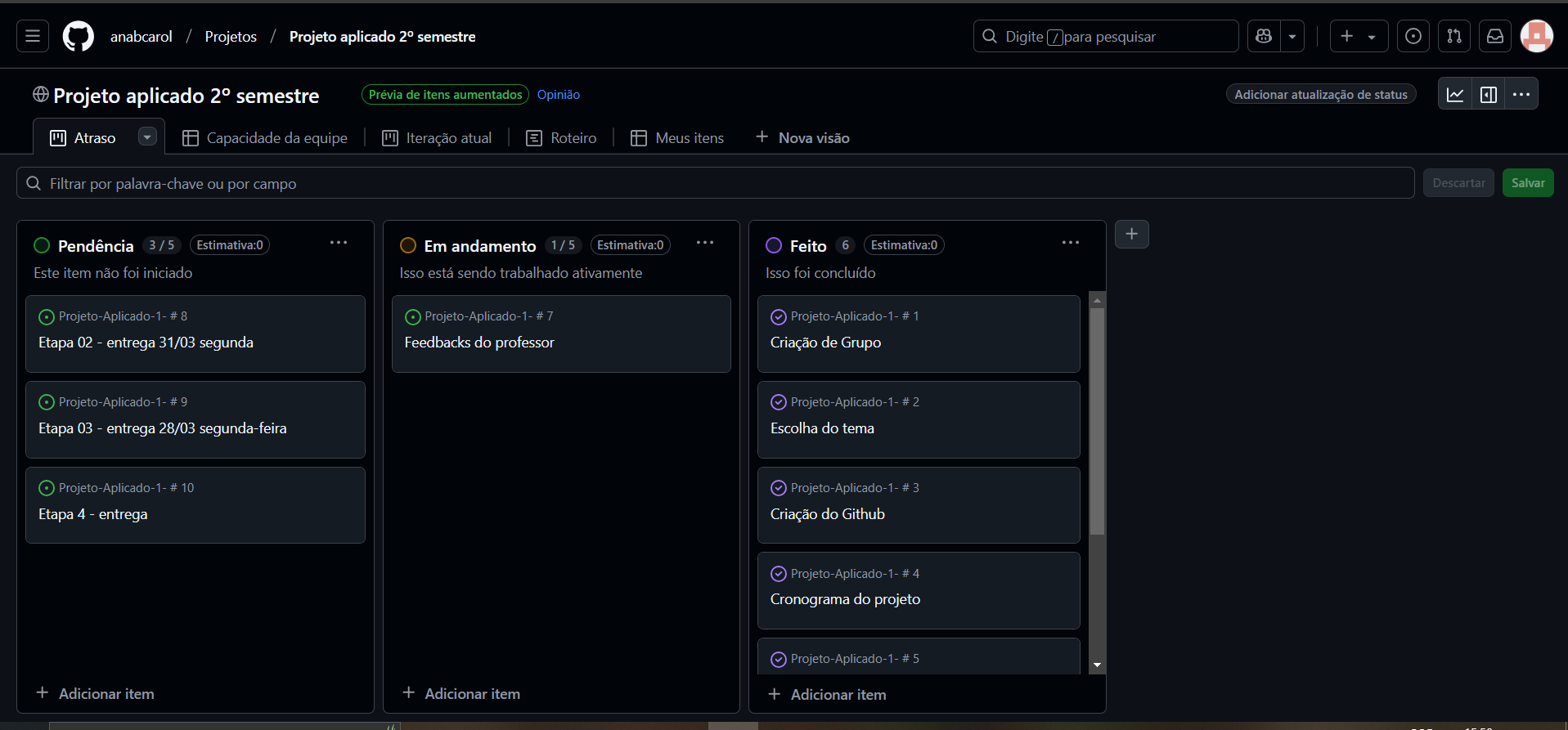
  
Fonte: Box Plot obtido no Google Colab.

# Capítulo 5. Github\*

Criação de página para gerenciamento de projetos, compartilhamento de documentos, repositório códigos e arquivos no [Github](https://github.com/anabcarol/Projeto-aplicado-2-semestre)\* (plataforma de compartilhamento e visualização de códigos e projetos)

## **5.1. Página do Projeto** Estruturação de etapas em formato “Kaban” de projeto para melhor visualização e distribuição de tarefas e colaboração dentro do grupo.

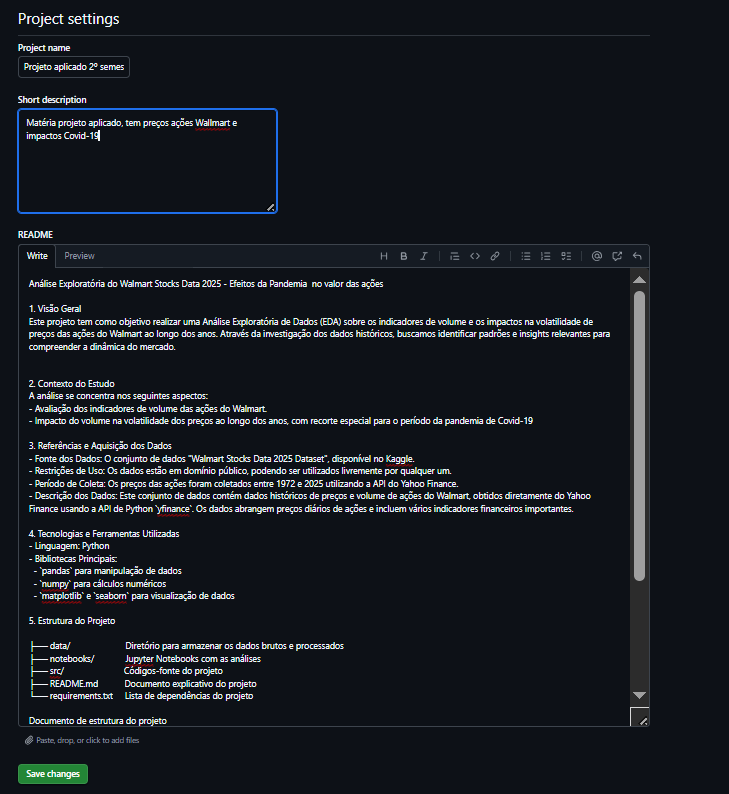
### Figura 12: Página projeto

  
Fonte: Print de tela do repositório na plataforma Github.

## **5.2. Readme**

Apresentação inicial da proposta e metodologia do trabalho. Para consulta geral dentro da plataforma Github\*, com link principais para acessos aos principais documentos e informações relevantes para o projeto.

### Figura 13: Página Readme

  
Fonte: Print de tela do repositório na plataforma Github.

## 5.3. Repositório [Link](https://github.com/anabcarol/Projeto-Aplicado-1-)

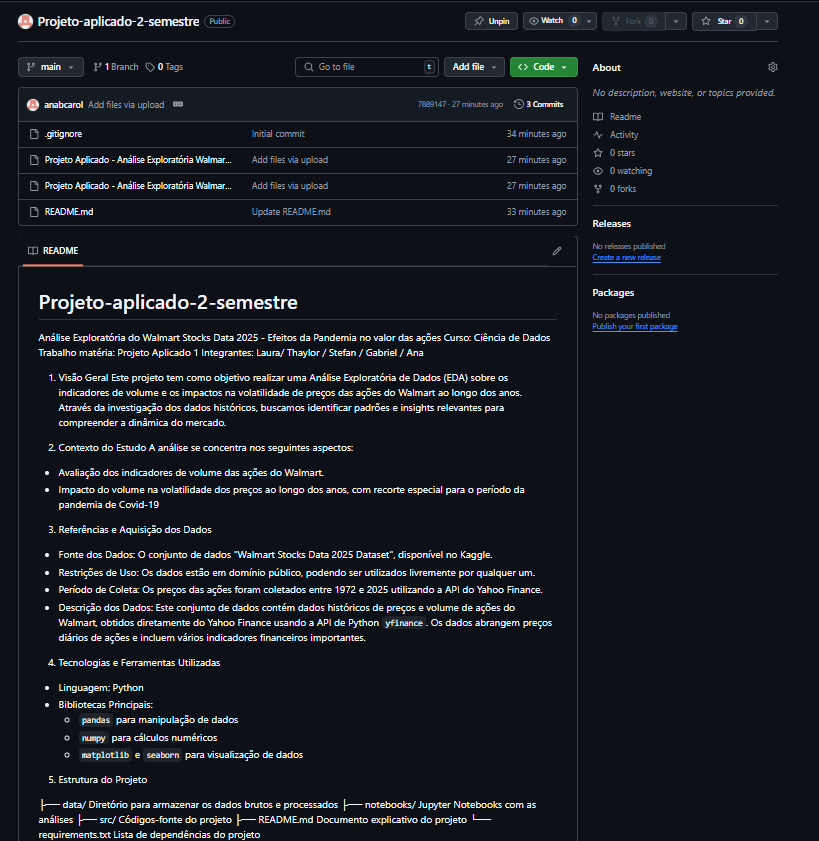
Repositório do projeto, onde todos os arquivos de documentos e códigos serão enviados. As interações com colaboração entre os integrantes do grupo ocorrem por esse espaço.

### Figura 14: Página Repositório

## 

Fonte: Print de tela do repositório na plataforma Github.

### Figura 15: Página Repositório Arquivos

  
Fonte: Print de tela do repositório na plataforma Github.

## 5.4. Google Colab\*

Criação de página compartilhada para criação do código para desenvolvimento do código em python. [Link](https://colab.research.google.com/drive/1BVU-DsnR8biC8RhUBw-ZMWF53eBP8wDr?usp=sharing)    
Plataforma Goggle que permite escrita e teste de códigos em diferentes linguagens. Criamos um espaço compartilhado dentro da plataforma para melhor troca e interação. com código formulados é possível transferi-los para nosso repositório no [Github](https://github.com/anabcarol/Projeto-aplicado-2-semestre)\*.

### Figura 16: Página Google Colab



Fonte: Print de tela do repositório na plataforma Google Colab.

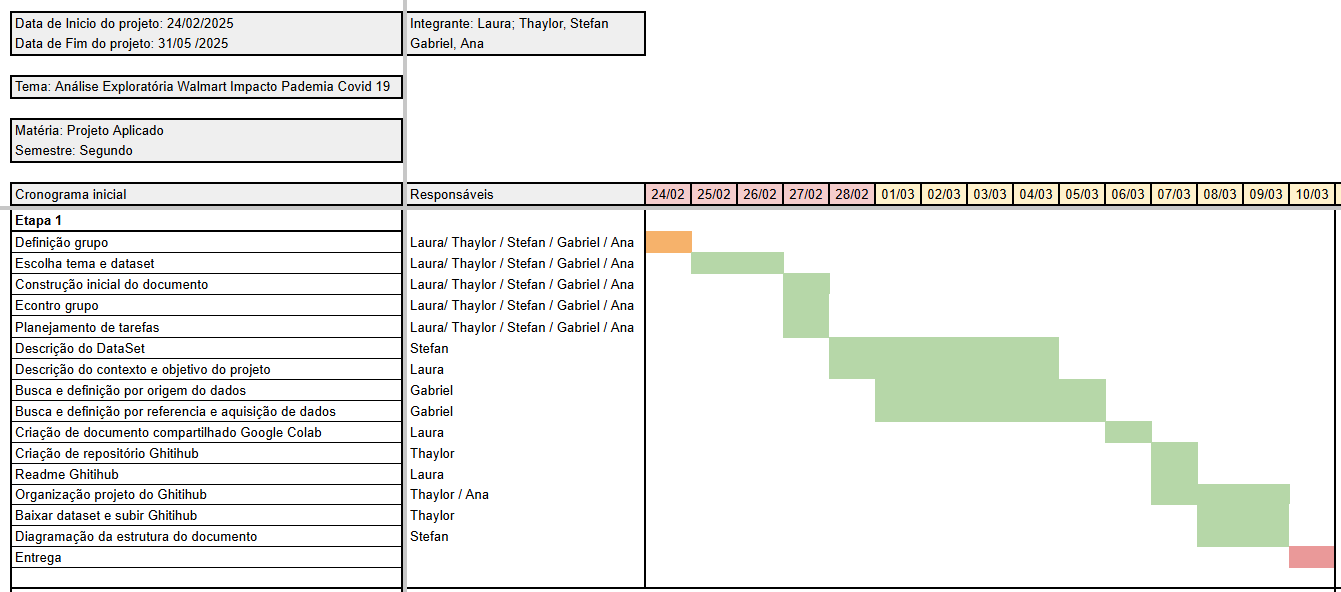
# Capítulo 7. Cronograma

Propostas de cronograma com previsão para cada etapa do projeto, com descritivo de tarefas e responsáveis com tema com definição de prazos:   
 [Link planilha](https://docs.google.com/spreadsheets/d/1gFoNW3ZzSZPmSNMNsbDJu5Zv3R5ipOifsELXW9G9KJk/edit?usp=sharing)

## 7.1. Etapa 1

Estrutura base do projeto com definição de grupos, integrantes, qual empresa e dataset serão usados e qual o tema da pesquisa, com divisão de tarefas iniciais para organização do documento.

### Figura 17: Cronograma etapa 1

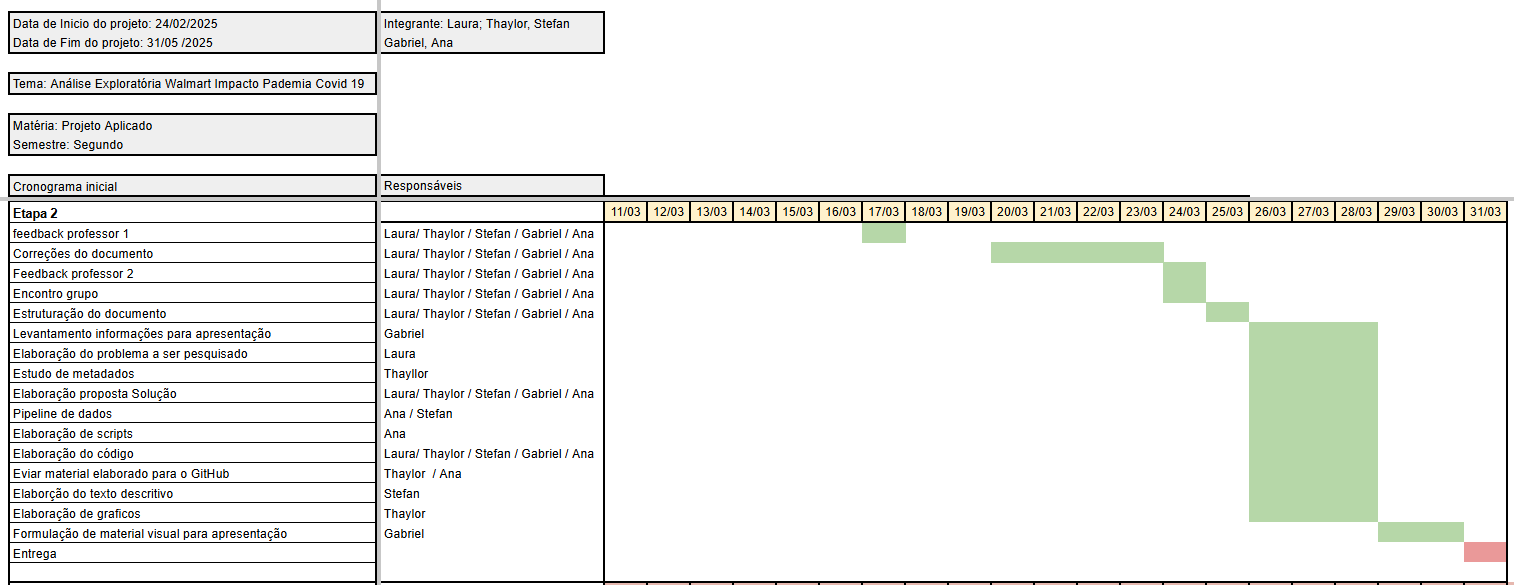


Fonte: Print de tela do Excel.

## 7.2. Etapa 2

Nessa etapa devemos aprofundar a organização do documento com adição de itens relacionados a apresentação da empresa e do problema, adicionar a proposta analítica inserida no documento, scripts\* da análise exploratória em python e apresentação de metadados.

### Figura 18: Cronograma etapa 2

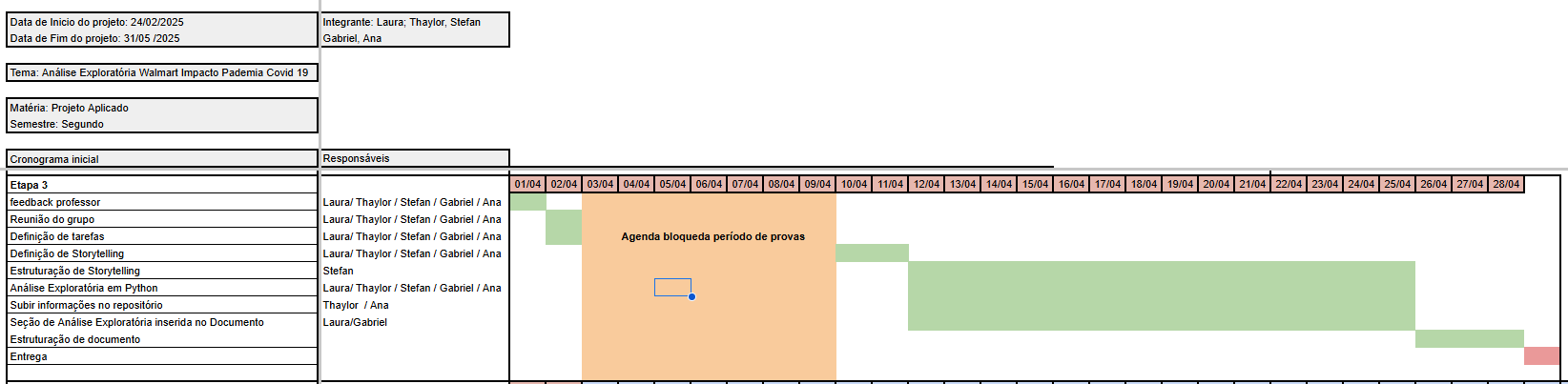


Fonte: Print de tela do Excel

## 7.3. Etapa 3

Na etapa 3 realizaremos um esboço do Storytelling\* descrevendo a narrativa que o grupo pretende fazer para a gravação final. Deve conter Scripts\* da Análise Exploratória em Python e O texto apresentando uma descrição das variáveis, complementando os metadados.

### Figura 19: Cronograma etapa 3



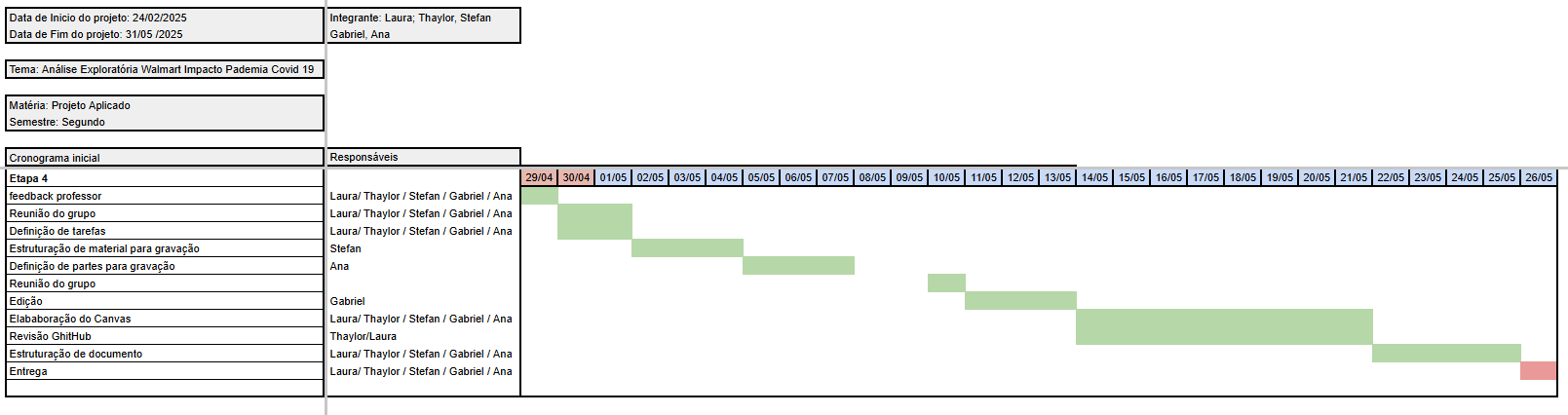
Fonte: Print de tela do Excel

## 

## 7.4. Etapa 4

Nessa etapa todas as informações são compiladas e organizadas em uma apresentação de vídeo feita pelo grupo e o documento final deve ser apresentado.

### Figura 20: Cronograma etapa 4



Fonte: Print de tela do Excel

# 

# Referências

[The Business Model Analiyst](https://businessmodelanalyst.com/pt/walmart-business-model/?srsltid=AfmBOor1-149dEFG_HPHXgFGKe7mC9VIsFyWMfiDP_JKte40dJmcauvw&utm_source=chatgpt.com)

[SEC](https://www.sec.gov/Archives/edgar/data/104169/000010416921000033/wmt-20210131.htm?utm_source=chatgpt.com)

[CNN Brasil](https://www.cnnbrasil.com.br/economia/macroeconomia/carrefour-gpa-e-via-varejo-confira-o-ranking-dos-maiores-varejistas-do-brasil/?utm_source=chatgpt.com)

<https://capitaloneshopping.com/research/walmart-statistics>

<https://www.statista.com/topics/1451/walmart/>

<https://www.supermarketnews.com/independents-regional-grocers/walmart-kroger-costco-make-top-three-grocery-retailers-list>

# Glossário

**Ações**: Parte do capital de uma empresa que é negociada na bolsa de valores. Quem compra uma ação se torna sócio da empresa e pode lucrar com sua valorização ou receber dividendos.

**Volatilidade**: Medida de quanto o preço de um ativo (como ações) varia ao longo do tempo. Alta volatilidade indica mudanças bruscas de preço, enquanto baixa volatilidade sugere estabilidade.

**Dividendos**: Pagamentos feitos por uma empresa aos seus acionistas, geralmente como parte dos lucros obtidos. Não são obrigatórios e variam conforme o desempenho da empresa.

**Desdobramento de Ações**: Processo em que uma empresa divide suas ações em mais unidades. Por exemplo, em um desdobramento 2:1, cada ação existente se transforma em duas, reduzindo o preço individual, mas mantendo o valor total investido.

**Lockdown**: Medida restritiva adotada durante a pandemia de COVID-19 para limitar a circulação de pessoas, visando reduzir a propagação do vírus. Impactou diretamente a economia e o comportamento do mercado.

**E-commerce**: Comércio eletrônico, ou seja, venda de produtos e serviços pela internet. Durante a pandemia, o Walmart ampliou suas operações online para atender à crescente demanda.

**Kaggle**: Plataforma online usada por cientistas de dados para compartilhar datasets, códigos e competições. Foi a fonte secundária dos dados deste estudo.

**API (Interface de Programação de Aplicações)**: Conjunto de regras que permite que diferentes softwares se comuniquem. Por exemplo, a API `yfinance` foi usada para extrair dados do Yahoo Finance\*.

**Desvio Padrão**: Indicador estatístico que mede a dispersão dos dados em relação à média. No contexto deste trabalho, foi usado para quantificar a volatilidade dos preços das ações.

**Boxplot**: Gráfico que resume a distribuição de dados usando quartis (mínimo, primeiro quartil, mediana, terceiro quartil e máximo). Ajuda a identificar outliers\* (valores extremos).

**Storytelling**: Técnica de narrativa que usa dados e análises para contar uma história coerente. Foi aplicada para explicar os padrões encontrados nas ações do Walmart.

**Pipeline**: Sequência de etapas estruturadas para processar dados, desde a coleta até a geração de insights. Neste projeto, incluiu pré-processamento, análise exploratória e visualização.

**Yahoo Finance**: Plataforma de informações financeiras que fornece dados históricos de ações, índices de mercado e notícias econômicas. Fonte primária dos dados utilizados.

**Google Colab**: Ferramenta gratuita baseada em nuvem para escrever e executar códigos em Python, usada neste projeto para análises e experimentos.

**GitHub**: Plataforma de hospedagem de código-fonte e colaboração em projetos. Foi utilizada para armazenar scripts, gráficos e documentação do trabalho.

**Inflação**: Aumento generalizado dos preços de bens e serviços ao longo do tempo. Durante a pandemia, a inflação nos EUA atingiu 9,1% em 2022, afetando custos operacionais e preços das ações.  
  
**Heatmap**: Representação visual de dados usando cores para indicar valores. No contexto financeiro, pode mostrar correlações entre variáveis (ex.: volume de negociação e preço).

**Outlier**: Valor atípico que se distancia significativamente da maioria dos dados. Por exemplo, quedas abruptas nas ações em março de 2020 foram consideradas outliers.

**Média Móvel**: Cálculo estatístico que suaviza flutuações de curto prazo, mostrando tendências de longo prazo. Usada para analisar a evolução suave dos preços das ações.

**Risco-Retorno**: Princípio financeiro que relaciona o potencial de ganho de um investimento com seu nível de risco. Durante a pandemia, o Walmart apresentou alta volatilidade (risco), mas também oportunidades de retorno elevado.

**Script**: É um arquivo de texto contendo um conjunto de instruções escritas em uma linguagem de programação (como Python, R ou JavaScript). Essas instruções são executadas sequencialmente por um interpretador ou compilador para realizar tarefas específicas, como manipular dados, gerar gráficos ou automatizar processos.   
 **Notebooks Jupyter:** É um documento interativo que combina código, texto explicativo, gráficos e resultados em um único ambiente. É amplamente usado em projetos de ciência de dados, análise financeira e pesquisa, pois permite executar tarefas, visualizar dados e documentar o processo de forma integrada.

# 