

**DESAFIO ORACLE NEXT EDUCATION & ALURA
DATA SCIENCE COM PYTHON**

Relatório de venda de uma das lojas para investimento em outro negócio

Helena Cardoso

BRASIL, 2025

Sumário

Objetivo.....	2
Introdução.....	3
Desenvolvimento.....	4
Fundamentos teórico.....	4
Bibliotecas Python.....	5
Prática do Desafio - Descrição do Código.....	5
1. Importação dos dados:.....	5
2. Análise do faturamento:.....	5
3. Vendas por Categoria:.....	6
4. Média de Avaliação das Lojas:.....	6
5. Produtos Mais e Menos Vendidos:.....	7
6. Frete Médio por Loja:.....	8
7. Análise dos Gráficos:.....	9
Observação.....	11
Relatório de Resultados e Conclusão.....	11
Fatores Adicionais a Considerar para Venda e Suposições Baseadas nos Resultados..	12
1. Potencial de Crescimento Futuro:.....	12
2. Valor Estratégico:.....	13
3. Custos Operacionais e Eficiência:.....	13
4. Concorrência Local:.....	13
5. Condições do Imóvel e Contrato de Locação:.....	13
6. Impacto nos Funcionários:.....	14
Conclusão.....	14
Referências.....	15

Objetivo

Este projeto tem como objetivo aplicar os conhecimentos adquiridos durante as aulas e atividades do programa **One Next Education** da Oracle & Alura, com foco em **Linguagem de Programação Python, análise de dados, funções, estruturas de controle e tratamento de exceções**. A partir dos dados fornecidos pela equipe da Alura, foi realizada uma análise de viabilidade de investimento para o Sr. João, visando a venda de uma de suas lojas.

Introdução

O Sr. João deseja abrir um novo negócio e, para isso, considera vender uma de suas lojas. A decisão será fundamentada em dados reais de vendas, que foram analisados utilizando a linguagem **Python** devido à sua simplicidade, versatilidade e ampla adoção no mercado de **Data Science**.

Durante este estudo, foram utilizadas ferramentas como **Google Colab**, **Git** e **GitHub**, além de bibliotecas essenciais de Python para análise de dados. A finalidade é fornecer uma análise precisa, embasada e visualmente clara para apoiar a decisão de venda.

Desenvolvimento

Fundamentos teórico

Lógica de programação

Lógica de programação é a base para desenvolver qualquer software. Trata-se da capacidade de estruturar e sequenciar instruções para que o computador execute tarefas de forma eficiente e correta.

Baseia - se em estruturas lógicas (sequência, decisão, repetição) e no pensamento algorítmico, permitindo transformar um problema em uma solução passo a passo.

Algoritmo de programação

É um conjunto finito de passos ordenados que descrevem como solucionar um problema ou executar uma determinada tarefa. Na programação, o algoritmo é a “receita” que orienta o computador, e pode ser escrito em linguagem natural, fluxogramas ou diretamente em código.

Linguagem de programação

É um conjunto de regras e sintaxe usado para escrever instruções para que um computador possa executar as tarefas. Ela serve como ponte entre o raciocínio humano e a execução pela máquina.

Linguagem de programação Python

O **Python** é uma linguagem de programação de alto nível, interpretada, de tipagem dinâmica e com sintaxe simples de fácil aprendizado.

Características principais:

- Código legível e conciso;
- Grande comunidade e bibliotecas diversificadas;
- Ideal para análise de dados, inteligência artificial, web, automação, etc.

Bibliotecas Python

pandas: manipulação e análise de dados.

matplotlib.pyplot: criação de gráficos.

numpy (se aplicável): cálculos numéricos eficientes.

Prática do Desafio - Descrição do Código

Abaixo vamos realizar uma análise exploratória dos dados de vendas de quatro lojas diferentes para identificar insights sobre o desempenho de cada uma.

1. Importação dos dados:

A primeira parte do código importa a **biblioteca pandas** e carrega os dados de quatro arquivos CSV hospedados em URLs diferentes em DataFrames separados (loja, loja2, loja3, loja4). Em seguida, exibe as primeiras linhas do DataFrame loja para uma visualização inicial dos dados.

```
▼ Importação dos dados

[ ] import pandas as pd

url1 = "https://raw.githubusercontent.com/alura-es-cursos/challenge1-data-science/refs/heads/main/base-de-dados-challenge-1/loja_1.csv"
url2 = "https://raw.githubusercontent.com/alura-es-cursos/challenge1-data-science/refs/heads/main/base-de-dados-challenge-1/loja_2.csv"
url3 = "https://raw.githubusercontent.com/alura-es-cursos/challenge1-data-science/refs/heads/main/base-de-dados-challenge-1/loja_3.csv"
url4 = "https://raw.githubusercontent.com/alura-es-cursos/challenge1-data-science/refs/heads/main/base-de-dados-challenge-1/loja_4.csv"

loja = pd.read_csv(url1)
loja2 = pd.read_csv(url2)
loja3 = pd.read_csv(url3)
loja4 = pd.read_csv(url4)

loja.head()
```

2. Análise do faturamento:

- Calcula o faturamento total de cada loja somando a coluna 'Preço' de cada DataFrame, convertendo-a para tipo float;
- Imprime o faturamento de cada loja formatado como moeda (R\$);
- Calcula e imprime o faturamento total de todas as lojas combinadas;

✓ 1. Análise do faturamento

```
[ ] # Somando apenas o preço das vendas, sem considerar as parcelas
faturamento_loja1 = loja['Preço'].astype(float).sum()
faturamento_loja2 = loja2['Preço'].astype(float).sum()
faturamento_loja3 = loja3['Preço'].astype(float).sum()
faturamento_loja4 = loja4['Preço'].astype(float).sum()

# Resultados
print(f"Faturamento Loja 1: R${faturamento_loja1:.2f}")
print(f"Faturamento Loja 2: R${faturamento_loja2:.2f}")
print(f"Faturamento Loja 3: R${faturamento_loja3:.2f}")
print(f"Faturamento Loja 4: R${faturamento_loja4:.2f}")

# Faturamento total de todas lojas
faturamento_total = faturamento_loja1 + faturamento_loja2 + faturamento_loja3 + faturamento_loja4
print(f"\nFaturamento total de todas as lojas: R${faturamento_total:.2f}")
```

3. Vendas por Categoria:

- Define uma função vendas_por_categoria que agrupa os dados de uma loja pela 'Categoria do Produto' e soma os 'Preço' para obter o total de vendas por categoria;
- Aplica esta função a cada DataFrame de loja e imprime as vendas por categoria para cada uma.

✓ 2. Vendas por Categoria

```
[ ] # Função para calcular e mostrar as vendas por categoria de uma loja
def vendas_por_categoria(loja, numero):
    vendas = loja.groupby('Categoria do Produto')['Preço'].sum()
    print(f"\nVendas por categoria na Loja {numero}:")
    print(vendas)

# Mostrar resultados para cada loja
vendas_por_categoria(loja, 1)
vendas_por_categoria(loja2, 2)
vendas_por_categoria(loja3, 3)
vendas_por_categoria(loja4, 4)
```

4. Média de Avaliação das Lojas:

- Define uma função avaliacao_media que calcula a média da coluna 'Avaliação da compra' para um determinado DataFrame de loja;
- Aplica esta função a cada DataFrame de loja e imprime a avaliação média formatada com duas casas decimais.

3. Média de Avaliação das Lojas

```
def avaliacao_media(loja, numero):  
    media = loja['Avaliação da compra'].mean()  
    print(f"Avaliação média na Loja {numero}: {media:.2f}")  
  
# Mostrar resultados para cada loja  
avaliacao_media(loja, 1)  
avaliacao_media(loja2, 2)  
avaliacao_media(loja3, 3)  
avaliacao_media(loja4, 4)
```

5. Produtos Mais e Menos Vendidos:

- Cria uma lista com as URLs dos arquivos CSV.
- Concatena todos os DataFrames das lojas em um único DataFrame (df_total), adicionando uma coluna 'Loja' para identificar a origem dos dados.
- Garante que as colunas 'Preço' e 'Quantidade de parcelas' tenham os tipos de dados corretos (float e int, respectivamente).
- Cria uma coluna 'Receita' que é uma cópia da coluna 'Preço'.
- Agrupa os dados por 'Loja' e 'Produto' e soma a 'Receita' para obter a receita por produto em cada loja.
- Define uma função top_produtos que, para cada loja, identifica e imprime os 3 produtos que geraram mais receita e os 3 que geraram menos receita.
- Itera sobre as lojas únicas no df_total e aplica a função top_produtos a cada uma.

✓ 4. Produtos Mais e Menos Vendidos

```
[ ] # Create a list of urls
    urls = [url1, url2, url3, url4]

    # Ler todos os CSVs e adicionar uma coluna 'Loja'
    df_total = pd.concat(
        [pd.read_csv(url).assign(Loja=f"Loja {i+1}") for i, url in enumerate(urls)],
        ignore_index=True
    )

    # Garantir os tipos corretos
    df_total['Preço'] = df_total['Preço'].astype(float)
    df_total['Quantidade de parcelas'] = df_total['Quantidade de parcelas'].astype(int)

    # A receita é simplesmente o Preço (não multiplicamos pela quantidade de parcelas)
    df_total['Receita'] = df_total['Preço']

    # Agrupar por Loja e Produto somando a receita
    vendas = df_total.groupby(['Loja', 'Produto'])['Receita'].sum().reset_index()

    # Função para mostrar os 3 produtos que mais e menos dinheiro geraram
    def top_produtos(loja):
        dados = vendas[vendas['Loja'] == loja]

        top_mais_vendidos = dados.sort_values(by='Receita', ascending=False).head(3)
        top_menos_vendidos = dados.sort_values(by='Receita', ascending=True).head(3)
```

```
    print(f"\n{loja} - TOP 3 produtos que MAIS dinheiro geraram:")
    for _, linha in top_mais_vendidos.iterrows():
        print(f"    {linha['Produto']} - Receita: R${linha['Receita']:.2f}")

    print(f"\n{loja} - TOP 3 produtos que MENOS dinheiro geraram:")
    for _, linha in top_menos_vendidos.iterrows():
        print(f"    {linha['Produto']} - Receita: R${linha['Receita']:.2f}")

    # Lista única de lojas
    lojas_unicas = df_total['Loja'].unique()

    # Mostrar resultados para cada loja
    for loja in lojas_unicas:
        top_produtos(loja)
```

6. Frete Médio por Loja:

- Garante que a coluna 'Frete' no df_total seja do tipo float.;
- Agrupa os dados por 'Loja' e calcula a média da coluna 'Frete' para obter o custo médio de envio por loja;
- Imprime o custo médio de envio para cada loja formatado como moeda (R\$).

✓ 5. Frete Médio por Loja

```
[ ] # Garantir que o custo de envio é numérico
    df_total['Frete'] = df_total['Frete'].astype(float)

# Agrupar por loja e calcular a média do custo de envio
envio_medio = df_total.groupby('Loja')['Frete'].mean()

# Mostrar resultados
print("\nEnvio médio por loja:")
for loja, medio in envio_medio.items():
    print(f"{loja}: R${medio:.2f}")
```

7. Análise dos Gráficos:

- Importa a biblioteca matplotlib.pyplot para a geração de gráficos.
- Garante que a coluna 'Preço' no df_total seja do tipo float e cria a coluna 'Receita' como uma cópia.
- Calcula as receitas totais por loja, a receita média por venda por loja e a distribuição percentual das receitas totais.
- Gera três gráficos:
- Um gráfico de barras verticais mostrando as receitas totais por loja.
- Um gráfico de linha mostrando a receita média por venda por loja.
- Um gráfico de pizza mostrando a participação de cada loja nas receitas totais.
- Exibe cada gráfico.

✓ Análise dos Gráficos

```
[ ] import matplotlib.pyplot as plt

df_total['Preço'] = df_total['Preço'].astype(float)

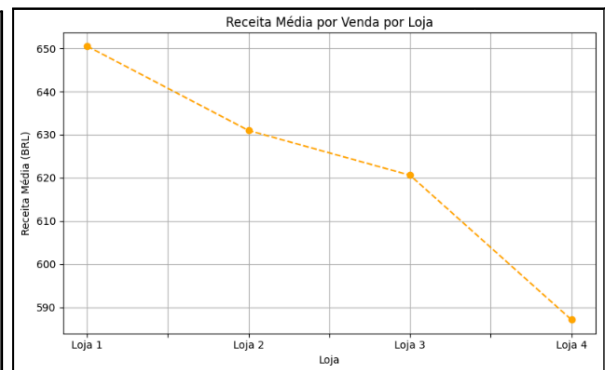
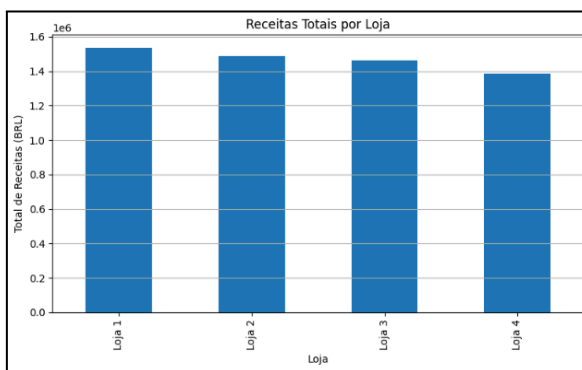
# Garantir tipo correto
df_total['Preço'] = df_total['Preço'].astype(float)

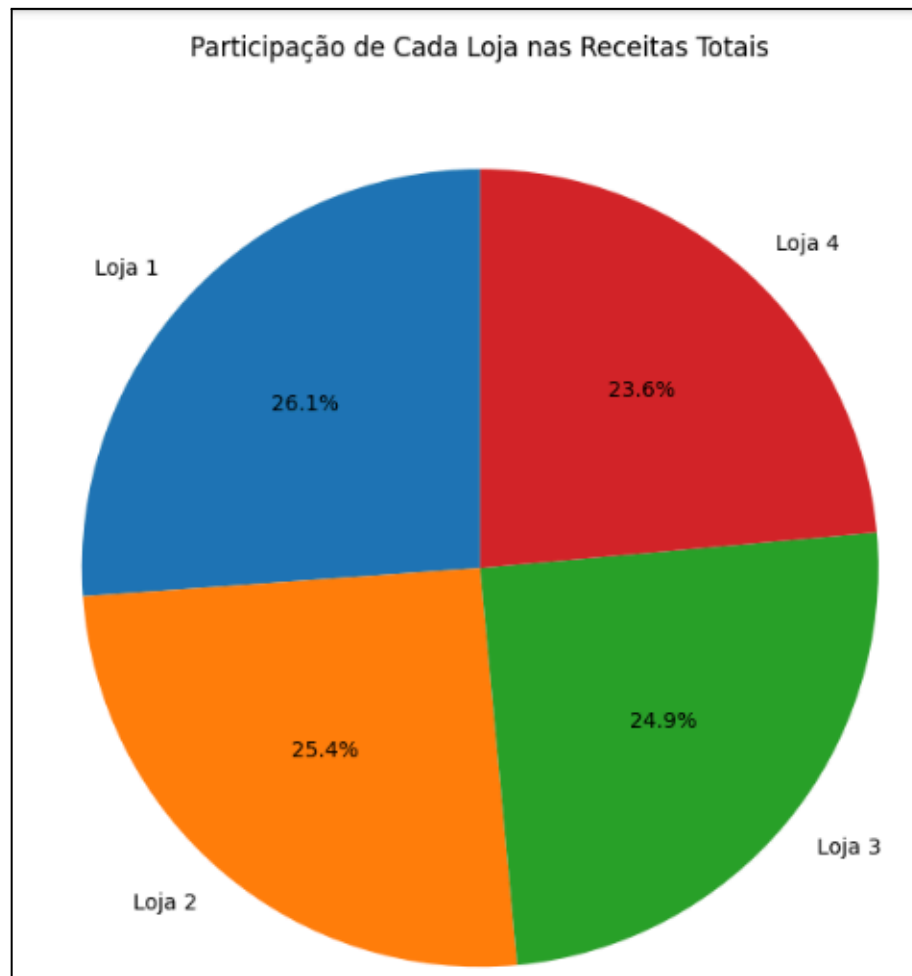
# A receita de cada venda é simplesmente o preço
df_total['Receita'] = df_total['Preço']

# Agrupar dados
receitas_totais = df_total.groupby('Loja')['Receita'].sum()
receita_media = df_total.groupby('Loja')['Receita'].mean()
distribuicao_receitas = receitas_totais

# 🍷 Gerar 3 gráficos realmente diferentes

# 1. Barra vertical: Total de receitas
plt.figure(figsize=(8,5))
receitas_totais.plot(kind='bar')
plt.title('Receitas Totais por Loja')
plt.ylabel('Total de Receitas (BRL)')
plt.xlabel('Loja')
plt.grid(axis='y')
plt.tight_layout()
plt.show()
```





Observação

Com base nas análises financeiras e nos fatores adicionais que discutimos, a Loja 4 seria a candidata mais provável a ser considerada para venda, pois apresenta o menor faturamento total e a menor receita média por venda. No entanto, é crucial lembrar que outros fatores estratégicos, operacionais e de mercado devem ser cuidadosamente avaliados antes de tomar uma decisão final

Relatório de Resultados e Conclusão

Com base nas análises realizadas e nos gráficos gerados, podemos observar os seguintes pontos:

- **Faturamento:** a Loja 1 apresenta o maior faturamento total, seguida pela Loja 2, Loja 3 e, por último, a Loja 4.

- **Vendas por Categoria:** a distribuição das vendas por categoria varia entre as lojas, mas 'eletronicos' e 'eletrodomesticos' geralmente representam uma parcela significativa das vendas em todas elas.
- **Avaliação Média:** as avaliações médias das lojas são bastante próximas, com todas apresentando uma média acima de 3.9. A Loja 3 possui a maior avaliação média, enquanto a Loja 1 tem a menor.
- **Produtos Mais e Menos Vendidos:** os produtos que mais geram receita ('TV Led UHD 4K', 'Geladeira', 'Celular Plus X42') são consistentes entre as lojas, assim como os que menos geram receita ('Corda de pular', 'Cubo mágico 8x8', 'Dinossauro Rex'). No entanto, a ordem e a receita exata variam um pouco.
- **Frete Médio por Loja:** a Loja 1 tem o frete médio mais alto, enquanto a Loja 4 tem o frete médio mais baixo. Isso pode influenciar a atratividade para os clientes.
- **Análise Gráfica:**
 - O gráfico de barras confirma que a Loja 1 lidera em receita total, seguida pelas outras em ordem decrescente.
 - O gráfico de linha da receita média por venda mostra que a Loja 1 também tem a maior receita média por venda, sugerindo que os produtos vendidos nela podem ter um preço médio mais alto ou uma combinação de produtos mais rentável.
 - O gráfico de pizza ilustra claramente a participação de cada loja no faturamento total, reforçando que a Loja 1 contribui com a maior fatia.

Fatores Adicionais a Considerar para Venda e Suposições Baseadas nos Resultados

Além das métricas de faturamento, receita média e participação no faturamento total, que já indicam a Loja 4 como a de menor desempenho financeiro, outros fatores cruciais devem ser considerados em uma decisão de venda. Com base nas análises já realizadas, podemos fazer algumas suposições informadas:

1. Potencial de Crescimento Futuro:

- **Suposição:** embora a Loja 4 tenha o menor faturamento atualmente, ela pode estar localizada em uma região com alto potencial de crescimento demográfico ou econômico no futuro. Isso poderia justificar mantê-la, pois seu desempenho pode melhorar significativamente com o tempo.

- **Como verificar:** seria necessário analisar dados demográficos da região, planos de desenvolvimento urbano, concorrência local e pesquisas de mercado sobre o potencial de consumo.

2. Valor Estratégico:

- **Suposição:** A Loja 4 pode ter um valor estratégico para a rede que vai além do seu faturamento atual. Por exemplo, ela pode servir como ponto de distribuição para outras lojas, ter uma base de clientes fiel em um nicho específico ou deter uma licença ou contrato valioso.
- **Como verificar:** Seria preciso avaliar o papel da Loja 4 na logística da empresa, analisar a composição e lealdade da sua base de clientes e revisar quaisquer contratos ou acordos comerciais específicos da loja.

3. Custos Operacionais e Eficiência:

- **Suposição:** Mesmo com menor receita, a Loja 4 pode ter custos operacionais significativamente mais baixos ou ser mais eficiente em termos de gestão de estoque e pessoal do que as outras lojas. Isso poderia torná-la mais rentável em uma base proporcional.
- **Como verificar:** Seria necessário analisar detalhadamente os relatórios de custos operacionais de cada loja (aluguel, salários, utilities, etc.) e métricas de eficiência.

4. Concorrência Local:

- **Suposição:** A Loja 4 pode enfrentar uma concorrência mais acirrada em sua localização, o que dificulta alcançar o mesmo desempenho das outras lojas. Vender a loja poderia ser uma forma de sair de um mercado saturado.
- **Como verificar:** Seria importante realizar uma análise da concorrência na área de atuação da Loja 4, identificando os principais concorrentes, seus produtos/serviços e estratégias de precificação.

5. Condições do Imóvel e Contrato de Locação:

- **Suposição:** As condições do imóvel onde a Loja 4 está localizada (necessidade de reformas, tamanho, acesso) ou os termos do contrato de locação (custo, duração, cláusulas de rescisão) podem influenciar a decisão de venda. Um contrato de locação desfavorável ou um imóvel que exige grandes investimentos pode tornar a venda mais atrativa.
- **Como verificar:** Seria necessário avaliar o estado físico do imóvel, os termos do contrato de locação e os custos associados a ele.

6. Impacto nos Funcionários:

- **Suposição:** A venda da Loja 4 teria um impacto nos funcionários. A empresa precisaria considerar opções como transferência para outras unidades, indenizações ou assistência na recolocação.
- **Como verificar:** Seria necessário avaliar o número de funcionários na Loja 4 e as implicações legais e éticas de uma possível demissão em massa.

Em resumo, enquanto a análise financeira aponta a Loja 4 como a candidata mais provável para venda, uma decisão final e estratégica requer a avaliação de fatores externos e internos que vão além dos dados de receita e lucro. Seria fundamental realizar uma diligência mais aprofundada nesses pontos para tomar a melhor decisão para o futuro da empresa.

Conclusão

Considerando o faturamento total, a receita média por venda e a participação no faturamento global, a **Loja 1** é a que demonstra o melhor desempenho financeiro. Embora a Loja 3 tenha a maior avaliação média, os indicadores de receita apontam a Loja 1 como a mais lucrativa. A Loja 4 apresenta o menor faturamento e a menor receita média por venda.

Com base exclusivamente nos dados financeiros e de desempenho de vendas apresentados, a **Loja 4** seria a candidata mais provável a ser considerada para venda, pois é a que contribui menos para o faturamento total e tem a menor receita média por venda. No entanto, uma decisão final sobre a venda de uma loja deve levar em consideração outros fatores estratégicos e de mercado que não estão presentes neste conjunto de dados.

Referências

GITHUB DOCS. *Clonar um repositório*. GitHub Docs. Disponível em: <https://docs.github.com/pt/repositories/creating-and-managing-repositories/cloning-a-repository>. Acesso em: 16 ago. 2025.

CHAGAS, C. *Como criar um README para seu perfil no GitHub*. Alura, 2022. Disponível em: <https://www.alura.com.br/artigos/como-criar-um-readme-para-seu-perfil-github>. Acesso em: 16 ago. 2025.

CHAGAS, C. *O que é Git e GitHub*. Alura, 2022. Disponível em: <https://www.alura.com.br/artigos/o-que-e-git-github>. Acesso em: 16 ago. 2025.

DIO. *Versionamento de código: conceitos básicos*. Digital Innovation One, 2023. Disponível em: <https://www.dio.me/articles/versionamento-de-codigo-conceitos-basicos>. Acesso em: 16 ago. 2025.

ASIMOV ACADEMY. *Google Colab: guia definitivo para iniciantes*. Hub Asimov, 2023. Disponível em: <https://hub.asimov.academy/blog/google-colab-guia-definitivo-para-iniciantes/>. Acesso em: 16 ago. 2025.

LOCAWEB. *Lógica de programação: o que é?*. Blog Locaweb, 2021. Disponível em: <https://www.locaweb.com.br/blog/temas/codigo-aberto/logica-de-programacao-o-que-e/>. Acesso em: 16 ago. 2025.

BUENO, G. *Estruturas de repetição*. GitHub Pages, 2021. Disponível em: https://gabrielbueno072.github.io/rea-aed/aula_rep.html. Acesso em: 16 ago. 2025.

ALURA. *Linguagem Python: por onde começar*. Alura, 2022. Disponível em: <https://www.alura.com.br/artigos/python>. Acesso em: 16 ago. 2025.

ALURA. *Análise de dados: entenda o que é e como aprender*. Alura, 2022. Disponível em: <https://www.alura.com.br/artigos/analise-de-dados>. Acesso em: 16 ago. 2025.

MIRO. *O que são relatórios?* Miro, 2024. Disponível em: <https://miro.com/pt/relatorios/o-que-sao-relatorios/>. Acesso em: 16 ago. 2025.