

PROJETO FINAL  
MÓDULO 4

# ABASTECENDO O CARRO

ANÁLISE EXPLORATÓRIA DOS VALORES DE  
COMBUSTÍVEIS NO BRASIL NOS MESES  
DE JULHO E AGOSTO DE 2022

RESILIA

# SQUAD



**Lucas  
Canella**

GESTOR DE  
CONHECIMENTO

LUCASCANELLA



**Carlos  
Augustin**

GESTOR DE GENTE E  
ENGAJAMENTO

CARLOS-AUGUSTIN

CXRLXX



**Bruno  
Silva**

COLABORADOR I

BRUNOGUEDESDSILVA

BRUNOSILL

**RESILIA**

# EXPECTATIVA

REALIZAR UMA ANÁLISE EXPLORATÓRIA RELACIONADA À SÉRIE HISTÓRICA DE PREÇO DE VENDA DA GASOLINA E DO ETANOL BASEADOS NOS DOIS ÚLTIMOS MESES DE ACORDO COM AGENCIA NACIONAL DE PETRÓLEO E GÁS NATURAL E BIOCOMBUSTÍVEIS (ANP).

**RESILIA**

# FERRAMENTAS UTILIZADAS PELO TIME



github



jupyter



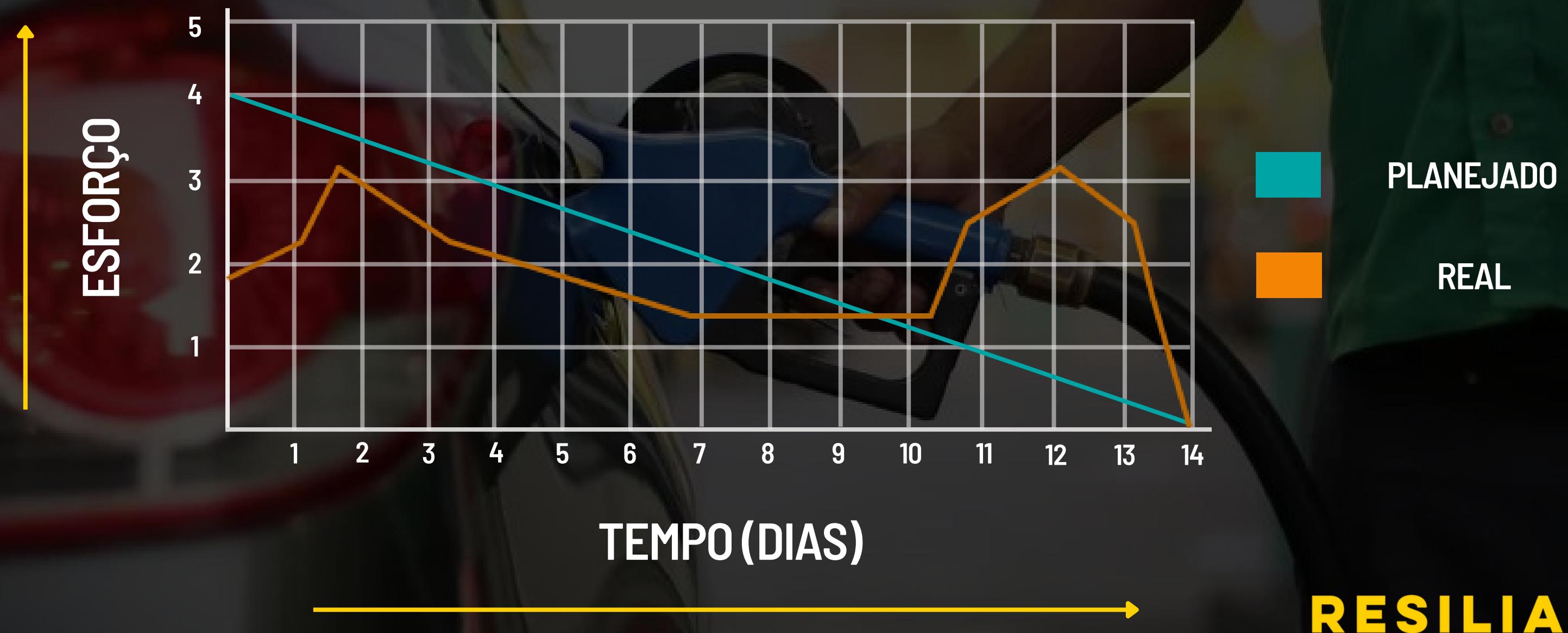
notion



DISCORD

RESILIA

# GRÁFICO DE BURNDOWN



# RETROSPECTIVA

1. DEFINIÇÃO DE FUNÇÕES E TAREFAS;
2. TRATAMENTO E ANÁLISE EXPLORATÓRIA DOS DADOS;
3. REUNIÕES DE APRESENTAÇÃO DOS RESULTADOS ALCANÇADOS;
4. ORGANIZAÇÃO DO ARQUIVO E INSIGHTS SOBRE AS PERGUNTAS;
5. ORGANIZAÇÃO DA APRESENTAÇÃO;

# PONTOS POSITIVOS E DE MELHORIA

## PONTOS POSITIVOS

1. PROATIVIDADE.
2. NÍVEIS DE CONHECIMENTOS VARIADOS EM RELAÇÃO AO CONTEÚDO.
3. EQUIPE FLEXÍVEL

## PONTOS DE MELHORIA

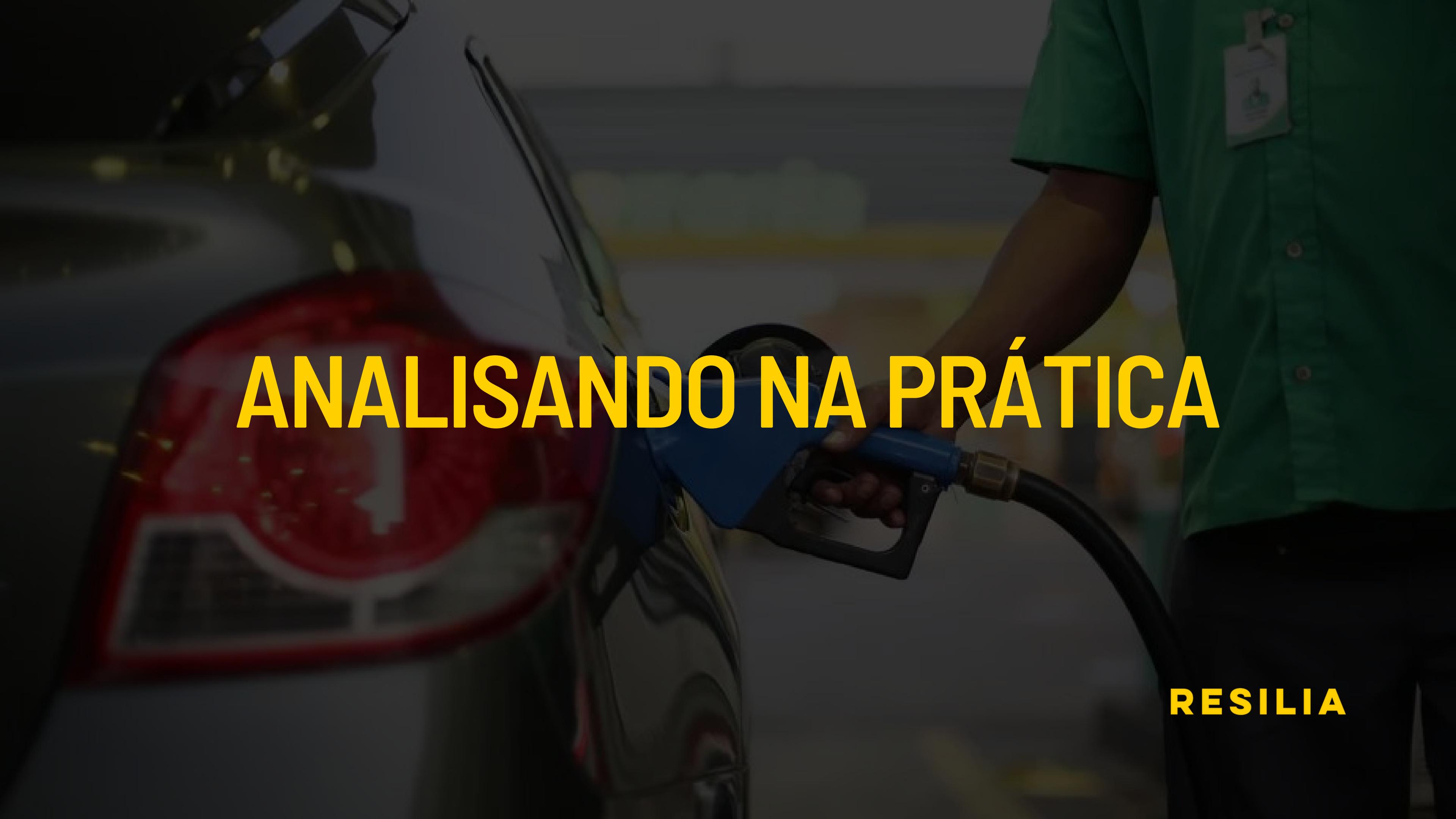
1. COMUNICAÇÃO;
2. CAPACIDADE DE "REALINHAR A ROTA";
3. SE PERMITIR SER COMANDADO PELA URGÊNCIA.

**RESILIA**

# AGÊNCIA NACIONAL DO PETRÓLEO, BIOCOMBUSTÍVEIS E GÁS NATURAL (ANP)

RESPONSÁVEL PELAS REGRAS E RESOLUÇÕES DE EXPLORAÇÃO, PRODUÇÃO, REFINO, COMERCIALIZAÇÃO E TRANSPORTE DO PETRÓLEO E SEUS DERIVADOS NO BRASIL. ALÉM DISSO, É VINCULADA AO MINISTÉRIO DE MINAS E ENERGIA E, TAMBÉM, MANTÉM O BANCO DE DADOS DE EXPLORAÇÃO E PRODUÇÃO (BDEP), NO QUAL SÃO PUBLICADOS DADOS SOBRE A INDÚSTRIAS DO PETRÓLEO, BIOCOMBUSTÍVEL E GÁS NATURAL, COMO A PRODUÇÃO NACIONAL E O PREÇO MÉDIA DE VENDA DE COMBUSTÍVEL SEMANALMENTE.

**RESILIA**

A photograph showing a person in a green uniform with a name tag, pumping fuel into a red car at a gas station. The person is wearing a blue cap and a dark t-shirt underneath the uniform. The car's fuel tank is visible, and the pump is connected to the vehicle.

**ANALISANDO NA PRÁTICA**

**RESILIA**

# IMPORTAÇÃO DAS BIBLIOTECAS

- ▼ BIBLIOTECA PARA ANÁLISE E MANIPULAÇÃO DE DADOS PODENDO TRABALHAR COM DADOS TABULARES, COMO UMA PLANILHA EXCEL OU UMA TABELA SQL.

```
import pandas as pd
```

- ▼ BIBLIOTECA QUE FORNECE UM GRANDE CONJUNTO DE FUNÇÕES E OPERAÇÕES QUE AJUDAM A EXECUTAR CÁLCULOS NUMÉRICOS COM MAIOR FACILIDADE

```
import numpy as np
```

- ▼ BIBLIOTECA ABRANGENTE PARA CRIAR VISUALIZAÇÕES ESTÁTICAS, ANIMADAS E INTERATIVAS. O MATPLOTLIB TORNA AS COISAS FÁCEIS FÁCEIS E AS DIFÍCEIS POSSÍVEIS.

```
import matplotlib  
import matplotlib.pyplot as plt
```

- ▼ BIBLIOTECA PARA FAZER GRÁFICOS ESTATÍSTICOS. ELE SE BASEIA NO MATPLOTLIB E SE INTEGRA DIRETAMENTE ÀS ESTRUTURAS DE DADOS DO PANDAS.

```
import seaborn as sns
```

# IMPORTAÇÃO DE DADOS

JULHO DE 2022

```
def julhoCsv():
    try:
        url_julho = 'https://www.gov.br/anp/pt-br/centrais-de-conteudo/dados-abertos/arquivos/shpc/dsan/2022/precos-gasolina-etanol-07.csv'
        dados_brutos_julho = pd.read_csv(url_julho, sep=';')
        return dados_brutos_julho
    except:
        url_julho = 'https://raw.githubusercontent.com/lucasCanella/Projeto-modulo-4/colab1/dados/dados_originais/precos-gasolina-etanol-julho.csv'
        dados_brutos_julho = pd.read_csv(url_julho, sep=';')
        return dados_brutos_julho

dados_brutos_julho = julhoCsv()
dados_brutos_julho.head(4)
```

	Regiao - Sigla	Estado - Sigla	Municipio	Revenda	CNPJ da Revenda	Nome da Rua	Numero Rua	Complemento	Bairro	Cep	Produto	Data da Coleta	Valor de Venda	Valor de Compra	Unidade de Medida	Bandeira
0	SE	ES	COLATINA	SAO FRANCISCO DE ASSIS COMERCIO DE COMBUSTIVEI...	08.519.545/0001-10	PRACA FIDELIS FERRARI	35	NaN	LACE	29703-030	GASOLINA ADITIVADA	01/07/2022	7,48	NaN	R\$ / litro	VIBRA ENERGIA
1	SE	ES	COLATINA	SAO FRANCISCO DE ASSIS COMERCIO DE COMBUSTIVEI...	08.519.545/0001-10	PRACA FIDELIS FERRARI	35	NaN	LACE	29703-030	GASOLINA	01/07/2022	7,38	NaN	R\$ / litro	VIBRA ENERGIA
2	SE	ES	COLATINA	POSTO MOTOCAP LTDA	04.814.652/0001-10	PRACA ALMIRANTE BARROSO	52	LETRA A	LACE	29703-055	ETANOL	01/07/2022	5,99	NaN	R\$ / litro	VIBRA ENERGIA
3	SE	ES	COLATINA	POSTO MOTOCAP LTDA	04.814.652/0001-10	PRACA ALMIRANTE BARROSO	52	LETRA A	LACE	29703-055	GASOLINA	01/07/2022	7,39	NaN	R\$ / litro	VIBRA ENERGIA

# IMPORTAÇÃO DE DADOS

AGOSTO DE 2022

```
def agostoCsv():
    try:
        url_agosto = 'https://www.gov.br/anp/pt-br/centrais-de-conteudo/dados-abertos/arquivos/shpc/dsan/2022/precos-gasolina-etanol-08.csv'
        dados_brutos_agosto = pd.read_csv(url_agosto, sep=';')
        return dados_brutos_agosto
    except:
        url_agosto = 'https://raw.githubusercontent.com/lucasCanella/Projeto-modulo-4/colab1/dados/dados_originais/precos-gasolina-etanol-agosto.csv'
        dados_brutos_agosto = pd.read_csv(url_agosto, sep=';')
        return dados_brutos_agosto

dados_brutos_agosto = agostoCsv()
dados_brutos_agosto.head(4)
```

	Regiao - Sigla	Estado - Sigla	Municipio	Revenda	CNPJ da Revenda	Nome da Rua	Numero Rua	Complemento	Bairro	Cep	Produto	Data da Coleta	Valor de Venda	Valor de Compra	Unidade de Medida	Bandeira
0	NE	AL	ARAPIRACA	AUTO POSTO M M GARROTE LTDA	08.738.994/0001-50	RODOVIA AL-220	5848	KM 96	BOM SUCESSO	57309-035	ETANOL	01/08/2022	4,79	NaN	R\$ / litro	VIBRA ENERGIA
1	NE	AL	ARAPIRACA	AUTO POSTO M M GARROTE LTDA	08.738.994/0001-50	RODOVIA AL-220	5848	KM 96	BOM SUCESSO	57309-035	GASOLINA	01/08/2022	5,59	NaN	R\$ / litro	VIBRA ENERGIA
2	NE	AL	ARAPIRACA	AUTO POSTO MASSARANDUBA LTDA.	07.248.398/0001-29	RODOVIA AL 220	4698	NaN	PLANALTO	57308-000	GASOLINA ADITIVADA	01/08/2022	5,99	NaN	R\$ / litro	BRANCA
3	NE	AL	ARAPIRACA	AUTO POSTO MASSARANDUBA LTDA.	07.248.398/0001-29	RODOVIA AL 220	4698	NaN	PLANALTO	57308-000	ETANOL	01/08/2022	4,99	NaN	R\$ / litro	BRANCA

# JUNÇÃO DOS DATAFRAMES

```
lista_dados= [dados_brutos_julho, dados_brutos_agosto]
```

```
dados_brutos= pd.concat([lista_dados, verify_integrity=True])
```

```
dados_brutos
```

	Regiao - Sigla	Estado - Sigla	Municipio	Revenda	CNPJ da Revenda	Nome da Rua	Numero Rua	Complemento	Bairro	Cep	Produto	Data da Coleta	Valor de Venda	Valor de Compra	Unidade de Medida	Bandeira
0	SE	ES	COLATINA	SAO FRANCISCO DE ASSIS COMERCIO DE COMBUSTIVEI...	08.519.545/0001-10	PRACA FIDELIS FERRARI	35	NaN	LACE	29703-030	GASOLINA ADITIVADA	01/07/2022	7,48	NaN	R\$ / litro	VIBRA ENERGIA
1	SE	ES	COLATINA	SAO FRANCISCO DE ASSIS COMERCIO DE COMBUSTIVEI...	08.519.545/0001-10	PRACA FIDELIS FERRARI	35	NaN	LACE	29703-030	GASOLINA	01/07/2022	7,38	NaN	R\$ / litro	VIBRA ENERGIA
2	SE	ES	COLATINA	POSTO MOTOCAP LTDA	04.814.652/0001-10	PRACA ALMIRANTE BARROSO	52	LETRA A	LACE	29703-055	ETANOL	01/07/2022	5,99	NaN	R\$ / litro	VIBRA ENERGIA
3	SE	ES	COLATINA	POSTO MOTOCAP LTDA	04.814.652/0001-10	PRACA ALMIRANTE BARROSO	52	LETRA A	LACE	29703-055	GASOLINA	01/07/2022	7,39	NaN	R\$ / litro	VIBRA ENERGIA
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
127171	NE	AL	ARAPIRACA	AUTO POSTO M M GARROTE LTDA	08.738.994/0001-50	RODOVIA AL-220	5848	KM 96	BOM SUCESSO	57309-035	ETANOL	01/08/2022	4,79	NaN	R\$ / litro	VIBRA ENERGIA
127172	NE	AL	ARAPIRACA	AUTO POSTO M M GARROTE LTDA	08.738.994/0001-50	RODOVIA AL-220	5848	KM 96	BOM SUCESSO	57309-035	GASOLINA	01/08/2022	5,59	NaN	R\$ / litro	VIBRA ENERGIA
127173	NE	AL	ARAPIRACA	AUTO POSTO MASSARANDUBA LTDA.	07.248.398/0001-29	RODOVIA AL 220	4698	NaN	PLANALT 0	57308-000	GASOLINA ADITIVADA	01/08/2022	5,99	NaN	R\$ / litro	BRANCA
127174	NE	AL	ARAPIRACA	AUTO POSTO MASSARANDUBA LTDA.	07.248.398/0001-29	RODOVIA AL 220	4698	NaN	PLANALT 0	57308-000	ETANOL	01/08/2022	4,99	NaN	R\$ / litro	BRANCA

# OBSERVANDO OS DADOS

```
# demonstração do formato da array
```

```
dados_brutos.shape
```

```
((127176, 16))
```

```
#retorno dos valores ausentes
```

```
dados_brutos.isna().sum()
```

Regiao - Sigla	0
Estado - Sigla	0
Municipio	0
Revenda	0
CNPJ da Revenda	0
Nome da Rua	0
Numero Rua	71
Complemento	100097
Bairro	412
Cep	0
Produto	0
Data da Coleta	0
Valor de Venda	0
Valor de Compra	127176
Unidade de Medida	0
Bandeira	0
dtype: int64	

```
dados_brutos.info()
```

#	COLUMN	NON-NULL COUNT	DTYPE
0	REGIAO SIGLA	127176 NON-NULL	OBJECT
1	ESTADO SIGLA	127176 NON-NULL	OBJECT
2	MUNICIPIO	127176 NON-NULL	OBJECT
3	REVENDA	127176 NON-NULL	OBJECT
4	CNPJ DA REVENDA	127176 NON-NULL	OBJECT
5	NOME DA RUA	127176 NON-NULL	OBJECT
6	NUMERO RUA	127105 NON-NULL	OBJECT
7	COMPLEMENTO	27079 NON-NULL	OBJECT
8	BAIRRO	126764 NON-NULL	OBJECT
9	CEP	127176 NON-NULL	OBJECT
10	PRODUTO	127176 NON-NULL	OBJECT
11	DATA DA COLETA	127176 NON-NULL	OBJECT
12	VALOR DE VENDA	127176 NON-NULL	FLOAT64
13	VALOR DE COMPRA	0 NON-NULL	FLOAT64
14	UNIDADE DE MEDIDA	127176 NON-NULL	OBJECT
15	BANDEIRA	127176 NON-NULL	OBJECT

# LIMPEZA DE DADOS

## 1. NOMENCLATURA DAS COLUNAS

```
dados_brutos.columns = dados_brutos.columns.str.replace('-', '')
dados_brutos.columns = dados_brutos.columns.str.replace(' ', '_')
dados_brutos.columns = dados_brutos.columns.str.lower()
dados_brutos.columns
```

```
Index(['regiao_sigla', 'estado_sigla', 'municipio', 'revenda',
       'cnpj_da_revenda', 'nome_da_rua', 'numero_rua', 'complemento', 'bairro',
       'cep', 'produto', 'data_da_coleta', 'valor_de_venda', 'valor_de_compra',
       'unidade_de_medida', 'bandeira'],
      dtype='object')
```

**RESILIA**

# LIMPEZA DE DADOS

## 2. MODIFICAÇÃO DO PADRÃO DAS DATAS

```
dados_brutos['Data da Coleta'] = pd.to_datetime(dados_brutos['Data da Coleta'], dayfirst = True)  
dados_brutos['Data da Coleta']
```

```
0    2022-07-01  
1    2022-07-01  
2    2022-07-01  
3    2022-07-01  
4    2022-07-01  
...  
127171 2022-08-31  
127172 2022-08-31  
127173 2022-08-31  
127174 2022-08-31  
127175 2022-08-31
```

NAME: DATA DA COLETA, LENGTH: 127176, DTTYPE: DATETIME64[NS]

RESILIA

# LIMPEZA DE DADOS

## 3. ADEQUAÇÃO DOS VALORES DA COLUNA VALOR\_DE\_VENDA

```
# TROCA DE VÍRGULA PARA PONTO E DE STRING PARA FLOAT
```

```
valores_virgula = dados_brutos['valor_de_venda'].tolist()
valores_ponto = [x.replace(',', '.') for x in valores_virgula]
dados_brutos['valor_de_venda'] = valores_ponto
dados_brutos['valor_de_venda'] = dados_brutos['valor_de_venda'].astype(float)
```

## 4. VERIFICANDO COLUNAS VAZIAS OU COM POUcos VALORES PARA APAGAR

```
dados_brutos.valor_de_compra.unique()
```

```
array([nan])
```

```
dados_brutos.drop(columns=['valor_de_compra'], inplace=True)
```



MESMO PROCEDIMENTO NA  
COLUNA 'COMPLEMENTO'

**RESILIA**

# DICIONÁRIO

**REGIÃO:** NOME DA REGIÃO PERTENCENTE AO BRASIL

**ESTADO:** NOME DO ESTADO PERTENCENTE AO BRASIL

**MUNICÍPIO:** NOME DO MUNICÍPIO

**REVENDA:** RAZÃO SOCIAL DO POSTO DE COMBUSTÍVEL

**CNPJ DA REVENDA:** CADASTRO NACIONAL DA PESSOA JURÍDICA DO POSTO DE COMBUSTÍVEL

**NOME DA RUA:** RUA ONDE SE LOCALIZA O POSTO DE COMBUSTÍVEL

**NÚMERO RUA:** NÚMERO ONDE SE LOCALIZA O POSTO DE COMBUSTÍVEL

**COMPLEMENTO:** COMPLEMENTO DO ENDEREÇO ONDE SE LOCALIZA O POSTO DE COMBUSTÍVEL

**BAIRRO:** NOME DO BAIRRO PERTENCENTE AO ENDEREÇO DO POSTO DE COMBUSTÍVEL

**CEP:** CÓDIGO DE ENDEREÇAMENTO POSTAL PERTENCENTE AO ENDEREÇO DO POSTO DE COMBUSTÍVEL

**PRODUTO:** TIPOS DE PRODUTOS COM QUE O POSTO DE COMBUSTÍVEL TRABALHA

**DATA DA COLETA:** QUANDO O DADO FOI COLETADO

**VALOR DE VENDA:** VALOR REPASSADO AO CLIENTE

**VALOR DE COMPRA:** VALOR PAGO PELA COMPRA

**UNIDADE DE MEDIDA:** UNIDADE DE MEDIDA DE RELAÇÃO TIPO DE MOEDA X TIPO DE VOLUME

**BANDEIRA:** DISTRIBUIDORA DOS COMBUSTÍVEIS

**RESILIA**

# VAMOS ÀS PERGUNTAS?!

MAS, ANTES DE RESPONDÊ-LAS, FOI CRIADA UMA FUNÇÃO ANÔNIMA 'LAMBDA' PARA APLICAR A FORMATAÇÃO DECIMAL NOS NOSSOS RESULTADOS:

```
formatar = lambda x: float("%.2f" % x)
```

RESILIA

# ANÁLISE EXPLORATÓRIA

- ▼ 1. COMO SE COMPORTARAM O PREÇO DOS COMBUSTÍVEIS DURANTE OS 2 MESES CITADOS? OS VALORES DO ETANOL E DA GASOLINA TIVERAM UMA DE QUEDA OU DIMINUIÇÃO?

Dataframe com a média dos valores de todos os tipos de combustíveis agrupados por data (gasolina, gasolina aditivada e etanol).

```
venda_por_data_gasolina_aditivada = dados.loc[dados['produto'] == 'GASOLINA ADITIVADA'].groupby(['data_da_coleta'])  
['valor_de_venda'].mean().reset_index()  
  
venda_por_data_gasolina = dados.loc[dados['produto'] == 'GASOLINA'].groupby(['data_da_coleta'])  
['valor_de_venda'].mean().reset_index()  
  
venda_por_data_etanol = dados.loc[dados['produto'] == 'ETANOL'].groupby(['data_da_coleta'])['valor_de_venda'].mean().reset_index()  
  
venda_por_data = dados.groupby(['data_da_coleta'])['valor_de_venda'].mean().reset_index()  
  
venda_por_data.head(10)
```

	data_da_coleta	valor_de_venda
0	2022-07-01	6.621690
1	2022-07-04	6.077862
2	2022-07-05	6.185041
3	2022-07-06	5.959738
4	2022-07-07	5.797347
5	2022-07-08	5.783333
6	2022-07-11	5.738744
7	2022-07-12	5.782261
8	2022-07-13	5.601108
9	2022-07-14	5.661235

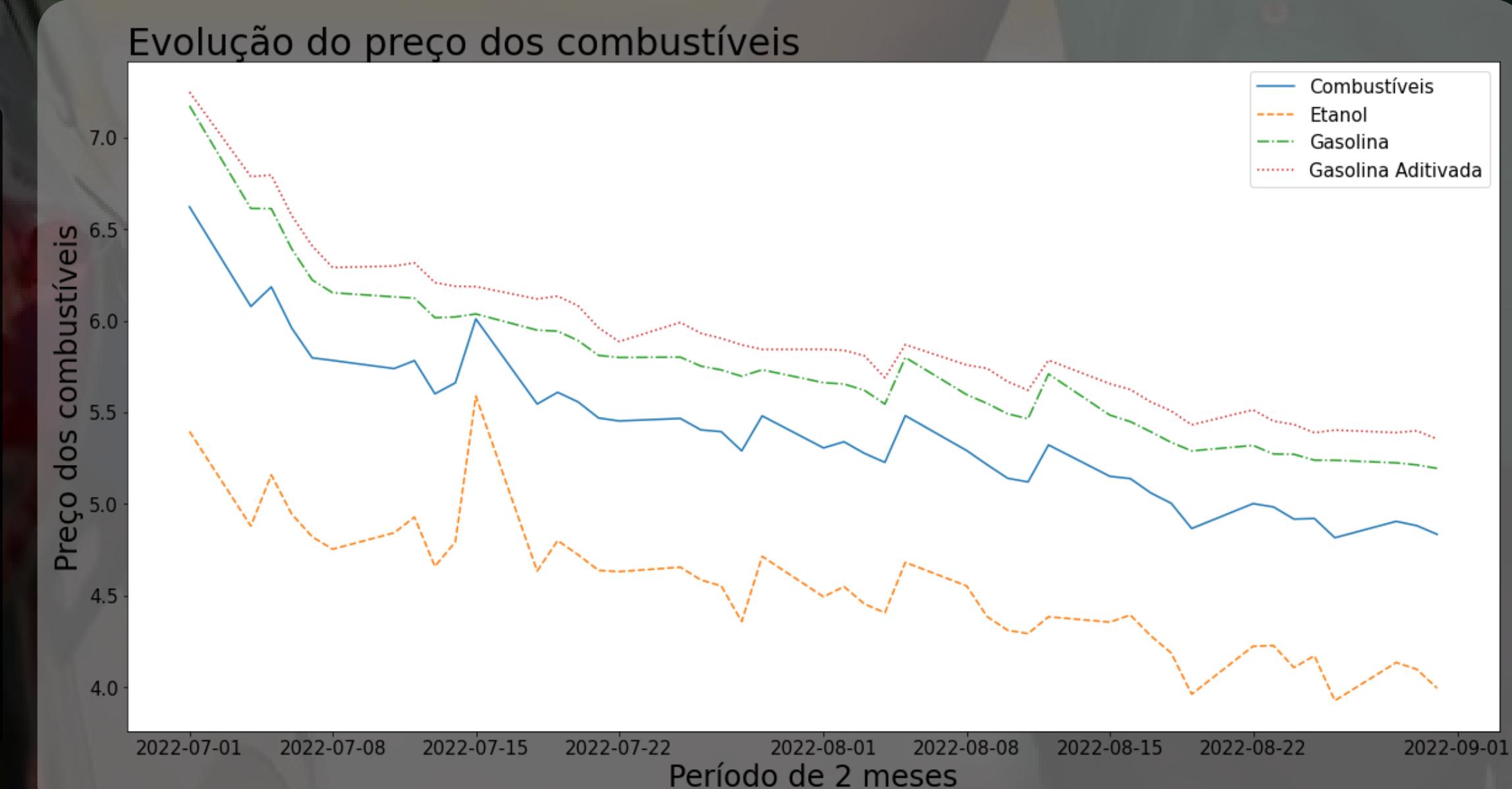
RESILIA

# ANÁLISE EXPLORATÓRIA

1. COMO SE COMPORTARAM O PREÇO DOS COMBUSTÍVEIS DURANTE OS 2 MESES CITADOS? OS VALORES DO ETANOL E DA GASOLINA TIVERAM UMA DE QUEDA OU DIMINUIÇÃO?

**Com esses dataframes criados, utilizamos o matplotlib para unir essas informações em um gráfico de linha:**

```
plt.figure(figsize=(20,10))
plt.rc('xtick', labelsize=15)
plt.rc('ytick', labelsize=15)
plt.rc('legend', fontsize=15)
plt.plot(venda_por_data['data_da_coleta'], venda_por_data['valor_de_venda'],
label='Combustíveis')
plt.plot(venda_por_data_etanol['data_da_coleta'], venda_por_data_etanol['valor_de_venda'],
label='Etanol', linestyle="--")
plt.plot(venda_por_data_gasolina['data_da_coleta'],
venda_por_data_gasolina['valor_de_venda'], label='Gasolina', linestyle="-.")
plt.plot(venda_por_data_gasolina_aditivada['data_da_coleta'],
venda_por_data_gasolina_aditivada['valor_de_venda'], label='Gasolina Aditivada', linestyle=":")
plt.xlabel("Período de 2 meses", size=24)
plt.ylabel("Preço dos combustíveis", size=24)
plt.title("Evolução do preço dos combustíveis", loc='left', size=30)
plt.legend()
plt.show()
```



# RESILIA

# ANÁLISE EXPLORATÓRIA

2. QUAL O PREÇO MÉDIO DA GASOLINA E DO ETANOL NESSES DOIS MESES?

```
media_gasolina = formatar(dados.loc[dados['produto'] == 'GASOLINA']['valor_de_venda'].mean())
media_gasolina
```

Média da gasolina: 5.71

```
media_etanol = formatar(dados.loc[dados['produto'] == 'ETANOL']['valor_de_venda'].mean())
media_etanol
```

Média do etanol: 4.51

**RESILIA**

# ANÁLISE EXPLORATÓRIA

## 3. QUAIS OS 5 ESTADOS COM O PREÇO MÉDIO DA GASOLINA E DO ETANOL MAIS CAROS?

#Gasolina:

```
gasolina_df = dados.loc[dados['produto'] == 'GASOLINA']
gasolina_df.groupby(['estado_sigla'])['valor_de_venda'].mean().apply(formatar).nlargest(5)
```

# Etanol:

```
etanol_df = dados.loc[dados['produto'] == 'ETANOL']
etanol_df.groupby(['estado_sigla'])['valor_de_venda'].mean().apply(formatar).nlargest(5)
```

GASOLINA

estado_sigla	valor_de_venda
AC	6.21
AM	6.19
PI	6.08
TO	6.05
BA	6.03

Name: valor\_de\_venda, dtype: float64

ETANOL

estado_sigla	valor_de_venda
AP	5.90
RS	5.66
PA	5.62
RO	5.60
RR	5.57

Name: valor\_de\_venda, dtype: float64

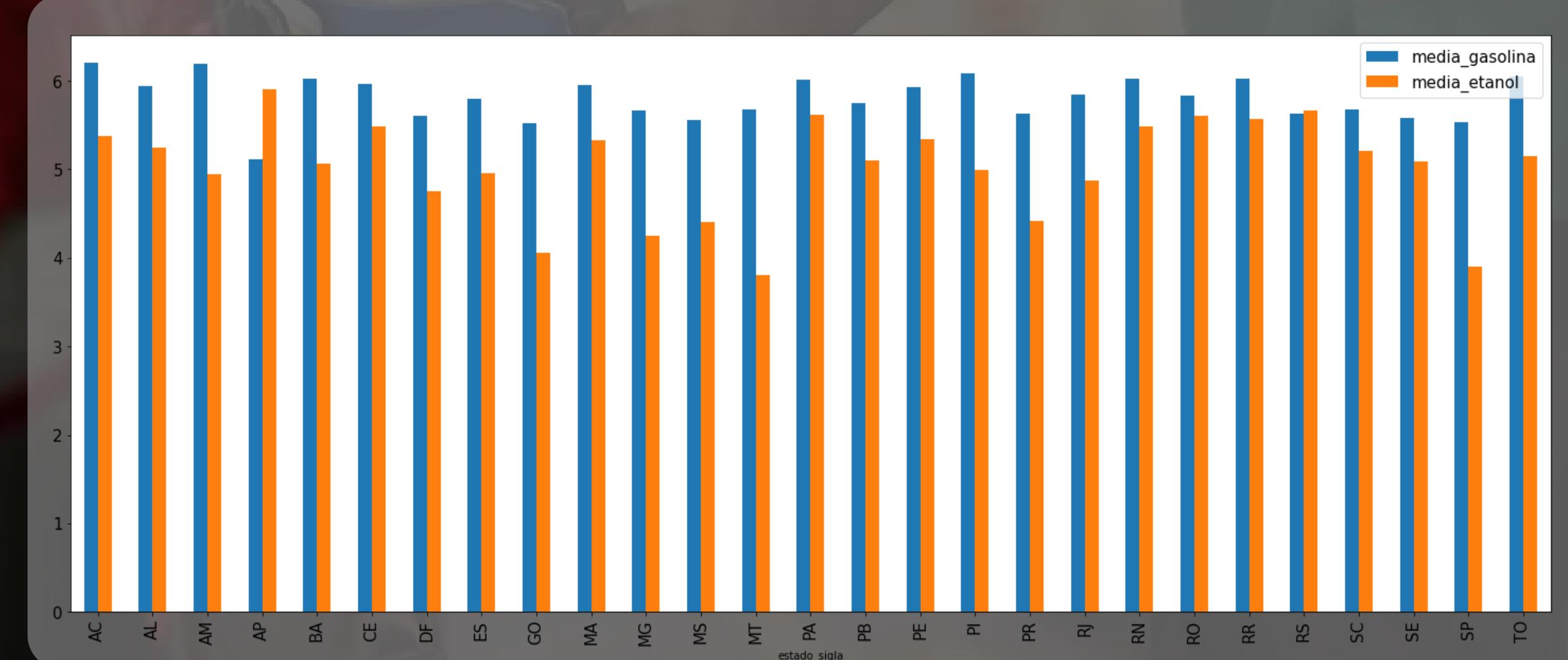
**RESILIA**

# ANÁLISE EXPLORATÓRIA

## 4. QUAL O PREÇO MÉDIO DA GASOLINA E DO ETANOL POR ESTADO?

```
estados_gasolina_etanol = pd.merge(estados_gasolina, estados_etanol, left_index=True, right_index=True)
estados_gasolina_etanol.rename(columns = {'valor_de_venda_x': 'media_gasolina', 'valor_de_venda_y': 'media_etanol'}, inplace = True)
estados_gasolina_etanol.head()
```

estado_sigla	media_gasolina	media_etanol
AC	6.21	5.38
AL	5.94	5.24
AM	6.19	4.95
AP	5.11	5.90
BA	6.03	5.07



# ANÁLISE EXPLORATÓRIA

5. QUAL O MUNICÍPIO QUE POSSUI O MENOR PREÇO PARA A GASOLINA E PARA O ETANOL?

Gasolina:

```
gasolina_df.loc[gasolina_df['valor_de_venda'] == gasolina_df['valor_de_venda'].min()]
```

Município com menor preço para gasolina: Passo Fundo

Etanol:

```
etanol_df.loc[etanol_df['valor_de_venda'] == etanol_df['valor_de_venda'].min()]
```

Município com menor preço etanol: Presidente Prudente

**RESILIA**

# ANÁLISE EXPLORATÓRIA

6. QUAL O MUNICÍPIO QUE POSSUI O MAIOR PREÇO PARA A GASOLINA E PARA O ETANOL?

Gasolina:

```
gasolina_df.loc[gasolina_df['valor_de_venda'] == gasolina_df['valor_de_venda'].max()]
```

Município com maior preço para gasolina: Gurupi

Etanol:

```
etanol_df.loc[etanol_df['valor_de_venda'] == etanol_df['valor_de_venda'].max()]
```

Município com maior preço para etanol: Bagé

**RESILIA**

# ANÁLISE EXPLORATÓRIA

## 7. QUAL A REGIÃO QUE POSSUI O MAIOR VALOR MÉDIO DA GASOLINA?

#descrever as siglas das regiões para melhor visualização das queries

```
dados_novo_arquivo.regiao_sigla=dados_brutos.regiao_sigla.map({'CO':'CO: Centro-Oeste', 'SE': 'SE: Sudeste', 'S':'S: Sul', 'NE': 'NE: Nordeste', 'N': 'Norte'})
```

```
regiao_maior_med_gas=dados_novo_arquivo[dados_novo_arquivo['produto']=='GASOLINA'].groupby(['regiao_sigla', 'produto']).aggregate({'valor_de_venda':'mean'}).round(3).sort_values('valor_de_venda', ascending=False)
regiao_maior_gas.style.background_gradient(cmap='Oranges')
```

regiao_sigla	produto	valor_de_venda
N: Norte	GASOLINA	5.972000
NE: Nordeste	GASOLINA	5.956000
S: Sul	GASOLINA	5.641000
SE: Sudeste	GASOLINA	5.616000
CO: Centro-Oeste	GASOLINA	5.582000



```
regiao_maior_gas.nlargest(1, columns = 'valor_de_venda')
```

regiao_sigla	produto	valor_de_venda
N: Norte	GASOLINA	5.972

**RESILIA**

# ANÁLISE EXPLORATÓRIA

## 8. QUAL A REGIÃO QUE POSSUI O MENOR VALOR MÉDIO DO ETANOL?

```
regiao_menor_med_etanol = dados_novo_arquivo[dados_novo_arquivo['produto']=='ETANOL'].groupby(['regiao_sigla',  
'produto']).aggregate( {"valor_de_venda":"mean"}).round(3).sort_values('valor_de_venda')  
regiao_menor_med_etanol.style.background_gradient(cmap='BuPu')
```

regiao_sigla	produto	valor_de_venda
CO: Centro-Oeste	ETANOL	4.107000
SE: Sudeste	ETANOL	4.150000
S: Sul	ETANOL	4.842000
NE: Nordeste	ETANOL	5.246000
N: Norte	ETANOL	5.341000



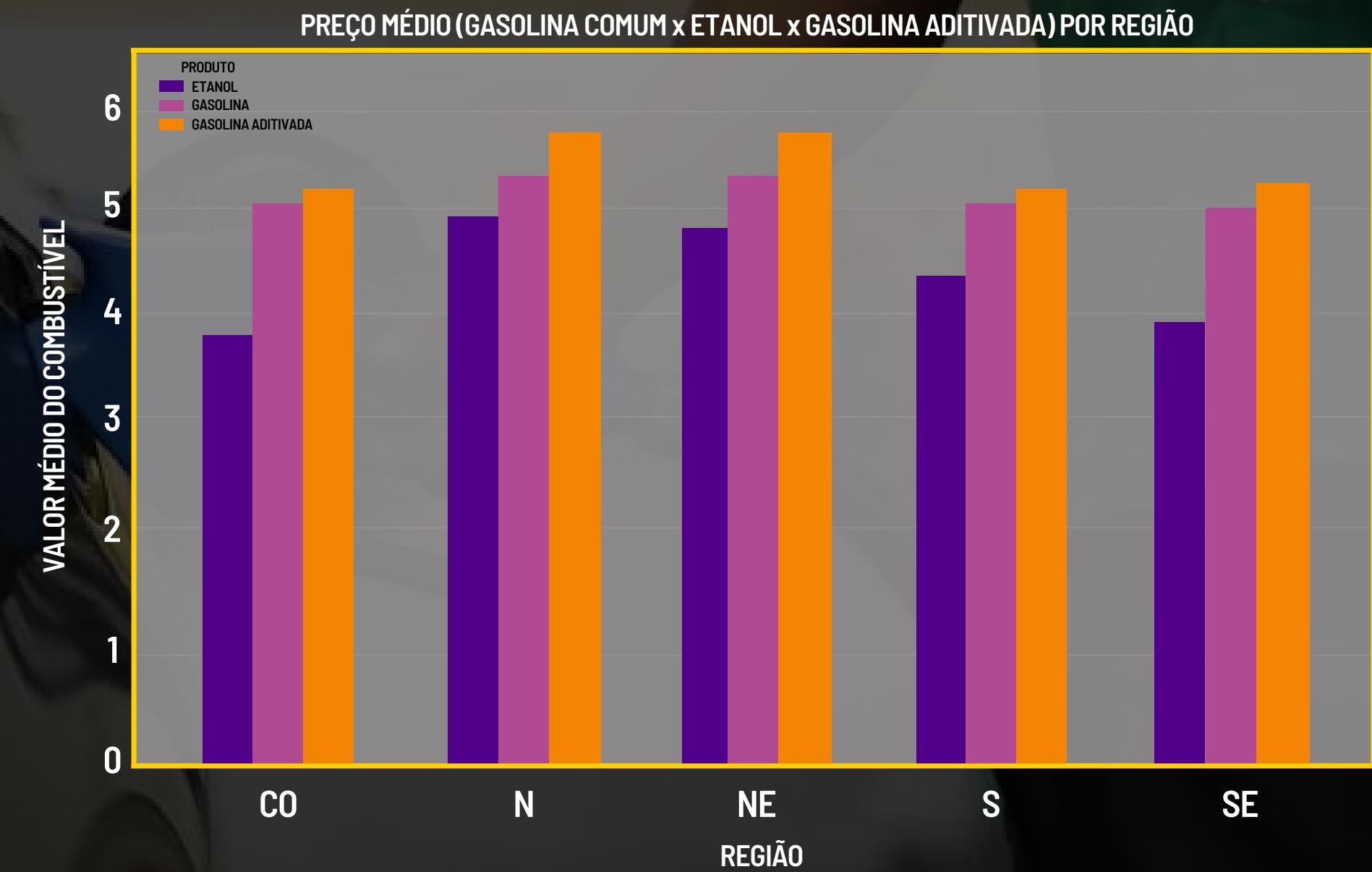
```
regiao_menor_med_etanol.nsmallest(1, columns = 'valor_de_venda')
```

regiao_sigla	produto	valor_de_venda
CO: Centro-Oeste	ETANOL	4.107

**RESILIA**

# ANÁLISE EXPLORATÓRIA

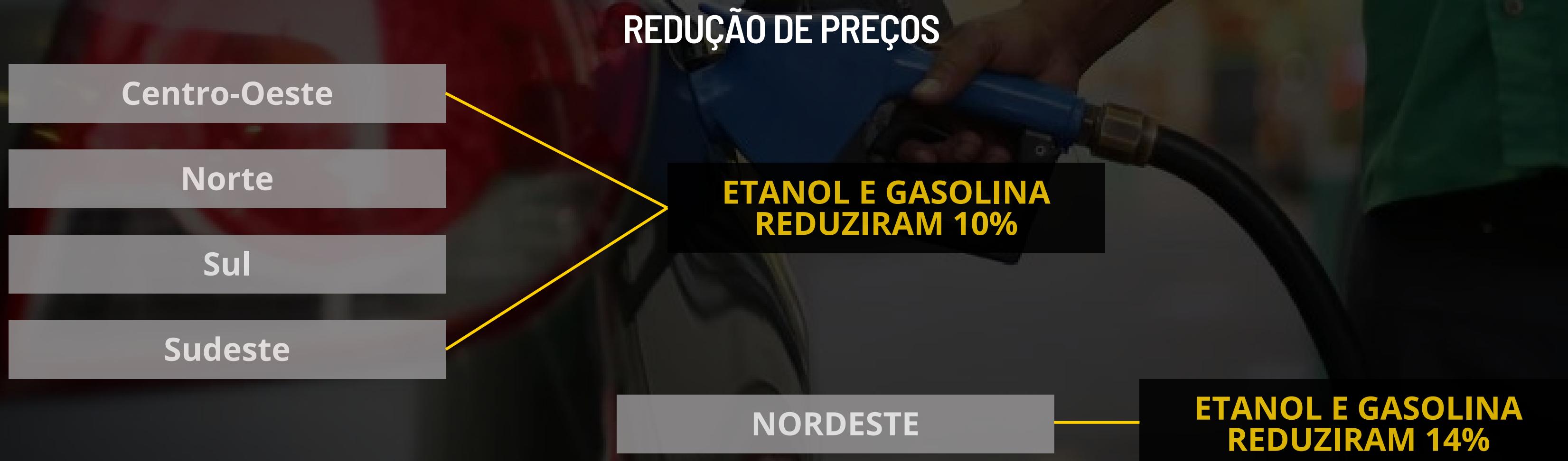
```
vl_medio_regiao = dados_novo_arquivo.groupby(['regiao_sigla',
'produto'])['valor_de_venda'].mean().reset_index()
grafico_val_med_reg = plt.figure(figsize=(60,30))
plt.title('PREÇO MÉDIO (GASOLINA COMUM x ETANOL x GASOLINA ADITIVADA) POR REGIÃO', fontsize=40)
sns.set(font_scale=3)
sns.barplot(y='valor_de_venda', x='regiao_sigla', hue='produto', data =
vl_medio_regiao, palette='plasma');
plt.ylabel('VALOR MÉDIO DA GASOLINA')
plt.xlabel('REGIÃO')
plt.legend(title="PRODUTO")
plt.tight_layout()
```



# ANÁLISE EXPLORATÓRIA

9. HÁ ALGUMA CORRELAÇÃO ENTRE O VALOR DO COMBUSTÍVEL (GASOLINA E ETANOL) E A REGIÃO ONDE ELE É VENDIDO?

```
dados.groupby(['regiao_sigla', 'produto', pd.Grouper(key='data_da_coleta', freq='M')]).agg({'valor_de_venda': 'mean'})
```



# ANÁLISE EXPLORATÓRIA

10. HÁ ALGUMA CORRELAÇÃO ENTRE O VALOR DO COMBUSTÍVEL  
(GASOLINA E ETANOL) E A BANDEIRA QUE VENDE ELE?

```
dados_brutos.groupby(['bandeira', 'produto', pd.Grouper(key='data_da_coleta', freq='M')]).agg({'valor_de_venda': 'mean'})
```

**BANDEIRAS QUE COBRAM MAIOR VALOR**  
\*2 DO NORTE E 1 DO NORDESTE

PETRONAC	6.190000
EQUADOR	5.935000
SP	5.934000

**BANDEIRAS QUE COBRAM MENOR VALOR**  
\*TODAS DE SÃO PAULO (SUDESTE)

TORRAO	5.008000
ROYAL FIC	4.983000
RUFF C.J.	4.643000

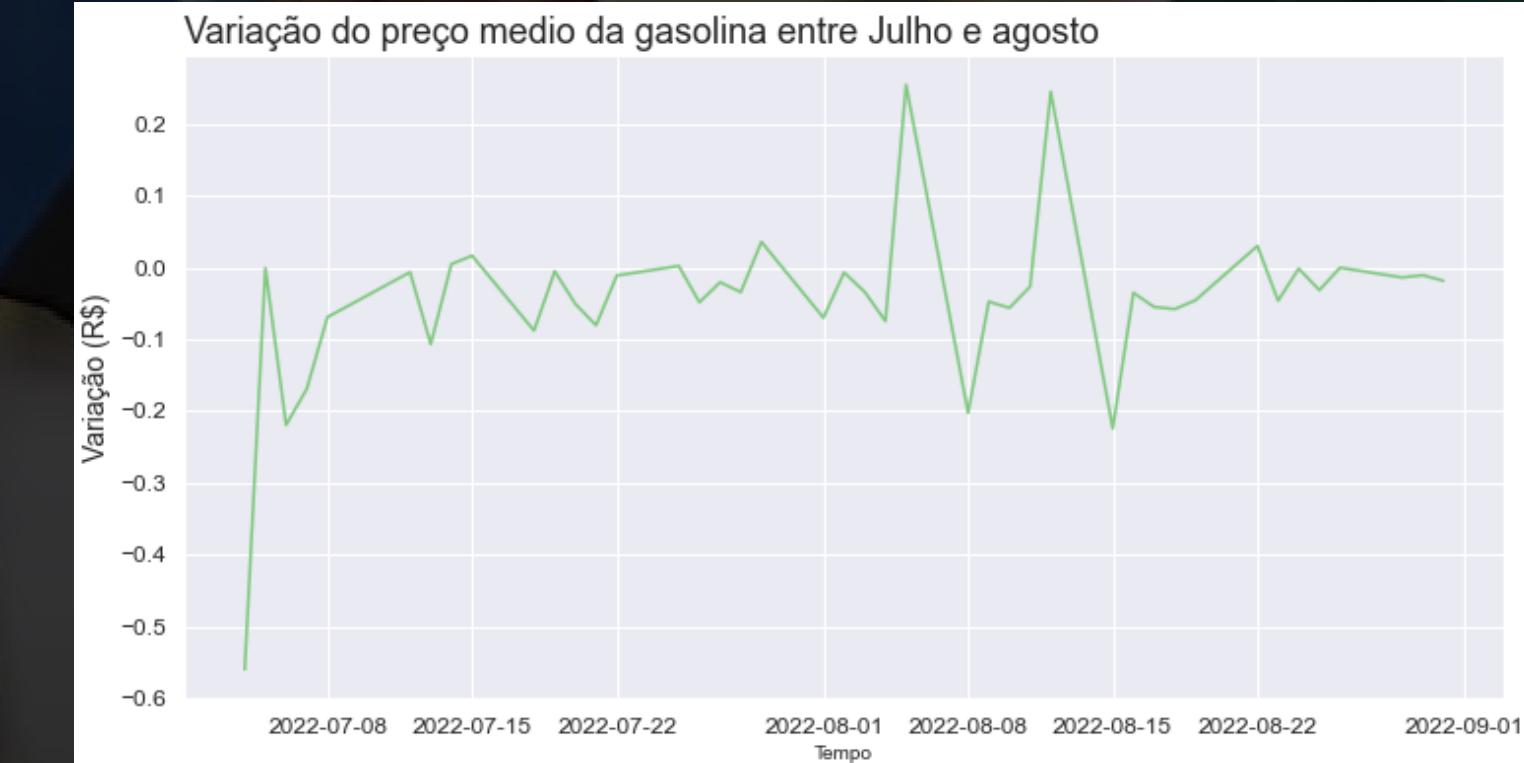
**RESILIA**

# ANÁLISE EXPLORATÓRIA

## 11. VARIAÇÃO DO PREÇO MÉDIO DA GASOLINA

```
dados_gasolina = dados.query("produto == 'GASOLINA')  
dados_valor_medio_gasolina = dados_gasolina.groupby("data_da_coleta").mean()  
dados_valor_medio_gasolina["variacao"] = dados_valor_medio_gasolina['valor_de_venda'].diff()  
dados_valor_medio_gasolina.head()
```

data_da_coleta	valor_de_venda	variacao
2022-07-01	7.48	NaN
2022-07-04	6.59	-0.560366
2022-07-05	6.59	-0.001112
2022-07-06	6.29	-0.001112
2022-07-07	5.99	-0.169634

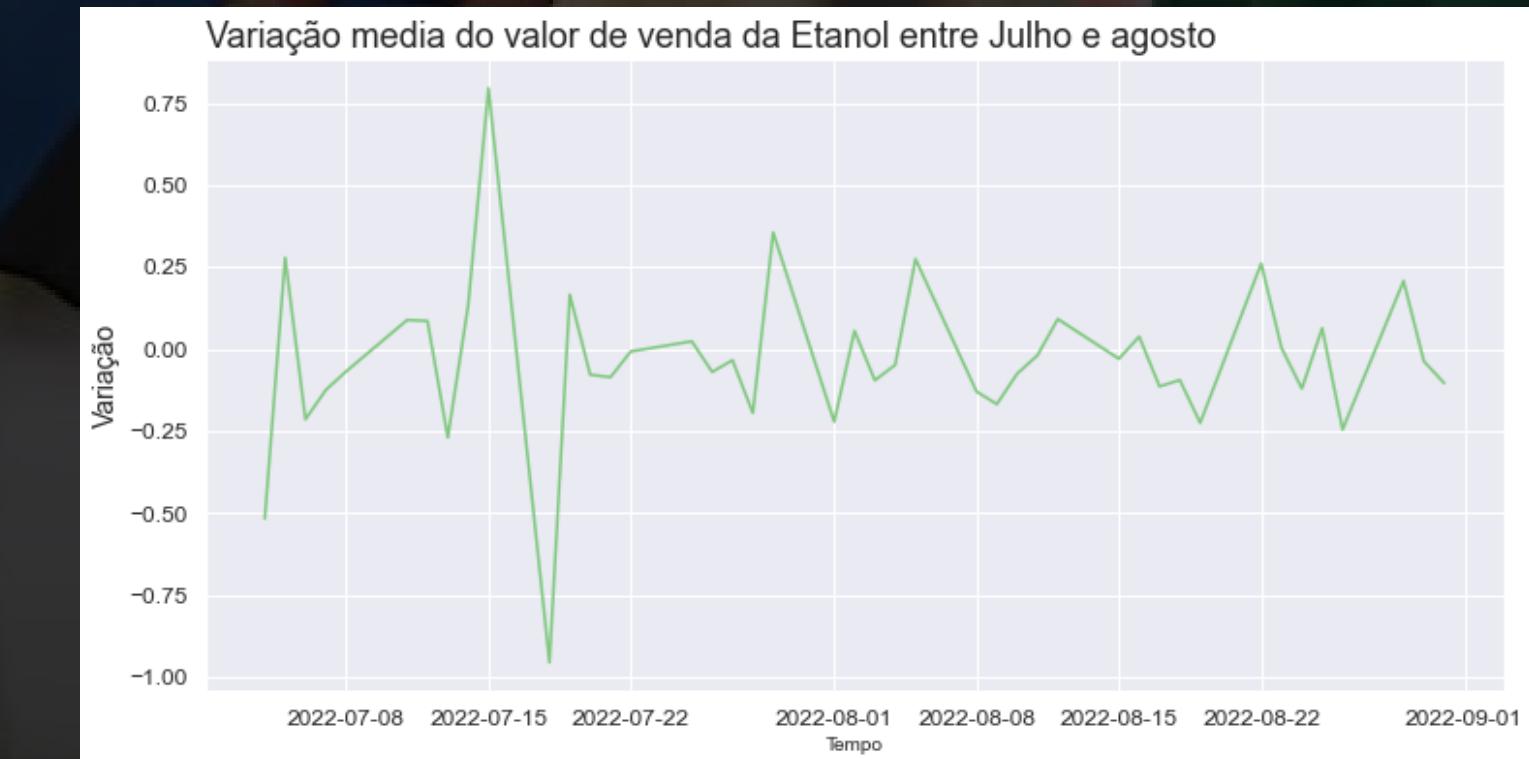


# ANÁLISE EXPLORATÓRIA

## 12. VARIAÇÃO DO PREÇO MÉDIO DO ETANOL

```
dados_etanol = dados.query("produto == 'ETANOL")
dados_valor_medio_etanol = dados_etanol.groupby("data_da_coleta").median()
dados_valor_medio_etanol
dados_valor_medio_etanol[["variacao"]] = dados_valor_medio_etanol['valor_de_venda'].diff()
dados_valor_medio_etanol.head()
```

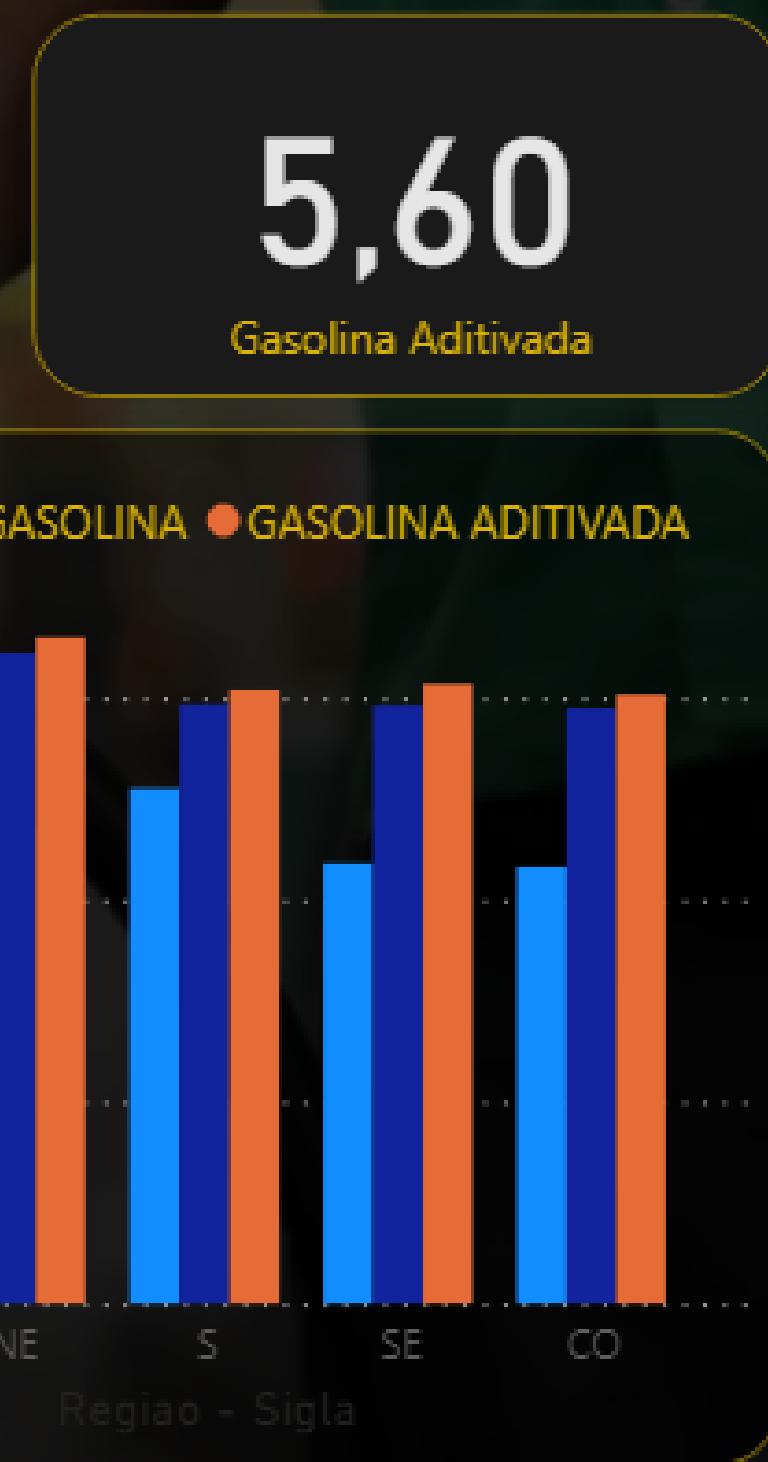
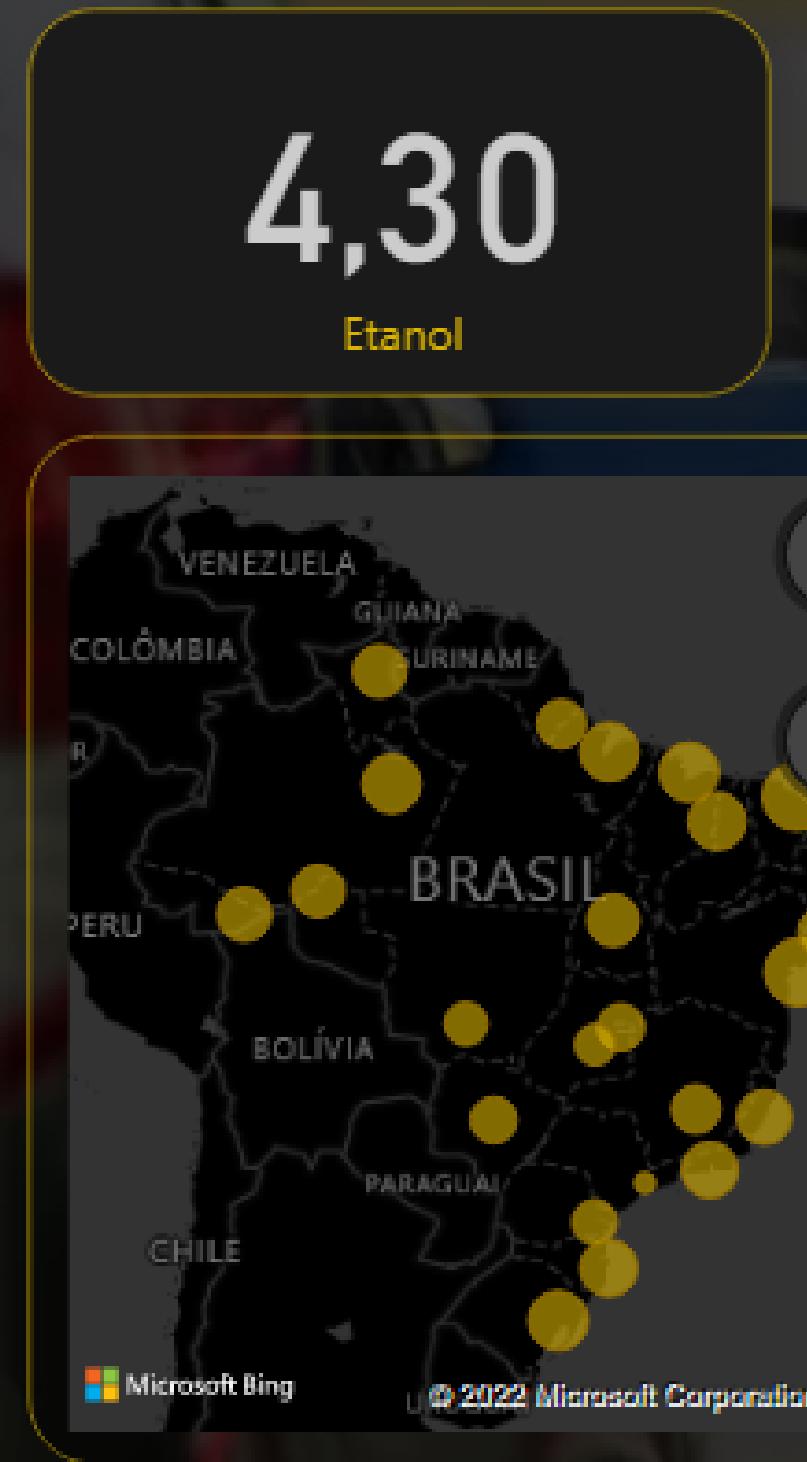
data_da_coleta	valor_de_venda	variacao
2022-07-01	5.570	NaN
2022-07-04	4.680	-0.516129
2022-07-05	5.185	0.278932
2022-07-06	4.750	-0.213570
2022-07-07	4.755	-0.124216



# BI

ESTADO
AC
AL
AM
AP
BA
CE
DF
ES
GO
MA
MG
MS
MT
PA

PRODUTO
GASOLINA ADITIVADA
GASOLINA



## RESILIA

# REFERÊNCIAS

[HTTPS://CBIE.COM.BR/ARTIGOS/O-QUE-E-A-ANP/](https://cbie.com.br/artigos/o-que-e-a-anp/)

[HTTPS://DADOS.GOV.BR/DATASET/SERIE-HISTORICA-DE-PRECOS-DE-COMBUSTIVEIS-POR-REVENDA](https://dados.gov.br/dataset/serie-historica-de-precos-de-combustiveis-por-revenda)

**RESILIA**



OBRIGADO(A)!

PROJETO  
CONCLUÍDO  
SQUAD 6