

CENTRO UNIVERSITÁRIO INTERNACIONAL- UNINTER
CURSO DE ENGENHARIA DE SOFTWARE

JOÃO PEDRO COSTA DE MENDONÇA – RU: 3618712
MOACIR DOMINGOS SILVA JÚNIOR – RU: 3539252

**ANÁLISE DE DADOS PÚBLICOS REFERENTES AOS
COMBUSTÍVEIS NOS ANOS DE 2020 E 2021**

PORTO ALEGRE-RS
2023

JOÃO PEDRO COSTA DE MENDONÇA – RU:3618712
MOACIR DOMINGOS SILVA JÚNIOR – RU: 3539252

ANALISANDO O IMPACTO DO PREÇO DOS COMBUSTÍVEIS NO DIA A DIA

Pesquisa apresentada ao curso
de Engenharia de Software,
apresentado na disciplina de
Atividade Extensionista II do
Centro Universitário
Internacional UNINTER.

PORTO ALEGRE-RS
2023

RESUMO

Este relatório trata-se de uma análise através dos dados disponibilizados no site do governo federal. O intuito é demonstrar os valores dos combustíveis nos anos de 2020 e 2021, anos que estão marcados na história como anos pandêmicos e de interferência da Guerra entre Rússia e Ucrânia. Pela influência econômica para um país e sua população, os combustíveis têm uma importância singular que o fazem ser explanados nesse documento.

SUMÁRIO

1. INTRODUÇÃO.....	3
2. CRONOGRAMA.....	4
3. PLANEJAMENTO INICIAL / METODOLOGIA.....	4
4. DESENVOLVIMENTO.....	5
4.1 OBTENÇÃO DOS DADOS E IMPORTAÇÃO DA BIBLIOTECA PANDAS.....	5
4.2 ANÁLISE E MANIPULAÇÃO DOS DADOS.....	6
5. CONCLUSÃO.....	13
6. REFERÊNCIAS.....	14

1. INTRODUÇÃO

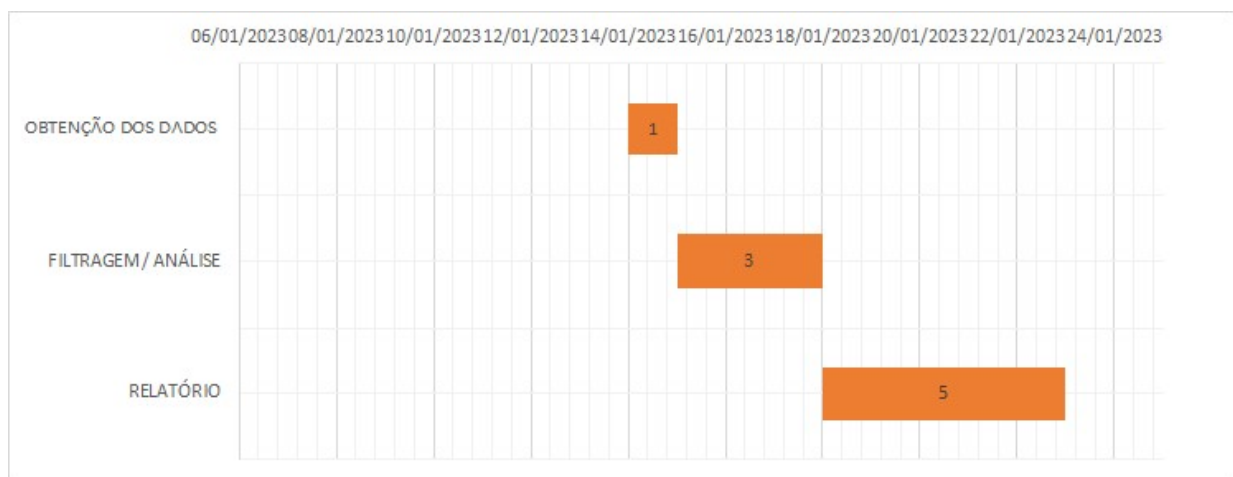
Este documento foi elaborado com o objetivo de relatar as atividades referentes a análise dos preços dos combustíveis nos anos de 2020 e 2021. Sua escolha se deu pelo fato de sua importância nacional e mundial. Dois fatores foram importantes para a escolha desse período de análise, o primeiro foi a pandemia da COVID-19 e o segundo foi a Guerra entre Rússia e Ucrânia, além da variação cambial do dólar. Fatores esses que aumentaram a inflação sobre os combustíveis, pois reduziram a oferta e prejudicaram as cadeias de distribuição.

2. CRONOGRAMA

As atividades desenvolvidas foram divididas em três fases: obtenção dos dados públicos, filtragem/ análise dos dados e relatório.

A obtenção dos dados do governo é a parte mais rápida do processo pois todos esses dados estão disponibilizados no site do governo federal, bastando apenas baixá-los. Na segunda fase, foi a parte mais demorada pois foi necessário gerar o código em linguagem *Python* para se obter o resultado desejado, esta fase teve duração de 01 (uma) semana. E por fim, na terceira fase, foi menos demorada que a segunda fase pois basicamente foi fazer uma comparação com os dados.

FIGURA 01 - GRÁFICO DE GANTT DAS ATIVIDADES



Fonte: O autor

3. PLANEJAMENTO INICIAL/ METODOLOGIA

O planejamento inicial estava voltado para analisar os dados públicos e gerar um estudo em cima dele, visto que era um dos exemplos de temas propostos para a atividade. A notória importância dos combustíveis sendo sempre o foco de grande parte dos noticiários e de abordagens políticas, despertou o interesse em realizar este estudo.

Utilizando a linguagem de programação *Python* de alto nível, ou seja, com sintaxe mais simplificada e próxima da linguagem humana, utilizada nas mais diversas aplicações, incluindo a ciência e análise de dados. Dentro das linguagens de programação existem as bibliotecas, sendo estas uma coleção de recursos e funções

usados por programa de computador e que foram desenvolvidas e disponibilizadas para o uso pelos desenvolvedores. A biblioteca utilizada nesta atividade foi a *Pandas*, biblioteca criada para a linguagem *Python* para a manipulação e análise de dados, oferecendo estruturas e operações para manipular tabelas numéricas e séries temporais.

4. DESENVOLVIMENTO

4.1 OBTENÇÃO DOS DADOS E IMPORTAÇÃO DA BIBLIOTECA PANDAS

Os dados em formato de tabelas (.csv) foram baixados a partir do site www.gov.br e após isso alocados no *Google Drive* para que pudessem ser manipulados através do *Google Colab*. As informações estavam separadas por semestres, sendo necessário a junção dos *Dataframes* em um único, para que a análise seja anual, conforme Figura 1.

Figura 1: Conectando com o Google Drive

▼ Importando bibliotecas

```
[ ] import pandas as pd
    from google.colab import drive
```

▼ Importando as tabelas dos combustíveis 2020

```
▶ drive.mount('/content/drive')
df_2020_1 = pd.read_csv('/content/drive/MyDrive/dados/combustivel/2020-01.csv', sep=';', encoding='ISO-8859-1')

[ ] drive.mount('/content/drive')
df_2020_2 = pd.read_csv('/content/drive/MyDrive/dados/combustivel/2020-02.csv', sep=';', encoding='ISO-8859-1')

[ ] df_Com2020 = pd.concat([df_2020_1, df_2020_2], ignore_index = True)
    df_Com2020
```

Fonte: O autor

Figura 2: Amostra dos dados contidos nos Dataframes

719296	SE	RJ	RIO DE JANEIRO	AUTO POSTO REDE PARCEIRO ANIL LTDA	30.865.289/0001-72	AVENIDA TENENTE CORONEL MUNIZ DE ARAGAO	0	LOT ESCOLA PAL 30470	ANIL 22765-006	DIESEL	29/12/2020
719297	SE	RJ	RIO DE JANEIRO	AUTO POSTO REDE PARCEIRO ANIL LTDA	30.865.289/0001-72	AVENIDA TENENTE CORONEL MUNIZ DE ARAGAO	0	LOT ESCOLA PAL 30470	ANIL 22765-006	ETANOL	29/12/2020
719298	SE	RJ	RIO DE JANEIRO	AUTO POSTO REDE PARCEIRO ANIL LTDA	30.865.289/0001-72	AVENIDA TENENTE CORONEL MUNIZ DE ARAGAO	0	LOT ESCOLA PAL 30470	ANIL 22765-006	GASOLINA ADITIVADA	29/12/2020
719299	SE	RJ	RIO DE JANEIRO	AUTO POSTO REDE PARCEIRO ANIL LTDA	30.865.289/0001-72	AVENIDA TENENTE CORONEL MUNIZ DE ARAGAO	0	LOT ESCOLA PAL 30470	ANIL 22765-006	DIESEL S10	29/12/2020

719300 rows x 16 columns

Fonte: O autor

4.2 ANÁLISE E MANIPULAÇÃO DOS DADOS

A manipulação dos dados foi feita com o auxílio da biblioteca Pandas, inicialmente as tabelas estavam separadas pelos semestres de cada ano, sendo necessário unir as tabelas dos mesmos anos, como foi feito nos anos de 2020 e 2021, conforme mostrado nas Figuras 1 e 3.

Figura 3: Conectando com o Google Drive e unindo as tabelas

```
[5] drive.mount('/content/drive')
df_2021_1 = pd.read_csv('/content/drive/MyDrive/dados/combustivel/2021-01.csv', sep=';', encoding='ISO-8859-1')

Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).

[6] drive.mount('/content/drive')
df_2021_2 = pd.read_csv('/content/drive/MyDrive/dados/combustivel/2021-01.csv', sep=';', encoding='ISO-8859-1')

Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).

[7] df_Com2021 = pd.concat([df_2021_1, df_2021_2], ignore_index = True)
df_Com2021
```

Fonte: O autor

Figura 4: Amostra dos dados contidos na tabela de 2021

669358	SE	ES	ARACRUZ	COMBUSTIVEL SHOW EIRELI	23.797.660/0001- 40	AVENIDA CORONEL VENANCIO FLORES	28	NaN	CENTRO	29190- 010	ETANOL	28/06/2021	4,79
669359	SE	ES	ARACRUZ	POSTO DE COMBUSTIVEL SHOW EIRELI	23.797.660/0001- 40	AVENIDA CORONEL VENANCIO FLORES	28	NaN	CENTRO	29190- 010	DIESEL	28/06/2021	4,19
669360	SE	ES	ARACRUZ	POSTO DE COMBUSTIVEL SHOW EIRELI	23.797.660/0001- 40	AVENIDA CORONEL VENANCIO FLORES	28	NaN	CENTRO	29190- 010	GASOLINA ADITIVADA	28/06/2021	5,67
669361	SE	ES	ARACRUZ	POSTO DE COMBUSTIVEL SHOW EIRELI	23.797.660/0001- 40	AVENIDA CORONEL VENANCIO FLORES	28	NaN	CENTRO	29190- 010	DIESEL S10	28/06/2021	4,24

669362 rows x 16 columns

Fonte: O autor

Após essa etapa de unificar as tabelas, foi feita uma filtragem nos dados para ser utilizado somente aqueles que fossem importantes para este relatório. Com isso, algumas informações foram descartas, como por exemplo: “Endereço” ou a “Bandeira do Posto”. Permanecendo a visualização das colunas: “Estado”, “Produto” e “Valor de Venda”, conforme Figura 5.

Figura 5: Filtrando os dados

Filtrando para utilizar as colunas que interessam para a análise

```
[8] df_Com2020.columns.values

array(['i»¿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)

[9] df_Com2020_Col = ['Estado - Sigla', 'Produto', 'Valor de Venda']

[10] df_Filtrado2020 = df_Com2020.filter(items = df_Com2020_Col)
df_Filtrado2020 = df_Filtrado2020.rename(columns = {'Estado - Sigla': 'Estado'})
df_Filtrado2020
```

	Estado	Produto	Valor de Venda
0	SP	GASOLINA	4,399
1	SP	ETANOL	3,199
2	SP	DIESEL S10	3,899

Fonte: O autor

Em seguida, foram separados os tipos de combustíveis para que a análise e a média de preço fossem feitas de acordo com seu tipo, sendo observado que o Etanol não foi utilizado neste estudo. Separando assim em “Gasolina”, “Gasolina Aditivada”, “Etanol”, “Diesel S10” e “Diesel”, conforme Figura 6.

Figura 6: Tipos de combustíveis

```
df_Filtrado2021 = df_Com2021.filter(items = df_Com2021.Col)
df_Filtrado2021 = df_Filtrado2021.rename(columns = {'Estado - Sigla':'Estado'})
df_Filtrado2021
```

	Estado	Produto	Valor de Venda
0	RS	GASOLINA	4,599
1	RS	ETANOL	4,199
2	RS	GASOLINA ADITIVADA	4,799
3	RS	DIESEL	3,499
4	RS	DIESEL S10	3,599
...
669357	ES	GASOLINA	5,67
669358	ES	ETANOL	4,79
669359	ES	DIESEL	4,19
669360	ES	GASOLINA ADITIVADA	5,67
669361	ES	DIESEL S10	4,24

669362 rows x 3 columns

Fonte: O autor

Para deixar mais detalhado este estudo, foi feita uma filtragem para inicialmente coletar as informações somente dos 2 (dois) tipos de gasolinas presentes nos Dataframes dos respectivos anos. Com isso, o resultado obtido foram as linhas referentes a “Gasolina” e “Gasolina Aditivada”, conforme Figuras 7 e 8. Para se obter esse resultado foi necessário utilizar o com “regex” utilizando o acento circunflexo, onde corresponde ao início da string, e no modo MULTILINE também corresponde imediatamente após cada nova linha. O mesmo processo foi feito para os tipos de diesel, conforme Figuras 9 e 10.

Figura 7: Filtrando pelo tipo de Gasolina em 2020

```
df_Com2020_Col = ['Estado - Sigla', 'Produto', 'Valor de Venda']
dfGasolina = df_Filtrado2020.filter(items=df_Com2020_Col)
regex = '^GASOLINA'
dfGasolina = dfGasolina.loc[dfGasolina['Produto'].str.contains(regex, regex=True)]
dfGasolina
```

	Produto	Valor de Venda
0	GASOLINA	4,399
4	GASOLINA	4,69
8	GASOLINA	4,46
11	GASOLINA	5,2
14	GASOLINA	5,2
...
719289	GASOLINA ADITIVADA	4,599
719292	GASOLINA	4,399
719293	GASOLINA	4,879
719295	GASOLINA	4,999
719298	GASOLINA ADITIVADA	5,199

229626 rows x 2 columns

Fonte: O autor

Figura 8: Filtrando pelo tipo de Gasolina em 2021

```
df_Com2021_Col = ['Estado - Sigla', 'Produto', 'Valor de Venda']
dfGasolina21 = df_Filtrado2021.filter(items=df_Com2021_Col)
regex = '^GASOLINA'
dfGasolina21 = dfGasolina21.loc[dfGasolina21['Produto'].str.contains(regex, regex=True)]
dfGasolina21
```

	Produto	Valor de Venda
0	GASOLINA	4,599
2	GASOLINA ADITIVADA	4,799
5	GASOLINA	4,459
7	GASOLINA ADITIVADA	4,499
9	GASOLINA	4,539
...
669352	GASOLINA ADITIVADA	5,559
669353	GASOLINA	5,459
669354	GASOLINA	5,59
669357	GASOLINA	5,67
669360	GASOLINA ADITIVADA	5,67

296900 rows x 2 columns

Fonte: O autor

Figura 9: Filtrando pelo tipo de Diesel em 2020

```
df_Com2020_Col = ['Estado - Sigla', 'Produto', 'Valor de Venda']
dfDiesel = df_Filtrado2020.filter(items=df_Com2020_Col)
regex = '^DIESEL'
dfDiesel = dfDiesel.loc[dfDiesel['Produto'].str.contains(regex, regex=True)]
dfDiesel
```

	Produto	Valor de Venda
2	DIESEL S10	3,899
6	DIESEL S10	3,78
10	DIESEL S10	3,69
12	DIESEL	4,85
13	DIESEL S10	4,95
...
719287	DIESEL S10	3,499
719288	DIESEL S10	3,869
719290	DIESEL	3,799
719296	DIESEL	3,699
719299	DIESEL S10	3,799

286695 rows x 2 columns

Fonte: O autor

Figura 10: Filtrando pelo tipo de Diesel em 2021

```
df_Com2021_Col = ['Estado - Sigla', 'Produto', 'Valor de Venda']
dfDiesel21 = df_Filtrado2021.filter(items=df_Com2021_Col)
regex = '^DIESEL'
dfDiesel21 = dfDiesel21.loc[dfDiesel21['Produto'].str.contains(regex, regex=True)]
dfDiesel21
```

	Produto	Valor de Venda
3	DIESEL	3,499
4	DIESEL S10	3,599
8	DIESEL	3,419
11	DIESEL S10	3,699
14	DIESEL S10	3,899
...
669349	DIESEL	4,329
669350	DIESEL S10	4,359
669356	DIESEL S10	4,48
669359	DIESEL	4,19
669361	DIESEL S10	4,24

204288 rows x 2 columns

Fonte: O autor

E por fim, obteve-se os valores referentes ao preço médio das gasolinas, conforme Figuras 11 e 12, e o dos tipos de diesel conforme Figuras 13 e 14.

Figura 11: Preço médio das Gasolinas em 2020

```
dfGasolina['Valor de Venda'] = dfGasolina['Valor de Venda'].str.replace(',', '.')
dfGasolina = dfGasolina.astype({'Valor de Venda': 'float32'})
dfGasolina = dfGasolina.groupby('Produto').mean()
dfGasolina
```

Valor de Venda	
Produto	
GASOLINA	4.280126
GASOLINA ADITIVADA	4.587793

Fonte: O autor

Figura 12: Preço Médio das Gasolinas em 2021

```
dfGasolina21['Valor de Venda'] = dfGasolina21['Valor de Venda'].str.replace(',', '.')
dfGasolina21 = dfGasolina21.astype({'Valor de Venda': 'float32'})
dfGasolina21 = dfGasolina21.groupby('Produto').mean()
dfGasolina21
```

Valor de Venda	
Produto	
GASOLINA	5.346717
GASOLINA ADITIVADA	5.495662

Fonte: O autor

Figura 13: Preço do Diesel em 2020

```
[ ] dfDiesel['Valor de Venda'] = dfDiesel['Valor de Venda'].str.replace(',','.')
dfDiesel = dfDiesel.astype({'Valor de Venda':'float32'})
dfDiesel = dfDiesel.groupby('Produto').mean()
dfDiesel
```



Valor de Venda	
Produto	
DIESEL	3.417827
DIESEL S10	3.508126

Fonte: O autor

Figura 14: Preço do Diesel em 2021

```
▶ dfDiesel21['Valor de Venda'] = dfDiesel21['Valor de Venda'].str.replace(',','.')
dfDiesel21 = dfDiesel21.astype({'Valor de Venda':'float32'})
dfDiesel21 = dfDiesel21.groupby('Produto').mean()
dfDiesel21
```



Valor de Venda	
Produto	
DIESEL	4.256016
DIESEL S10	4.309581

Fonte: O autor

6. CONCLUSÃO:

O valor dos combustíveis se mantém alto, fato que implica diretamente na economia do país, visto que 75% do transporte país é feito pelo modal rodoviário, conforme uma pesquisa divulgada pela Fundação Dom Cabral. Devido a essa dependência, a economia do país é impactada, a alimentação fica mais cara, o transporte, a saúde e a educação. Com isso, a inflação também aumenta e o salário mínimo não acompanha esse aumento, fazendo com que a população perca parte do poder de compra. Outra questão importante foi a média de desemprego no país que chegou aos 14,2% em 2020, a maior já registrada desde de 2012, conforme noticiado na página de economia do site Uol.

Já em 2021, a média anual de desemprego foi de 13,2% ainda sofrendo pelos impactos do aumento dos combustíveis, Guerra na Ucrânia e pandemia. Segundo o economista Flauzino Antunes o setor agrícola é o que mais utiliza combustível, portanto foi fortemente impactado.

Destarte, o controle dos preços dos combustíveis é de suma importância para a economia do país e a revisão e diminuição da dependência do modal rodoviário é outro ponto a ser observado com mais cuidado, para que no futuro não se tenha implicações tão relevantes no dia a dia devido ao aumento dos combustíveis.

BIBLIOGRAFIA:

Série Histórica de Preços de Combustíveis e de GLP. **GOV.BR**,2023. Disponível em: <<https://www.gov.br/anp/pt-br/centrais-de-conteudo/dados-abertos/serie-historica-de-precos-de-combustiveis>>. Acesso em: 14 de jan. de 2023.

Operações com expressões regulares. **DOCS.PYTHON**,2023. Disponível em: <<https://docs.python.org/pt-br/3/library/re.html>>. Acesso em: 20 de jan. de 2023.

Salário mínimo desvaloriza ‘um tanque de gasolina’ desde 2019.

NOTÍCIAS.R7.COM,2021. Disponível em: <<https://noticias.r7.com/economia/salario-minimo-desvaloriza-um-tanque-de-gasolina-desde-2019-06072022>>. Acesso em: 21 de jan. de 2023.

Modal rodoviário é o mais utilizado no Brasil.

SUMMITMOBILIDADE.ESTADAO.COM.BR,2022. Disponível em: <[Alta dos preços, subida do diesel e as consequências na sociedade de aumentos seguidos em todos os setores. **LIVEFMCAMPOS.COM**,2021. Disponível em:](https://summitmobilidade.estadao.com.br/ir-e-vir-no-mundo/modal-rodoviario-e-o-mais-utilizado-no-brasil/#:~:text=Uma%20pesquisa%20dos%20Custos%20Log%C3%ADsticos,dos%20governos%20federais%20e%20estaduais.>”. Acesso em: 21 de jan. de 2023.</u></p></div><div data-bbox=)

<[Desemprego no país cai para 11,1% no quarto trimestre de 2021. **GOV.BR**,2022.](https://www.livefmcampos.com/alta-dos-precos-subida-do-diesel-e-as-consequencias-na-sociedade-de-aumentos-seguidos-em-todos-os-setores/#:~:text=O%20MEGA%20AUMENTO%20NO%20PRE%C3%87O,AT%C3%89%20O%20FINAL%20DO%20ANO.>”. Acesso em: 21 de jan. de 2023.</u></p></div><div data-bbox=)

Disponível em: <<https://www.gov.br/pt-br/noticias/trabalho-e-previdencia/2022/02/desemprego-no-pais-cai-para-11-1-no-quarto-trimestre-de-2021>>. Acesso em: 21 de jan. de 2023.

Como aumento dos combustíveis afeta preço de outros produtos no dia a dia ?

Especialistas explicam. **G1.GLOBO.COM**,2022. Disponível em: <<https://g1.globo.com/df/distrito-federal/noticia/2022/03/26/como-aumento-dos-combustiveis-afeta-preco-de-outros-produtos-no-dia-a-dia-especialistas-explicam.ghtml>>. Acesso em: 21 de jan. de 2023.