



UFRJ

UNIVERSIDADE FEDERAL
DO RIO DE JANEIRO

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

TRABALHO 1 | TÓP. ESP. EM ENG. DE DADOS
RELATÓRIO DE ANÁLISE DE DADOS: PERFIL DO ELEITORADO E DADOS
METEOROLÓGICOS

DHAYSE DE LIMA TITO DRE: 120019062
GIOVANNI ANSELMO FERNANDES COELHO DE ORNELLAS DRE: 120136123
GUILHERME GUEDES GUIMARÃES MERÇON DRE: 117098748
EDUARDO DAEMON TEIXEIRA DOS SANTOS DRE: 120154798

Rio de Janeiro-RJ

2024

SUMÁRIO

1	ANÁLISE DO PERFIL DO ELEITORADO	3
1.1	INTRODUÇÃO	3
1.2	DESCRIÇÃO DO DATASET	3
1.2.1	Origem dos Dados	3
1.3	CARGA	3
1.3.0.1	Contagem de linhas, colunas e células	4
1.4	TRANSFORMAÇÕES NOS DADOS	6
1.4.1	Projeção de Colunas	6
1.4.2	Eliminação de Linhas Inválidas	6
1.5	ANÁLISES REALIZADAS	7
1.5.1	Tabelas Dinâmicas	7
1.5.2	Pacote Suplementar do Excel	9
1.5.3	Filtros Aplicados	10
1.5.4	Resultados da Análise do Dataset01	15
2	ANÁLISE DE DADOS METEOROLÓGICOS	18
2.1	INTRODUÇÃO	18
2.2	DESCRIÇÃO DO DATASET	18
2.2.1	Origem dos Dados	18
2.2.2	Carga	18
2.2.2.1	Importação dos Dados	18
2.2.2.2	Contagem de linhas, colunas e células	18
2.2.3	Variáveis Seleccionadas	19
2.3	TRANSFORMAÇÕES NOS DADOS	19
2.3.1	Eliminação de Colunas	19
2.3.2	Eliminação de Linhas Inválidas	19
2.4	ANÁLISES REALIZADAS	22
2.4.1	Análise Trimestral da Temperatura do Ar	22
2.4.2	Cálculos das Estatísticas Básicas	22
2.5	ANÁLISE COMPARATIVA ENTRE 2022 E 2023	25
2.5.1	Precipitação Total	25
2.5.1.1	Tendência Geral ao Longo do Ano	26
2.5.1.2	Comparação Mensal e Horária	26
2.5.1.3	Padrão Horário de 2023	26
2.5.1.4	Pontos de Destaque	26
2.5.2	Pressão Atmosférica	27
2.5.3	Temperatura do Ar	29
2.5.3.1	Média de Temperatura Horária (2022 vs. 2023)	30

2.5.3.1.1	<i>Tendência:</i>	30
2.5.3.1.2	<i>Comparação:</i>	30
2.5.3.1.3	<i>Pontos de Destaque:</i>	30
2.5.3.2	Média de Temperatura Mensal (2022 vs. 2023)	31
2.5.3.2.1	<i>Tendência:</i>	31
2.5.3.2.2	<i>Comparação:</i>	31
2.5.3.2.3	<i>Pontos de Destaque:</i>	31
2.5.4	Vento: Direção e Velocidade	32
2.5.4.1	Análise da Média de Vento e Rajada Máxima	32
2.6	ANÁLISE DE TEMPERATURA POR AGRUPAMENTO TEMPORAL .	34
2.7	FÓRMULAS UTILIZADAS	34
3	RESULTADO DAS ANÁLISE	35
3.1	LINK PARA O REPOSITÓRIO	35

1 ANÁLISE DO PERFIL DO ELEITORADO

1.1 INTRODUÇÃO

Neste capítulo, é apresentada a análise do perfil do eleitorado para as eleições municipais de 2024, comparando São Paulo e Rio de Janeiro. O objetivo é explorar características demográficas e fornecer insights que possam ser utilizados por campanhas políticas.

1.2 DESCRIÇÃO DO DATASET

1.2.1 Origem dos Dados

Os dados foram obtidos no site de dados abertos do Tribunal Superior Eleitoral - Assessoria de Gestão Eleitoral (TSE/AGEL: https://dadosabertos.tse.jus.br/pt_PT/dataset/eleitorado-atual), contendo informações sobre os eleitores de todo o Brasil, se encontrando, por exemplo, separados por Município e por local de votação (todos UFs). Aqui tomamos apenas os datasets avulsos do Rio de Janeiro e de São Paulo.

1.3 CARGA

Conforme as orientações do TRE em anexo aos datasets, o arquivo se encontra com os caracteres codificados em "Latin 1" e os campos estão entre aspas e separados por ponto e vírgula. Assim, realizamos a importação de dados no Excel através do PowerQuery conforme as seguintes etapas:

1. Abrimos o Excel e acessamos a aba **Dados**.
2. Clicamos em **Obter Dados Externos** e, em seguida, selecionamos a opção **De Texto/CSV**.
3. Na janela aberta, escolhemos a opção **Delimitador** e definimos o delimitador como **ponto e vírgula (;)**.
4. No campo **Origem do arquivo**, selecionamos **Windows (ANSI)**.
5. No campo **Deteção do Tipo de Dados**, selecionamos **Com base nas primeiras 200 linhas**.
6. Para preparar os dados como se encontram, selecionamos a opção **Carregar para...** e optamos por **Conexão**.

Dessa forma, preparamos os dados para poder posteriormente transformá-los e carregá-los para tabelas dinâmicas mais compactas, evitando, assim, atingir o limite de carga do programa que encontramos ao tentar importar os dados para planilhas simples, visto que os datasets estão na ordem de centenas de milhões de células.

1.3.0.1 Contagem de linhas, colunas e células

Inicialmente, utilizando ferramentas embutidas do windows podemos fazer uma contagem sequencial de linhas e colunas presentes em nossos datasets, incluindo o cabeçalho e colunas indesejadas. Verificado após a carga que não há células vazias ou nulas, podemos tomar o número de linhas à menos do cabeçalho multiplicado pelo número de colunas para obter o total de células com dados.

```

Administrador: Windows PowerShell

PS C:\Windows\system32> cd D://Downloads
PS D:\Downloads> Get-Content .\perfil_eleitor_secao_ATUAL_RJ\perfil_eleitor_secao_ATUAL_RJ.csv | Measure-Object -Line
>>
    Lines Words Characters Property
    -----
6455238

PS D:\Downloads>
PS D:\Downloads>
PS D:\Downloads> Get-Content .\perfil_eleitor_secao_ATUAL_RJ\perfil_eleitor_secao_ATUAL_RJ.csv | Measure-Object -Line
>>
    (Get-Content .\perfil_eleitor_secao_ATUAL_RJ\perfil_eleitor_secao_ATUAL_RJ.csv -First 1) -split ';'
>>
"DT_GERACAO"
"HH_GERACAO"
"ANO_ELEICAO"
"SG_UF"
"CD_MUNICIPIO"
"NM_MUNICIPIO"
"NR_ZONA"
"NR_SECAO"
"NR_LOCAL_VOTACAO"
"NM_LOCAL_VOTACAO"
"CD_GENERO"
"DS_GENERO"
"CD_ESTADO_CIVIL"
"DS_ESTADO_CIVIL"
"CD_FAIXA_ETARIA"
"DS_FAIXA_ETARIA"
"CD_GRAU_ESCOLARIDADE"
"DS_GRAU_ESCOLARIDADE"
"CD_RACA_COR"
"DS_RACA_COR"
"CD_IDENTIDADE_GENERO"
"DS_IDENTIDADE_GENERO"
"CD_QUILOMBOLA"
"DS_QUILOMBOLA"
"CD_INTERPRETE_LIBRAS"
"DS_INTERPRETE_LIBRAS"
"QT_ELEITORES_PERFIL"
"QT_ELEITORES_BIOMETRIA"
"QT_ELEITORES_DEFICIENCIA"
"QT_ELEITORES_INC_NM_SOCIAL"
PS D:\Downloads>

```

(a) Linhas de comando para o dataset RJ

```

Administrador: Windows PowerShell

PS D:\Downloads> Get-Content .\perfil_eleitor_secao_ATUAL_SP\perfil_eleitor_secao_ATUAL_SP.csv | Measure-Object -Line
>>
    Lines Words Characters Property
    -----
17423905

PS D:\Downloads> (Get-Content .\perfil_eleitor_secao_ATUAL_SP\perfil_eleitor_secao_ATUAL_SP.csv -First 1) -split ';'
>>
"DT_GERACAO"
"HH_GERACAO"
"ANO_ELEICAO"
"SG_UF"
"CD_MUNICIPIO"
"NM_MUNICIPIO"
"NR_ZONA"
"NR_SECAO"
"NR_LOCAL_VOTACAO"
"NM_LOCAL_VOTACAO"
"CD_GENERO"
"DS_GENERO"
"CD_ESTADO_CIVIL"
"DS_ESTADO_CIVIL"
"CD_FAIXA_ETARIA"
"DS_FAIXA_ETARIA"
"CD_GRAU_ESCOLARIDADE"
"DS_GRAU_ESCOLARIDADE"
"CD_RACA_COR"
"DS_RACA_COR"
"CD_IDENTIDADE_GENERO"
"DS_IDENTIDADE_GENERO"
"CD_QUILOMBOLA"
"DS_QUILOMBOLA"
"CD_INTERPRETE_LIBRAS"
"DS_INTERPRETE_LIBRAS"
"QT_ELEITORES_PERFIL"
"QT_ELEITORES_BIOMETRIA"
"QT_ELEITORES_DEFICIENCIA"
"QT_ELEITORES_INC_NM_SOCIAL"
PS D:\Downloads>

```

(b) Linhas de comando para o dataset SP

Ignorando o cabeçalho, encontramos a seguinte quantidade de dados:

Contagem de dados do Rio de Janeiro:

- **Linhas:** 6,455,237
- **Colunas:** 30

- **Células:** 193,657,110

Contagem de dados de São Paulo:

- **Linhas:** 17,423,904
- **Colunas:** 30
- **Células:** 522,717,120

Verificamos após a carga dos dados no Excel que há coerência com a contagem feita pelo PowerShell do Windows.

1.4 TRANSFORMAÇÕES NOS DADOS

Ainda, para remover dados indesejados do dataset, vamos fazer uma filtragem de colunas e linhas em duas etapas através do PowerQuery, antes mesmo da carga dos dados.

1.4.1 Projeção de Colunas

Selecionando a conexão feita com os datasets e clicando em "Editar", abrimos o editor do PowerQuery, onde podemos escolher que colunas desejamos manter ou remover.

Colunas selecionadas para análise:

- NM_MUNICIPIO
- CD_GENERO, DS_GENERO
- CD_ESTADO_CIVIL, DS_ESTADO_CIVIL
- CD_FAIXA_ETARIA, DS_FAIXA_ETARIA
- CD_GRAU_ESCOLARIDADE, DS_GRAU_ESCOLARIDADE
- QT_ELEITORES_PERFIL

Total de células remanescentes para RJ e SP, respectivamente: 64,552,370 e 174,239,040. Assim, removemos 20 colunas indesejadas e reduzimos cada dataset para 33% do tamanho original, facilitando significativamente o processamento.

1.4.2 Eliminação de Linhas Inválidas

Ainda no editor do PowerQuery, clicando no ícone de filtro das colunas restantes e desmarcando as entradas "NÃO INFORMADO" ou, no caso da coluna de faixa etária, "inválido" eliminamos dados que não nos interessam, aperfeiçoando a consistência da análise. Após essa remoção, restam 6,421,115 linhas na planilha do RJ e 17,405,241 linhas na de SP.

Total de células remanescentes para RJ e SP, respectivamente: 64,211,150 e 174,052,410. Concluídas ambas transformações, reduzimos os datasets à 33.16% e 33.30% dos tamanhos originais, vastamente agilizando o processamento.

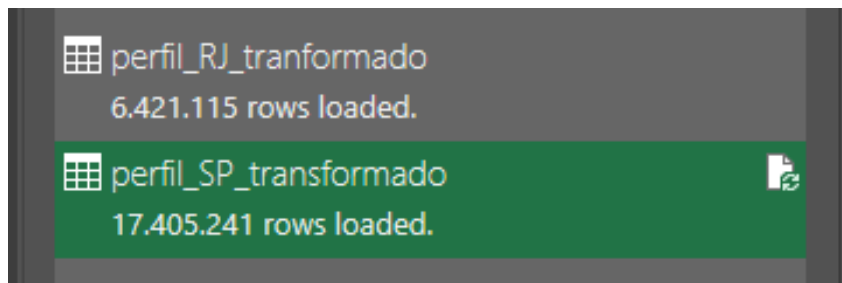


Figura 2 – Cargas dos datasets finais apresentando o total de linhas carregadas

1.5 ANÁLISES REALIZADAS

Para as análises seguintes, selecionamos as conexões estabelecidas com os datasets transformados e, finalmente, carregamos eles para tabelas dinâmicas, onde podemos apresentar os dados em diferentes agrupamentos, de forma geral sumarizados pelo número total de eleitores por grupo.

1.5.1 Tabelas Dinâmicas

À partir das Conexões geradas dos datasets transformados, podemos finalmente fazer cargas pontuais dos dados com diferentes agrupamentos através da função de Tabela Dinâmica do Excel. Selecionando uma Conexão e clicando em "Carregar Para..." podemos importar e exibir os dados como uma tabela Dinâmica.

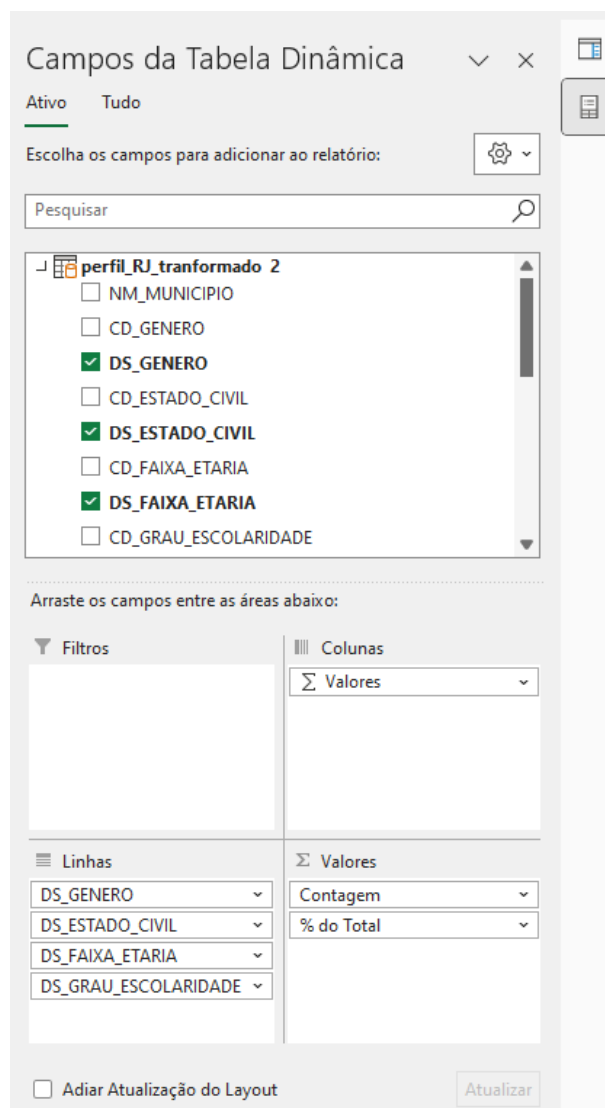


Figura 3 – Escolhendo categorias e valores de agrupamento

Após a inserção de uma Tabela Dinâmica, escolhemos na aba "Campos da Tabela Dinâmica" as categorias que desejamos apresentar nas fileiras e colunas bem como os valores a serem representados no corpo da tabela. À título de exemplo, apresentamos na tabela a seguir eleitores de ambos os estados agrupados por gênero, faixa etária e escolaridade.

Eleitorado Rio de Janeiro			Eleitorado São Paulo		
	Contagem	% do Total		Contagem	% do Total
FEMININO	3,571,115	55.62%	FEMININO	9,440,417	54.24%
MASCULINO	2,850,000	44.38%	MASCULINO	7,964,824	45.76%
CASADO	995,545	15.50%	CASADO	3,149,535	18.10%
16 anos	3	0.00%	16 anos	23	0.00%
17 anos	23	0.00%	17 anos	71	0.00%
18 anos	46	0.00%	18 anos	281	0.00%
19 anos	115	0.00%	19 anos	612	0.00%
20 anos	216	0.00%	20 anos	1,033	0.01%
21 a 24 anos	2,314	0.04%	21 a 24 anos	10,515	0.06%
25 a 29 anos	12,443	0.19%	25 a 29 anos	60,424	0.35%
30 a 34 anos	31,536	0.49%	30 a 34 anos	152,208	0.87%
35 a 39 anos	50,842	0.79%	35 a 39 anos	237,830	1.37%
ANALFABETO	90	0.00%	ANALFABETO	534	0.00%
ENSINO FUNDAMENTAL COMPLETO	2,965	0.05%	ENSINO FUNDAMENTAL COMPLETO	17,636	0.10%
ENSINO FUNDAMENTAL INCOMPLETO	6,806	0.11%	ENSINO FUNDAMENTAL INCOMPLETO	23,590	0.14%
ENSINO MÉDIO COMPLETO	18,183	0.28%	ENSINO MÉDIO COMPLETO	80,567	0.46%
ENSINO MÉDIO INCOMPLETO	6,423	0.10%	ENSINO MÉDIO INCOMPLETO	29,244	0.17%
LÊ E ESCRIVE	256	0.00%	LÊ E ESCRIVE	958	0.01%
SUPERIOR COMPLETO	10,605	0.17%	SUPERIOR COMPLETO	61,160	0.35%
SUPERIOR INCOMPLETO	5,514	0.09%	SUPERIOR INCOMPLETO	24,141	0.14%
40 a 44 anos	71,756	1.12%	40 a 44 anos	305,864	1.76%
45 a 49 anos	83,645	1.30%	45 a 49 anos	331,933	1.91%
50 a 54 anos	89,846	1.40%	50 a 54 anos	336,826	1.94%
55 a 59 anos	98,004	1.53%	55 a 59 anos	349,919	2.01%
60 a 64 anos	113,723	1.77%	60 a 64 anos	347,518	2.00%
65 a 69 anos	113,669	1.77%	65 a 69 anos	319,720	1.84%
70 a 74 anos	96,613	1.50%	70 a 74 anos	260,528	1.50%
75 a 79 anos	72,496	1.13%	75 a 79 anos	181,675	1.04%
80 a 84 anos	51,031	0.79%	80 a 84 anos	114,011	0.66%
85 a 89 anos	33,530	0.52%	85 a 89 anos	62,810	0.36%
90 a 94 anos	22,772	0.35%	90 a 94 anos	31,443	0.18%
95 a 99 anos	28,233	0.44%	95 a 99 anos	25,941	0.15%
100 anos ou mais	22,689	0.35%	100 anos ou mais	18,350	0.11%
DIVORCIADO	135,161	2.10%	DIVORCIADO	546,277	3.14%
SEPARADO JUDICIALMENTE	45,450	0.71%	SEPARADO JUDICIALMENTE	158,789	0.91%
SOLTEIRO	1,631,547	25.41%	SOLTEIRO	3,973,298	22.83%
VIÚVO	42,297	0.66%	VIÚVO	136,925	0.79%
Total Geral	6,421,115	100.00%	Total Geral	17,405,241	100.00%

Figura 4 – Tabelas Dinâmicas agrupadas por gênero, faixa etária, estado civil e escolaridade, resumando a contagem total de eleitores por grupo

Com a Tabela Dinâmica construída, podemos facilmente inserir um Gráfico Dinâmico, que responde às mudanças da Tabela dinamicamente, com a opção "Gráfico Dinâmico" na aba "Análise de Tabela Dinâmica".

1.5.2 Pacote Suplementar do Excel

Através da adição do suplemento "Analysis Toolpak" ou "Ferramentas de Análise", encontrado nativamente na instalação do Excel, na aba de suplementos das opções do Excel, podemos calcular rapidamente estatísticas básicas como medidas de tendências centrais e de dispersão, intervalo de valores, valores máximos e mínimos dentre outros. Bastando selecionar a opção "Análise de Dados" na aba "Dados" do Excel, dentre as análises possíveis escolher "Estatística Descritivas" e confirmar a coluna de dados dos quais se deseja obter as estatísticas, gerando tabelas como as [16a], [16b], [17a] e [17b] apresentadas na seção de interpretação dos dados.

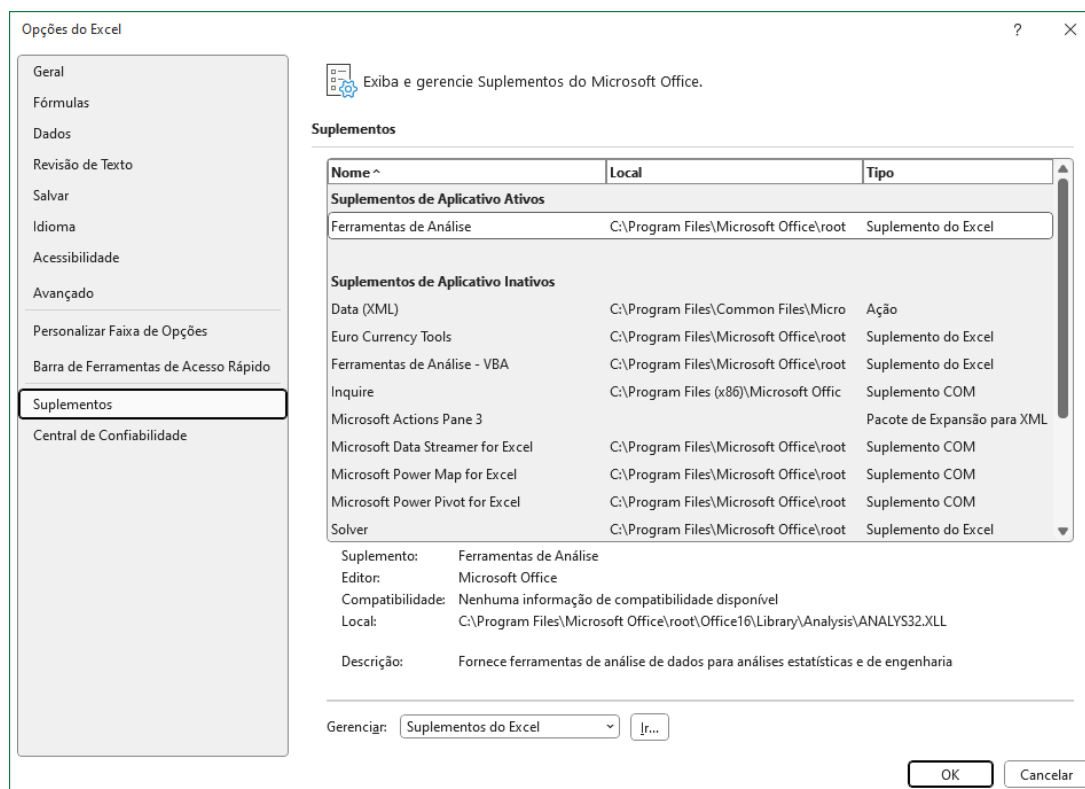


Figura 5 – Aba de suplementos no Excel

1.5.3 Filtros Aplicados

Foram realizadas, para cada estado, seleções com base em gênero, faixa etária, estado civil, município e combinações relevantes de certos critérios, para identificar perfis demográficos de interesse. As tabelas e gráficos dinâmicos gerados se encontram à seguir:

RJ- Faixa Etária	Eleitores
16 anos	16,433
17 anos	39,056
18 anos	100,611
19 anos	126,671
20 anos	128,382
21 a 24 anos	302,456
25 a 29 anos	358,359
30 a 34 anos	408,562
35 a 39 anos	466,825
40 a 44 anos	517,683
45 a 49 anos	522,674
50 a 54 anos	523,344
55 a 59 anos	578,416
60 a 64 anos	581,556
65 a 69 anos	512,911
70 a 74 anos	400,247
75 a 79 anos	282,351
80 a 84 anos	193,718
85 a 89 anos	129,083
90 a 94 anos	84,442
95 a 99 anos	85,558
100 anos ou mais	61,777
Total Geral	6,421,115

(a) Base eleitoral do RJ agrupada por Faixa Etária

SP - Faixa Etária	Eleitores
16 anos	49,722
17 anos	108,699
18 anos	289,945
19 anos	354,955
20 anos	353,458
21 a 24 anos	797,666
25 a 29 anos	1,050,333
30 a 34 anos	1,221,001
35 a 39 anos	1,396,990
40 a 44 anos	1,646,943
45 a 49 anos	1,672,134
50 a 54 anos	1,609,104
55 a 59 anos	1,623,588
60 a 64 anos	1,512,214
65 a 69 anos	1,287,284
70 a 74 anos	976,322
75 a 79 anos	631,012
80 a 84 anos	378,391
85 a 89 anos	213,653
90 a 94 anos	108,539
95 a 99 anos	76,765
100 anos ou mais	46,523
Total Geral	17,405,241

(b) Base eleitoral de SP agrupada por Faixa Etária

RJ- Gênero	Eleitores	%
FEMININO	3,571,115	55.62%
MASCULINO	2,850,000	44.38%
Total Geral	6,421,115	100.00%

RJ- Estado Civil	Eleitores	%
SOLTEIRO	3,384,894	52.72%
CASADO	2,258,556	35.17%
DIVORCIADO	407,339	6.34%
VIÚVO	256,239	3.99%
SEPARADO JUDICIALMENTE	114,087	1.78%
Total Geral	6,421,115	100.00%

(a) Base eleitoral do RJ agrupada por Gênero e por Estado Civil

SP - Gênero	Eleitores	%
FEMININO	9,440,417	54.24%
MASCULINO	7,964,824	45.76%
Total Geral	17,405,241	100.00%

SP - Estado Civil	Eleitores	%
SOLTEIRO	8,028,759	46.13%
CASADO	6,648,196	38.20%
DIVORCIADO	1,550,139	8.91%
VIÚVO	794,649	4.57%
SEPARADO JUDICIALMENTE	383,498	2.20%
Total Geral	17,405,241	100.00%

(b) Base eleitoral de SP agrupada por Gênero e por Estado Civil

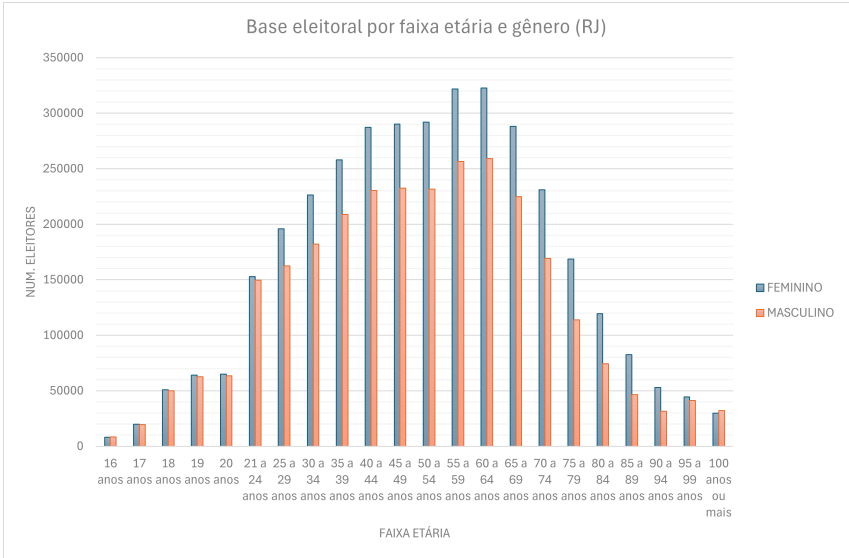


Figura 8 – Histograma da base eleitoral do RJ agrupada por Faixa Etária e Gênero

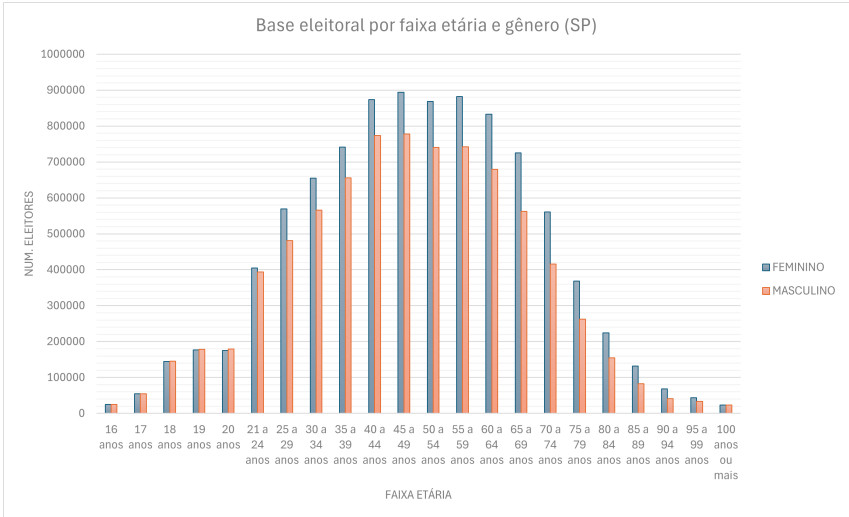


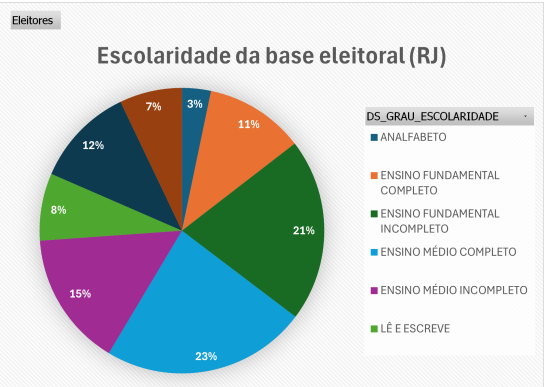
Figura 9 – Histograma da base eleitoral de SP agrupada por Faixa Etária e Gênero

RJ- Escolaridade	Eleitores
ANALFABETO	210,742
ENSINO FUNDAMENTAL COMPLETO	724,358
ENSINO FUNDAMENTAL INCOMPLETO	1,331,486
ENSINO MÉDIO COMPLETO	1,495,354
ENSINO MÉDIO INCOMPLETO	981,585
LÊ E ESCRIVE	490,228
SUPERIOR COMPLETO	731,138
SUPERIOR INCOMPLETO	456,224
Total Geral	6,421,115

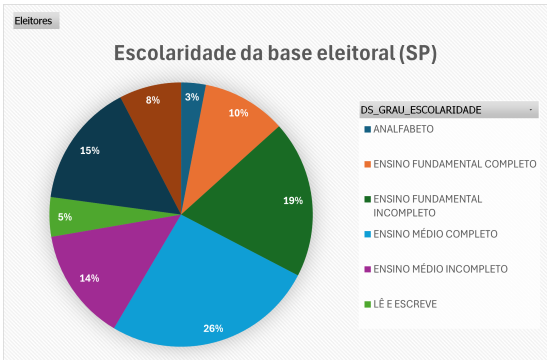
(a) Base eleitoral do RJ agrupada por Escolaridade

SP - Escolaridade	Eleitores
ANALFABETO	527,134
ENSINO FUNDAMENTAL COMPLETO	1,797,226
ENSINO FUNDAMENTAL INCOMPLETO	3,352,838
ENSINO MÉDIO COMPLETO	4,507,123
ENSINO MÉDIO INCOMPLETO	2,399,829
LÊ E ESCRIVE	839,999
SUPERIOR COMPLETO	2,654,123
SUPERIOR INCOMPLETO	1,326,969
Total Geral	17,405,241

(b) Base eleitoral de SP agrupada por Escolaridade



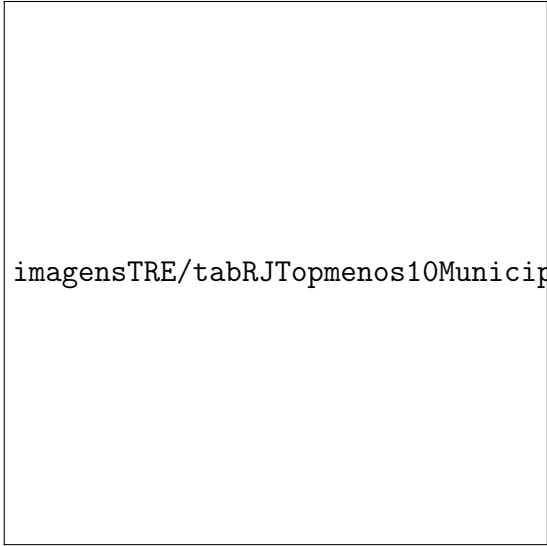
(a) Gráfico em pizza para base eleitoral do RJ agrupada por Escolaridade



(b) Gráfico em pizza para base eleitoral de SP agrupada por Escolaridade

RJ- Municípios	Eleitores	%
RIO DE JANEIRO	2,331,918	36.32%
DUQUE DE CAXIAS	332,372	5.18%
SÃO GONÇALO	318,879	4.97%
NOVA GUARUJÁ	278,324	4.33%
NITERÓI	214,083	3.33%
SÃO JOÃO DE MERITI	194,289	3.03%
CAMPOS DOS GOYTACAZES	181,780	2.83%
BELFORD ROXO	159,886	2.49%
PETRÓPOLIS	127,328	1.98%
VOLTA REDONDA	111,936	1.74%
Total Geral *	6,421,115	100.00%

(a) Base eleitoral do RJ agrupada por Município apresentando os 10 primeiros por quantidade de eleitor



(b) Base eleitoral do RJ agrupada por Município apresentando os 10 últimos por quantidade de eleitor

SP - Municípios	Eleitores	%
SÃO PAULO	4,636,558	26.64%
GUARULHOS	465,520	2.67%
CAMPINAS	425,199	2.44%
SÃO BERNARDO DO CAMPO	313,042	1.80%
SANTO ANDRÉ	291,735	1.68%
OSASCO	291,645	1.68%
SOROCABA	262,971	1.51%
SÃO JOSÉ DOS CAMPOS	255,702	1.47%
RIBEIRÃO PRETO	225,077	1.29%
SANTOS	183,104	1.05%
Total Geral *	17,405,241	100.00%

(a) Base eleitoral do SP agrupada por Município apresentando os 10 primeiros por quantidade de eleitor

imagenesTRE/tabSPTopmenos10Municipios.png

(b) Base eleitoral do SP agrupada por Município apresentando os 10 últimos por quantidade de eleitor

Base Eleitoral RJ	Estado Civil	SEPARADO					Gênero
	CASADO	DIVORCIADO	JUDICIALMENTE	SOLTEIRO	VÍVULO	Total	FEMININO
Faixa Etária							MASCULINO
16 anos	0.00%	0.00%	0.00%	0.23%	0.00%	0.23%	
17 anos	0.00%	0.00%	0.00%	0.50%	0.00%	0.50%	
18 anos	0.00%	0.00%	0.00%	1.41%	0.00%	1.42%	
19 anos	0.00%	0.00%	0.00%	1.78%	0.00%	1.79%	
20 anos	0.02%	0.00%	0.00%	1.79%	0.00%	1.82%	
21 a 24 anos	0.28%	0.01%	0.01%	3.66%	0.01%	4.28%	
25 a 29 anos	1.11%	0.07%	0.00%	4.30%	0.01%	5.48%	
30 a 34 anos	1.87%	0.19%	0.01%	4.26%	0.01%	6.34%	
35 a 39 anos	2.46%	0.36%	0.02%	4.36%	0.02%	7.23%	
40 a 44 anos	3.02%	0.61%	0.04%	4.31%	0.08%	8.06%	
45 a 49 anos	3.17%	0.78%	0.08%	3.98%	0.12%	8.13%	
50 a 54 anos	3.28%	0.90%	0.12%	3.66%	0.21%	8.17%	
55 a 59 anos	3.54%	1.05%	0.18%	3.86%	0.38%	9.01%	
60 a 64 anos	3.68%	1.15%	0.28%	3.30%	0.63%	9.03%	
65 a 69 anos	3.37%	1.01%	0.33%	2.52%	0.85%	8.07%	
70 a 74 anos	2.60%	0.79%	0.19%	1.78%	0.90%	6.26%	
75 a 79 anos	2.18%	0.41%	0.22%	1.17%	0.75%	4.72%	
80 a 84 anos	1.64%	0.21%	0.14%	0.73%	0.62%	3.35%	
85 a 89 anos	1.18%	0.10%	0.09%	0.46%	0.48%	2.31%	
90 a 94 anos	0.77%	0.04%	0.05%	0.27%	0.35%	1.48%	
95 a 99 anos	0.16%	0.02%	0.04%	0.26%	0.33%	1.25%	
100 anos ou mais	0.37%	0.01%	0.02%	0.18%	0.27%	0.85%	
TOTAL	7.65%	7.65%	0.52%	48.10%	5.44%	69.36%	

(a) Mapa de calor da base eleitoral do RJ
Feminina agrupada por Estado Civil e
Faixa etária

Base Eleitoral RJ	Estado Civil	SEPARADO					Gênero
	CASADO	DIVORCIADO	JUDICIALMENTE	SOLTEIRO	VIVO	Grand Total	
Paiva Estária							FEMININO
16 anos	0.00%	0.00%	0.00%	0.29%	0.00%	0.29%	
17 anos	0.00%	0.00%	0.00%	0.68%	0.00%	0.68%	
18 anos	0.00%	0.00%	0.00%	1.74%	0.00%	1.75%	
19 anos	0.00%	0.00%	0.00%	2.19%	0.00%	2.20%	
20 anos	0.01%	0.00%	0.01%	2.21%	0.00%	2.23%	
21 a 24 anos	0.08%	0.00%	0.01%	5.15%	0.15%	5.25%	
25 a 29 anos	0.44%	0.01%	0.00%	5.24%	0.00%	5.70%	
30 a 34 anos	1.11%	0.06%	0.01%	5.21%	0.00%	6.39%	
35 a 39 anos	1.78%	0.16%	0.01%	5.30%	0.01%	7.33%	
40 a 44 anos	2.52%	0.30%	0.03%	5.22%	0.03%	8.08%	
45 a 49 anos	2.93%	0.43%	0.06%	4.70%	0.03%	8.16%	
50 a 54 anos	3.15%	0.56%	0.09%	4.28%	0.05%	8.13%	
55 a 59 anos	3.44%	0.68%	0.13%	4.66%	0.09%	9.00%	
60 a 64 anos	3.99%	0.77%	0.20%	3.99%	0.15%	9.09%	
65 a 69 anos	3.99%	0.73%	0.27%	2.86%	0.22%	7.89%	
70 a 74 anos	3.39%	0.71%	0.27%	1.54%	0.23%	5.94%	
75 a 79 anos	2.54%	0.27%	0.18%	0.83%	0.18%	4.00%	
80 a 84 anos	1.79%	0.13%	0.11%	0.44%	0.13%	2.60%	
85 a 89 anos	1.18%	0.06%	0.06%	0.24%	0.09%	1.63%	
90 a 94 anos	0.80%	0.03%	0.04%	0.17%	0.07%	1.10%	
95 a 99 anos	0.09%	0.03%	0.06%	0.25%	0.11%	1.44%	
100 anos ou mais	0.40%	0.05%	0.04%	0.10%	0.14%	1.12%	
	54.34%	4.85%	0.72%	47.96%	1.48%	100.00%	

(b) Mapa de calor da base eleitoral do RJ
Masculina agrupada por Estado Civil e
Faixa etária

Base Eleitoral SP	Estado Civil		SEPARADO			Gênero	
	CASADO	DIVORCIADO	JUDICIALMENTE	SOLTEIRO	VÍVULO	Total Geral	
Faixa Etária							FEMININO
16 anos	0.00%	0.00%	0.00%	0.25%	0.00%	0.25%	MASCULINO
17 anos	0.00%	0.00%	0.00%	0.58%	0.00%	0.58%	
18 anos	0.01%	0.00%	0.00%	1.52%	0.00%	1.53%	
19 anos	0.02%	0.00%	0.00%	1.85%	0.00%	1.87%	
20 anos	0.04%	0.00%	0.00%	1.80%	0.00%	1.85%	
21 a 24 anos	0.42%	0.02%	0.01%	3.83%	0.00%	4.28%	
25 a 29 anos	1.52%	0.15%	0.01%	4.35%	0.01%	6.03%	
30 a 34 anos	2.46%	0.40%	0.02%	4.04%	0.02%	6.94%	
35 a 39 anos	3.12%	0.73%	0.05%	3.90%	0.04%	7.85%	
40 a 44 anos	3.73%	1.09%	0.12%	4.22%	0.09%	9.25%	
45 a 49 anos	3.90%	1.27%	0.20%	3.90%	0.18%	9.47%	
50 a 54 anos	3.87%	1.38%	0.28%	3.33%	0.33%	9.20%	
55 a 59 anos	3.94%	1.52%	0.35%	2.94%	0.59%	9.34%	
60 a 64 anos	3.78%	1.47%	0.39%	2.28%	0.91%	8.82%	
65 a 69 anos	3.33%	1.20%	0.35%	1.63%	1.16%	7.68%	
70 a 74 anos	2.61%	0.78%	0.19%	1.09%	1.20%	5.84%	
75 a 79 anos	1.79%	0.38%	0.16%	0.65%	0.93%	3.90%	
80 a 84 anos	1.12%	0.15%	0.09%	0.36%	0.65%	2.37%	
85 a 89 anos	0.68%	0.06%	0.05%	0.19%	0.42%	1.39%	
90 a 94 anos	0.37%	0.02%	0.02%	0.10%	0.22%	0.72%	
95 a 99 anos	0.24%	0.01%	0.01%	0.07%	0.13%	0.46%	
100 anos ou mais	0.12%	0.00%	0.01%	0.04%	0.08%	0.25%	
Total Geral	37.05%	10.53%	2.33%	42.36%	6.73%	100.00%	

(a) Mapa de calor da base eleitoral do SP
Feminina agrupada por Estado Civil e
Faixa etária

Base Eleitoral SP	Estado Civil		SEPARADO				Gênero
	CASADO	DIVORCIADO	JUDICIALMENTE	SOLTEIRO	VIVO	Total Geral	
Faixa Etária							FEMININO
16 anos	0.00%	0.00%	0.00%	0.31%	0.00%	0.31%	MASCULINO
17 anos	0.00%	0.00%	0.00%	0.68%	0.00%	0.68%	
18 anos	0.06%	0.00%	0.00%	1.82%	0.00%	1.83%	
19 anos	0.01%	0.00%	0.00%	2.22%	0.00%	2.24%	
20 anos	0.01%	0.00%	0.00%	2.22%	0.00%	2.24%	
21 a 24 anos	0.13%	0.00%	0.01%	4.76%	0.1%	4.94%	
25 a 29 anos	0.76%	0.04%	0.00%	5.24%	0.00%	6.04%	
30 a 34 anos	1.91%	0.16%	0.01%	5.01%	0.01%	7.10%	
35 a 39 anos	2.99%	0.30%	0.04%	4.62%	0.01%	8.23%	
40 a 44 anos	3.14%	0.14%	0.01%	5.15%	0.02%	8.46%	
45 a 49 anos	4.17%	0.78%	0.15%	4.63%	0.04%	9.77%	
50 a 54 anos	4.23%	0.90%	0.23%	3.66%	0.7%	9.30%	
55 a 59 anos	4.39%	1.05%	0.31%	3.43%	0.13%	9.31%	
60 a 64 anos	4.36%	1.06%	0.34%	2.55%	0.21%	8.53%	
65 a 69 anos	4.01%	0.89%	0.32%	1.50%	0.29%	7.06%	
70 a 74 anos	3.27%	0.58%	0.23%	0.82%	0.13%	5.02%	
75 a 79 anos	2.28%	0.26%	0.12%	0.30%	0.24%	3.20%	
80 a 84 anos	1.43%	0.10%	0.06%	0.18%	0.16%	1.94%	
85 a 89 anos	0.79%	0.03%	0.03%	0.08%	0.10%	1.03%	
90 a 94 anos	0.39%	0.01%	0.01%	0.04%	0.05%	0.51%	
95 a 99 anos	0.33%	0.01%	0.01%	0.04%	0.03%	0.42%	
100 anos ou mais	0.23%	0.00%	0.01%	0.02%	0.03%	0.29%	
Total Geral	39.54%	6.86%	1.89%	49.88%	1.72%	100.00%	

(b) Mapa de calor da base eleitoral do SP
Masculina agrupada por Estado Civil e
Faixa etária

1.5.4 Resultados da Análise do Dataset01

Aqui está uma explicação concisa de moda, média e mediana: Moda: é o valor que aparece com mais frequência em um conjunto de dados. Exemplo: em 1, 2, 2, 3, a moda é 2. Média: é a soma de todos os valores dividida pela quantidade de dados. Exemplo: para 1, 2, 3, a média é $(1 + 2 + 3) / 3 = 2$. Mediana: é o valor central quando os dados estão organizados em ordem crescente. Se houver um número ímpar de dados, é o valor no meio; se houver um número par, é a média dos dois valores centrais. Exemplo: em 1, 2, 3, 4, a mediana é $(2 + 3) / 2 = 2,5$.

<i>Num. Eleitores por Faixa Etária (RJ)</i>	
Mean	291868.8636
Standard Errc	42598.75111
Median	292403.5
Mode	#N/D
Standard Dev	199805.8535
Sample Varia	39922379113
Kurtosis	-1.665058667
Skewness	0.100386577
Range	565123
Minimum	16433
Maximum	581556
Sum	6421115
Count	22

(a) Estatística Descritiva obtida para a base eleitoral do RJ agrupada por faixa etária

<i>Num. Eleitores por Faixa Etária (SP)</i>	
Mean	791147.3182
Standard Errc	131579.7382
Median	714339
Mode	#N/D
Standard Dev	617163.6779
Sample Varia	3.80891E+11
Kurtosis	-1.666415095
Skewness	0.195320453
Range	1625611
Minimum	46523
Maximum	1672134
Sum	17405241
Count	22

(b) Estatística Descritiva obtida para a base eleitoral de SP agrupada por faixa etária

<i>Num. Eleitores por Município (RJ)</i>	
Mean	69794.72826
Standard Error	25775.13393
Median	18839
Mode	#N/D
Standard Deviation	247226.3996
Sample Variance	61120892679
Kurtosis	79.21429673
Skewness	8.633796845
Range	2327547
Minimum	4371
Maximum	2331918
Sum	6421115
Count	92

(a) Estatística Descritiva obtida para a base eleitoral do RJ agrupada por Municípios

<i>Num. Eleitores por Município (SP)</i>	
Mean	26984.86977
Standard Error	7366.982392
Median	5496
Mode	3076
Standard Deviation	187098.1482
Sample Variance	35005717069
Kurtosis	574.5657972
Skewness	23.35779448
Range	4635915
Minimum	643
Maximum	4636558
Sum	17405241
Count	645

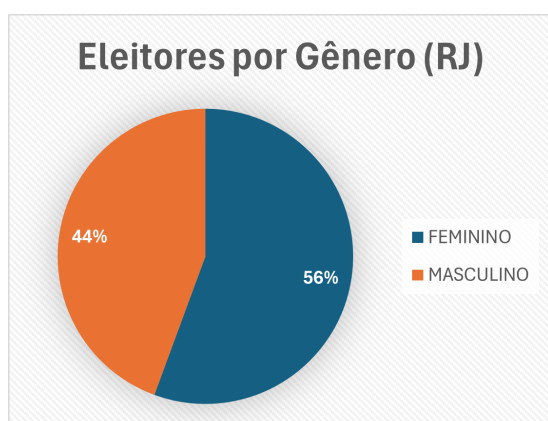
(b) Estatística Descritiva obtida para a base eleitoral de SP agrupada por Municípios

Podemos fazer algumas comparações entre ambos estados à partir desses valores estatísticos. A primeira, referente as médias de eleitores por município. Vemos

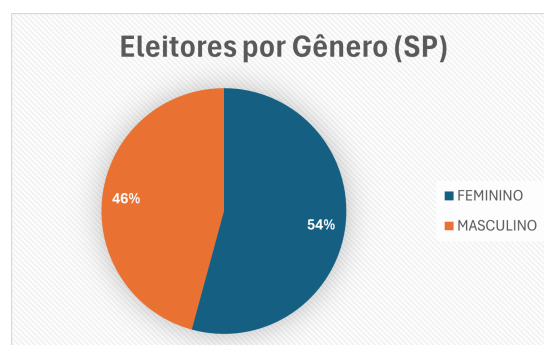
o valor de 26.984 em SP e de 69.795 no RJ (diferença de 2,58 vezes), podemos inferir a partir disso que, apesar do estado de São Paulo ter uma população maior que a do Rio de Janeiro, a população do último é maior em menos municípios (ver item 1.3.0.1). Esse fato, combinado com o desvio padrão do RJ, que chega a ser 247.000, enquanto o de SP é de 187.000, ou seja, há uma variabilidade significativamente maior no primeiro do que no segundo, nos pinta uma imagem da distribuição da população em ambos estados. Em SP, temos mais municípios, um estado maior, porém uma população menos concentrada, mais espalhada entre eles, enquanto o RJ tem mais pessoas em seus municípios, com uma concentração maior em poucos municípios (sabemos que há uma aglomeração maior em sua capital). Temos também o valor da mediana de cada um dos estados, que é mais representativo de um município "normal" nesse caso, do que a média (que pode ser muito impactada pelos outliers e portanto não dá uma imagem acurada da cidade comum no estado). Para o RJ, temos uma mediana de 18.839 e para SP, de 5.496, uma diferença entre ambos de 3,47 vezes, bem maior que a diferença de 2,58 para a média. Além disso, esses são valores bem menores que aqueles achados para as médias. Podemos extrapolar a partir disso que não só o Rio tem uma população mais concentrada na sua capital, como também as cidades no estado tendem a ter uma população consideravelmente maior, mesmo sem considerarmos a diferença do número de municípios e o vies gerado pelos outliers (que afetam a média). Finalmente, temos a moda, uma medida para datasets mais qualitativos que não é válida nessa aplicação, e portanto é inexistente para o RJ, significando que não repetição o número de habitantes em mais de um município. Já em SP, vemos sim, um valor de moda, devido às cidades de Santa Alvertina e Lagoinha que tem a mesma população.

Quando pensamos em estatísticas referentes ao estado interessantes para um assessor de campanha eleitoral, o coeficiente de variação de eleitores é uma delas. Para o Rio de Janeiro, esse valor é de 3,52%, enquanto para São Paulo, ele é de 6,93%, um valor quase duas vezes maior. Vemos que a relação se inverteu em relação ao desvio padrão. Isso ocorreu porque o valor do CV é ajustado pela divisão do DP pela média e a média no RJ é 3,47 vezes maior que a de SP. Ou seja, um candidato precisaria levar em consideração, uma assimetria mais acentuada em SP, já que a dispersão relativa (mostrada no CV) é uma medida mais eficaz para representar a concentração em um estado do que a dispersão absoluta. Num ponto de vista estratégico, isso implicaria por exemplo no uso de táticas mais concentradas em SP do que no RJ, por exemplo, um tour das 10 cidades mais populosas do estado de São Paulo. Para o Rio, devido a população mais homogênea, talvez uma tática como investimento maior em horário eleitoral (propaganda eleitoral, tanto televisionada quanto por rádio) seja mais efetivo, pois atingiria todos municípios mais ou menos igualmente. Em relação ao CV de gênero em ambos estados, vemos uma diferença desprezível entre ambos,

mas se tomarmos ele em consideração para a formulação de uma tática, podemos apontar o CV maior no RJ como representativo de uma maior variação em número entre homens e mulheres nos diferentes municípios. Isso por sua vez pode fazer com que seja mais interessante criar táticas eleitorais com propaganda voltada a um gênero específico (como a defesa de propostas feministas) em certos municípios do Rio de Janeiro, onde essa maior concentração de mulheres é mais acentuada. Além do CV de Gênero, podemos usar também a própria proporção geral entre homens e mulheres no estado, para contruir uma tática. No caso, vemos que o RJ tem 56% de eleitoras enquanto SP tem 54%, essa diferença pequena, porém relevante, pode significar a margem entre a vitória e uma derrota em muitas eleições, e portanto, podemos simplesmente considerar que é mais favorável uma estratégia mais voltada ao público feminino no RJ. Uma outra possível análise, seria a comparação básica das taxas de escolaridade em ambos estados (ver 1.5.2), porém podemos ver que as diferenças são desprezíveis entre ambos, e portanto essa análise não é pertinente.



(a) Gráfico em pizza para base eleitoral do RJ por gênero



(b) Gráfico em pizza para base eleitoral de SP por gênero

É imediato perceber nos Mapas de Calor construídos ([13b],[14b],[15a],[15b]) para ambos estados e gêneros, agrupados por faixa etária e estado civil, que se assemelham muito entre si, sugerindo que o estado e o gênero não afeta significativamente a distribuição de eleitores por faixa etária versus estado civil, predominando eleitores solteiros da faixa dos 21 aos 60 anos seguido por casados dos 40 aos 65.

2 ANÁLISE DE DADOS METEOROLÓGICOS

2.1 INTRODUÇÃO

Neste capítulo, é realizada a análise comparativa de dados meteorológicos de dois anos consecutivos para uma estação meteorológica específica. O objetivo é identificar variações climáticas significativas.

Observação sobre o UTC:

O UTC não é 0000 durante os períodos matutinos e vespertinos analisados, pois os horários apresentados (como 0300 UTC a 1500 UTC) correspondem à conversão do horário local de Curitiba, que está em UTC-3.

2.2 DESCRIÇÃO DO DATASET

2.2.1 Origem dos Dados

Os dados meteorológicos foram obtidos no site do Instituto Nacional de Meteorologia (INMET), acessível via <https://portal.inmet.gov.br/dadoshistoricos>.

2.2.2 Carga

2.2.2.1 Importação dos Dados

Para realizarmos a importação de dados no Excel, fizemos o seguinte:

1. Abrimos o Excel e acessamos a aba **Dados**.
2. Clicamos em **Obter Dados Externos** e, em seguida, selecionamos a opção **De Texto**.
3. Na janela **Assistente de Importação de Texto**, escolhemos a opção **Delimitado** e definimos o delimitador como **ponto e vírgula (;)**.
4. No campo **Origem do arquivo**, selecionamos **Windows (ANSI)**.
5. No qualificador de texto, optamos pela opção **Nenhum**.
6. Na janela **Configurações Avançadas de Importação de Texto**, configuramos o separador decimal como **vírgula (,)** e o separador de milhar como **Nenhum**.

2.2.2.2 Contagem de linhas, colunas e células

Planilha de 2022 - Sem Tratamento:

- **Linhas:** 8760
- **Colunas:** 19
- **Células:** 125411

Planilha de 2023 - Sem Tratamento:

- **Linhas:** 8760
- **Colunas:** 19
- **Células:** 161478

Observação: Pode-se notar que o número de células não corresponde a multiplicação das linhas com as colunas, isso se deve ao fato da existência de alguns valores vazios em ambas as tabelas.

2.2.3 Variáveis Seleccionadas

As variáveis quantitativas seleccionadas para a análise incluem:

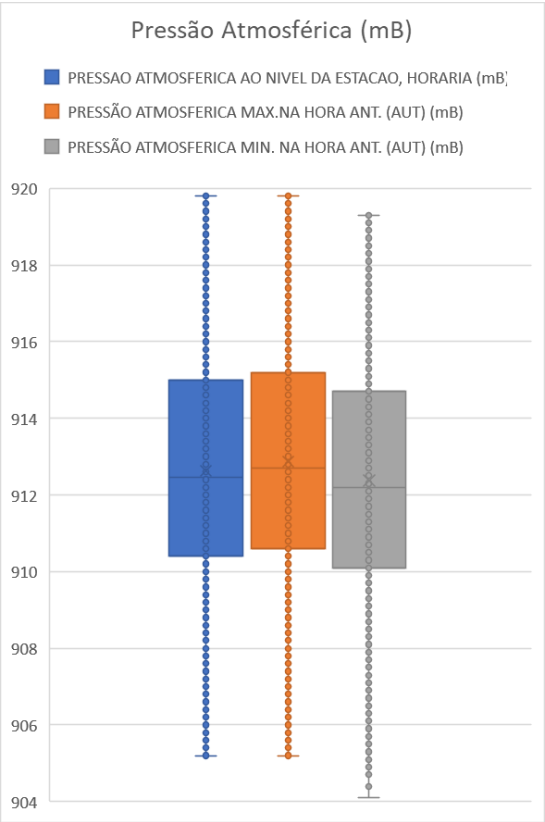
- DATA e HORA UTC
- PRECIPITAÇÃO TOTAL (mm)
- PRESSÃO ATMOSFÉRICA (mB)
- TEMPERATURA DO AR (°C)
- VENTO (direção e velocidade)

2.3 TRANSFORMAÇÕES NOS DADOS**2.3.1 Eliminação de Colunas**

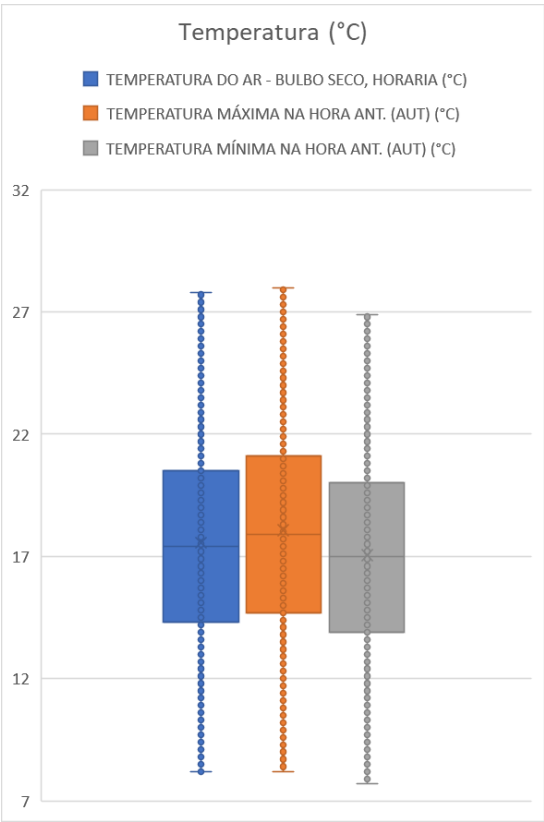
Foram removidas colunas que não eram relevantes para a análise, o que otimizou o tamanho do dataset. Após a retirada, o número de células ficaram: 101.560 e 104.363, dos anos de 2022 e 2023 respectivamente.

2.3.2 Eliminação de Linhas Inválidas

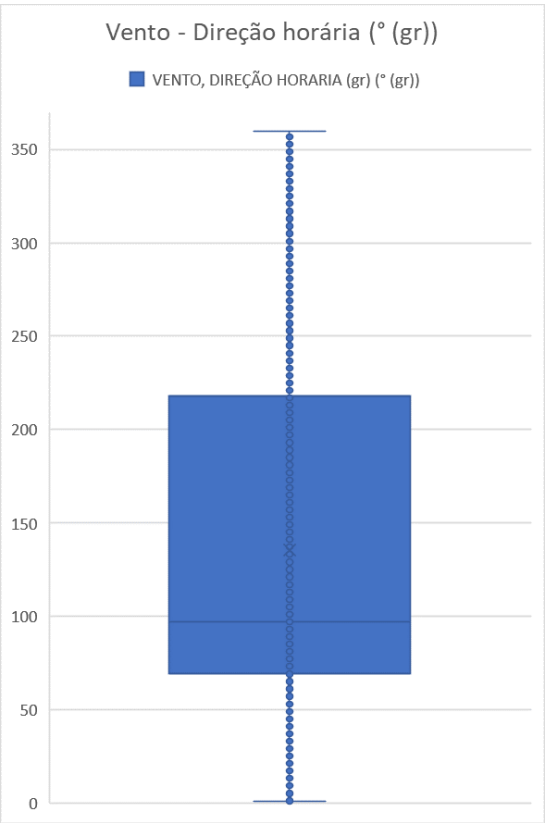
Linhas com valores nulos ou outliers foram eliminadas para garantir a precisão da análise. Após as eliminações, os dados de 2022 ficaram com 7428 linhas, já os de 2023 ficaram com 7468. Achemos por bem fazer alguns Gráficos de Caja, tanto de 2022 quanto de 2023, para verificação dos outliers. Utilizando o conceito de primeiro e terceiro quartil, foram calculados os limites superiores e inferiores de cada coluna. Assim, como podemos observar nos gráficos da figura 19 e 20 os outliers foram removidos.



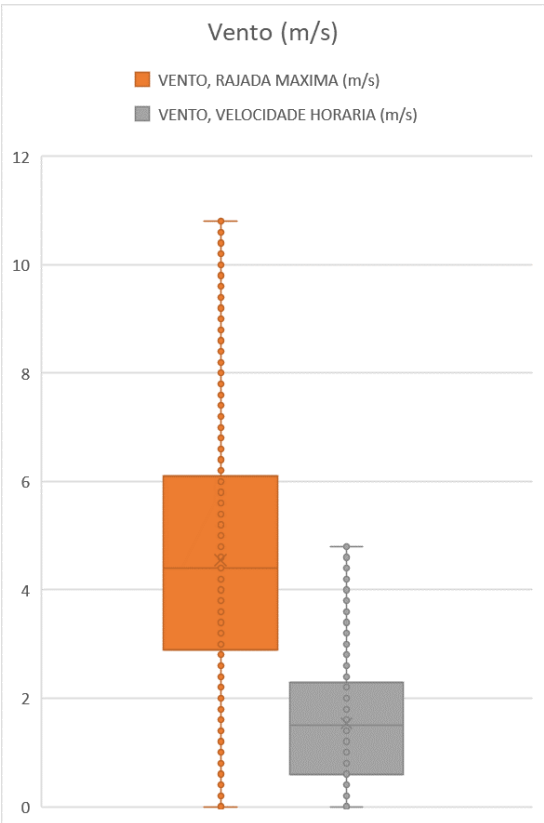
(a) Pressão Atmosférica (mB) - 2022



(b) Temperatura (°C) - 2022

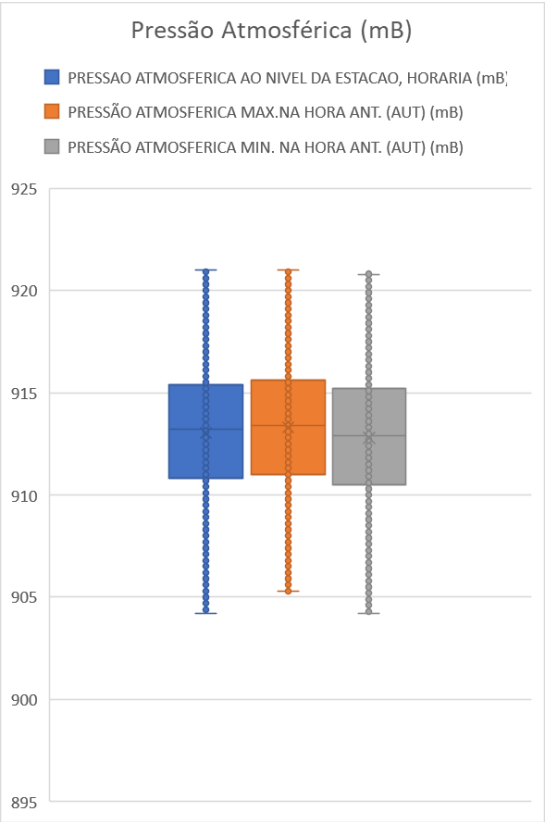


(c) Vento (°) - 2022

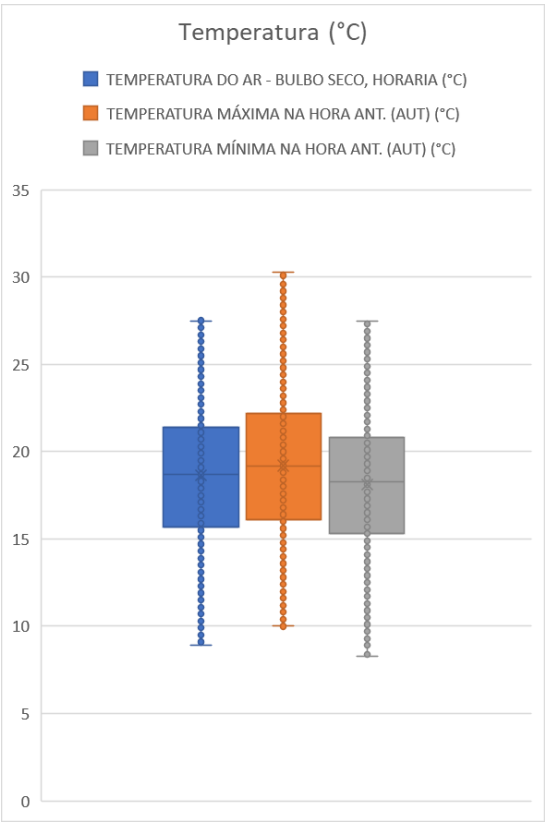


(d) Vento (m_s) – 2022

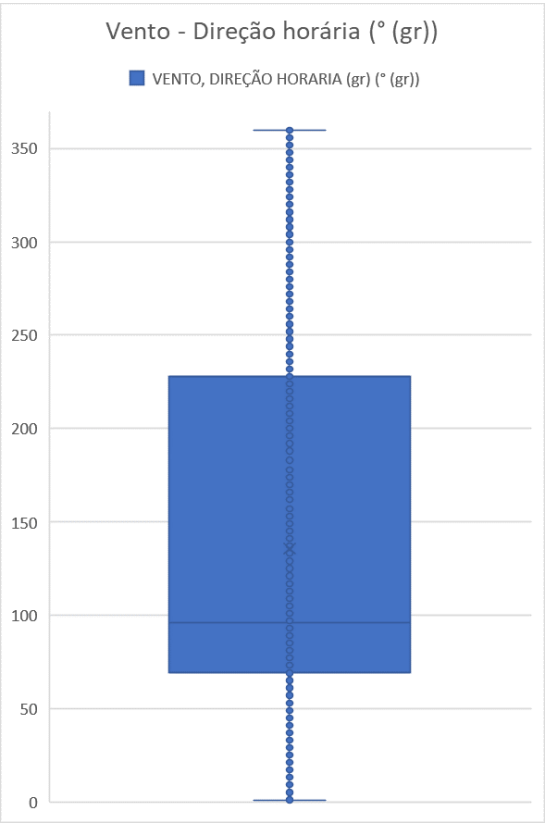
Figura 19 – Organização dos gráficos de 2022



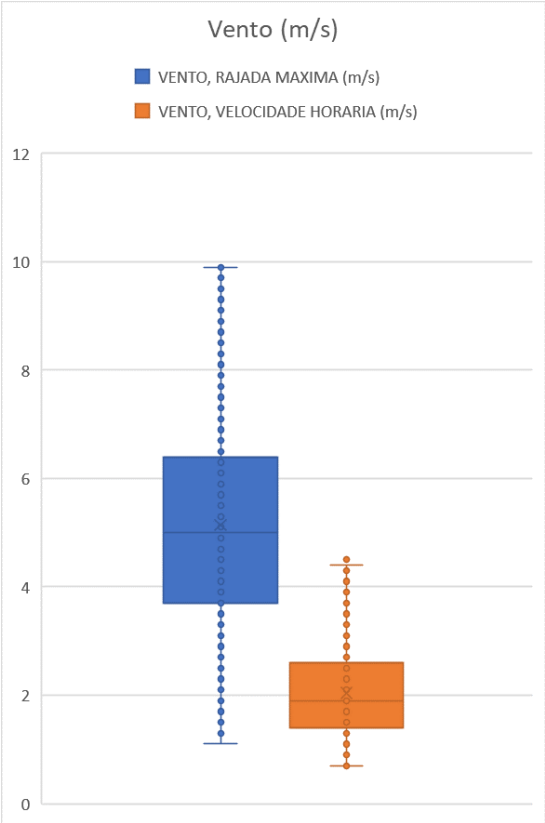
(a) Pressão Atmosférica (mB) - 2023



(b) Temperatura (°C) - 2023



(c) Vento (°) - 2023



(d) Vento (m_s) - 2023

Figura 20 – Organização dos gráficos de 2023

2.4 ANÁLISES REALIZADAS

2.4.1 Análise Trimestral da Temperatura do Ar

Trimestres	Máximo	Mínimo	Média
Trim1	27.8	10.8	20.5
Trim2	26.9	8.3	16.0
Trim3	27.1	8.2	15.4
Trim4	27.8	9.0	18.2
Total Geral	27.8	8.2	17.5

Tabela 1 – Máximos, Mínimos e Médias de Temperatura por Trimestre de 2022

Trimestres	Máximo	Mínimo	Média
Trim1	27.5	12.9	20.5817
Trim2	27.5	8.9	16.9677
Trim3	27.5	9.3	16.3976
Trim4	27.5	9.7	18.9964
Total Geral	27.5	8.9	18.2769

Tabela 2 – Máximos, Mínimos e Médias de Temperatura por Trimestre de 2023

2.4.2 Cálculos das Estatísticas Básicas

2022	Média	Moda	Mediana
PRECIPITAÇÃO TOTAL, HORÁRIO (mm)	0.0981	0	0
P.Atm AO NÍVEL DA ESTAÇÃO, HORÁRIA (mB)	912.625	911.1	912.45
P.Atm MÁX. NA HORA ANT. (AUT) (mB)	912.866	911.1	912.7
P.Atm MÍN. NA HORA ANT. (AUT) (mB)	912.369	911.1	912.2
TEMP DO AR - BULBO SECO, HORÁRIA (°C)	17.5540	17.3	17.4
TEMP MÁXIMA NA HORA ANT. (AUT) (°C)	18.0779	17.6	17.9
TEMP MÍNIMA NA HORA ANT. (AUT) (°C)	17.0573	17.2	17.0
VENTO, DIREÇÃO HORÁRIA (gr) (° (gr))	135.3681	73	97
VENTO, RAJADA MÁXIMA (m/s)	4.5447	0	4.4
VENTO, VELOCIDADE HORÁRIA (m/s)	1.5355	0	1.5

Tabela 3 – Estatísticas Descritivas para 2022

2022	Variância	Desvio Padrão	Amplitude
PRECIPITAÇÃO TOTAL, HORÁRIO (mm)	0.1830	0.4278	4.6
P.Atm AO NÍVEL DA ESTAÇÃO, HORÁRIA (mB)	9.8686	3.1414	14.6
P.Atm MÁX. NA HORA ANT. (AUT) (mB)	9.8436	3.1375	14.6
P.Atm MÍN. NA HORA ANT. (AUT) (mB)	9.9965	3.1617	15.2
TEMP DO AR - BULBO SECO, HORÁRIA (°C)	17.2970	4.1590	19.6
TEMP MÁXIMA NA HORA ANT. (AUT) (°C)	18.5729	4.3096	19.8
TEMP MÍNIMA NA HORA ANT. (AUT) (°C)	15.7465	3.9682	19.2
VENTO, DIREÇÃO HORÁRIA (gr) (° (gr))	9581.20	97.8836	359
VENTO, RAJADA MÁXIMA (m/s)	5.0855	2.2551	10.9
VENTO, VELOCIDADE HORÁRIA (m/s)	1.2389	1.1131	4.8

Tabela 4 – Estatísticas Descritivas para 2022

2022	IQR	Máximo	Mínimo
PRECIPITAÇÃO TOTAL, HORÁRIO (mm)	0	4.6	0
P.Atm AO NÍVEL DA ESTAÇÃO, HORÁRIA (mB)	4.5	919.8	905.2
P.Atm MÁX. NA HORA ANT. (AUT) (mB)	4.8	919.8	905.2
P.Atm MÍN. NA HORA ANT. (AUT) (mB)	897.3	919.3	904.1
TEMP DO AR - BULBO SECO, HORÁRIA (°C)	6.0	27.8	8.2
TEMP MÁXIMA NA HORA ANT. (AUT) (°C)	6.8	28	8.2
TEMP MÍNIMA NA HORA ANT. (AUT) (°C)	3.3	26.9	7.7
VENTO, DIREÇÃO HORÁRIA (gr) (° (gr))	213.6	360	1
VENTO, RAJADA MÁXIMA (m/s)	4.8	10.9	0
VENTO, VELOCIDADE HORÁRIA (m/s)	1.7	4.8	0

Tabela 5 – Estatísticas Descritivas para 2022

2023	Média	Moda	Mediana
PRECIPITAÇÃO TOTAL, HORÁRIO (mm)	0.1665	0	0
P.Atm AO NÍVEL DA ESTAÇÃO, HORÁRIA (mB)	913.097	912.7	913.3
P.Atm MÁX. NA HORA ANT. (AUT) (mB)	913.340	913.9	913.5
P.Atm MÍN. NA HORA ANT. (AUT) (mB)	912.836	913.6	913
TEMP DO AR - BULBO SECO, HORÁRIA (°C)	18.296	19.5	18.3
TEMP MÁXIMA NA HORA ANT. (AUT) (°C)	18.831	19.5	18.8
TEMP MÍNIMA NA HORA ANT. (AUT) (°C)	17.787	18.6	17.9
VENTO, DIREÇÃO HORÁRIA (gr) (° (gr))	135.634	84	96
VENTO, RAJADA MÁXIMA (m/s)	4.657	3.3	4.5
VENTO, VELOCIDADE HORÁRIA (m/s)	1.762	1.7	1.7

Tabela 6 – Estatísticas Descritivas para 2023

2023	Variância	Desvio Padrão	Amplitude
PRECIPITAÇÃO TOTAL, HORÁRIO (mm)	1.1491	1.0720	30.2
P.Atm AO NÍVEL DA ESTAÇÃO, HORÁRIA (mB)	10.5108	3.2420	17
P.Atm MÁX. NA HORA ANT. (AUT) (mB)	10.3709	3.2204	15.7
P.Atm MÍN. NA HORA ANT. (AUT) (mB)	10.7188	3.2740	17.3
TEMP DO AR - BULBO SECO, HORÁRIA (°C)	16.1990	4.0248	18.6
TEMP MÁXIMA NA HORA ANT. (AUT) (°C)	17.5354	4.1875	20.3
TEMP MÍNIMA NA HORA ANT. (AUT) (°C)	14.9661	3.8686	19.2
VENTO, DIREÇÃO HORÁRIA (gr) (° (gr))	9831.97	99.1563	359
VENTO, RAJADA MÁXIMA (m/s)	4.0719	2.0179	9.4
VENTO, VELOCIDADE HORÁRIA (m/s)	1.0005	1.0002	4.6

Tabela 7 – Estatísticas Descritivas para 2023

2023	IQR	Máximo	Mínimo
PRECIPITAÇÃO TOTAL, HORÁRIO (mm)	0	30.2	0
P.Atm AO NÍVEL DA ESTAÇÃO, HORÁRIA (mB)	4.6	921	904
P.Atm MÁX. NA HORA ANT. (AUT) (mB)	4.9	921	905.3
P.Atm MÍN. NA HORA ANT. (AUT) (mB)	897	920.8	903.5
TEMPERATURA DO AR - BULBO SECO, HORÁRIA (°C)	5.5	27.5	8.9
TEMP MÁXIMA NA HORA ANT. (AUT) (°C)	6.4	30.3	10
TEMP MÍNIMA NA HORA ANT. (AUT) (°C)	3	27.5	8.3
VENTO, DIREÇÃO HORÁRIA (gr) (° (gr))	219.5	360	1
VENTO, RAJADA MÁXIMA (m/s)	4.5	9.9	0.5
VENTO, VELOCIDADE HORÁRIA (m/s)	1.5	4.6	0

Tabela 8 – Estatísticas Descritivas para 2023

2.5 ANÁLISE COMPARATIVA ENTRE 2022 E 2023

2.5.1 Precipitação Total

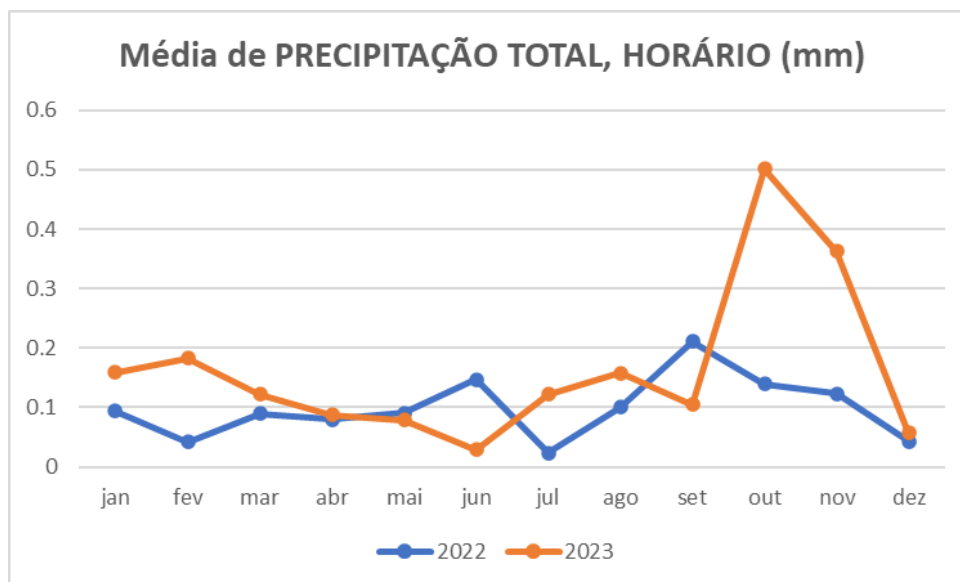


Figura 21 – Precipitação Ano

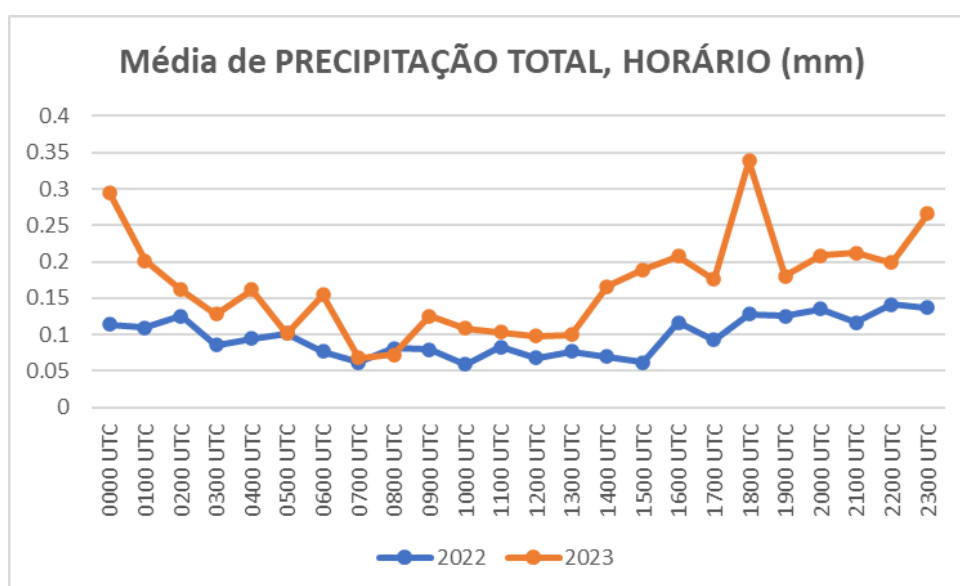


Figura 22 – Precipitação Dia

Os dois gráficos apresentam a média de precipitação total (mm) para os anos de 2022 e 2023, com o primeiro gráfico focado na variação mensal e o segundo na variação horária (UTC). A análise conjunta revela um padrão de variações mais acentuadas em 2023 em ambas as escalas temporais.

2.5.1.1 Tendência Geral ao Longo do Ano

Em **2022**, a precipitação se manteve relativamente estável ao longo dos meses, sem grandes picos, refletindo um comportamento mais uniforme. Em **2023**, por outro lado, apresentou-se uma variação significativa ao longo do ano, especialmente a partir de **setembro**, com um grande **pico de precipitação em outubro**, seguido por uma queda brusca em novembro e dezembro.

2.5.1.2 Comparação Mensal e Horária

O primeiro gráfico mostra que, enquanto nos primeiros meses de ambos os anos (de janeiro a agosto) as médias mensais de precipitação são bastante semelhantes, **2023** se destaca com um **aumento acentuado em outubro**. Esse comportamento é reforçado pelo segundo gráfico, que mostra a variação diária. Nele, em **2023**, há dois picos principais de precipitação por hora: um no início da madrugada e outro mais significativo no final da tarde, que não aparecem em 2022.

2.5.1.3 Padrão Horário de 2023

Em termos de precipitação horária, **2023** apresenta variações mais pronunciadas ao longo do dia. O aumento mais expressivo ocorre no final da tarde, indicando que, além de maior volume anual, há também uma concentração da precipitação em horários específicos. **2022** mantém um comportamento muito mais estável em todas as horas do dia, com pouca variação entre os diferentes horários.

2.5.1.4 Pontos de Destaque

O comportamento de **2023**, tanto em termos de meses quanto de horas, sugere eventos de chuva mais intensos e concentrados, especialmente no **outono** e em horários específicos do dia. O pico de **outubro de 2023** no primeiro gráfico é particularmente notável e coincide com o aumento de precipitação nas horas da tarde, conforme evidenciado no segundo gráfico. Além disso a média anual de aumentou em 70% de 2022 para 2023.

O ano de **2023** foi marcado por padrões de precipitação mais variáveis e intensos, tanto ao longo do ano quanto do dia, com destaque para o mês de **outubro** e as **tardes**. Já **2022** apresentou uma precipitação mais estável e previsível, sem grandes oscilações.

2.5.2 Pressão Atmosférica

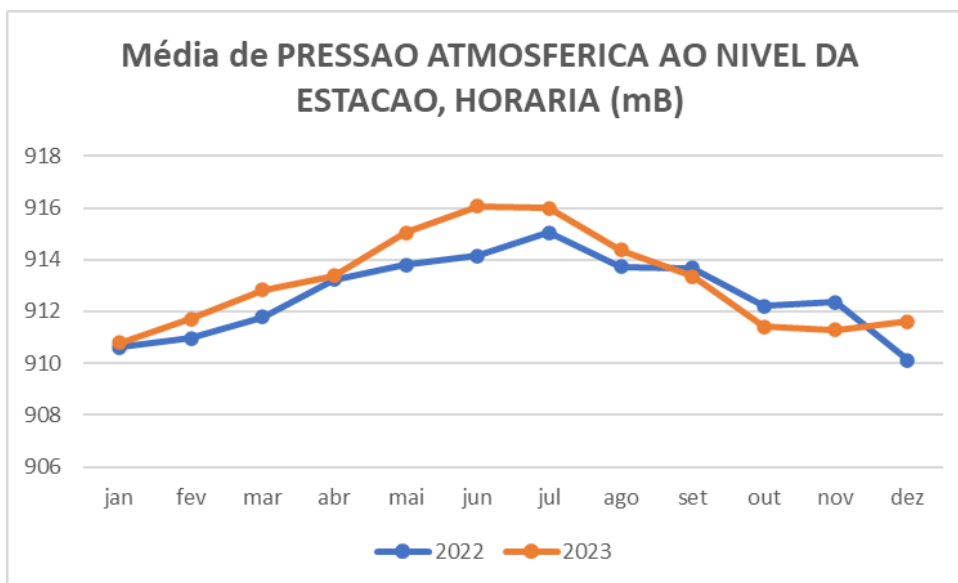


Figura 23 – Pressão Ano

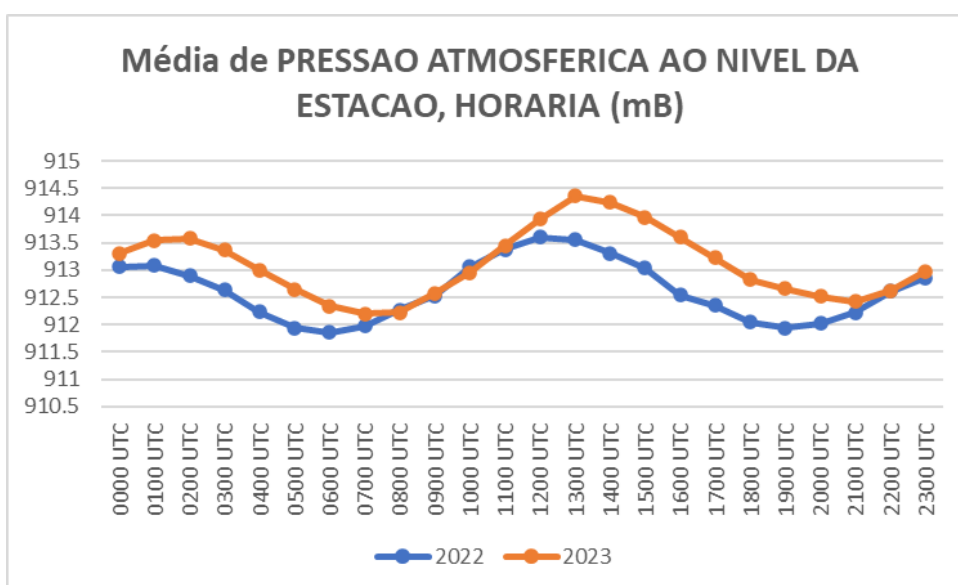


Figura 24 – Pressão Dia

Variação Horária

Tendências Observadas:

- Em ambos os anos, a pressão atmosférica é mais baixa durante a manhã (por volta das 05:00 BRT) e atinge seu pico à tarde (entre 11:00 e 12:00 BRT).
- Após o pico à tarde, a pressão diminui novamente até a meia-noite.

Comparação:

- A pressão atmosférica em 2023 tende a ser ligeiramente mais alta do que em 2022, especialmente durante as horas da tarde, onde a diferença chega a ser de aproximadamente 1 mB.
- A amplitude das variações diárias é maior em 2023, com um pico mais acentuado e uma depressão menos pronunciada comparada a 2022.

Pontos de Destaque:

- Em 2022, a pressão atmosférica atinge seu ponto mais baixo (~ 911.5 mB) por volta das 05:00 BRT e seu pico (~ 913.5 mB) às 12:00 BRT.
- Em 2023, a pressão atmosférica atinge seu ponto mais baixo (~ 912 mB) por volta das 05:00 BRT e seu pico (~ 914.5 mB) às 11:00 BRT.
- A diferença mais notável é durante o pico da tarde, onde 2023 mostra uma pressão mais alta comparada a 2022.

Variação Mensal**Tendências Observadas:**

- Em ambos os anos, há um aumento gradual na pressão atmosférica de janeiro a meio do ano (junho/julho), seguido por uma diminuição na segunda metade do ano.
- O pico anual ocorre em julho para 2022 e junho para 2023.

Comparação:

- A pressão atmosférica média mensal em 2023 é geralmente mais alta do que em 2022, com uma diferença notável no meio do ano (junho/julho), onde a pressão atinge cerca de 916 mB em 2023 comparado a 914 mB em 2022.
- A queda na pressão atmosférica após o pico de meio ano é mais acentuada em 2023 do que em 2022, resultando em valores similares no final do ano.

Pontos de Destaque:

- Em 2022, a pressão atinge seu ponto mais baixo (~ 910 mB) em janeiro e dezembro, e seu pico (~ 914 mB) em julho.
- Em 2023, a pressão atinge seu ponto mais baixo (~ 910 mB) em janeiro e dezembro, e seu pico (~ 916 mB) em junho.
- Os valores de pressão atmosférica em 2023 são consistentemente mais altos que em 2022, especialmente no primeiro semestre do ano.
- A pressão atmosférica em 2023 apresenta um aumento mais rápido no início do ano e uma queda mais pronunciada no final do ano.

Resumo Final

Os dois gráficos mostram padrões consistentes de variação horária e mensal da pressão atmosférica entre os anos de 2022 e 2023, com 2023 exibindo valores ligeiramente mais altos em ambos os períodos. As tendências diárias e sazonais são similares, mas 2023 apresenta uma maior amplitude nas variações horárias e um pico mais elevado no meio do ano. Essas observações indicam que, apesar das variações diárias e sazonais, os padrões gerais se mantêm consistentes entre os anos.

2.5.3 Temperatura do Ar

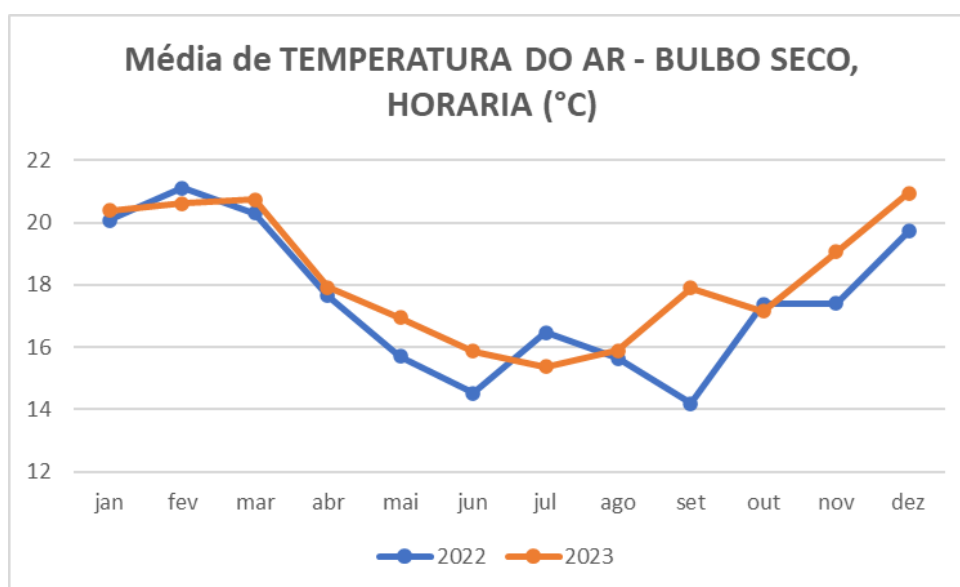


Figura 25 – Temp Ano

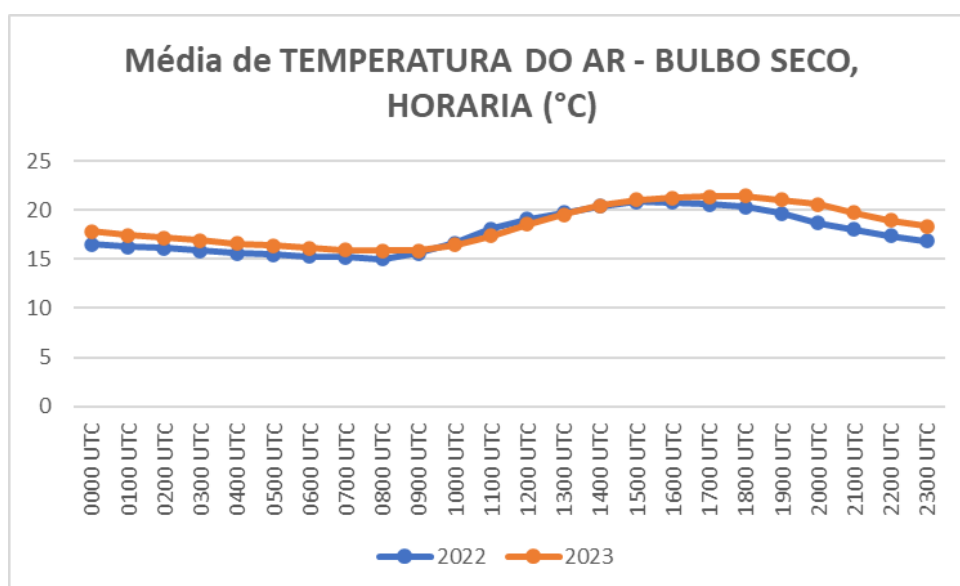


Figura 26 – Temp Dia

Os gráficos fornecidos comparam as médias de temperatura horária e mensal dos anos de 2022 e 2023. A seguir, são discutidas as tendências observadas, as comparações entre os dois anos e os pontos de destaque.

2.5.3.1 Média de Temperatura Horária (2022 vs. 2023)

2.5.3.1.1 *Tendência:*

- A variação de temperatura ao longo do dia segue o padrão típico de um ciclo diurno, com as menores temperaturas ocorrendo durante as primeiras horas da manhã (**UTC 05:00 - 07:00**, equivalente a **02:00 - 04:00** no horário de Curitiba) e os picos de temperatura ao redor do meio-dia e início da tarde (**UTC 14:00 - 17:00**, que corresponde a **11:00 - 14:00** em Curitiba).
- Há um aumento gradual da temperatura nas primeiras horas da manhã, atingindo o máximo no início da tarde, seguido por um declínio até a noite.

2.5.3.1.2 *Comparação:*

- A linha de **2023** (representada em laranja) está **consistentemente acima** da linha de **2022** (representada em azul) durante a maior parte do dia, indicando que 2023 foi um ano mais quente em termos de temperatura média horária.
- A maior diferença entre os dois anos ocorre no período entre **10:00 UTC e 17:00 UTC** (correspondente a **07:00 - 14:00** no horário de Curitiba), onde as temperaturas de 2023 são significativamente mais altas.

2.5.3.1.3 *Pontos de Destaque:*

- O **pico de temperatura** em ambos os anos ocorre entre **14:00 - 17:00 UTC** (equivalente a **11:00 - 14:00** em Curitiba).
- As menores variações de temperatura entre os dois anos ocorrem nas primeiras horas do dia (**00:00 - 06:00 UTC**, ou **21:00 - 03:00** no horário de Curitiba).
- As temperaturas de 2023 se mantêm **mais elevadas** ao longo do dia, especialmente durante o período da tarde.

2.5.3.2 Média de Temperatura Mensal (2022 vs. 2023)

2.5.3.2.1 Tendência:

- Ambos os anos apresentam o padrão sazonal esperado, com temperaturas **mais altas no verão** (início e fim do ano) e **mais baixas no inverno** (meio do ano).
- A curva de 2022 apresenta uma descida mais suave, enquanto a curva de 2023 mostra uma maior variação mês a mês, especialmente nos meses de transição de estações.

2.5.3.2.2 Comparação:

- O ano de **2023** tem temperaturas médias **consistentemente mais elevadas** que 2022, exceto nos meses de **abril** e **novembro**, onde as diferenças são menores.
- As maiores discrepâncias ocorrem em **setembro** e **dezembro**, onde 2023 registra temperaturas significativamente mais altas do que 2022.

2.5.3.2.3 Pontos de Destaque:

- A **temperatura mínima** é registrada em **julho** para ambos os anos, sendo mais baixa em 2022 (cerca de 14 °C) do que em 2023.
- O mês de **setembro de 2023** se destaca por apresentar um aumento significativo nas temperaturas em relação a 2022.
- A **temperatura máxima** é registrada em **janeiro** e **dezembro** em ambos os anos, com 2023 apresentando temperaturas mais altas, especialmente em dezembro (aproximadamente 22 °C).

2.5.4 Vento: Direção e Velocidade

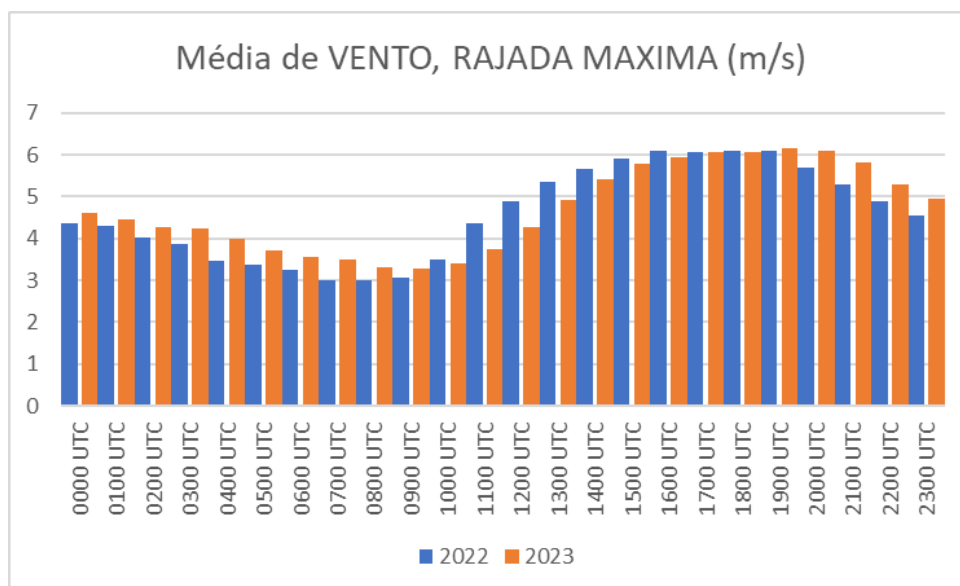


Figura 27 – Média de Rajada Máxima ao Longo de um dia.

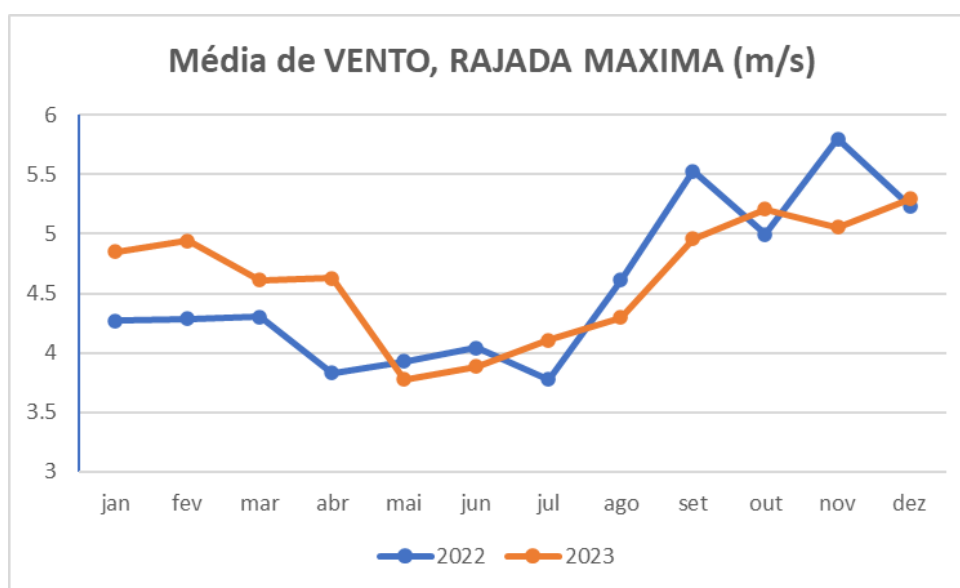


Figura 28 – Média de Rajada Máxima ao Longo do Ano.

2.5.4.1 Análise da Média de Vento e Rajada Máxima

- 1. Tendência Geral:** Pela assimetria do histograma à direita observada na figura 27, a rajada máxima mostra uma tendência de aumento ao longo do dia para ambos os anos. Como observado no gráfico 30, ao longo dos meses a rajada máxima apresenta flutuações para ambos os anos, mas há um aumento de cerca de 2% na média de 2023 em comparação com 2022.
- 2. Comparação por Horário:**

- **Períodos Matutinos (0300 UTC a 1500 UTC):** Durante as primeiras horas da manhã, a média de vento em 2023 é consistentemente mais alta do que em 2022, indicando um aumento na intensidade do vento nesse período. Esse fenômeno pode estar relacionado a características locais de microclima e influências geográficas da cidade.
- **Períodos Vespertinos (1500 UTC a 0200 UTC):** Após o meio-dia, a diferença entre os dois anos é ainda mais pronunciada, com 2023 mostrando médias superiores, especialmente nas horas do final da tarde e início da noite. Isso pode estar ligado a alterações nos padrões de temperatura e pressão atmosférica em Curitiba.

3. Resumo da Comparação Trimestral

- **Primeiro Trimestre (Janeiro a Março):** Em 2023, a média de vento destacou-se em relação a 2022, com picos mais altos, especialmente em fevereiro e março.
- **Segundo Trimestre (Abril a Junho):** Após uma queda acentuada em maio para ambos os anos, a média de 2023 ainda permaneceu um pouco acima de 2022. Em junho, a média se recuperou, superando a de 2022.
- **Terceiro Trimestre (Julho a Setembro):** Em julho, 2023 atingiu um pico significativo, indicando rajadas mais intensas. Em agosto, a média apresentou uma leve queda, mas ainda permaneceu acima de 2022.
- **Quarto Trimestre (Outubro a Dezembro):** No último quadrimestre, as médias de 2023 se destacaram, principalmente em setembro e novembro, com picos acentuados. Em dezembro, a média estabilizou, mas continuou acima da de 2022.

2.6 ANÁLISE DE TEMPERATURA POR AGRUPAMENTO TEMPORAL

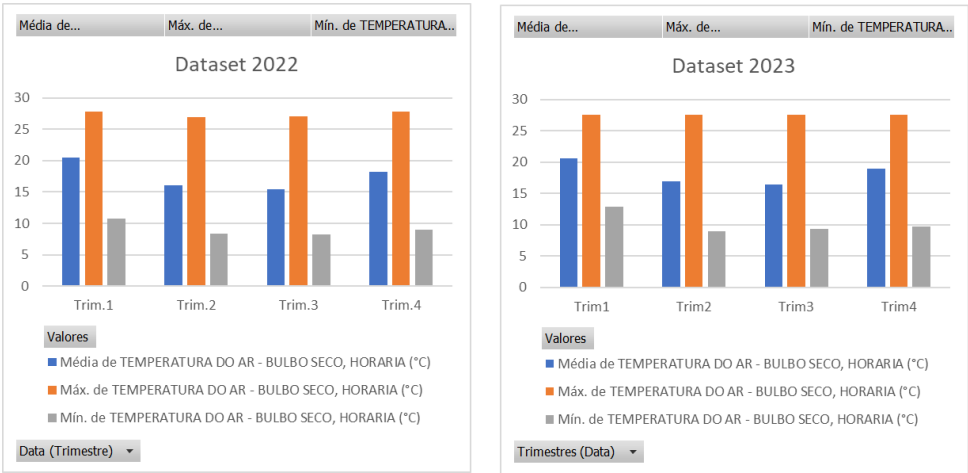


Figura 29 – Análise Trimestral

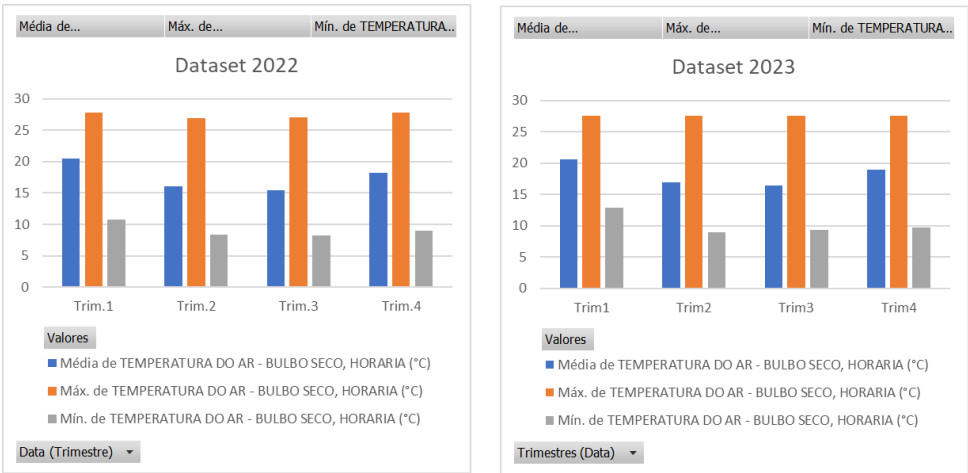


Figura 30 – Análise Anual

2.7 FÓRMULAS UTILIZADAS

Para encontrar a média, máximo e mínimo foi utilizado tabela dinâmica. Para as demais foram utilizadas: MOD.ÚNICO, MED, VAR e DESVPAD.P. Além disso foram utilizadas fórmulas mais complexas para cálculo temporais, exemplo:

=MODO.ÚNICO(SE(MÊS('2022 (2)'!A2:A7429)=B101;'2022 (2)'!K2:K7429))

3 RESULTADO DAS ANÁLISE

3.1 LINK PARA O REPOSITÓRIO

Repositório Git <https://github.com/dhaysetito/TSE-INMET-Analysis>.