PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS

Tecnologia em Banco de Dados

**Análise da influência da variação da capacidade de geração e demanda consumidora no aumento da tarifa de energia elétrica**

Bruno Silva Devesa

Sanderson Esteves Vieira

Anselmo Berriel de Lira

Douglas Gonçalves Guglielmelli

Belo Horizonte

2022

Sumário

[1. INTRODUÇÃO 2](#_Toc103773742)

[2. OBJETIVO 3](#_Toc103773743)

[3. METODOLOGIA 3](#_Toc103773744)

[3.1. Descrição da Base de Dados 3](#_Toc103773745)

[3.2. Arquitetura 3](#_Toc103773746)

[3.2.1. Padrão *Pipe-Filter* 4](#_Toc103773747)

[3.2.2. Padrão *Layered Pattern* 4](#_Toc103773748)

[3.2.3. Diagramas 5](#_Toc103773749)

[3.2.4. Descrição Metodológica 7](#_Toc103773750)

[4. DESENVOLVIMENTO 7](#_Toc103773751)

[4.1. Extração e limpeza 7](#_Toc103773752)

[4.2. Limpeza e Transformação 8](#_Toc103773753)

[4.3. Visualização dos Dados 8](#_Toc103773754)

[5. APRESENTAÇÃO E CONCLUSÃO 8](#_Toc103773755)

[REFERÊNCIAS 9](#_Toc103773756)

[APÊNDICE A - Roteiro de Tratamento de Dados 11](#_Toc103773757)

[APÊNDICE B - Roteiro de Análise de Dados 12](#_Toc103773758)

# INTRODUÇÃO

A energia brasileira é cara, tendo vários fatores apontados como causa, sendo os principais a tributação, falhas e perdas no fornecimento, e a grande dependência das hidrelétricas. (CAMPAGNOLO, 2022; DEISTER, 2021; INSP, 2021; SABER ELÉTRICA, 2022)

Conforme o Balanço Energético Nacional de 2021, a matriz elétrica brasileira é composta em 65% da geração hídrica. O Brasil sempre foi visto como país com grande potencial energético, apesar disso, a grande dependência de hidrelétricas traz desafios com relação à distribuição e aos períodos de seca. Em agosto de 2021, a Agência Nacional de Energia Elétrica (Aneel) criou mais uma bandeira tarifária, a bandeira tarifária “escassez hídrica”, 50% mais cara que a bandeira vermelha patamar 2, a mais alta existente até então. (MINISTÉRIO DE MINAS E ENERGIA, 2021a; SILVEIRA et.al, 2021)

Apesar de hidrelétricas serem predominantes no Brasil, o país possui regiões com grande potencial para geração de energia por outras fontes. Além disso, em 10 anos, o custo da energia eólica ficou três vezes mais barata, e a energia solar teve redução de dez vezes. O Tribunal Regional Federal da 5ª Região, junto com a Companhia Energética de Pernambuco (Celpe) inaugurou sua terceira usina fotovoltaica em agosto de 2021, e, de acordo com *International Energy Agency* (2022), projetos fotovoltaicos representam 70% de todas as adições programadas para os próximos anos a partir de 2022. (ASSOCIAÇÃO BRASILEIRA DE ENERGIA SOLAR FOTOVOLTAICA, 2021; LAZARD, 2021)

Seja o contribuidor do aumento de tarifas a má gestão dos recursos hídricos, ou os períodos de estiagem, investimentos em outras matrizes energéticas amenizariam estes aumentos ao reduzir a dependência de hidrelétricas. (REDAÇÃO RBA, 2022; MINISTÉRIO DE MINAS E ENERGIA, 2021b)

Analisando dados de consumo e geração dos últimos anos, conseguiríamos avaliar a proporção do efeito da escassez de oferta e pressão de demanda sobre tarifas, investigar se a diversificação da matriz elétrica já influencia positivamente no preço ao consumidor, ou se as tributações e outros fatores ainda impedem a observação desses efeitos.

# OBJETIVO

Neste trabalho visamos explorar se o preço da energia para o consumidor varia proporcionalmente com o gap entre demanda e capacidade de geração, segmentando a análise por fontes de geração e região, no período de 2018 a 2021.

# METODOLOGIA

## Descrição da Base de Dados

Para obter informações de consumo, usaremos os dados de Carga de Energia do Operador Nacional do Sistema Elétrico (ONS).

Para informações de geração de energia disponível, utilizaremos os dados de capacidade de geração, divulgados no Sistema de Informações de Geração da ANEEL (Agência Nacional de Energia Elétrica) (SIGA), em conjunto com os dados de Disponibilidade do Sistema de Geração da ONS. O primeiro nos traz dados da potência nominal dos empreendimentos geradores, o segundo informa o tempo de disponibilidade do sistema de energia. Com ambos, podemos inferir a oferta real de energia em um período. A base do SIGA também fornece tipos de usina e o combustível utilizado na unidade geradora.

Para analisar o valor da tarifa de energia elétrica, utilizaremos a base de “Tarifas De Aplicação Das Distribuidoras De Energia Elétrica”, mantida pela Agência Nacional de Energia Elétrica (ANEEL).

Além disso, usaremos dados consolidados e tabelas fornecidas pela Empresa de Pesquisa Energética (EPE) para eventuais conferências de informações obtidas ao longo do trabalho.

## Arquitetura

O padrão ETL compreende as etapas que o profissional que está trabalhando com os dados deverá cumprir para que esses dados resolvam problemas previamente definidos. Essas etapas compreendem todo o processo de extração dos dados brutos de fontes diferentes, transformação para torná-los inteligíveis e, posteriormente, carregados em um sistema de destino para sua análise.

Para o desenvolvimento de nossa solução, tomamos por referências as arquiteturas *Pipe-filter* e *Layered Pattern*.

### Padrão *Pipe-Filter*

O padrão de projeto *Pipe-Filter Pattern* é um padrão já conhecido para execução de processos em fila, típico de processos de processamento de dados, onde a execução é iniciada a partir de um gatilho, e concluída após o processamento sequencial das etapas definidas (filtros).

Figura 1 – Esquema ilustrativo padrão *Pipe-Filter*

[Diagrama

Descrição gerada automaticamente](https://static.packt-cdn.com/products/9781789809770/graphics/assets/ab9be4ee-22f4-4f04-982f-3b04a754468c.png)

**Fonte:** LAVIERI, 2019

Neste trabalho, lidaremos com os dados de forma sequencial partindo do arquivamento de .*csv* em pasta específica, tratamento dos .*csv* com regras em *Python*, carga dos dados para banco em *SQL*, e cruzamento de dados em *SQL* e VizQL.

### Padrão *Layered Pattern*

O padrão *Layered Pattern* possui tipicamente 4 camadas: Base de Dados (*Database Layer*), Persistência (*Persistence Layer*), Negócios (*Business Layer*), Apresentação (*Presentation Layer*).

Na primeira, camada de Base de Dados, tem-se o banco de dados com tabelas populadas. O banco é acessado pelo script de tratamento de dados. A camada de Persistência de dados contém a parte do código que realiza as operações de criação, atualização, consultas e exclusões de dados (CRUD). A camada de Negócios contém a execução de regras necessárias para o cruzamento de dados e obtenção de informações relativas ao objetivo do projeto. Por fim, a camada de Apresentação coloca os dados tratados de forma visual*.*

Figura 2 - Esquema ilustrativo padrão arquitetural Layered Pattern

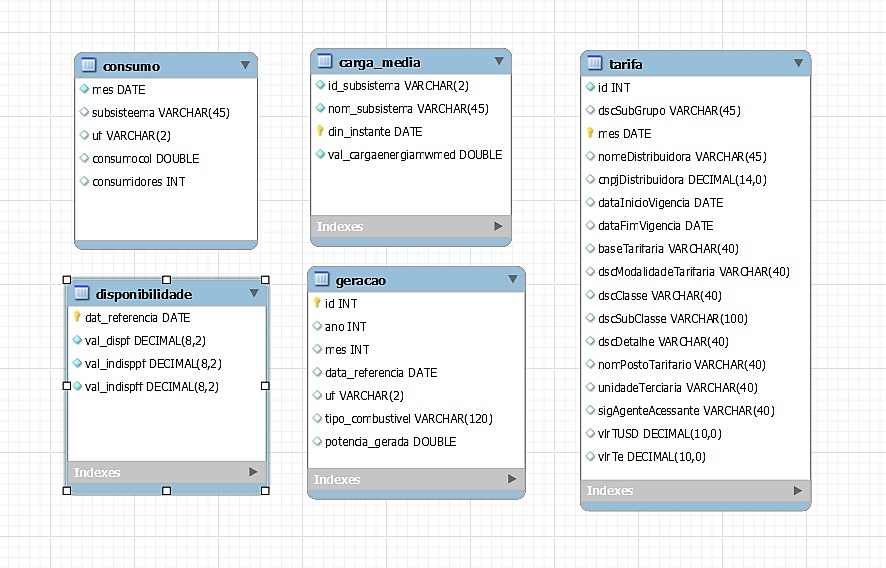
[](https://www.oreilly.com/library/view/software-architecture-patterns/9781491971437/ch01.html)

**Fonte**: RICHARDS, 2015

Neste trabalho, com adaptações, a camada de Base de Dados se tratará dos arquivos *.csv* originados após tratamento dos *.csv* originais por scripts *Python*. A camada de persistência será representada pelos scripts de *SQL* que fazem a carga dos dados para o *MySQL*. A camada de negócios, representada pelos scripts em *SQL* que realizam o cruzamento dos dados de diferentes tabelas para obtenção de informações relacionadas ao objetivo. E, finalmente, a camada de Apresentação representada pela ferramenta de BI, *Tableau*, incluindo eventuais códigos em *VizQL* nele incluídos.

### Diagramas

**Figura 3** – Diagrama de Classes



**Fonte:** Autoria própria.

Figura 4 – Diagrama de atividades

Linha do tempo

Descrição gerada automaticamente

**Fonte:** Autoria própria

Figura 5 – Diagrama de Infraestrutura

Diagrama

Descrição gerada automaticamente

**Fonte:** Autoria própria

### Descrição Metodológica

Seguindo o *ETL*, no desenvolvimento deste trabalho, a coleta de dados e criação da camada de Base de Dados (*Database Layer)* se iniciará pelo armazenamento local dos arquivos coletados das diferentes fontes. Os dados de origem encontram-se em arquivos .*csv*. Os arquivos serão lidos e processados por scripts em *Python* para tratamento dos dados de acordo com o **APÊNDICE A** - Roteiro Tratamento de Dados. Os dados tratados serão salvos em novos arquivos .*csv*.

Estes arquivos por sua vez, serão carregados e unificados em um único banco no *MySQL* a partir de script em *SQL*. O modelo de classes exposto na Figura 3 demonstra as relações e o arranjo que darão origem ao nosso banco de dados com os dados agrupados e tratados. Segue dicionário de dados para complementar e auxiliar a compreensão do diagrama de classes

* Consumo: Dados sobre consumo de energia elétrica.

|  |  |  |
| --- | --- | --- |
| **Coluna** | **Tipo** | **Definição** |
| mes | Date | Data de registro do dado |
| uf | String | Sigla da unidade federativa |
| consumo | Float | Consumo em MWh |

* Disponibilidade: Indicadores de disponibilidade de geração de energia

|  |  |  |
| --- | --- | --- |
| **Coluna** | **Tipo** | **Definição** |
| dat\_referencia | Date | Data de referência |
| val\_dispf | Float | Indicador de Disponibilidade das Funções Geração |
| val\_insdisppf | Float | Indicador de Indisponibilidade Programada das Funções Geração |
| val\_insdispff | Float | Indicador de Indisponibilidade Forçada das Funções Geração |

* Carga\_media: Carga de energia média por unidade federativa

|  |  |  |
| --- | --- | --- |
| **Coluna** | **Tipo** | **Definição** |
| id\_subsistema | String | Sigla da unidade federativa |
| nome\_subsistema | String | Nome da unidade federativa |
| dn\_instante | Date | Data de referência |
| val\_cargaenergiamwmed | Float | Valor da carga de energia média medido em MW |

* Geracao: Dados sobre capacidade de geração de energia

|  |  |  |
| --- | --- | --- |
| **Coluna** | **Tipo** | **Definição** |
| id | Integer | Identificador único do registro no sistema |
| ano | Integer | Ano de referência |
| mes | Integer | Mês de referência |
| data\_referencia | Date | Data de registro do dado |
| uf | String | Sigla da unidade federativa |
| tipo\_combustivel | String | Combustível da unidade geradora |
| potencia\_gerada | Float | Potência nominal da unidade geradora em MW |

* Tarifa: Tarifas homologadas das distribuidoras de energia elétrica

|  |  |  |
| --- | --- | --- |
| **Coluna** | **Tipo** | **Definição** |
| id | Integer | Identificador único do registro no sistema |
| dscSubGrupo | String | Subgrupos tarifários definidos pela Resolução Normativa nº 1000/2021. |
| rres | Date | Data de referência |
| nomeDistribuidora | String | Nome da unidade distribuidora |
| cnpjDistribuidora | Decimal | Cadastro Nacional de Pessoas Jurídicas do Ministério da Fazenda |
| dataInidoVigencia | Date | Informação sobre a data de início da vigência das tarifas |
| dataFimVigencia | Date | Informação sobre a data de fim da vigência das tarifas |
| baseTrifaria | String | Tarifa de Aplicação e Base Econômica |
| DscModalidadeTarifaria | String | Conjunto de tarifas aplicáveis às componentes de consumo de energia elétrica demanda |
| DscClasse | String | Classificação do tipo de unidade consumidora conforme a finalidade de utilização da energia elétrica |
| DscSubClasse | String | Classificação do tipo de unidade consumidora, como subdivisão das classes |
| DscDetalhe | String | Conjunto complementar de variáveis utilizadas na definição das tarifas quanto a critérios de aplicação ou universo de acessantes elegíveis |
| NomPostoTarifario | String | Identificação do posto tarifário, |
| unidadeTerciaria | String | Unidade da tarifa, conforme a grandeza elétrica aplicável |
| SigAgenteAcessante | String | Aplicável nos casos de tarifas nominais, aplicáveis especificamente a um usuário (distribuidora, unidade consumidora ou gerador) |
| VlrTusd | Decimal | Apresenta o valor da Tarifa de Uso do Sistema de Distribuição -TUSD em valor monetário (R$/MWh ou R$/kW) |
| VlrTe | Decimal | Apresenta o valor da Tarifa de Energia –TE em valor monetário (R$/MWh). |

Com o banco de dados populado, é criado um Roteiro de Análise de dados, incluído no **APÊNDICE B** - Roteiro de Análise de Dados, o qual será referenciado para criação de regras de negócio, que são as regras pertinentes especificamente ao objetivo do trabalho. Essas regras poderão ser implementadas tanto no próprio banco, a partir de consultas *SQL*, quanto no software Tableau, na criação dos painéis de visualização.

# DESENVOLVIMENTO

## Extração e limpeza

Os algoritmos para extração e limpeza dos dados bruto, desenvolvidos de acordo com o Apêndice A, e os arquivos *.csv* resultantes, com dados tratados, foram subidos para o *GitHub* com acesso no link:

[*https://bit.ly/GitHub-DataBase-ExtracaoTratamento*](https://bit.ly/GitHub-DataBase-ExtracaoTratamento)

## Limpeza e Transformação

Os algoritmos para carga dos dados contidos nos novos .csv ao *MySQL* foram desenvolvidos e subidos para o *GitHub*, com acesso no link:

[*https://bit.ly/GitHub-PersistenceBusiness-CarregamentoCruzamento*](https://bit.ly/GitHub-PersistenceBusiness-CarregamentoCruzamento)

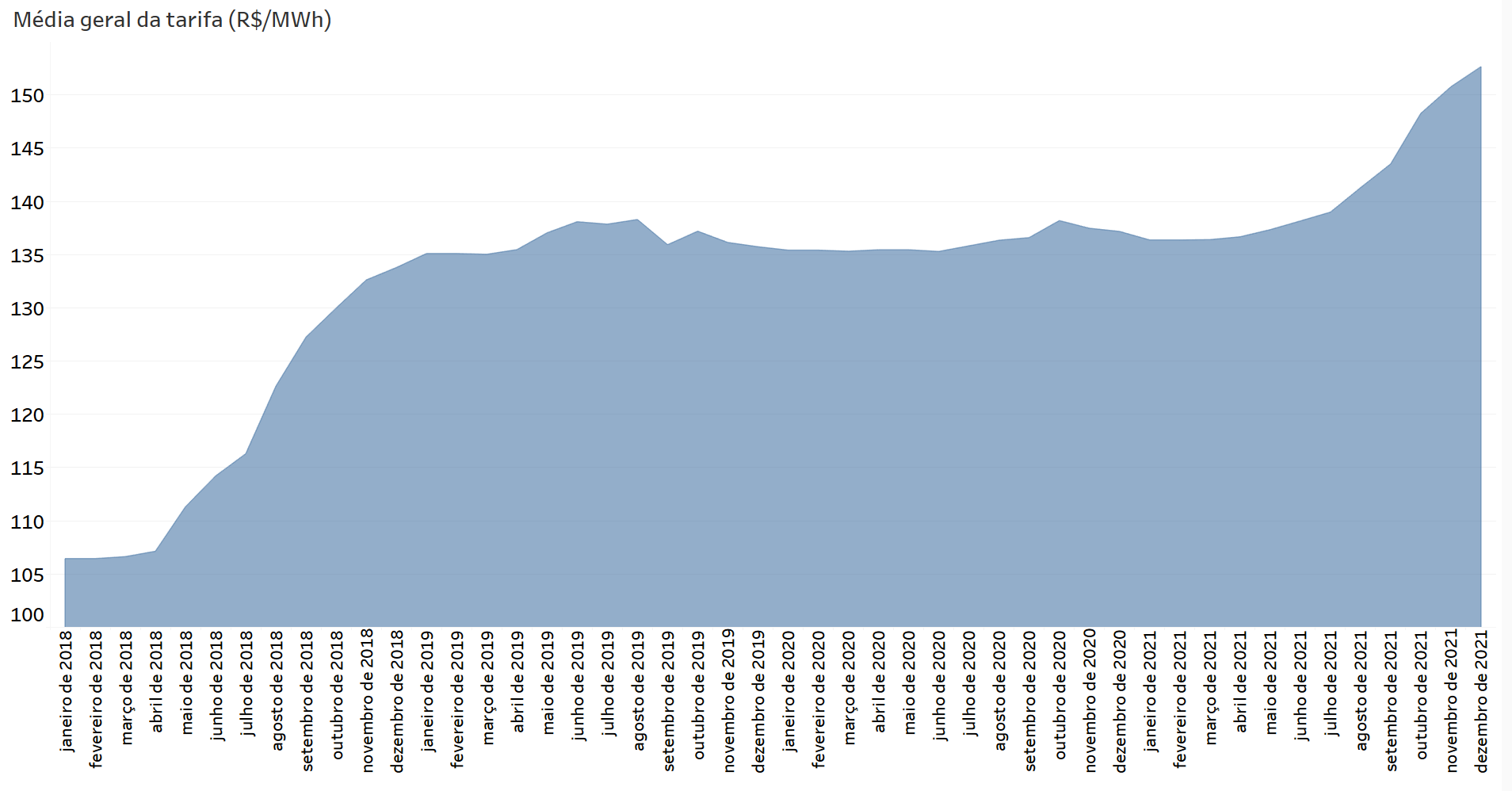
## Visualização dos Dados

O banco de dados populado no *MySQL* foi importado pelo *Tableau*, onde diferentes visualizações foram criadas tomando por referência o APÊNDICE B - Roteiro de Análise de Dados, para obtenção de informações. O painel desenvolvido foi subido para o GitHub com acesso no link.

**INSERIR LINK**

# APRESENTAÇÃO E ANÁLISE

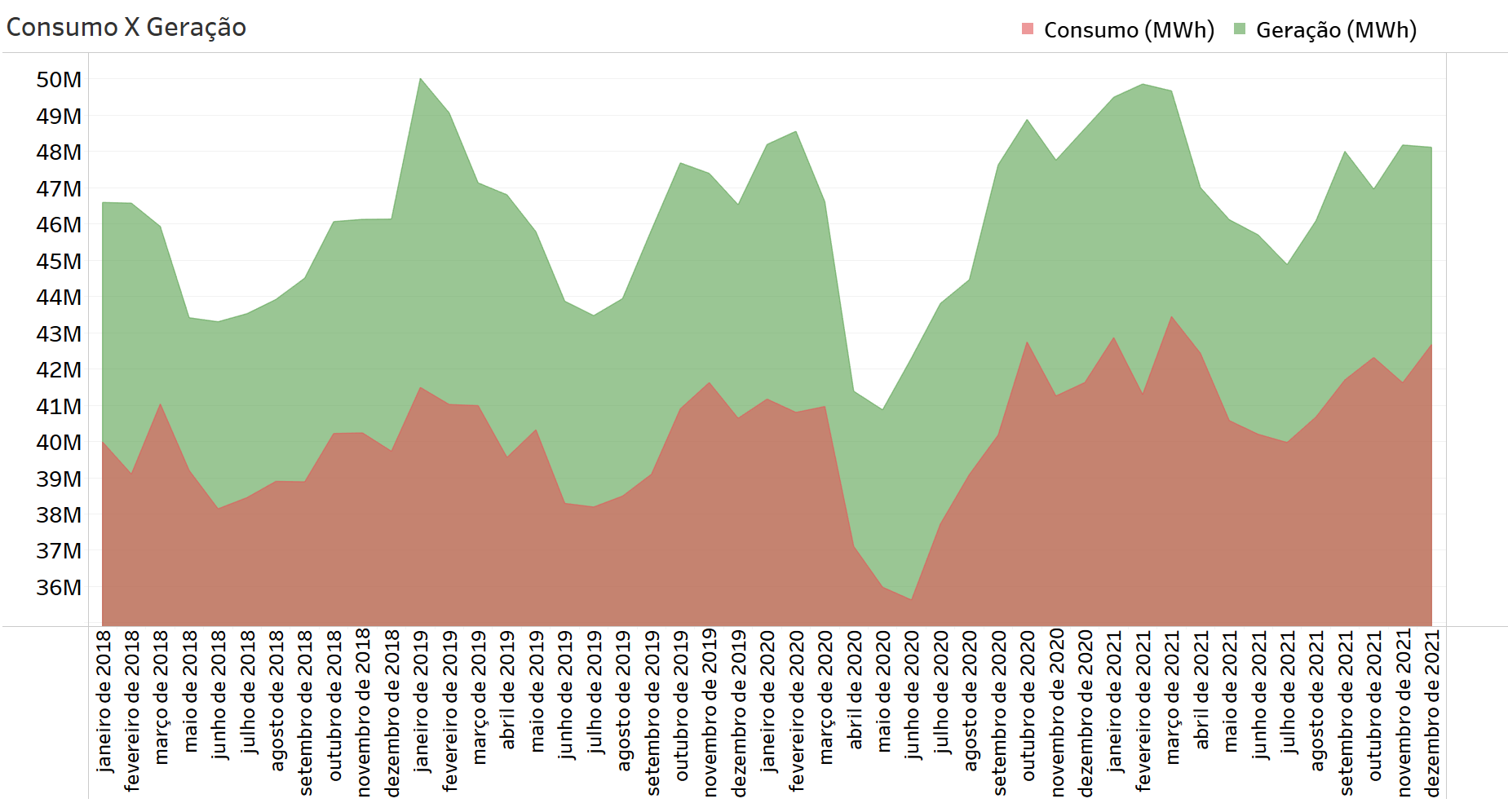
Figura 6 – Gráfico do valor médio da tarifa de energia ao longo do tempo



**Fonte:** Autoria própria

Nesse gráfico buscamos ilustrar com detalhe a variação da tarifa de energia elétrica ao longo do nosso horizonte de estudo (2018-2021). Percebe-se uma tendência contínua de crescimento com destaque para dois períodos de forte alta. A partir do segundo trimestre de 2018 houve um aumento muito significativo nos valores, assim como no período a partir de agosto de 2021. No intervalo entre esses períodos a variação na tarifa ocorreu em menor escala. É plausível que fatores externos tenham sido determinantes para a variação tão drástica nesses dois momentos, o que pode ser melhor explorado nas próximas etapas.

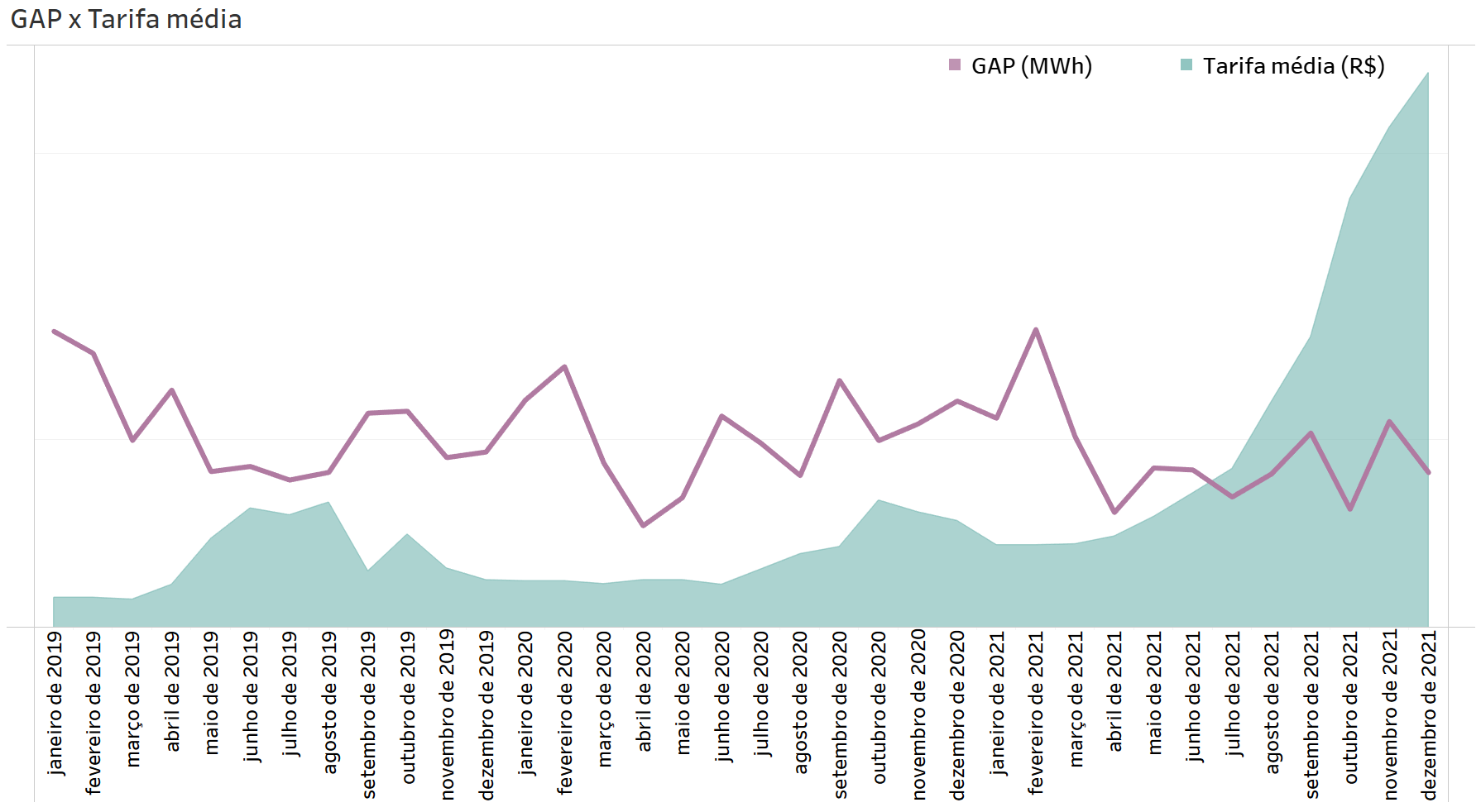
Figura 7 – Gráfico de comparação da potência gerada com consumo médio



**Fonte:** Autoria própria

Comparamos nesse gráfico a potência elétrica gerada em todo o país com o consumo de energia médio para o mesmo período. O que nos chamou mais atenção foi como as duas medidas estão correlacionadas. Com destaque para o período de fevereiro até setembro de 2020, onde fica evidente como as duas variáveis caminham juntas. Em uma análise futura, testes estatísticos podem ser realizados para medir o grau de correlação existente.

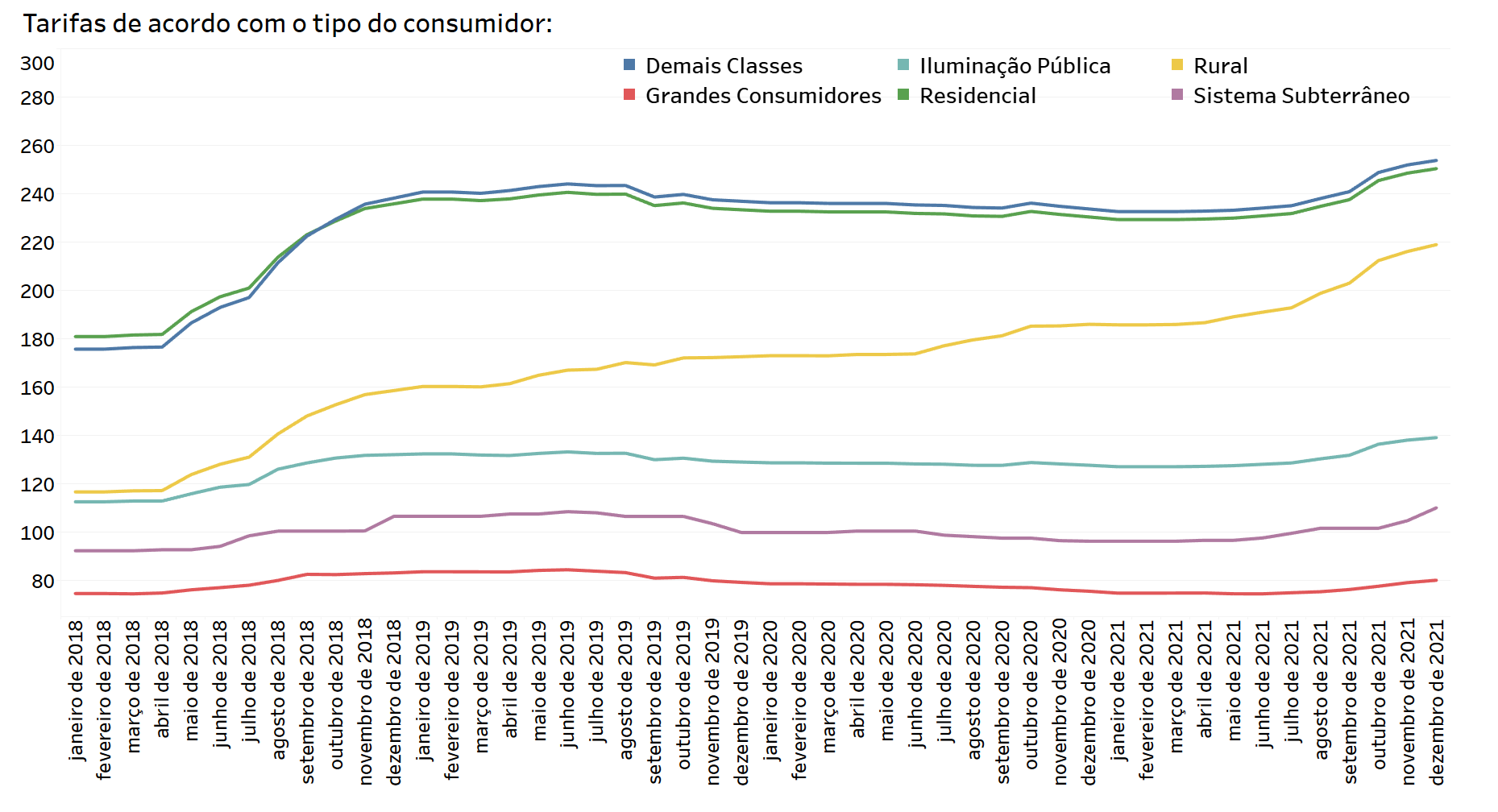
Figura 8 – Gráfico comparação do GAP com a tarifa



**Fonte:** Autoria própria

Uma das principais questões que pretendemos analisar com esse trabalho é a relação entre o gap (geração de energia – consumo) e a tarifa elétrica. Com essa ilustração fica evidente que essa relação é mais complexa do que antecipávamos. Não há uma correlação óbvia, e em alguns momentos como no início de 2019, o gap diminuiu enquanto a tarifa subia. Já no período entre abril e outubro de 2020 as duas medidas subiram concomitantemente. Além disso a tarifa teve um aumento muito grande a partir de junho de 2021, enquanto a variação do gap se manteve relativamente dentro do mesmo intervalo ao longo de todo o período estudado.

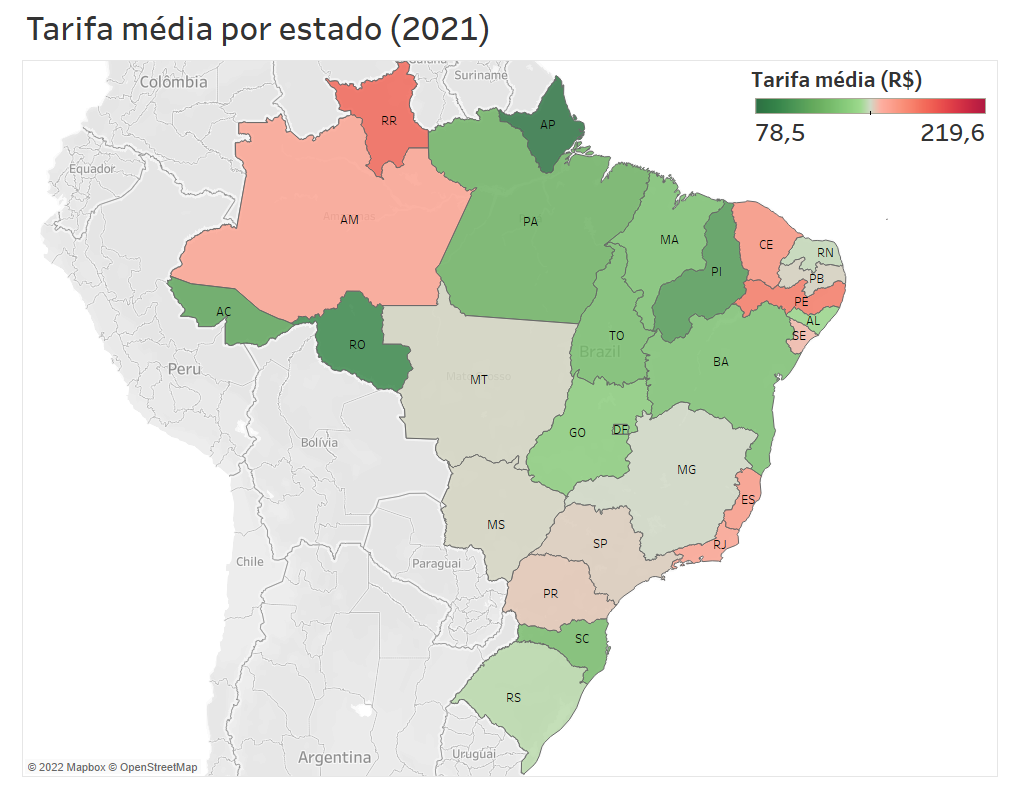
Figura 9 – Tarifa média de acordo com o tipo de consumidor



**Fonte:** Autoria própria

Com esse gráfico pretendemos elucidar a diferença da tarifa paga pelos grupos de consumidores específicos. Primeiramente, destaca-se a discrepância entre o valor cobrado do grupo residencial e dos grandes consumidores, e que perpassa todo o período destacado. Sabíamos de antemão que os valores residenciais seriam elevados, mas a diferença existente é de uma magnitude surpreendente. Além disso a variação da tarifa cobrada de consumidores rurais sofreu um grande aumento ao longo dos anos, proporcionalmente maior do que qualquer outro grupo. Tal questão deverá ser melhor explorada na próxima etapa.

Figura 10 – Tarifa média de acordo com o estado em 2021



**Fonte:** Autoria própria

Esse gráfico ilustra a variação da tarifa média de energia para cada estado do Brasil no ano de 2021. Pretendemos oferecer um panorama geral da disparidade do custo da energia para os consumidores de diferentes localidades. Destaca-se o estado de Roraima que apresenta os valores mais elevados, e o estado da Amapá que possui o menor valor médio. Por fim, ressalto que essa visualização quando observada no Tableau permite acompanhar a evolução ao longo dos anos, e para esse documento optamos pelo ano de 2021 por ser mais recente.

# CONCLUSÃO

# REFERÊNCIAS

|  |
| --- |
| AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. **Tarifas de aplicação das distribuidoras de energia elétrica**: 2022. Disponível em: https://dadosabertos.aneel.gov.br/pt\_BR/dataset/tarifas-distribudoras-energia-eletrica  ASSOCIAÇÃO BRASILEIRA DE ENERGIA SOLAR FOTOVOLTAICA. **TRF5 vai instalar mais uma usina fotovoltaica:** 2022. Disponível em: <https://www.absolar.org.br/noticia/trf5-vai-instalar-mais-uma-usina-fotovoltaica/>. Acesso em 11 ab. 2022. |
| CAMPAGNOLO, Edson. **5 razões para o Brasil ter uma das energias mais caras do mundo**. Paraná: Agência Sistema FIEP, 2 abr. 2018. Disponível em: https://agenciafiep.com.br/2018/04/02/5-razoes-para-o-brasil-ter-uma-das-energias-mais-caras-do-mundo/. Acesso em: 11 abr. 2022. |
| DEISTER, Jaqueline. **Aumento da conta de luz**: o que está por trás da "crise hídrica" instaurada no Brasil?. Rio de Janeiro: Brasil de Fato, 9 jul. 2021. Disponível em: https://www.brasildefato.com.br/2021/07/09/aumento-da-conta-de-luz-o-que-esta-por-tras-da-crise-hidrica-instaurada-no-brasil. Acesso em: 11 abr. 2022. |
| EMPRESA DE PESQUISA ENERGÉTICA. **Dados do anuário estatístico de energia elétrica**: 2018. Disponível em: https://www.epe.gov.br/pt/publicacoes-dados-abertos/dados-abertos/dados-do-anuario-estatistico-de-energia-eletrica |
| INSP – THERM TRAFO SERVICE ENGENHARIA ELÉTRICA. **Por que a energia elétrica no brasil é tão cara?**. [*S. l.*], 9 fev. 2021. Disponível em: https://www.insp-therm.com.br/blog/por-que-a-energia-eletrica-no-brasil-e-tao-cara/. Acesso em: 11 abr. 2022. |
| INTERNATIONAL ENERGY AGENCY. Overview. *In*: **Brazil**. [*S. l.*], abr 2022. Disponível em: https://www.iea.org/countries/brazil. Acesso em: 8 abr. 2022. |
| LAVIERI, Edward. **Hands-On Design Patterns with Java.** Packt, abr. 2019. ISBN: 9781789809770. Disponível em:< *https://subscription.packtpub.com/book/programming/9781789809770/10/ch10lvl1sec83/understanding-the-pipe-filter-pattern*>. Acesso em: 25/04/2022 |
| LAZARD. **Lazard’s levelized cost of energy analysis**: Version 15.0. [*S. l.*], out 2021. Disponível em: https://www.lazard.com/perspective/levelized-cost-of-energy-levelized-cost-of-storage-and-levelized-cost-of-hydrogen/. Acesso em: 8 abr. 2022. |
| MINISTÉRIO DE MINAS E ENERGIA. Agência Nacional de Energia Elétrica . **FAQ**: Tire suas dúvidas sobre a Bandeira Escassez Hídrica. 31 ago. 2021b. Disponível em: https://www.gov.br/aneel/pt-br/assuntos/noticias/2022/faq-tire-suas-duvidas-sobre-a-bandeira-escassez-hidrica. Acesso em: 11 abr. 2022. |
| MINISTÉRIO DE MINAS E ENERGIA. Empresa de Pesquisa Energética. **Balanço Energético Nacional 2021**: Ano base 2020. Brasil, 2021a. Disponível em: https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/balanco-energetico-nacional-2021. Acesso em: 11 abr. 2022. |
| OPERADOR NACIONAL DE SISTEMA ELÉTRICO. **Capacidade geração:** 8 de abril de 2022. Disponível em: https://dados.ons.org.br/dataset/capacidade-geracao/resource/a6412542-f2ce-408e-b51d-19a48cc50b62?inner\_span=True |
| REDAÇÃO RBA (ed.). **Culpa da chuva?**: Especialista mostra que conta de luz mais cara é resultado da inação do governo. São Paulo: Rede Brasil Atual, 1 set. 2021. Disponível em: https://www.redebrasilatual.com.br/ambiente/2021/09/culpa-da-chuva-especialista-mostra-que-conta-de-luz-mais-cara-e-resultado-da-inacao-do-governo/. Acesso em: 11 abr. 2022. |
| RICHARDS, Mark. **Software Architecture Patterns.** O'Reilly Media, Inc**, f**ev. 2015. ISBN: 9781491924242. Disponíve em: https://www.oreilly.com/library/view/software-architecture-patterns/9781491971437/ch01.html. Acesso em25 abr.2022 |
| SABER ELÉTRICA. **5 Razões para o Brasil ter uma das Energias mais Caras do Mundo**. [*S. l.*], 2022. Disponível em: https://www.sabereletrica.com.br/energias-mais-caras-do-mundo/. Acesso em: 11 abr. 2022. |
| SILVEIRA, Daniel; ALVARENGA , Darlan; GERBELLI, Luiz. **Conta de luz está cada vez mais cara**: Entenda por que ela sobe e quais os problemas dessa escalada de preços. **G1**, Rio de Janeiro e São Paulo, 27 ago. 2021. Disponível em: https://g1.globo.com/economia/noticia/2021/08/27/conta-de-luz-esta-cada-vez-mais-cara-entenda-por-que-ela-sobe-e-quais-os-problemas-dessa-escalada-de-precos.ghtml. Acesso em: 8 abr. 2022. |

# APÊNDICE A - Roteiro de Tratamento de Dados

Passo a passo do que o algoritmo de tratamento dos dados deverá executar. As transformações devem ser salvas em novo arquivo .*csv*, com *encoding* em UTF-8.

**A ) Base Consumo (Responsável: Douglas)**

*Fonte: EPE Anuário Estatístico de Energia*

1. Descartar colunas: 2, 6 a 11.
2. Reformatar coluna 1 (Data) e 4 (SetorN1).
3. Agregar linhas repetidas somando para Consumo e Consumidores.

**B) Base Tarifa (Responsável: Bruno)**

*Fonte ANEEL Tarifas Homologadas Distribuidoras*

1. Converter todos os preços para R$/MWh
2. Converter datas para formato apropriado
3. Converter divisores decimais
4. Descartar colunas: 1, 2.
5. Criar relação mês a mês de acordo com as datas de início e fim das tarifas.

Ex: Mês n, percorrer linhas da base, caso mês n estiver entre DataInicio e DataFim, inserir dados na linha. Próximo mês. Loop até fim.

**C) Base Capacidade de Geração (Responsável: Anselmo)**

*Fonte: ANEEL SIGA Empreendimentos Geração*

* Descartar colunas: 1, 3, 6, 7, 8, 10, 13, 14, 15, 18, 19, 20
* Selecionar apenas linhas onde DscFaseUsina : “Operação”
* Descartar coluna DscFaseUsina
* Criar colunas “Início considerado” e “Fim considerado”. Utilizar de árvore de regras para definir quais datas utilizar a partir das colunas de Data Entrada Operação, Data Inicio Vigencia e Data Fim Vigencia.
* Tratar Datas restantes (Colunas início e fim considerado podem ter trazido erros).
* Criar relação mês a mês de acordo com as datas de início e fim consideradas, com script.

Ex: Mês n, percorrer linhas da base, caso mês n estiver entre DataInicio e DataFim, inserir dados na linha. Próximo mês. Loop até fim.

**D) Base Disponibilidade de Geração (Responsável: Sanderson)**

*Fonte: ONS Disponibilidade*

1. Anos em arquivos separados, percorrer os arquivos e juntá-los em um só adicionando as linhas de um sob o outro.
2. Base utiliza “.” como separador decimal, caso necessário, converter.

**E) Base Carga de Energia (Responsável: Sanderson)**

*Fonte: ONS Carga de Energia*

* Anos em arquivos separados, percorrer os arquivos e juntá-los em um só adicionando as linhas de um sob o outro.

# APÊNDICE B - Roteiro de Análise de Dados

Objetivo do trabalho: “Explorar se o preço da energia para o consumidor varia proporcionalmente com o gap entre demanda e capacidade de geração, segmentando a análise por fontes de geração e região, no período de 2018 a 2021.”

Em outras palavras: a variação de tarifa no tempo é proporcional à variação do gap de fornecimento e demanda no tempo?

Matematicamente: **d/dt(Tarifa) α d/dt(GapDemanda) ?**

**Gráficos:**

Todo gráfico deve responder a alguma pergunta que ajuda a chegar mais perto do objetivo.

|  |  |  |
| --- | --- | --- |
|  | **Pergunta** | **Gráfico** |
| 1 | Como varia a Tarifa no tempo? | **Gráfico de linha** “Tarifa” vs “Mês”, com filtro de região. |
| 2 | Qual o Gap de geração e consumo e como varia no tempo? | **Gráfico de barras duplas** “Consumo” e “Geração” vs “Mês”, com filtro de matriz geradora e região. |
| 3 | Como variam o gap e a tarifa comparativamente? A geração, disponibilidade, consumo, fonte, influenciam na variação da Tarifa? | **Key Influencers**, analisar “Tarifa” explicada por “Geração”, “Disponibilidade”, “Consumo”, “Fonte”. |
| 4 | A Tarifa varia em função do Gap (y=f(x))? O quão próxima é essa relação?(R2) | **Regressão** “Tarifa” vs “GapDemanda”. |
| 5 | d/dt(Tarifa) α d/dt(GapDemanda) ? | **Gráfico de linha** das derivadas mensais dos dois. (Talvez o gráfico 3 já responda, se não responder e decidirmos fazer este, então vamos precisar ou criar o cálculo no BI (Ferramenta “Medidas” do Power BI), ou fazer no *MySQL* mesmo, criando uma coluna. |