

1. Objetivo da segunda entrega

Esta segunda entrega teve como objetivo aprofundar a análise da relação entre:

- número de interrupções diárias no fornecimento de energia elétrica em Brasília;
- variáveis climáticas (temperatura e precipitação diária);
- consumo mensal de energia elétrica da Neoenergia Brasília.

Atendendo às solicitações foram incluídas:

- agregações **semanais** e **mensais** das interrupções versus clima;
 - uso de **médias móveis** nas séries diárias;
 - construção de modelos de **previsão do número diário de interrupções** com separação temporal clara entre treino e teste;
 - análise conjunta de **falhas x consumo** e **consumo x temperatura**.
-

2. Bases de dados e pré-processamento

2.1. Dados de interrupções

- Arquivo base: `dados_completos_brasilia.csv`
- Período coberto (após integração com clima): **01/01/2017 a 31/05/2025**
- Unidade de análise diária:
 - data: derivada de `DatInicioInterrupcao` ;
 - variável de interesse: **número diário de interrupções**, calculado como o número de valores distintos de `NumOrdemInterrupcao` por dia.

Resumo da série diária de interrupções:

- número de dias: **3073**;
- média diária: **243,59 interrupções**;

- mínimo diário: **63 interrupções**;
- máximo diário: **1424 interrupções**.

2.2. Dados climáticos (INMET)

- Fonte: estação do INMET para Brasília, em base horária, de **2017 a 2025**.
- Arquivos horários foram agregados para gerar o arquivo diário:
 - `clima_diario_brasilia.csv`.

Variáveis usadas:

- **temperatura média diária** (`temperatura_media`), em °C;
- **precipitação total diária** (`precipitacao_total_mm`), em mm.

Período climático (base completa): **01/01/2017 a 31/07/2025**.

Resumo das séries diárias de clima (base INMET):

- **temperatura média diária**
 - período: 01/01/2017 a 31/07/2025;
 - média: **21,53 °C**;
 - mínimo: **11,31 °C**;
 - máximo: **29,92 °C**.
- **precipitação total diária**
 - período: 01/01/2017 a 31/07/2025;
 - média: **3,90 mm/dia**;
 - mínimo: **0,0 mm**;
 - máximo: **113,0 mm**.

Após integrar interrupções + clima (apenas dias com ambos os dados), o período efetivo de análise passou a ser **01/01/2017 a 31/05/2025**.

Na base diária integrada:

- **temperatura média diária**
 - média: **21,57 °C** (mínimo 11,31; máximo 29,92);
- **precipitação diária**
 - média: **3,97 mm/dia** (mínimo 0,0; máximo 113,0).

3. Interrupções x clima (diário, semanal e mensal)

3.1. Correlações diárias

Na base diária integrada (3073 dias), foi calculada a correlação de Pearson entre interrupções e clima:

- interrupções x temperatura média diária:
 - $r \approx 0,10$ (associação linear **fraca e positiva**; temperatura sozinha explica pouco da variabilidade diária de falhas);
- interrupções x precipitação diária:
 - $r \approx 0,35$ (associação linear **moderada**; dias com maior chuva tendem a apresentar mais interrupções, mas ainda com bastante dispersão).

3.2. Agregação semanal

A base diária integrada foi agregada por semana (frequência W), gerando séries semanais de:

- interrupções totais semanais;
- temperatura média semanal;
- precipitação total semanal.

Resumo:

- período semanal: **01/01/2017 a 01/06/2025**;
- interrupções semanais:
 - média: **1701,23 interrupções/semana**;
 - mínimo: 116;
 - máximo: 4914;
- temperatura média semanal:
 - média: **21,58 °C**;
 - faixa: 15,98 °C a 26,77 °C;

- precipitação total semanal:
 - média: **27,75 mm/semana**;
 - faixa: 0,0 mm a 173,4 mm.

Os gráficos semanais mostram que, ao nível semanal, o ruído diário diminui e alguns períodos chuvosos aparecem associados a patamares mais elevados de interrupções, embora a relação não seja perfeitamente linear.

3.3. Agregação mensal

Também foram geradas séries **mensais** de:

- interrupções mensais;
- temperatura média mensal;
- precipitação total mensal.

Período mensal: **01/01/2017 a 01/05/2025** (101 meses).

Resumo:

- interrupções mensais:
 - média: **7411,31 interrupções/mês**;
 - mínimo: 3570;
 - máximo: 15 336;
- temperatura média mensal:
 - média: **21,58 °C**;
 - faixa: 17,63 °C a 24,98 °C;
- precipitação total mensal:
 - média: **120,90 mm/mês**;
 - mínimo: 0,0 mm;
 - máximo: 539,4 mm.

Os gráficos mensais reforçam a sazonalidade do clima (estações chuvosa e seca) e sugerem que períodos de maior chuva tendem a concentrar maior número de interrupções, embora a relação ainda não seja perfeitamente explicada apenas pela precipitação.

4. Séries com média móvel diária

Para tornar mais visíveis os padrões de tendência e sazonalidade, foram calculadas médias móveis de **7 e 14 dias** para:

- número diário de interrupções;
- temperatura média diária;
- precipitação diária.

Colunas criadas:

- `interrupcoes_mm_7` , `interrupcoes_mm_14` ;
- `temperatura_media_mm_7` , `temperatura_media_mm_14` ;
- `precipitacao_total_mm_mm_7` , `precipitacao_total_mm_mm_14` .

Os gráficos correspondentes mostram:

- a **sazonalidade anual** da temperatura;
- ciclos sazonais de **chuva x seca** refletidos na precipitação;
- uma série de interrupções com flutuações mais suaves, onde se tornam mais claros:
 - picos em determinados períodos (possivelmente associados à estação chuvosa);
 - períodos de patamar mais baixo, compatíveis com épocas mais secas.

Essas médias móveis foram usadas apenas para análise descritiva e para construir um baseline de previsão (média dos 7 últimos dias).

5. Previsão do número diário de interrupções (treino x teste temporal)

5.1. Divisão temporal de treino e teste

A série diária de interrupções (integrada com clima) foi dividida de forma **estritamente temporal**, sem embaralhamento:

- **treino:**
 - 2458 dias, até **24/09/2023**;
- **teste:**
 - 615 dias, a partir de **25/09/2023**.

Para cada dia t , as previsões foram geradas utilizando **apenas dados até $t-1$** , evitando vazamento de informação do futuro.

5.2. Modelos de referência (baselines)

Foram avaliados dois modelos simples de referência:

1. Persistência

- Fórmula: $y_{\text{hat}}(t) = y(t-1)$
- A previsão do dia t é o número de interrupções observado no dia anterior.

2. Média móvel de 7 dias

- Fórmula: $y_{\text{hat}}(t) = \text{média dos 7 valores de } y(t-1) \text{ a } y(t-7)$
- A previsão do dia t é a média das interrupções dos 7 dias anteriores.

As métricas usadas foram:

- **MAE** – erro absoluto médio;
- **RMSE** – raiz do erro quadrático médio.

5.3. Resultados numéricos

Persistência ($y[t-1]$)

- **Treino** (2457 observações válidas):
 - MAE $\approx 57,41$ interrupções/dia;
 - RMSE $\approx 91,83$ interrupções/dia.
- **Teste** (615 observações válidas):
 - MAE $\approx 75,12$ interrupções/dia;
 - RMSE $\approx 111,83$ interrupções/dia.

Média móvel de 7 dias

- **Treino** (2457 observações válidas):
 - MAE ≈ **54,96** interrupções/dia;
 - RMSE ≈ **87,05** interrupções/dia.
- **Teste** (615 observações válidas):
 - MAE ≈ **68,44** interrupções/dia;
 - RMSE ≈ **107,23** interrupções/dia.

5.4. Interpretação

- Os erros no conjunto de teste são, como esperado, maiores do que no treino, indicando que os baselines não estão superajustados, mas enfrentam a variabilidade natural da série.
 - O modelo de **média móvel de 7 dias** apresenta desempenho ligeiramente superior ao da persistência, tanto em MAE quanto em RMSE, no treino e no teste.
 - Isso sugere que um simples alisamento das observações recentes já captura parte da dinâmica de curto prazo das interrupções, mas ainda com erros absolutos da ordem de **70 interrupções diárias**, o que indica espaço para modelos mais sofisticados no futuro (por exemplo, modelos de séries temporais ou regressão com variáveis climáticas e de consumo).
-

6. Interrupções x consumo de energia (base mensal)

6.1. Construção do consumo mensal

Para representar o consumo mensal do sistema de distribuição em Brasília, foi utilizada a base regulatória da ANEEL:

- Arquivo: `samp-balanco.csv`.

Filtro aplicado:

- `NumCPFCNPJ = 07522669000192` (Neoenergia Brasília);
- `DscModalidadeBalanco = "Energia Injetada"`;

- `DscFluxoEnergia` = "Disponibilidades" ;
- `DscCctBalanco` = "Energia Injetada Total" ;
- `DscDetalheBalanco` = "Energia Medida (kWh)" .

Após o filtro, os valores de `VlrEnergia` foram agregados por ano e mês, resultando em uma série mensal de **consumo total em kWh**.

Esse consumo foi integrado à base mensal de interrupções + clima, produzindo a base:

- `base_mensal_interrupcoes_clima_consumo.csv` ;
- período: **01/01/2017 a 01/05/2025** (101 meses).

6.2. Estatísticas mensais

- **Interrupções mensais:**
 - média: **7411,31 interrupções/mês**;
 - mínimo: 3570;
 - máximo: 15 336.
- **Consumo mensal total (kWh):**
 - média: **642 324 909 kWh/mês** (aprox. 642 GWh/mês);
 - mínimo: 538 711 696 kWh/mês;
 - máximo: 752 769 315 kWh/mês.

6.3. Correlação interrupções x consumo

A correlação de Pearson entre **interrupções mensais** e **consumo mensal** foi:

- `r ≈ 0,4764` .

Interpretação:

- Correlação **moderada e positiva**: meses com maior consumo tendem a apresentar maior número de interrupções.
- Isso sugere que a carga do sistema (volume de energia entregue) pode estar associada ao estresse da rede e, consequentemente, a uma maior frequência de falhas, ainda que outros fatores (clima, manutenção, eventos específicos) também desempenhem papel relevante.

7. Consumo x temperatura (base mensal)

Na mesma base mensal integrada, foi analisada a relação entre **consumo total de energia** e **temperatura média mensal**.

- temperatura média mensal:
 - média: **21,58 °C**;
 - mínimo: 17,63 °C;
 - máximo: 24,98 °C.

A correlação de Pearson calculada foi:

- consumo mensal (kWh) x temperatura média mensal:
 - $r \approx 0,5550$.

Interpretação:

- Correlação **moderada a forte e positiva**: meses mais quentes tendem a apresentar maior consumo de energia elétrica.
 - Esse comportamento é coerente com o aumento do uso de **refrigeração/condicionamento de ar** em períodos mais quentes, o que incrementa a demanda do sistema.
-

8. Síntese das principais conclusões da segunda entrega

1. Integração robusta de dados

Foi construída uma base diária integrada de interrupções + clima, cobrindo 2017–2025, com:

- 3073 dias;
- temperatura média diária em torno de 21,6 °C;
- precipitação média diária em torno de 4 mm;

- média de aproximadamente 244 interrupções diárias em Brasília.

2. Agregações semanais e mensais

As agregações semanais e mensais de interrupções com temperatura e precipitação deixaram os padrões sazonais mais claros:

- maior concentração de chuva em determinados períodos do ano;
- associação visível entre períodos chuvosos e maior número de interrupções.

3. Relação interrupções x clima (diário)

As correlações diárias foram:

- fraca entre interrupções e temperatura ($r \approx 0,10$);
- moderada entre interrupções e precipitação ($r \approx 0,35$).

Isso indica que a chuva está mais associada a falhas do que a temperatura, mas ainda há forte componente de variabilidade não explicada apenas pelo clima.

4. Médias móveis

A aplicação de médias móveis de 7 e 14 dias suavizou as séries de interrupções e clima, facilitando a visualização de tendências e sazonalidades, sem criar dados artificiais (apenas transformações da própria série).

5. Previsão com treino/teste temporal

A separação treino/teste foi estritamente temporal, com:

- treino até 24/09/2023 (cerca de 80% da série);
- teste a partir de 25/09/2023.

Baselines avaliados:

- persistência ($y[t-1]$);
- média móvel de 7 dias.

A média móvel de 7 dias superou a persistência em MAE e RMSE no conjunto de teste, mas com erros ainda na faixa de 68–75 interrupções diárias, sugerindo que modelos mais complexos podem ser investigados posteriormente.

6. Integração com consumo de energia

A partir do `samp-balanco.csv` da ANEEL, foi derivada uma série de consumo mensal (kWh) da Neoenergia Brasília, integrada à base mensal de interrupções + clima.

- correlação interrupções x consumo mensal $r \approx 0,48$, indicando que meses com maior consumo tendem a ter mais falhas, possivelmente por maior carregamento da rede.

7. Consumo x temperatura

A correlação consumo mensal x temperatura média mensal foi $r \approx 0,56$, sugerindo que o consumo aumenta em meses mais quentes, compatível com maior uso de ar-condicionado e refrigeração.