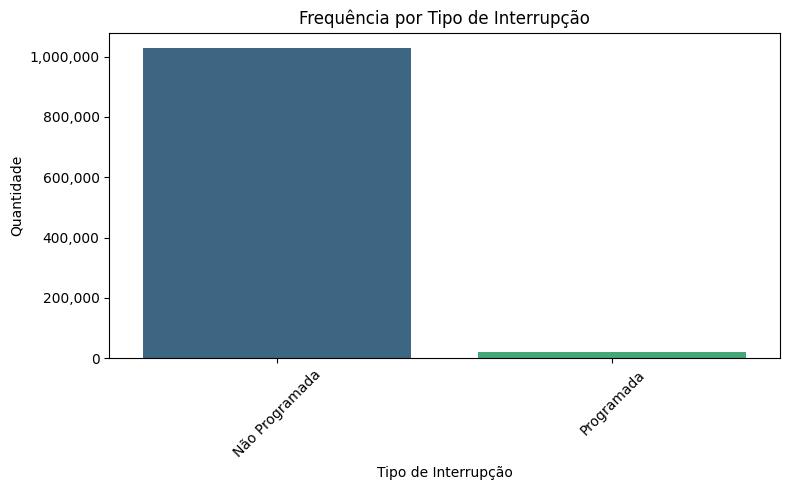
O projeto envolve a análise de dados de interrupções de energia elétrica, reclamações de consumidores e indicadores de qualidade de serviços em municípios. O objetivo é explorar a relação entre diferentes variáveis, como a quantidade de interrupções, o impacto nas populações locais, e como as reclamações se distribuem ao longo do tempo. Além disso, busca-se investigar a correlação entre indicadores de satisfação, qualidade e confiança dos consumidores.

O projeto foi realizado em várias etapas:

1. **Análise das interrupções de energia**: Foram exploradas as interrupções não programadas em diferentes municípios, incluindo a densidade dessas interrupções por número de habitantes. Gráficos como barras horizontais e linhas foram usados para visualizar a distribuição e a frequência dessas interrupções.
2. **Análise das reclamações**: Foi analisada a distribuição das tipologias de reclamações mais frequentes, além do tempo de resposta por agente de atendimento. Gráficos de barras e boxplots foram utilizados para visualizar a distribuição e o tempo médio de resposta.
3. **Correlação entre interrupções e reclamações**: Foi investigada a correlação entre o número de interrupções de energia e o número de reclamações feitas pelos consumidores nos municípios, usando gráficos de dispersão e linhas de tendência.
4. **Análise de qualidade, satisfação e confiança**: Foram explorados os indicadores de qualidade, satisfação e confiança, utilizando correlação e gráficos de dispersão para observar as relações entre esses indicadores.
5. **Modelo preditivo**: Um modelo de regressão foi treinado utilizando dados de interrupções para prever a duração das interrupções com base em características como o nível de tensão e o tipo de interrupção, utilizando o Random Forest.

Relatório de Análise de Interrupções de Energia Elétrica

O objetivo desta análise foi examinar a frequência das interrupções de energia elétrica, categorizadas entre **não programadas** e **programadas**, para entender melhor o impacto de cada tipo de interrupção no fornecimento de energia e possíveis áreas para melhoria.  
 **Frequência por Tipo de Interrupção**

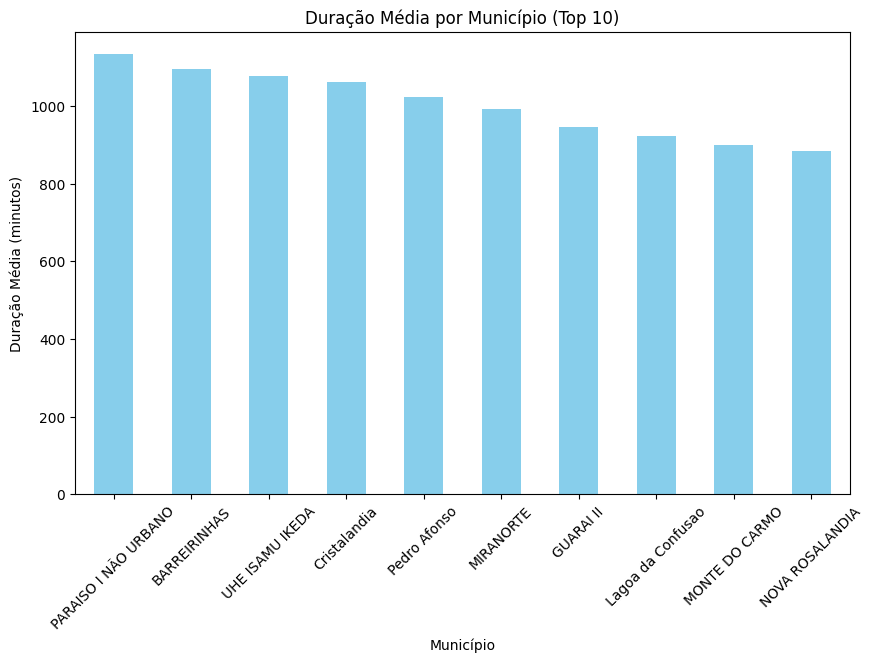
O Gráfico a seguir apresenta a frequência total das interrupções de energia, divididas por tipo.  


A **maioria das interrupções** são do tipo **não programada**, com um número significativamente maior (1.027.477) em comparação com as interrupções **programadas** (21.071).

**Interrupções não programadas** geralmente acontecem de forma inesperada e podem ser causadas por diversos fatores, como falhas técnicas, acidentes, ou eventos climáticos. Esse grande número de interrupções não programadas sugere que há uma instabilidade no fornecimento de energia, que poderia ser um ponto de melhoria para as concessionárias.

**Interrupções programadas** são aquelas previamente agendadas para manutenção ou upgrades do sistema, com um número muito menor, o que é esperado, já que esses eventos são planejados para minimizar o impacto sobre os consumidores.

Análise da Duração Média das Interrupções por Município (Top 10)

Esta análise tem como objetivo avaliar a **duração média das interrupções** de energia elétrica nos **10 municípios com as maiores durações**. Com isso, busca-se entender os padrões de interrupção e identificar as áreas que podem necessitar de maior atenção no que diz respeito à **duração das interrupções**.  
  
O gráfico a seguir apresenta a **duração média das interrupções** (em minutos) para os **10 municípios** com as maiores durações:   
  
A análise da duração média das interrupções de energia elétrica mostrou que alguns municípios enfrentam interrupções muito mais longas do que outros. O município de **PARAISO I NÃO URBANO** teve a maior duração média de interrupção, com 1133,32 minutos, seguido por **BARREIRINHAS** e **UHE ISAMU IKEDA**, ambos com durações acima de 1000 minutos. Esses locais parecem ter maiores dificuldades com a infraestrutura elétrica, o que pode justificar os longos períodos sem energia.

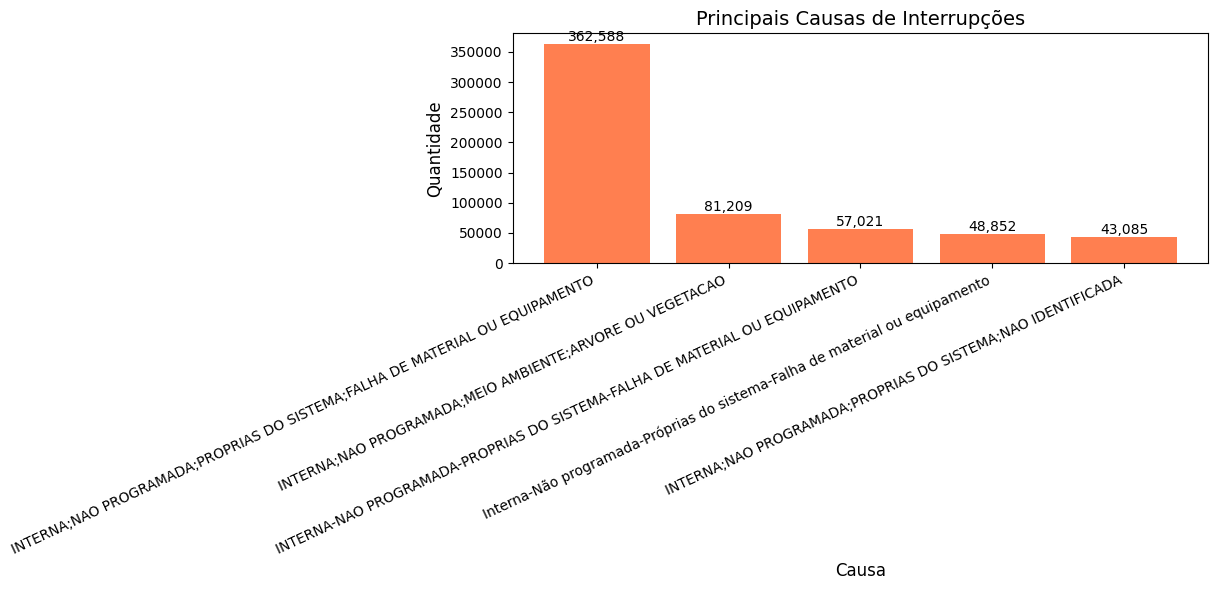
Além desses, outros municípios como **Cristalândia**, **Pedro Afonso** e **MIRANORTE** também apresentaram interrupções longas, acima de 900 minutos. Isso indica que, em algumas regiões, a duração das interrupções pode estar relacionada a fatores como **dificuldade de acesso** ou **infraestrutura mais frágil**, que tornam a manutenção e a recuperação do fornecimento de energia mais lenta.

Por outro lado, **NOVA ROSALANDIA** e **MONTE DO CARMO**, embora com tempos acima de 800 minutos, têm as durações mais curtas entre os 10 municípios mais afetados.

De forma geral, a análise sugere que os municípios com interrupções mais longas podem estar lidando com **problemas de infraestrutura**, **acesso difícil para manutenção** ou até mesmo **falta de modernização** na rede elétrica. Para melhorar a situação, seria importante investir em **melhorias na infraestrutura**, **manutenção preventiva** e **tecnologias de monitoramento**, especialmente nessas áreas, para reduzir o tempo de interrupção e melhorar o serviço de energia elétrica oferecido aos consumidores.

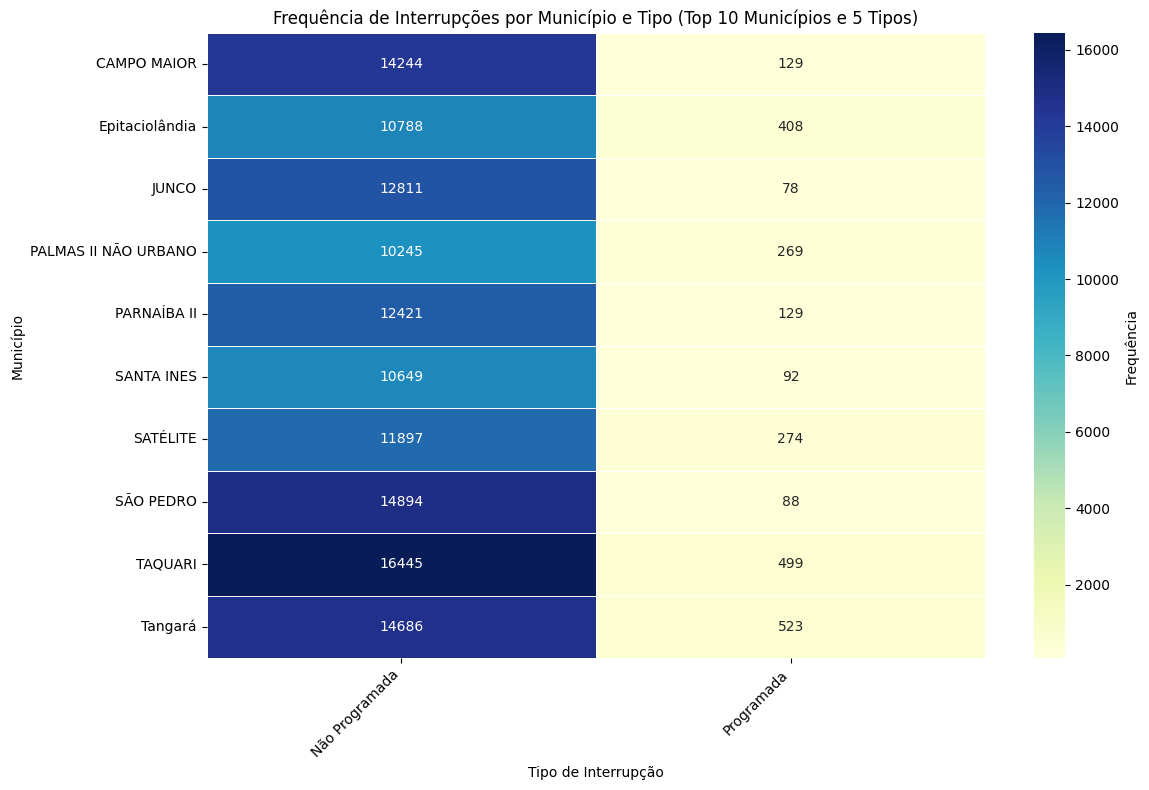
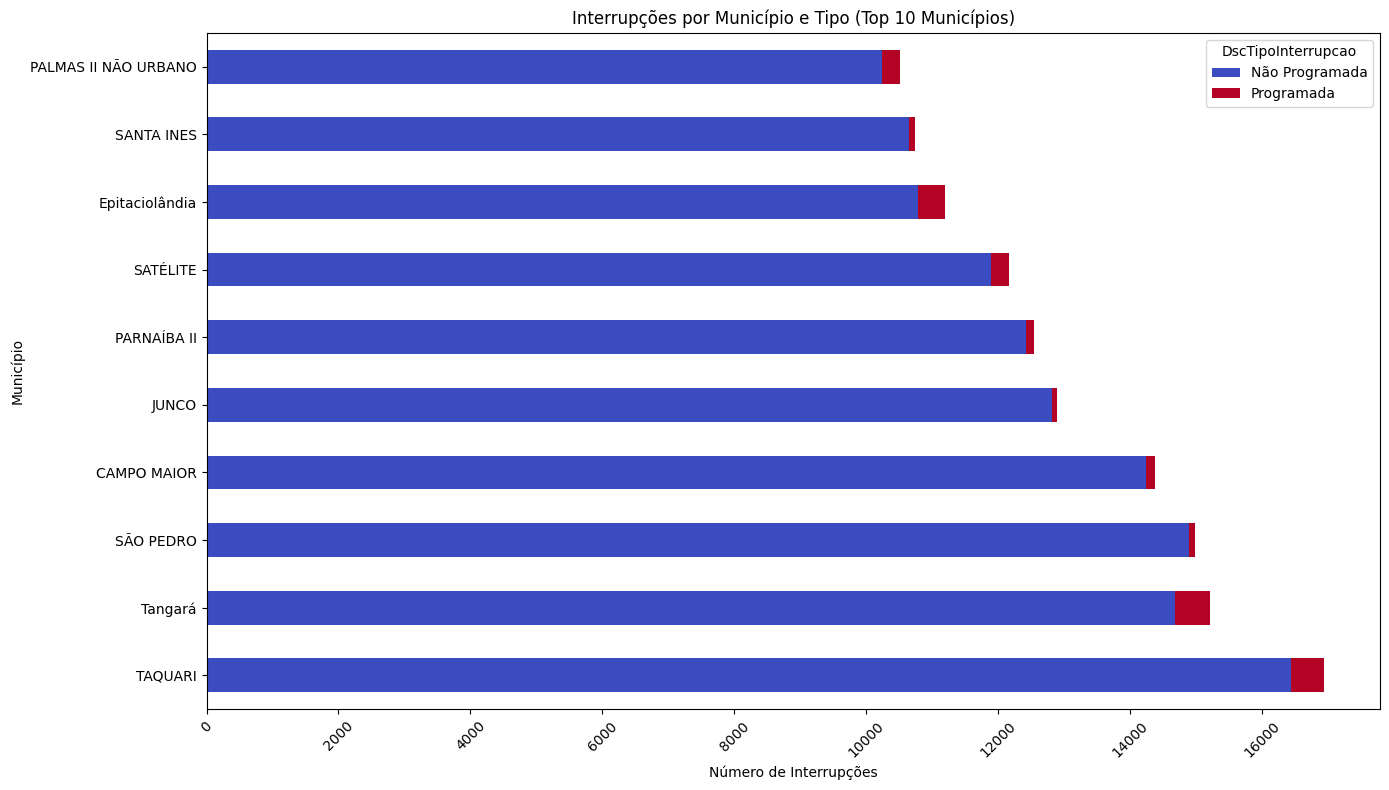
Análise das Principais Causas de Interrupções (Top 5)

O objetivo desta análise foi identificar as principais causas das interrupções no fornecimento de energia elétrica, com foco nas falhas mais frequentes que impactam os consumidores. A partir da categorização das causas das interrupções, foi possível entender as razões mais recorrentes para as falhas no sistema, destacando aspectos como problemas internos no sistema e interferências ambientais.

O gráfico a seguir apresenta **as Principais Causas de Interrupções (Top 5).** Por quantidade de ocorrências.  
  
  
  
A análise das causas das interrupções revelou que as falhas internas no sistema são as responsáveis pela grande maioria das interrupções de energia elétrica. A causa mais comum foi **falha de material ou equipamento**, com um total de 362.588 ocorrências. Esse tipo de falha é frequentemente relacionado à deterioração de componentes da rede elétrica ou problemas nos equipamentos usados para o fornecimento de energia.

Em segundo lugar, a **interferência ambiental**, especialmente com **árvores ou vegetação**, também é uma causa significativa de interrupções, com 81.209 registros. Esse problema é comum em áreas onde as redes elétricas estão localizadas em regiões de grande vegetação ou onde o manejo das árvores não é feito adequadamente.

Outras causas incluem variações na descrição de falhas **internas e não programadas**, com falhas de material ou equipamento, que somaram 57.021 e 48.852 ocorrências, respectivamente. A falta de clareza na identificação exata das falhas, como as chamadas **"não identificadas"** (43.085 interrupções), também contribui para uma parte das interrupções.  
  
Em resumo, as falhas **internas no sistema** e **problemas com a vegetação** são os principais fatores responsáveis pelas interrupções, e a variação na descrição de falhas sugere que é necessário aprimorar o **processo de identificação** e **documentação** dos eventos para uma gestão mais eficaz e para a implementação de **ações preventivas**.  
  
Distribuição e Impacto por Município e Tipo

O objetivo desta análise foi compreender o número total de interrupções no fornecimento de energia elétrica, separando-as por tipo (programada e não programada) para os 10 municípios com o maior número de ocorrências. Ao identificar esses municípios e os tipos de interrupções mais prevalentes, a análise busca fornecer insights sobre os padrões de falhas e planejar ações específicas para mitigar problemas recorrentes, além de otimizar a gestão e a alocação de recursos para manutenção.  
  
O gráfico a seguir apresenta Distribuição e Impacto por Município e Tipo:  
   
  
A análise dos dados de interrupções de energia revelou os 10 municípios com o maior número de registros, categorizados entre interrupções programadas e não programadas. Os municípios de **Taquari**, **Tangará** e **São Pedro** se destacam como os mais afetados por interrupções, com Taquari apresentando um total de 16.944 ocorrências. A maioria das interrupções nesses municípios são de natureza **não programada**, o que sugere que falhas imprevistas no sistema têm um impacto significativo. A interrupção programada, geralmente associada a manutenção planejada, também representa uma porção das falhas, mas de forma bem menos expressiva.

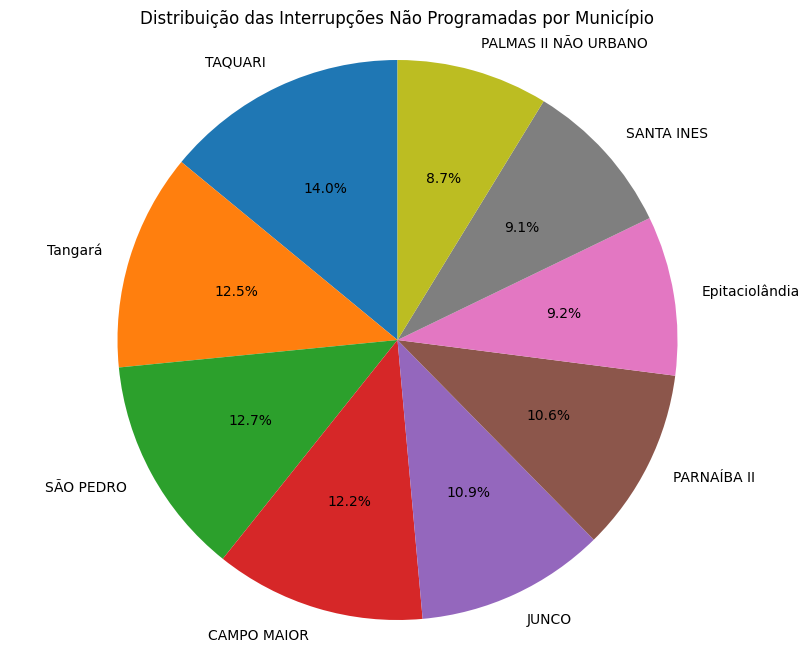
Esses dados fornecem uma visão detalhada das localidades com maior necessidade de atenção, permitindo que a gestão da energia elétrica seja mais eficaz, focando na resolução de problemas imprevistos e na manutenção preventiva, especialmente nas áreas com altas taxas de interrupções não programadas.

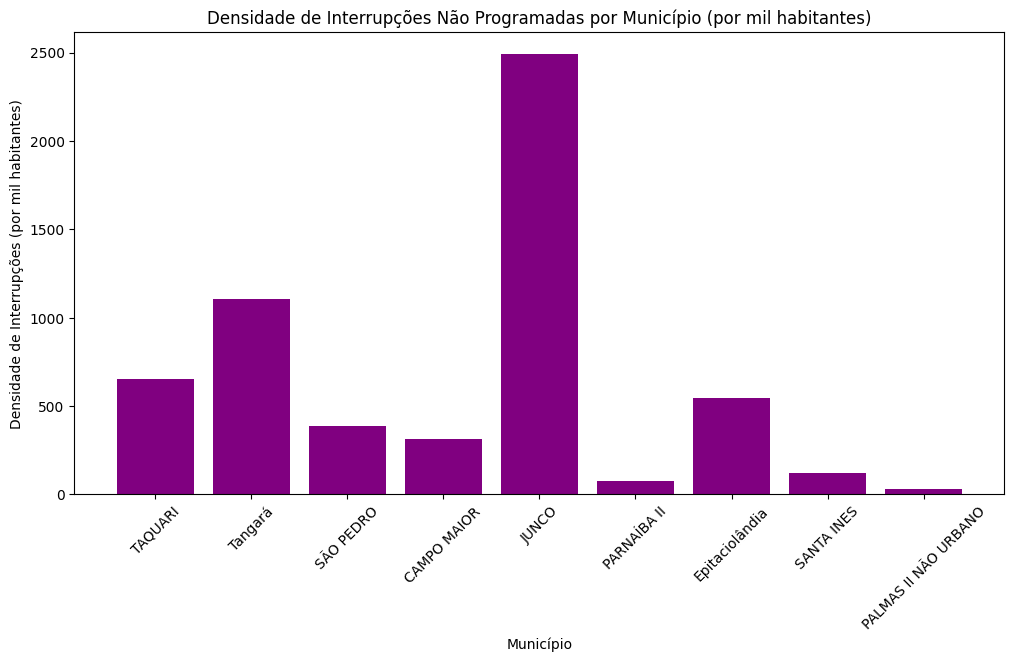
Análise de Interrupções Não Programadas por Município e sua Densidade em Função da População

A análise focou na comparação das interrupções de energia elétrica não programadas nos municípios selecionados, levando em consideração a densidade dessas interrupções em relação à população local. Para isso, foi calculada a densidade de interrupções não programadas por 1.000 habitantes, oferecendo uma visão mais precisa/real sobre a incidência dessas falhas em cada município.

Ao examinar os dados, observamos que municípios com populações menores, como **JUNCO** e **PARNAÍBA II**, apresentam uma densidade de interrupções significativamente mais alta, indicando que, apesar de possuírem menos habitantes, enfrentam um maior número de interrupções por habitante. Em contraste, municípios com populações maiores, como **PALMAS II NÃO URBANO** e **PARNAÍBA II**, têm uma densidade de interrupções mais baixa, o que pode ser resultado de um maior número de consumidores diluindo as interrupções.

Através do gráfico de pizza, podemos visualizar a distribuição das interrupções entre os municípios e entender como os eventos de interrupção não programada são distribuídos entre eles, enquanto o gráfico de barras ajuda a comparar a densidade das interrupções não programadas entre as diferentes localidades.



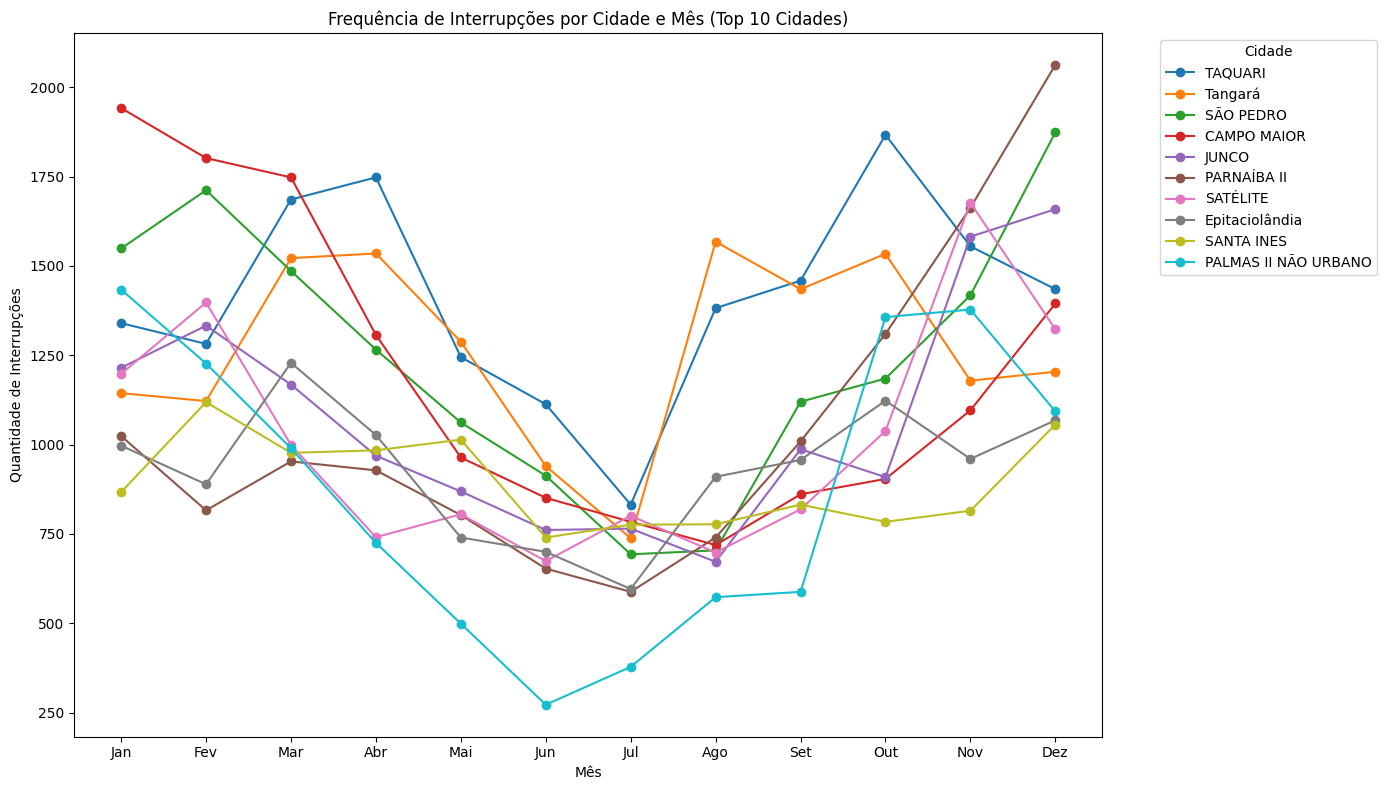


Análise da Frequência Mensal de Interrupções nas 10 Cidades com Mais Ocorrências

Esta análise busca entender a distribuição e variação mensal das interrupções de energia elétrica nas 10 cidades com mais ocorrências de falhas não programadas. O objetivo é observar os padrões de interrupção ao longo dos meses e identificar picos ou quedas significativas.

Ao examinar as frequências mensais das interrupções, algumas tendências se destacam:

* **TAQUARI** apresenta um padrão relativamente constante, com picos mais altos nos meses de 4 (1748) e 10 (1867), o que sugere que, durante o segundo semestre, o município enfrenta um número elevado de falhas.
* **Tangará** também segue uma tendência semelhante, com o maior número de interrupções nos meses de 8 (1568) e 10 (1534). Este comportamento pode indicar uma sazonalidade ou fatores específicos de clima ou demanda que aumentam as falhas nesses meses.
* **SÃO PEDRO** mostra uma variação significativa, com um pico nas interrupções em dezembro (1874), o que pode estar relacionado a eventos climáticos ou condições de rede mais desafiadoras no final do ano.
* **JUNCO** apresenta picos no início (1º e 2º meses) e no final (11º e 12º meses), o que pode sugerir uma maior ocorrência de falhas relacionadas a condições climáticas extremas ou manutenção mais intensiva nesses períodos.
* **PARNAÍBA II** possui uma tendência crescente de interrupções no segundo semestre, especialmente entre os meses 7 e 12, onde os números de interrupção aumentam drasticamente (de 588 para 2063), o que indica possíveis problemas recorrentes na rede durante o final do ano.

****

O gráfico de linhas mensal revela os meses mais críticos para interrupções, enquanto a análise de cada cidade permite a identificação de padrões que podem ser usados para priorizar ações de manutenção e melhorias nas infraestruturas. A variação nas interrupções também pode refletir condições externas, como eventos climáticos, demandas de energia ou falhas recorrentes de equipamento. Essa análise é crucial para otimizar o planejamento de reparos e a alocação de recursos para reduzir a frequência e impacto das interrupções.  
  
Insights

**Tipo de interrupção mais frequente:** As **interrupções não programadas** são as mais comuns, com um total de **1.027.477 ocorrências**. Este tipo de interrupção, que geralmente ocorre devido a falhas inesperadas no sistema, apresenta um desafio maior para as concessionárias em termos de planejamento e resposta rápida.

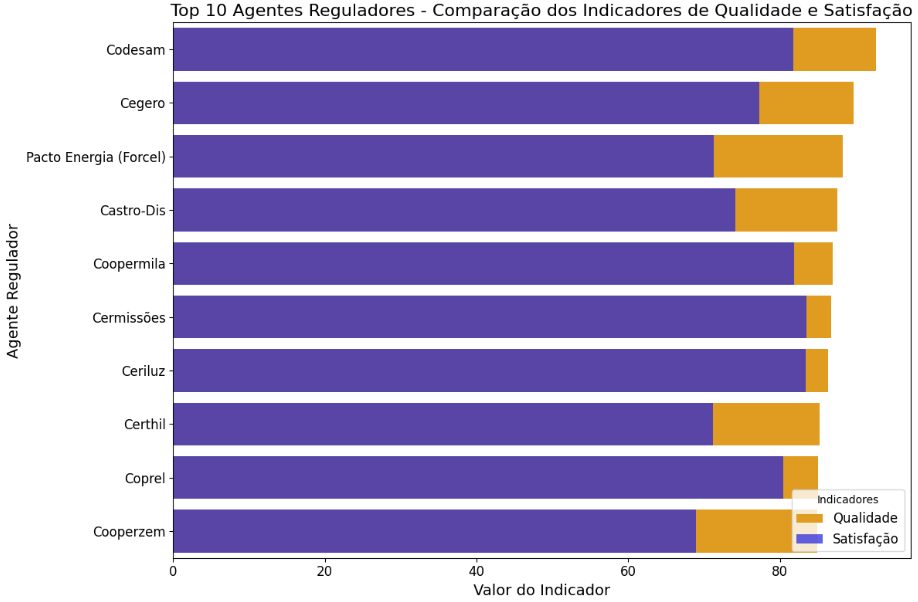
**Município com maior duração média de interrupção:** O município de **PARAISO I NÃO URBANO** apresenta a maior duração média de interrupção, com **1.133,32 minutos** por incidente. Isso sugere que, além da frequência, as interrupções nessas áreas são mais prolongadas, possivelmente devido a desafios logísticos ou dificuldades no restabelecimento do serviço.

**Causa mais comum das interrupções:** A principal causa das interrupções é **falha de material ou equipamento** no sistema, com **362.588 ocorrências**. Esta falha interna, muitas vezes relacionada a falhas em componentes da rede elétrica, destaca a necessidade de uma manutenção preventiva e da substituição de equipamentos mais suscetíveis a falhas.

Indicadores de Qualidade e Satisfação por Agente Regulador

A análise dos indicadores de qualidade e satisfação para os diferentes agentes reguladores revela um panorama de variação no desempenho das concessionárias de energia.

1. **Média de Indicadores de Qualidade e Satisfação por Agente Regulador**:
   * A **qualidade** dos serviços de energia, medida pelo indicador de qualidade, varia significativamente entre as concessionárias. O valor médio desse indicador para as empresas analisadas é de **57,61**, com algumas concessionárias apresentando números bem acima da média, como a **Pacto Energia (Forcel)** com **88,32** e a **RGE** com **76,39**.
   * Já o **indicador de satisfação** do cliente, que mede a percepção dos consumidores sobre a qualidade do serviço, tem uma média de **58,93**. Algumas concessionárias conseguem manter a satisfação dos clientes mais elevada, como a **CEA Equatorial** e a **Sulgipe**, que atingem **70,60**.
2. **Top 10 Agentes Reguladores por Qualidade e Satisfação**:
   * As concessionárias com os melhores desempenhos em qualidade e satisfação incluem:
     + **Codesam**: com a melhor pontuação em qualidade (92,62) e satisfação (81,72).
     + **Cegero** e **Pacto Energia (Forcel)** também se destacam, alcançando bons índices em ambas as categorias.
     + Outras concessionárias como **Coopermila**, **Cermissões**, e **Ceriluz** também estão entre as melhores, com índices de qualidade acima de **86** e altos índices de satisfação, refletindo uma entrega consistente de serviços de qualidade e uma boa percepção por parte dos consumidores.



**Resumo da Análise**: A análise dos dados revela que, embora haja uma variação considerável entre os agentes reguladores, algumas empresas têm se destacado por manter altos níveis de qualidade no fornecimento de energia e satisfação dos consumidores. Essas concessionárias, como **Codesam** e **Cegero**, mostram um bom equilíbrio entre qualidade do serviço e satisfação, o que sugere boas práticas de gestão e atendimento ao cliente.

Análise do Percentual de Reclamações por Município

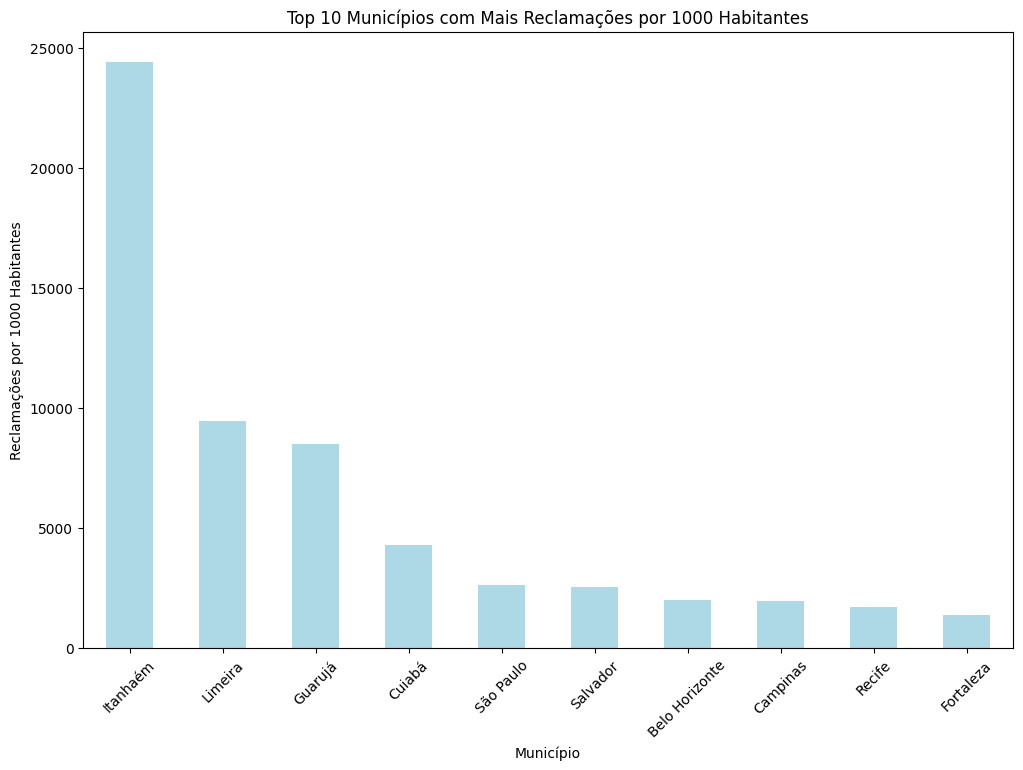
Nesta análise, avaliamos as reclamações recebidas nos 10 municípios com mais reclamações de consumidores. O objetivo é entender melhor o impacto das reclamações no contexto de cada cidade, levando em consideração o número de habitantes e o total de reclamações recebidas.

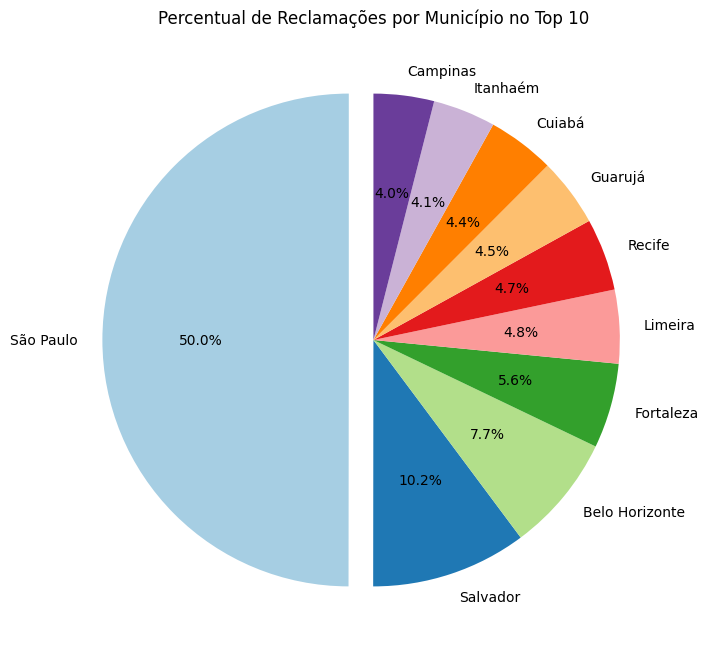
**Principais Resultados:**

* **Município com maior percentual de reclamações**: São Paulo se destaca, com **49,98%** das reclamações do Top 10, o que pode ser um reflexo tanto do tamanho da cidade quanto de um volume maior de interações com os serviços de energia.
* **Outros municípios**: Cidades como **Salvador** (10,21%) e **Belo Horizonte** (7,69%) também aparecem entre os mais impactados. Contudo, o percentual de reclamações para essas cidades diminui consideravelmente em comparação com São Paulo.

**Indicadores adicionais:**

* **Reclamações por 1000 habitantes**: Ao observar as reclamações per capita, percebemos que algumas cidades com menor população, como **Itanhaém** e **Limeira**, possuem uma maior densidade de reclamações por habitante. Isso indica uma maior insatisfação relativa em comparação com as grandes cidades.





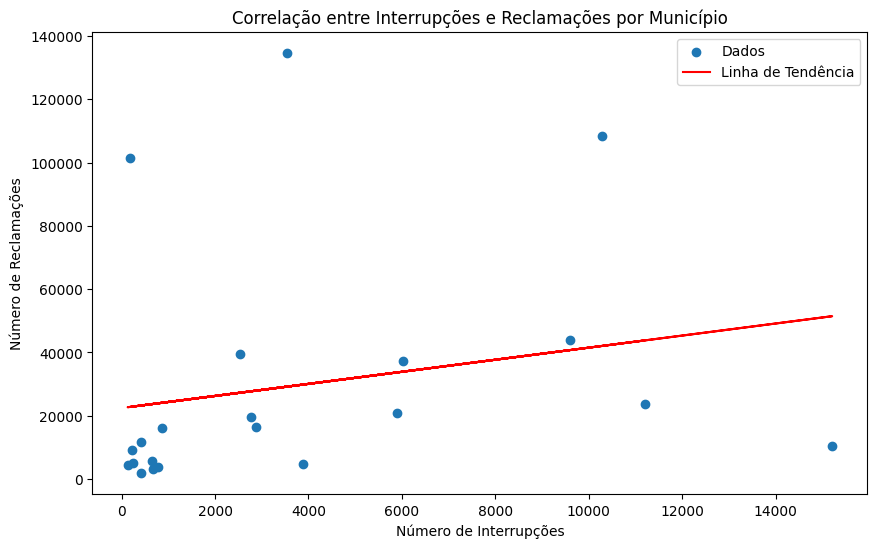
A cidade de **São Paulo** lidera em termos absolutos e percentuais de reclamações, com uma proporção significativa do total de reclamações recebidas. Enquanto isso, cidades com menor população, como **Itanhaém**, têm um índice elevado de reclamações per capita, sugerindo uma percepção mais forte de insatisfação. Essa análise fornece uma visão detalhada das áreas onde as concessionárias de energia podem precisar focar esforços para melhorar o atendimento e a infraestrutura.

#### Correlação entre Interrupções e Reclamações por Município

A análise de correlação entre **interrupções de energia** e **reclamações** por município revelou um valor de **0.22** para a correlação, o que indica uma relação **positiva fraca** entre os dois fatores. Ou seja, há uma tendência de que, à medida que o número de interrupções de energia aumenta, o número de reclamações também tende a aumentar, mas a relação não é forte.

**Correlação de 0.22**: Esse valor sugere que, embora haja uma associação entre interrupções e reclamações, ela não é muito forte. Pode indicar que outros fatores, além das interrupções, também influenciam o número de reclamações, como a qualidade do atendimento ao cliente, a percepção do consumidor ou a frequência e a duração das interrupções.

O **gráfico de dispersão** mostra como as interrupções e as reclamações se distribuem para cada município. A **linha de tendência** em vermelho sugere que, de forma geral, municípios com mais interrupções tendem a ter mais reclamações, mas a dispersão dos pontos indica que a relação não é perfeitamente linear.



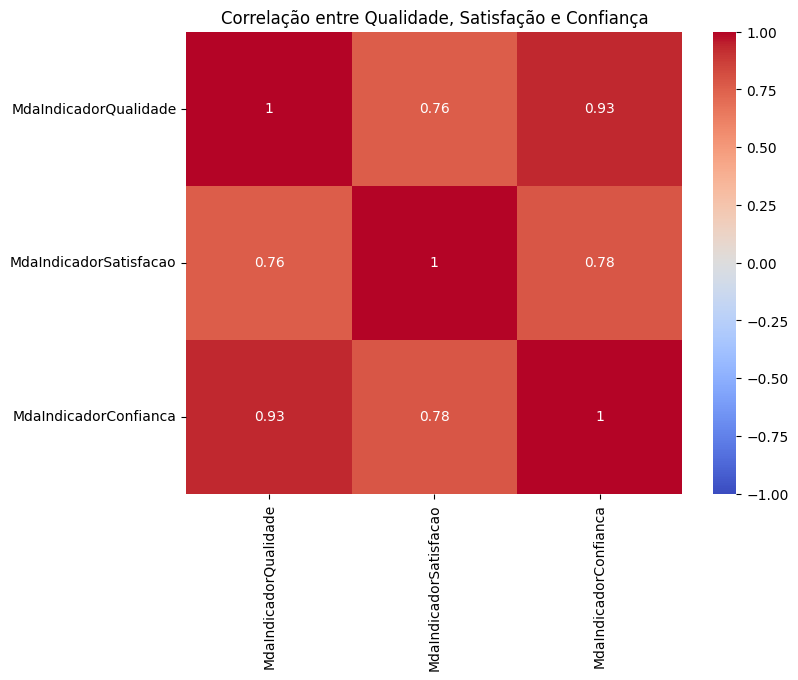
Apesar da correlação fraca, a análise ainda oferece insights importantes. Uma maior frequência de interrupções pode ser um indicativo de um aumento nas reclamações, mas outros fatores também devem ser considerados para uma análise mais completa e para tomar decisões estratégicas sobre como melhorar o serviço de energia e a satisfação dos consumidores.

#### Correlação entre Qualidade, Satisfação e Confiança

A análise de correlação entre os indicadores **Qualidade**, **Satisfação** e **Confiança** revela os seguintes resultados:

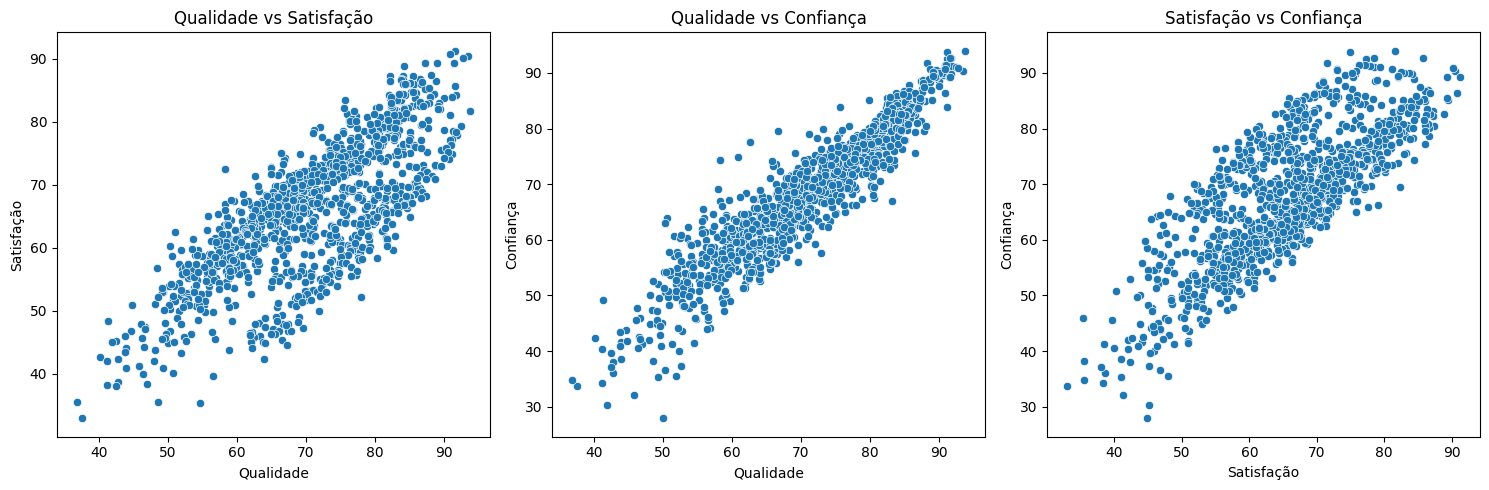
* **Qualidade vs Satisfação**: Correlação de **0.76**, o que indica uma relação bastante forte entre a qualidade do serviço e a satisfação do cliente. Isso sugere que, quanto melhor a qualidade do serviço, maior é a satisfação dos consumidores.
* **Qualidade vs Confiança**: Correlação de **0.93**, indicando uma relação muito forte entre a qualidade e a confiança. Isso sugere que consumidores mais satisfeitos com a qualidade do serviço tendem a confiar mais na empresa fornecedora de energia.
* **Satisfação vs Confiança**: Correlação de **0.78**, indicando uma relação forte entre satisfação e confiança. Isso reflete o fato de que consumidores satisfeitos geralmente têm mais confiança na empresa, o que pode levar a uma lealdade maior.

**Mapa de calor da correlação**: O mapa de calor apresenta as correlações entre os três indicadores, destacando a forte correlação entre **Qualidade** e **Confiança** (0.93), bem como entre **Qualidade** e **Satisfação** (0.76).



**Gráficos de dispersão**:

* **Qualidade vs Satisfação**: A dispersão dos pontos mostra que há uma tendência clara de que consumidores satisfeitos também percebem uma boa qualidade.
* **Qualidade vs Confiança**: A dispersão aqui também indica que a confiança aumenta com a melhoria da qualidade do serviço.
* **Satisfação vs Confiança**: Similarmente, consumidores satisfeitos parecem confiar mais na empresa de energia.



**Relação forte entre qualidade, satisfação e confiança**: A análise sugere que, para melhorar a satisfação e a confiança dos consumidores, a qualidade do serviço precisa ser uma prioridade. Investir na melhoria da qualidade dos serviços de energia provavelmente terá impactos positivos em outros indicadores de satisfação do cliente.

A técnica de **inferência aplicada** nas suas análises está centrada na utilização de modelos preditivos para entender e estimar comportamentos de dados históricos, com o objetivo de fornecer insights sobre as interrupções de energia elétrica e suas características. O processo de inferência pode ser explicado de forma mais detalhada em relação aos seus objetivos:

O objetivo é entender e prever a **duração das interrupções de energia elétrica** com base em variáveis como:

* **NumNivelTensao** (nível de tensão da rede de distribuição),
* **NumConsumidorConjunto** (número de consumidores afetados),
* **NumAno** (ano de ocorrência da interrupção),
* **DscTipoInterrupcao** (tipo da interrupção, como programada ou não programada).

### Técnica de Inferência Aplicada:

1. **Pré-processamento dos Dados**:
   * **Conversão de datas**: A primeira etapa no processo foi converter as colunas de datas (DatInicioInterrupcao e DatFimInterrupcao) para um formato adequado, o que possibilita o cálculo da **duração das interrupções**.
   * **Cálculo de duração**: A duração da interrupção foi calculada como a diferença entre as duas datas, expressa em minutos. Este é o valor que o modelo tentará prever.
2. **Engenharia de Variáveis**:
   * **One-hot encoding**: Para tratar variáveis categóricas (como o tipo de interrupção), foi utilizado o **one-hot encoding**. Essa técnica transforma variáveis categóricas em colunas binárias, facilitando a aplicação de modelos de aprendizado de máquina.
   * **Criação de novas features**: Colunas derivadas do tipo de interrupção foram incluídas como features adicionais, ajudando o modelo a capturar diferentes padrões relacionados ao tipo de evento.
3. **Divisão de Dados**:
   * O dataset foi dividido em conjuntos de **treinamento** e **teste** usando a função train\_test\_split. Isso garante que o modelo seja treinado em uma parte dos dados e avaliado em outra, evitando que o modelo se ajuste excessivamente aos dados de treinamento (overfitting).
   * **X** contém as variáveis independentes (features), e **y** contém a variável dependente (duração da interrupção).
4. **Modelo Preditivo**:
   * Foi escolhido o **Random Forest Regressor**, que é um modelo de aprendizado supervisionado robusto. Este modelo funciona bem com dados de alta dimensionalidade (como seu conjunto de dados com várias variáveis) e pode capturar interações complexas entre as variáveis.
   * O modelo foi treinado nos dados de treinamento (X\_train e y\_train), aprendendo a relação entre as características dos dados e a duração das interrupções.
5. **Avaliação do Modelo**:
   * Após treinar o modelo, o desempenho foi avaliado usando o **Erro Médio Absoluto (MAE)**, que mede a diferença média entre as previsões do modelo e os valores reais. No seu caso, o erro médio absoluto foi de **374,08 minutos**, ou seja, em média, o modelo está errando por cerca de 374 minutos (aproximadamente 6 horas) ao prever a duração das interrupções.

### Interpretação dos Resultados:

* O modelo de **Random Forest** foi capaz de capturar a relação entre as variáveis selecionadas e a duração das interrupções, embora o erro médio ainda seja relativamente alto. Isso pode ser atribuído à complexidade dos dados e a possível necessidade de mais features ou ajustamento do modelo.
* A técnica de **inferência** aqui aplicada envolve estimar como diferentes fatores (como o tipo de interrupção e o número de consumidores afetados) influenciam a duração das interrupções. Com isso, a análise oferece uma previsão quantitativa da duração média das interrupções com base em dados históricos.

A análise abordou uma dor crítica no setor de energia elétrica: a instabilidade e a imprevisibilidade das interrupções, que geram grandes desconfortos para os consumidores e um volume expressivo de reclamações. A incapacidade de prever e gerenciar essas interrupções de forma eficaz frequentemente resulta em baixa satisfação, aumento das queixas e perda de confiança por parte dos consumidores. Com base nas análises de dados, foi possível identificar padrões de interrupção, entender a relação entre qualidade do serviço e satisfação do consumidor, e desenvolver um modelo preditivo para estimar a duração das falhas.

Essa abordagem ajuda a resolver a dor do mercado relacionada à falta de previsibilidade e planejamento para as interrupções no fornecimento de energia, permitindo que as concessionárias programem estratégias mais eficazes de gerenciamento de crises e comunicação com os consumidores. Com um modelo que antecipa falhas e define melhor os períodos de manutenção, as empresas podem reduzir os impactos das interrupções, melhorar a confiança do cliente e, por consequência, diminuir o número de reclamações. Assim, a análise não só ajuda a mitigar o problema das interrupções inesperadas, mas também fortalece a relação entre as concessionárias e seus consumidores, resolvendo uma das maiores dores do setor.