



## AVALIAÇÃO DE MODELOS DE PREVISÃO DE SÉRIES TEMPORAIS, APLICADO ÀS DESPESAS DE UMA PREFEITURA NO INTERIOR DO CEARÁ

EVALUATION OF TIME SERIES FORECASTING MODELS APPLIED TO MUNICIPAL EXPENSES IN A CITY IN CEARÁ

EVALUACIÓN DE MODELOS DE PRONÓSTICO DE SERIES TEMPORALES APLICADOS A LOS GASTOS MUNICIPALES EN UNA CIUDAD DEL CEARÁ

José Wivo Gomes<sup>1</sup>, João Evangelista Dantas dos Santos<sup>2</sup>, & Rodolfo de Sousa Santos<sup>3</sup>

<sup>1 2 3</sup> Universidade Regional do Cariri

<sup>1</sup>wivo.gomes@urca.br <sup>2</sup>joao.evangelista@urca.br <sup>3</sup>rodolfo.sousa@urca.br

### ARTIGO INFO.

Recebido: 14.01.2025

Aprovado: 12.03.2025

Disponibilizado: 25.04.2025

**PALAVRAS-CHAVE:** Machine Learning; Previsão; Modelos estatísticos; Planejamento de recursos.

**KEYWORDS:** Machine Learning; Forecasting; Statistical Models; Resource Planning.

**PALABRAS CLAVE:** Aprendizaje Automático; Pronóstico; Modelos estadísticos; Planificación de Recursos.

\*Autor Correspondente: Gomes, J. W.

### RESUMO

Para atender às complexas necessidades de planejamento financeiro de municípios com alta volatilidade nos gastos, como os do interior do Ceará, é necessário possuir modelos preditivos capazes de projetar despesas e otimizar recursos de maneira precisa. O comparativo entre os modelos Auto ETS e Auto ARIMA no desempenho de previsões de despesas foi o questionamento central deste estudo, que verificou os resultados usando dados históricos de 2018 a 2023. As métricas de avaliação correspondem a MAE, RMSE, MSE, MAPE,  $R^2$  e correlação, associação a análises qualitativa de modelos. Os resultados, nesse sentido, apontam a melhor performance do Auto ETS em erros absolutos, com MAE de 21.510.499,88 e RMSE de 27.275.560,60 e explicação da variância com  $R^2$  de 0,9139. Enquanto isso, o Auto ARIMA apresenta a melhor precisão relativa, com MAPE = 23,63%. No entanto, os dois modelos têm correlação extremamente elevada, acima de 0,97, entre os valores reais e previstos. Ao testar as previsões, os modelos capturaram os padrões históricos, demonstrando alta aderência aos valores reais. Para o intervalo de previsão de 2024 a 2025, as estimativas mantiveram os padrões e variações sazonais em conformidade com o passado; o Auto ETS suavizou os picos mais extremos e se mostrou mais confiável; ao mesmo tempo que o Auto ARIMA lidou melhor com as pequenas variações contribuindo para o planejamento financeiro da cidade.

### ABSTRACT

To meet the complex financial planning needs of municipalities with high spending volatility, such as those in the interior of Ceará, it is necessary to have predictive models capable of accurately projecting expenses and optimizing resources. The comparison between the Auto ETS and Auto ARIMA models in the performance of expense forecasts was the central question of this study, which

verified the results using historical data from 2018 to 2023. The evaluation metrics correspond to MAE, RMSE, MSE, MAPE,  $R^2$  and correlation, association with qualitative analysis of models. The results, in this sense, indicate the best performance of Auto ETS in absolute errors, with MAE of 21,510,499.88 and RMSE of 27,275,560.60 and explanation of variance with  $R^2$  of 0.9139 ENREDO NDO. Meanwhile, Auto ARIMA presents the best relative accuracy, with MAPE = 23.63%. However, the two models have an extremely high correlation, above 0.97, between the actual and predicted values. PLOT. What is more, when testing the forecasts in 2023, the models capture historical patterns, demonstrating high adherence to the actual values. For the forecast range from 2024 to 2025, the estimates-maintained patterns and seasonal variations in line with the past; Auto ETS smoothed out the most extreme peaks and proved to be more reliable; at the same time, Auto ARIMA handled better the small variations contributing to the city's financial planning.

### RESUMEN

Para atender las complejas necesidades de planificación financiera de municipios con alta volatilidad de gastos, como los del interior de Ceará, es necesario contar con modelos predictivos capaces de proyectar gastos y optimizar recursos con precisión. La comparación entre los modelos Auto ETS y Auto ARIMA en el desempeño de las previsiones de gastos fue la cuestión central de este estudio, que verificó los resultados utilizando datos históricos de 2018 a 2023. Las métricas de evaluación corresponden a MAE, RMSE, MSE, MAPE,  $R^2$  y correlación, asociación con el análisis cualitativo de modelos. Los resultados, en este sentido, apuntan al mejor desempeño de Auto ETS en errores absolutos, con MAE de 21.510.499,88 y RMSE de 27.275.560,60 y explicación de la varianza con  $R^2$  de 0,9139 PLOT NDO. Mientras tanto, Auto ARIMA presenta la mejor precisión relativa, con MAPE = 23,63%. Sin embargo, ambos modelos tienen una correlación extremadamente alta, superior a 0,97, entre los valores reales y los previstos. TRAMA NDO. Es más, al probar los pronósticos en 2023, los modelos capturan patrones históricos, demostrando una alta adherencia a los valores reales. Para el rango de pronóstico 2024 a 2025, las estimaciones mantuvieron patrones estacionales y variaciones acordes con el pasado; El ETS automático suavizó los picos más extremos y demostró ser más confiable; Al mismo tiempo, Auto ARIMA gestionó mejor las pequeñas variaciones, contribuyendo a la planificación financiera de la ciudad.

## INTRODUÇÃO

A análise preditiva se estabeleceu como um instrumento crucial na administração pública, possibilitando a previsão de cenários e a melhoria do planejamento orçamentário. No âmbito das administrações municipais, a previsão de despesas empenhadas é fundamental para a eficácia da administração financeira, assegurando maior controle sobre os fundos e auxiliando na tomada de decisões precisas. De acordo com Hyndman e Athanasopoulos (2021), as séries temporais têm um papel crucial em questões de previsão, sendo frequentemente empregadas em cenários financeiros e operacionais para a avaliação de tendências e padrões sazonais.

Neste estudo, sugere-se o uso de métodos de aprendizado de máquina, particularmente modelos de séries temporais, para antecipar as despesas empenhadas da prefeitura de Juazeiro do Norte entre os anos de 2018 e 2023. Métodos como o Auto ETS têm mostrado maior efetividade em relação a abordagens convencionais, como o Auto ARIMA, em situações de alta variabilidade e sazonalidade (Hyndman & Khandakar, 2008). Pesquisas recentes indicam que os modelos ETS, especialmente nas configurações aditivas e multiplicativas (AAA e MMM), são capazes de representar com mais exatidão as dinâmicas complexas de séries temporais (Holt, 2004; Winters, 1960).

A análise dos modelos preditivos será baseada em indicadores amplamente utilizados, tais como MSE, MAE, RMSE, MAPE,  $R^2$  e correlação, assegurando a solidez e a fiabilidade dos resultados alcançados (Makridakis et al., 2018). Essa metodologia tem como objetivo não só confirmar a eficácia dos modelos, mas também proporcionar um alicerce para comparações com outras técnicas já existentes.

Assim, esta pesquisa auxilia na implementação prática de técnicas de previsão no âmbito da administração pública, ressaltando a capacidade da ciência de dados e da engenharia de produção como instrumentos estratégicos para aprimorar o planejamento financeiro e a administração de recursos públicos.

## METODOLOGIA

### Classificação da Pesquisa

A metodologia do estudo é caracterizada como sendo uma pesquisa aplicada, com o intuito de gerar conhecimento para fins práticos. Dentro deste trabalho, o objetivo é atender às necessidades locais, mais especificamente da cidade de Juazeiro do Norte, e buscar maneiras que possam ajudar a melhorar a gestão dos recursos públicos e seus gastos.

Do ponto de vista da abordagem, esta pesquisa é categorizada como uma pesquisa quantitativa, pois pressupõe que tudo pode ser mensurado, fazendo uso de recursos estatísticos, como porcentagem, média, moda, mediana, desvio padrão, coeficiente de correlação e análise de regressão. Dessa forma, traduz opiniões e informações em números, que serão classificados e organizados para análise. A abordagem quantitativa é uma forma de testar teorias objetivas, onde as relações entre variáveis são examinadas.

Essas variáveis podem, no entanto, ser aplicadas em instrumentos que permitam sua quantificação, de modo que os números resultantes possam ser submetidos a análises estatísticas. O relatório final segue um esquema estruturado, incluindo introdução, revisão de literatura e teoria, metodologia, achados e discussão, respectivamente (Creswell, 2014).

## Coleta de Dados

Os dados financeiros da Prefeitura de Juazeiro do Norte (2018-2023) foram obtidos no Portal da Transparência, o que garante a veracidade dos resultados obtidos (Silva et al., 2020). Este trabalho utilizou Python (versão 3.10.12) com bibliotecas eficientes, como Pandas e Numpy, para manipulação e limpeza dos dados, além de Matplotlib e Seaborn para visualização desses dados.

## Etapas de Filtragem e Processamento

- **Portal da Transparência:** Os dados foram coletados diretamente do portal da transparência, onde foram aplicados filtros para restringir a amostra aos dados de despesas no período de 2018 a 2023, que era o intervalo estudado. Escolheu-se esse período para obter uma amostra do tempo mais abrangente e relevante, contendo todos os anos completos. Outros filtros aplicados foram a escolha apenas das despesas públicas de sua natureza orçamentária contendo os gastos em infraestrutura e custeio da máquina pública.
- **Python:** A limpeza e organização dos dados compreenderam a eliminação de valores faltantes, além do *design* e reestruturação de categorias e datas, garantindo que os dados estivessem limpos e aptos para análise em série temporal.

## Preparação dos Dados

No caso das despesas brutas da Cidade de Juazeiro do Norte, os dados em primeiro lugar foram coletados e, posteriormente transformados através dos Pandas e Numpy, bibliotecas fortemente recomendadas ao trabalhar com data em série cronológica.

## Principais Etapas:

- **Importação e Pré-processamento:** Com base no modelo de organização de Hyndman e Athanasopoulos (2021), os dados foram estruturados com a coluna de data convertida em *date-time* e definida como índice, um aspecto importante para a construção de modelos de séries temporais.
- **Resampling:** Os dados diários foram consolidados em números mensais, facilitando a avaliação e revelando padrões sazonais e tendências (Tabela 1).

**Tabela 1.** Histórico mensal das despesas da cidade de 2018 a 2023

| Data       | Valor     |
|------------|-----------|
| 2018-01-31 | 143423867 |
| 2018-02-28 | 11194790  |
| 2018-03-31 | 60428223  |
| 2018-04-30 | 19569705  |
| 2018-05-31 | 57192770  |
| ...        | ...       |
| 2023-08-31 | 35130386  |
| 2023-09-30 | 64821865  |
| 2023-10-31 | 45679730  |
| 2023-11-30 | 54776005  |
| 2023-12-31 | 78381352  |

Fonte: Autores (2024).

Durante essa mudança, seus modelos foram adaptados segundo Shumway e Stoffer (2017). Esses processos também asseguraram que as transformações aplicadas garantissem a estacionariedade da série, sem necessidade de transformação logarítmica o que possibilitou aumentar a eficácia dos modelos de previsão.

### Treinamento de Modelos Preditivos

O treinamento dos modelos concentrou-se na previsão das despesas municipais de Juazeiro do Norte, usando-se técnicas de séries temporais. Os modelos a seguir foram usados:

- ARIMA – Este modelo captura tendências e padrões de autocorrelação e é, portanto, ideal para séries temporais com dependências temporais complexas. Box, Jenkins e Reinsel (2015), defendem que o ARIMA seja amplamente eficaz para a modelagem univariada, particularmente em séries temporais com tendência e variações cíclicas.
- ETS – Este modelo decompõe a série em erro, tendência e sazonalidade, de acordo com Hyndman e Athanasopoulos (2021). Ele seria adequado para as séries com sazonalidades acentuadas, como aumentos em determinados períodos em detrimento de outros devido a decisões de políticas públicas ou eventos locais. Os modelos foram treinados com base nos anais históricos do município, capturando, dessa forma, os padrões de longo prazo e sazonalidades.

### Testes Estatísticos

Os testes estatísticos são ferramentas fundamentais para avaliar a eficiência de modelos preditivos, especialmente em problemas de regressão e séries temporais. Eles permitem medir a precisão das previsões feitas pelos modelos ao comparar os valores reais com os valores previstos. A seguir, são apresentadas as principais métricas utilizadas.

MSE (Mean Squared Error): O Erro Médio Quadrático (MSE) mede o erro médio ao quadrado entre os valores reais ( $y_i$ ) e os valores preditos ( $\hat{y}_i$ ):

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (14)$$

Essa métrica penaliza erros grandes, pois os eleva ao quadrado, sendo sensível a *outliers*. Um MSE menor indica maior precisão do modelo. MAE (Mean Absolute Error): O Erro Médio Absoluto (MAE) avalia o erro médio em termos absolutos, calculado como:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (15)$$

Essa métrica oferece uma interpretação mais intuitiva, pois é expressa na mesma unidade da variável predita. Além disso, é menos sensível a *outliers* em comparação ao MSE. RMSE (Root Mean Squared Error): O Erro Médio Quadrático da Raiz (RMSE) é a raiz quadrada do MSE:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (16)$$

O RMSE combina as vantagens do MSE e facilita a interpretação, já que o resultado é expresso na mesma unidade da variável. Ele é útil quando grandes desvios são mais impactantes. MAPE (Mean Absolute Percentage Error): O Erro Médio Absoluto Percentual (MAPE) mede o erro médio em relação ao valor real, expresso como porcentagem:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \quad (17)$$

Essa métrica compara modelos em diferentes escalas, fornecendo uma visão relativa do erro. No entanto, é sensível a valores reais iguais a zero ( $y_i = 0$ ).  $R^2$  (Coeficiente de Determinação): Mede a proporção da variância dos valores reais que é explicada pelo modelo:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (18)$$

Valores próximos de 1 indicam um bom ajuste.

Correlação ( $r$ ): A correlação avalia a força e a direção da relação linear entre os valores reais e previstos:

$$r = \frac{\sum_{i=1}^n (y_i - \hat{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \hat{y})^2 \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}} \quad (19)$$

Um valor de  $r$  próximo de 1 indica uma forte correlação positiva, enquanto valores próximos de -1 indicam correlação negativa.

Ao avaliar o desempenho de modelos preditivos, é fundamental utilizar métricas complementares que proporcionem uma visão abrangente da eficácia do modelo. As principais métricas incluem o Erro Médio Absoluto (MAE), o Erro Quadrático Médio (MSE), a Raiz do Erro Quadrático Médio (RMSE), o Erro Percentual Absoluto Médio (MAPE), o coeficiente de determinação ( $R^2$ ) e a correlação.

O **MAE** mede a média das diferenças absolutas entre os valores previstos e os observados, fornecendo uma indicação clara da magnitude média dos erros. Já o **MSE** calcula a média dos quadrados dos erros, penalizando mais severamente erros maiores, o que pode ser útil para identificar modelos que cometem grandes erros com menor frequência (Hastie, Tibshirani, & Friedman, 2009). O **RMSE**, sendo a raiz quadrada do MSE, mantém a unidade dos dados originais, facilitando a interpretação dos resultados (James, Witten, Hastie, & Tibshirani, 2013).

O **MAPE** expressa o erro médio absoluto em termos percentuais, permitindo uma interpretação mais intuitiva da precisão relativa do modelo. No entanto, o MAPE pode ser problemático quando os valores reais são próximos de zero, levando a valores extremamente altos ou indefinidos (Makridakis, Wheelwright, & Hyndman, 2003).

O coeficiente de determinação,  **$R^2$** , indica a proporção da variabilidade dos dados que é explicada pelo modelo. Valores de  $R^2$  próximos de 1 sugerem que o modelo explica bem a variabilidade dos dados, enquanto valores próximos de 0 indicam o contrário (Draper & Smith, 1998). A **correlação**, como observado por Field (2013), avalia a força e a direção da relação linear entre as previsões do modelo e os valores observados, sendo uma métrica complementar que fornece *insights* adicionais sobre o desempenho do modelo.

Em resumo, ao utilizar essas métricas de forma conjunta, é possível obter uma avaliação mais robusta e detalhada do desempenho dos modelos preditivos, permitindo identificar pontos fortes e limitações específicas de cada abordagem.

## REFERENCIAL TEÓRICO

### Gestão pública

Segundo Vergara (2003), o sucesso na Gestão Pública resulta no desenvolvimento da comunidade política através do aprimoramento contínuo dos processos de resolução de problemas sociais. Em contraste, o fracasso pode condenar a comunidade ao subdesenvolvimento e a um ciclo persistente de problemas não resolvidos. O desafio da Gestão Pública é desenvolver mecanismos eficazes que permitam à comunidade política resolver seus problemas de maneira adequada, mobilizando recursos, coordenando as ações dos membros, corrigindo instituições que não atingem os resultados esperados e promovendo formas mais eficientes de cooperação social. Esse é um desafio técnico, político, intelectual e prático, mas, acima de tudo, essencial para o progresso.

Os princípios da gestão pública asseguram a eficácia e a legitimidade da administração e das políticas de recursos públicos. O objetivo desses princípios é garantir que os serviços governamentais e o que é feito pelo governo atendam às necessidades das pessoas de forma eficaz e transparente.

A eficiência é um dos pilares básicos da gestão pública, de acordo com Bresser-Pereira (1998); significa o uso ótimo de recursos para produzir ou maximizar resultados. Para o autor, a eficiência não é simplesmente redução de custos; em vez disso, é atingir os maiores resultados possíveis com os recursos disponíveis dentro das restrições orçamentárias do setor público.

Outro princípio consubstancial à gestão pública moderna é a eficácia, como capacidade de atingir os fins que foram estabelecidos nas políticas públicas (Denhardt, 2003). Para Ferreira et al. (2024), a eficácia se refere diretamente à capacidade de se implementar ações que, efetivamente, proporcionem um impacto positivo de vida da população. É um princípio voltado a resultados, que busca transformar as intenções do governo em ações concretas para o benefício público.

### **Pesquisa operacional aplicado às despesas**

A aplicação de métodos quantitativos e analíticos é essencial na solução de problemas complexos, como aqueles relacionados à otimização e alocação de recursos na administração pública. Esses métodos, amplamente utilizados na pesquisa operacional, combinam ferramentas matemáticas, estatísticas e computacionais para fornecer suporte decisório eficaz. Quando integrados à análise de séries temporais, permitem uma compreensão aprofundada de padrões e sazonalidades, contribuindo para o planejamento orçamentário estratégico e a distribuição eficiente de recursos ao longo do tempo.

Ao utilizar essas técnicas na gestão de despesas públicas, é possível prever desequilíbrios financeiros e implementar ações preventivas de forma assertiva. Por exemplo, a identificação de períodos de gastos elevados por meio de modelos preditivos auxilia na formulação de estratégias para otimizar o orçamento e evitar desperdícios. Assim, a sinergia entre abordagens analíticas e a análise de séries temporais promove uma gestão financeira mais transparente, eficiente e orientada a resultados.

### **Séries Temporais**

Uma série temporal é um conjunto de informações que são observadas em períodos regulares, podendo variar entre horas, dias, meses ou anos. Desta forma, os dados se organizam em uma série temporal, o que possibilita verificar o quanto uma determinada informação se alterou em diferentes pontas do tempo. As explicações de Morettin e Toloí (2006) reconhecem o valor prático que as séries temporais possuem, pois são amplamente aplicadas em diversas áreas como economia, meteorologia, etc., e para o propósito deste trabalho, em estudos relacionados a despesas do governo. A temporalidade implicada nos dados permite perceber algumas das tendências e das características sazonais disponíveis, com a finalidade de fazer interpretações visando ao futuro.

Com os dados de gastos anuais fornecidos de 2018 até 2023, da Prefeitura de Juazeiro do Norte, os gastos poderão ser analisados e a evolução estudada, com o objetivo principal de prever como as despesas podem vir a se comportar no futuro. Por meio da análise de séries temporais é possível visualizar períodos de altos gastos e melhor planejar os recursos da administração (Figura 1).



**Figura 1.** Série Temporal dos gastos mensais da Prefeitura de Juazeiro do Norte

Fonte: Autores (2024).

Essa abordagem, segundo Chatfield (2003), é aplicável nos ambientes onde a regularidade dos dados ao longo do tempo fornece informações valiosas para previsões e decisões. Isso é muito relevante na gestão pública, onde a prevenção de gastos e receitas futuras poderá vir em última instância afetar diretamente o orçamento e a prestação de serviços para a comunidade.

## RESULTADOS E DISCUSSÃO

Nesta parte, realizaremos uma análise descritiva dos dados das despesas mensais da prefeitura de Juazeiro do Norte e, em seguida, o modelo ARIMA será utilizado para prever as despesas.

### Análise Descritiva dos Gastos Mensais

A Tabela 2 a seguir apresenta as estatísticas descritivas da coluna "valor" que indica o gasto realizado durante os períodos de 2018-2023. Na análise foram determinadas as informações de número de despesas, média, desvio padrão, valor mínimo gasto, quartis e o valor máximo gasto.

**Tabela 2.** Análise estatística dos gastos da Prefeitura de Juazeiro do Norte

| Métrica | Valor           |
|---------|-----------------|
| mean    | R\$ 63 milhões  |
| std     | R\$ 63 milhões  |
| min     | R\$ 6 milhões   |
| 25%     | R\$ 34 milhões  |
| 50%     | R\$ 46 milhões  |
| 75%     | R\$ 65 milhões  |
| max     | R\$ 384 milhões |

Fonte: Autores (2024)

A análise revela que:

- **Média (Mean):** A média dos gastos é de aproximadamente 63 milhões, indicando o gasto médio ao longo do período.
- **Desvio Padrão (Std):** Com um desvio padrão de 63 milhões, observa-se uma alta variação dos gastos, sugerindo que os valores não são homogêneos. Em outras palavras, o desvio padrão alto, indica que os gastos apresentam valores que se afastam significativamente da média, apontando para uma distribuição heterogênea, ou "não homogênea". Esse desvio pode refletir picos sazonais ou despesas excepcionais em alguns meses.
- **Mínimo (Min):** O gasto mínimo registrado foi de cerca de 6 milhões, destaca-se o nível mais baixo de despesa.

- Quartis (25% e 75%): O primeiro quartil (34 milhões) indica que 25% dos gastos estão abaixo desse valor, enquanto o terceiro quartil (65 milhões) mostra que 75% dos gastos são iguais ou inferiores a esse montante.
- Máximo (Max): O pico de gasto foi de aproximadamente 384 milhões, revelando períodos de despesas significativamente mais altas.

A Figura 2, complementa a análise das métricas descritivas ao ilustrar o comportamento dos gastos ao longo do ano, com média mensal dos valores gastos. Observa-se uma variação acentuada nos valores, que corresponde ao desvio padrão elevado apresentado anteriormente, indicando que os gastos não são homogêneos ao longo do período.

**Figura 2.** Média das despesas mensais da Prefeitura de Juazeiro do Norte



Fonte: Autores (2024).

Em janeiro, os gastos médios atingem o valor mais alto do ano, refletindo o pico observado na análise estatística. Esse gasto elevado pode ser resultado de despesas concentradas no início do ano, como compromissos financeiros sazonais ou ajustes administrativos. Em contraste, o mês de fevereiro registra uma queda acentuada, com os valores mensais se estabilizando em um patamar mais baixo a partir de então.

Ao longo dos meses subsequentes, de março a novembro, os gastos mensais permanecem relativamente estáveis, condizendo com o valor mediano e os quartis observados nas métricas descritivas. Em dezembro, no entanto, há um aumento nos gastos, que pode ser associado a despesas de fechamento de ano, como cumprimento de metas orçamentárias ou ajustes fiscais, que elevam as despesas novamente.

Esse comportamento sazonal, com picos concentrados em janeiro e dezembro, reforça a ideia de que os gastos têm uma natureza cíclica e sugerem que fatores externos, como políticas e eventos administrativos, podem impactar o volume de despesas de forma significativa em determinados meses do ano.

Essas métricas sugerem uma alta dispersão nos gastos, possivelmente reflexo de políticas ou eventos sazonais que impactam o volume de despesas. A seguir, apresentamos os resultados obtidos através da aplicação do modelo ARIMA, que complementam esta análise inicial.



### Aplicação do Modelo ARIMA

A implementação do modelo ARIMA (Autoregressive Integrated Moving Average) para a análise das despesas da prefeitura de Juazeiro do Norte começa com uma decomposição da série temporal. Esta etapa permite separar os componentes que compõem a série: tendência, sazonalidade e ruído (ou resíduo).

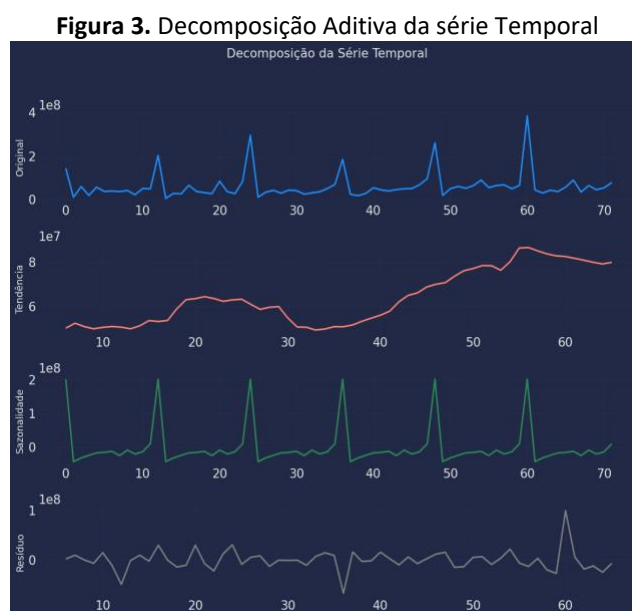
Matematicamente, isso pode ser representado como:

$$Y_t = T_t + S_t + R_t$$

onde:

- $Y_t$  é a série observada no tempo  $t$ .
- $T_t$  é a componente de tendência, refletindo a direção geral da série ao longo do tempo.
- $S_t$  é a componente sazonal, que captura as flutuações que ocorrem em intervalos regulares (como meses ou estações do ano).
- $R_t$  é a componente residual, representando variações aleatórias não explicadas pelos outros componentes.

O gráfico da Figura 3, esclarece as tendências e sazonalidades existentes na série temporal.



Fonte: Autores (2024).

### Componente Original

A linha azul representa as despesas mensais totais ao longo dos 72 meses conforme o apêndice B. Um pico significativo é observado em janeiro, onde os gastos são tradicionalmente mais altos. Esse aumento pode ser atribuído a despesas sazonalmente elevadas, como pagamentos de impostos, férias e outros custos extraordinários que costumam ocorrer no início do ano.

### Tendência

A linha vermelha ilustra uma tendência de crescimento nas despesas ao longo do tempo. Notamos um aumento exponencial a partir do 40º mês (abril de 2021) até o 56º mês (agosto de 2022), possivelmente devido à recuperação econômica durante e pós-pandemia, aumento dos custos operacionais e expansão de serviços públicos. Após esse período de crescimento, as despesas apresentaram uma estabilização em um nível mais alto, sugerindo que esses novos gastos se tornaram parte integral do orçamento.

### Sazonalidade

A linha verde, que representa o componente sazonal, se comporta de maneira semelhante à série original, evidenciando a presença de padrões sazonais claros nas despesas. Os picos que ocorrem em meses específicos refletem gastos recorrentes que seguem um padrão previsível ao longo do ano. Isso destaca que a série possui características sazonais bem definidas.

### Resíduo

A linha cinza mostra o componente residual, que representa variações que não são explicadas pela tendência ou sazonalidade. A proximidade dos resíduos a zero indica que as despesas mensais estão bem previstas pelos componentes de tendência e sazonalidade. Isso sugere que não há muitos desvios significativos ou eventos extraordinários que tenham impactado as despesas de maneira notável, evidenciando uma boa adequação do modelo utilizado para decompor a série temporal.

### Conclusão Geral

A análise da série temporal das despesas de Juazeiro do Norte revela uma tendência de crescimento consistente e padrões sazonais claros. A sazonalidade, que se comporta de forma semelhante à série original, destaca picos em meses específicos, como janeiro, e um crescimento exponencial nas despesas entre abril de 2021 e agosto de 2022. A estabilização subsequente e a continuidade dos padrões sazonais ressaltam a adaptação dos gastos às novas realidades econômicas. A proximidade dos resíduos a zero reforça a eficiência do modelo de decomposição utilizado, indicando que as despesas mensais são previsíveis e gerenciáveis, o que pode facilitar o planejamento financeiro e orçamentário.

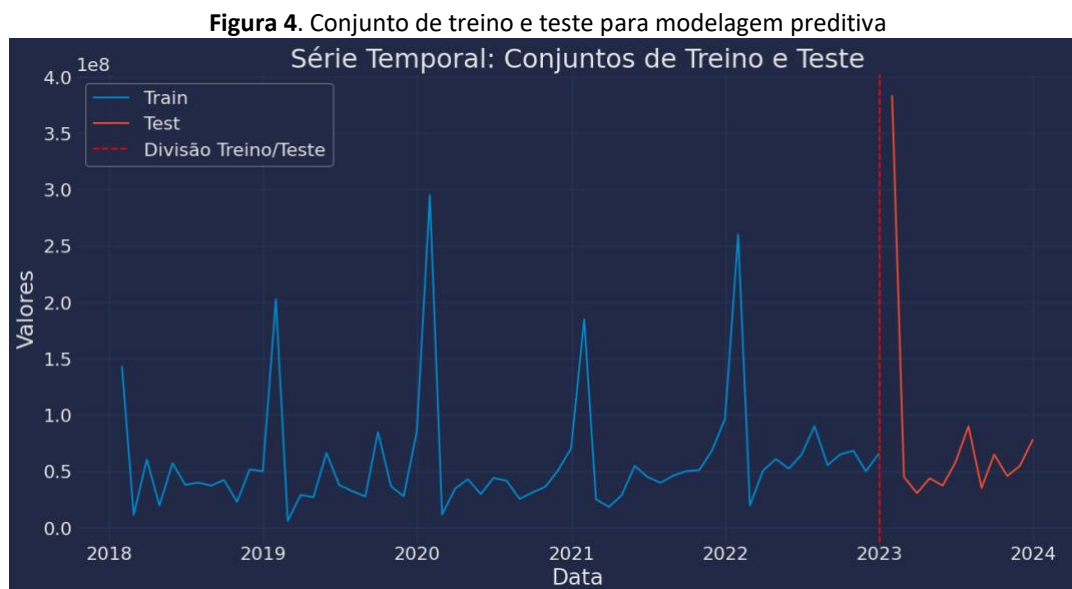
### Divisão dos dados em Treinamento e Teste

Para implementar o modelo ARIMA, os dados foram divididos em dois conjuntos: conjunto de treinamento e conjunto de teste. O conjunto de treinamento contém as observações até 1º de janeiro de 2023, enquanto o conjunto de teste abrange as observações posteriores a essa data.

A divisão dos dados é fundamental para avaliar a capacidade de um modelo de generalizar e prever corretamente novos dados. O conjunto de treinamento é utilizado para ajustar o modelo, enquanto o conjunto de teste permite medir seu desempenho em dados que não foram utilizados durante o aprendizado, assegurando que as previsões não sejam simplesmente uma repetição das informações já conhecidas.

A razão para essa divisão é fundamental na validação e avaliação do modelo. O conjunto de treinamento é utilizado para ajustar o modelo ARIMA, permitindo que ele aprenda os padrões subjacentes presentes nos dados. O conjunto de teste, por sua vez, é reservado para avaliar a capacidade preditiva do modelo em dados que ele não viu anteriormente.

Para séries temporais com sazonalidade, os autores sugerem que o ponto de corte entre treino e teste ocorra em uma posição que mantenha o padrão cíclico, evitando mudanças bruscas na série. Isso ajuda a garantir que o modelo possa capturar a sazonalidade sem extrapolações forçadas no conjunto de teste (Brockwell & Davis, 2002) (Figura 4).



O gráfico mostra a série temporal das despesas empenhadas da prefeitura, com uma divisão entre os conjuntos de treino e teste.

Análise do Gráfico:

- Linha Azul (Train): Representa o conjunto de dados de treino, que vai de 2018 até o final de 2022.
- Linha Vermelha (Test): Representa o conjunto de teste, que começa no início de 2023.
- Linha Tracejada Vermelha (Divisão Treino/Teste): Marca o ponto de separação entre os dados de treino e teste. Após essa linha, os dados do teste são usados para avaliar a capacidade do modelo de prever despesas futuras.

### Instanciando o modelo

Na modelagem de séries temporais, a escolha dos parâmetros  $p$  (autorregressão),  $d$  (diferenciação) e  $q$  (média móvel) no modelo ARIMA é uma etapa crucial para garantir previsões precisas. Tradicionalmente, esses parâmetros podem ser determinados através da análise visual das funções de autocorrelação (ACF) e autocorrelação parcial (PACF), como proposto por Box e Jenkins (2015). No entanto, neste trabalho, foi utilizado o Auto ARIMA, uma abordagem automatizada que seleciona os melhores parâmetros com base na minimização do Critério de Informação de Akaike (AIC), conforme discutido por Hyndman e Athanasopoulos (2021), eliminando a necessidade de inspeção manual dos gráficos de ACF e PACF.

O Auto ARIMA, descrito por Hyndman e Khandakar (2008), ajusta diferentes combinações de  $p$ ,  $d$  e  $q$ , selecionando a que apresenta o menor valor de AIC, melhorando o ajuste e a complexidade do modelo. Isso automatiza o processo de modelagem, evitando subjetividades na escolha dos parâmetros e garantindo que o modelo final seja o mais adequado para capturar os padrões da série temporal. Dessa forma, o uso do Auto ARIMA não só agiliza o processo de ajuste do modelo, como também aumenta a confiabilidade dos resultados, ao utilizar um critério quantitativo como o AIC para determinar os parâmetros ideais. Abaixo a Figura 5 do código de instanciamento do modelo ARIMA realizada no ambiente Python.

**Figura 5.** Código de instância do modelo preditivo

```

Carregar bibliotecas

[ ] from statsforecast import StatsForecast
    from statsforecast.models import AutoARIMA
    from statsforecast.arma import arma_string

Instanciando o modelo

[ ] season_length = 12 # Monthly data
    horizon = len(Y_test_df) # number of predictions

    models = [AutoARIMA(season_length=season_length)]

[ ] sf = StatsForecast(df=Y_train_df,
                      models=models,
                      freq='MS',
                      n_jobs=-1)

```

Fonte: Autores (2024).

Para configurar o modelo Auto ARIMA e adaptá-lo à natureza sazonal dos dados, definimos a variável `season_length` como 12, indicando que estamos lidando com dados mensais e que, portanto, a série apresenta um ciclo sazonal completo a cada 12 meses. O `horizon` foi definido como o comprimento do conjunto de dados de teste (`Y_test_df`), o que determina o número de previsões que o modelo fará.

Ao definir essa configuração, incluímos o modelo Auto ARIMA (`season_length = season_length`) na lista de modelos a serem ajustados e avaliados. Esse modelo otimiza automaticamente os parâmetros  $p$ ,  $d$ , e  $q$ , além de considerar a sazonalidade nos dados, identificando o padrão cíclico de 12 meses. A implementação dessa configuração nos permite gerar previsões mais precisas e alinhadas com a periodicidade observada na série temporal.

### Ajuste do Modelo

Após configurar o objeto `StatsForecast` com o conjunto de dados de treinamento e o modelo Auto ARIMA, a próxima etapa é ajustar o modelo aos dados. Na Figura 6 é mostrado o ajuste do modelo com o comando `sf.fit()`, iniciamos o processo de treinamento, permitindo que o modelo Auto ARIMA analise o conjunto de dados de treinamento (`Y_train_df`). Durante o ajuste, o modelo identifica os padrões e características da série, incluindo tendências, sazonalidade e componentes de ruído, conforme especificado.

**Figura 6.** Ajuste do modelo preditivo

```

Ajuste o modelo

▶ sf.fit()

/usr/local/lib/python3.10/dist-packages/statsforecast/core.py:476:
  warnings.warn(
StatsForecast(models=[AutoARIMA])

▶ StatsForecast(models=[AutoARIMA], freq='MS')

StatsForecast(models=[<class 'statsforecast.models.AutoARIMA'>])

[ ] arma_string(sf.fitted_[0,0].model_)

'ARIMA(1,0,0)(0,1,1)[12]'

[ ] result=sf.fitted_[0,0].model_
    print(result.keys())
    print(result['arma'])

dict_keys(['coef', 'sigma2', 'var_coef', 'mask', 'loglik', 'aic',
(1, 0, 0, 1, 12, 0, 1)

```

Fonte: Autores (2024).

Esse processo de treinamento é fundamental, pois permite ao Auto ARIMA encontrar os melhores parâmetros de autorregressão, diferenciação e média móvel, adaptando o modelo ao comportamento específico dos dados históricos. A partir desse ajuste, o modelo estará preparado para realizar previsões com base nos padrões aprendidos.

Após o ajuste do modelo com `sf.fit()`, obtemos uma representação textual do modelo ajustado com `arima_string(sf.fitted_[0,0].model_)`, que, neste caso, é `ARIMA(1,0,0)(0,1,1)[12]`. Essa notação descreve as especificações do modelo ARIMA, que foram ajustadas automaticamente para capturar as características da série temporal.

**I.(1,0,0): Essa primeira parte representa os parâmetros do modelo ARIMA, onde:**

- 1 é o parâmetro de autorregressão ( $p$ ), indicando que uma defasagem da série foi usada para prever o valor atual.
- 0 é o parâmetro de diferenciação ( $d$ ), sugerindo que os dados não precisaram ser diferenciados para se tornarem estacionários.
- 0 é o parâmetro de média móvel ( $q$ ), indicando que não houve necessidade de um termo de média móvel na modelagem da série.

**II.(0,1,1) [12]: Essa parte representa a sazonalidade do modelo:**

- 0 é o termo de autorregressão sazonal.
- 1 é o termo de diferenciação sazonal, sugerindo que uma diferenciação foi aplicada para capturar o padrão sazonal.
- 1 é o termo de média móvel sazonal.
- [12] indica que a sazonalidade ocorre em ciclos de 12 períodos, correspondente a uma série mensal com um ciclo anual.

Esse modelo ARIMA foi ajustado para capturar tanto a estrutura temporal quanto a sazonalidade dos dados, garantindo que a previsão leve em conta os padrões sazonais anuais.

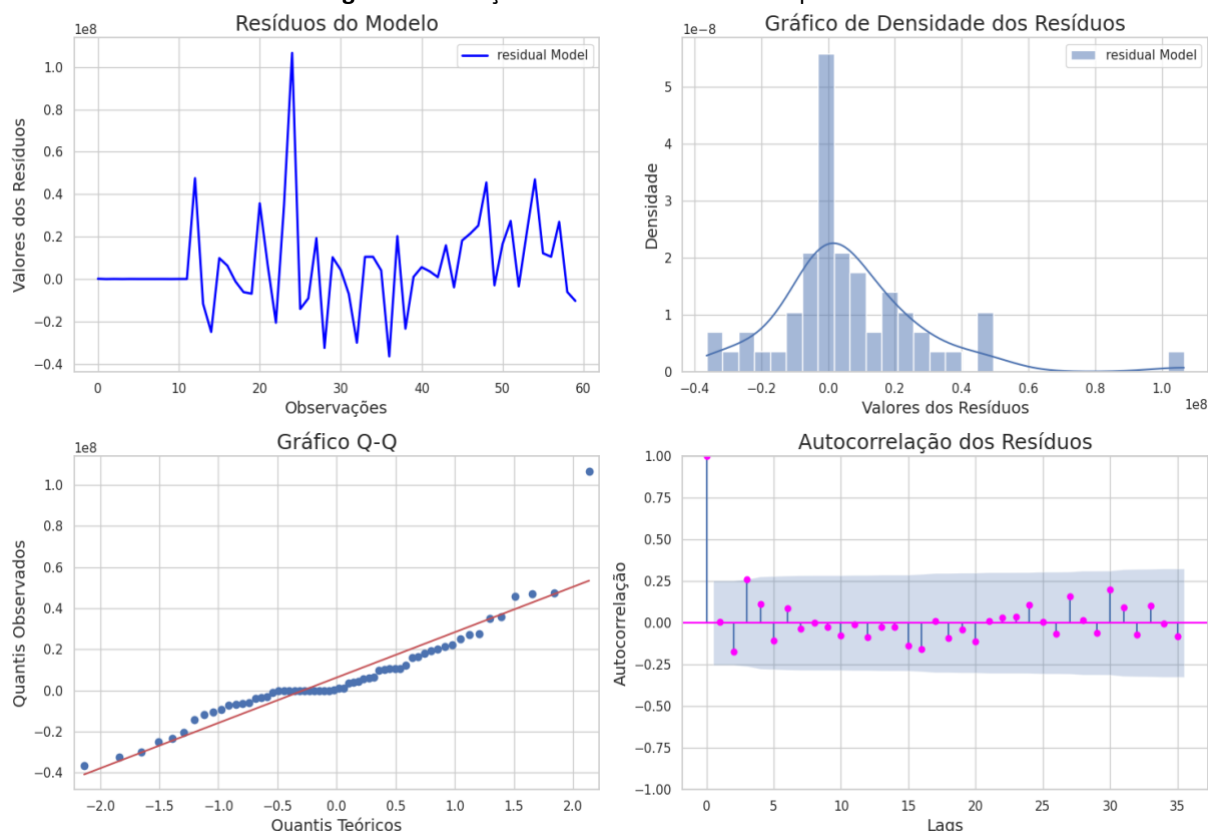
## Resíduos

Após ajustar um modelo ARIMA, precisamos verificar os resíduos para avaliar a qualidade e a adequação do modelo. Os resíduos são as diferenças entre os valores que o modelo previu e os valores que realmente ocorreram. Idealmente, se o modelo estivesse perfeito, todos os resíduos seriam iguais a zero. Portanto, ao analisar os resíduos, buscamos entender se o modelo está fazendo um bom trabalho. Segundo Box e Jenkins (2015), um dos fundamentos da modelagem ARIMA é que, após o ajuste, o comportamento da série deve ser totalmente explicado pelo modelo, deixando apenas ruído imprevisível, que é uma espécie de "barulho" aleatório.

Além disso, é importante verificar a presença de autocorrelação nos resíduos. Autocorrelação é quando os valores de uma série temporal estão relacionados entre si em diferentes momentos. Por exemplo, se um mês teve gastos altos, é provável que o próximo mês também tenha. Se os resíduos mostram autocorrelação significativa, isso sugere que o modelo ARIMA não capturou completamente a estrutura dos dados, o que pode exigir ajustes adicionais. Testes estatísticos, são frequentemente utilizados para avaliar essa dependência e garantir que os resíduos se comportem de forma independente, ou seja, sem relações previsíveis entre eles.

Outro aspecto relevante na análise dos resíduos é a normalidade e homoscedasticidade. A normalidade significa que os resíduos devem seguir uma distribuição que se assemelha a uma forma de sino, enquanto a homoscedasticidade indica que a variância dos resíduos deve ser constante ao longo do tempo. Através da Figura 7 abaixo, podemos perceber se os resíduos apresentam essas características, caso contrário, isso pode sugerir problemas na modelagem, como heterocedasticidade, que significa que a variabilidade dos dados muda ao longo do tempo (Hyndman & Athanasopoulos, 2021). Portanto, garantir que os resíduos sejam normais e homogêneos é fundamental para a validade das inferências feitas a partir do modelo.

**Figura 7.** Avaliação dos resíduos do modelo preditivo



Fonte: Autores (2024).

Os gráficos dos resíduos mostram resultados positivos que indicam que o modelo ARIMA capturou bem a estrutura dos dados em vários aspectos:

1. Resíduos ao longo do tempo (Gráfico de Resíduos do Modelo): A oscilação dos resíduos em torno de zero indica que o modelo conseguiu, em geral, prever a tendência principal dos dados. Isso é importante, pois sugere que o modelo capturou bem as flutuações médias dos dados ao longo do tempo.
2. Distribuição dos resíduos (Gráfico de Densidade): A densidade dos resíduos se concentra em torno de zero, o que é um sinal de que as previsões do modelo são, na maior parte das vezes, próximas dos valores reais. A curva se aproxima de uma distribuição normal, o que reforça a adequação do modelo e sugere que os erros de previsão são distribuídos de maneira razoavelmente aleatória.
3. Gráfico Q-Q: A maioria dos pontos no gráfico Q-Q está alinhada com a linha de referência, especialmente no centro, o que indica que os resíduos seguem uma distribuição próxima da normal. Isso é um bom sinal, pois a normalidade dos resíduos sugere que o modelo consiga capturar as tendências principais, reduzindo o risco de erros.

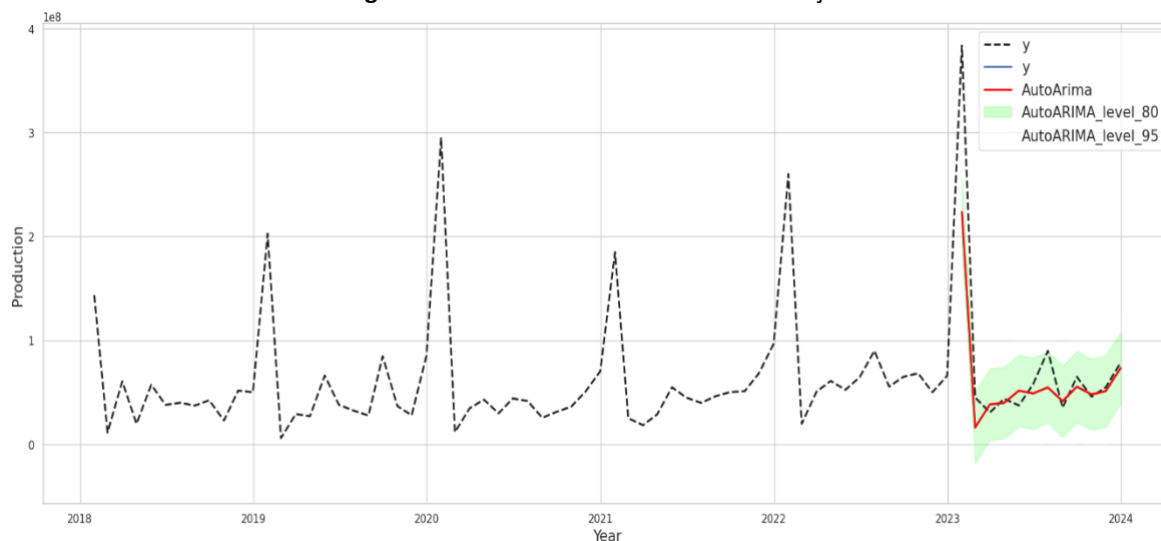


4. Autocorrelação dos Resíduos (ACF): A ausência de autocorrelação significativa na maioria dos *lags* sugere que o modelo ARIMA conseguiu capturar a estrutura temporal dos dados. Quando os resíduos não apresentam autocorrelação, indica que o modelo conseguiu extrair os padrões temporais dos dados, deixando apenas ruído aleatório.

Esses resultados indicam que o modelo ARIMA capturou de forma satisfatória os principais padrões dos dados, com resíduos que parecem estar próximos de uma distribuição normal e com pouca ou nenhuma autocorrelação. São sinais interessantes em uma análise de resíduos, pois demonstram que o modelo se ajusta bem aos dados e que suas previsões são consistentes.

Diante disso, criou-se um modelo de previsão ARIMA (Figura 8), onde seus componentes tentam se adaptar a sazonalidade do conjunto de dados.

**Figura 8.** Previsão com intervalo de confiança



Fonte: Autores (2024).

Neste gráfico, temos uma previsão das despesas mensais com o modelo Auto ARIMA, considerando os seguintes elementos e o nível de confiança das previsões:

1. Linha Tracejada Preta (Observado): Representa os valores reais das despesas mensais da prefeitura de Juazeiro do Norte ao longo do período de 2018 a 2023. Observa-se que os maiores picos ocorrem em janeiro de cada ano, o que indica um padrão sazonal, ligado a despesas específicas no início do ano.

2. Linha Vermelha (Previsão Auto ARIMA): Indica as previsões do modelo Auto ARIMA para o período mais recente. A linha vermelha tenta seguir a tendência geral da série temporal, considerando os dados históricos. Ela representa o valor previsto, mas sem captar completamente os picos altos observados em janeiro.

3. Faixa Verde (Intervalos de Confiança): O gráfico inclui uma área sombreada em verde, que mostra os intervalos de confiança de 80% e 95% em torno das previsões do modelo Auto ARIMA. O nível de confiança de 80% é o mais interno e o de 95% é o mais externo. Esses intervalos fornecem uma medida de incerteza das previsões. A faixa de 95% é mais ampla, sugerindo maior incerteza e menos precisão na previsão, especialmente nas áreas que cercam os picos.

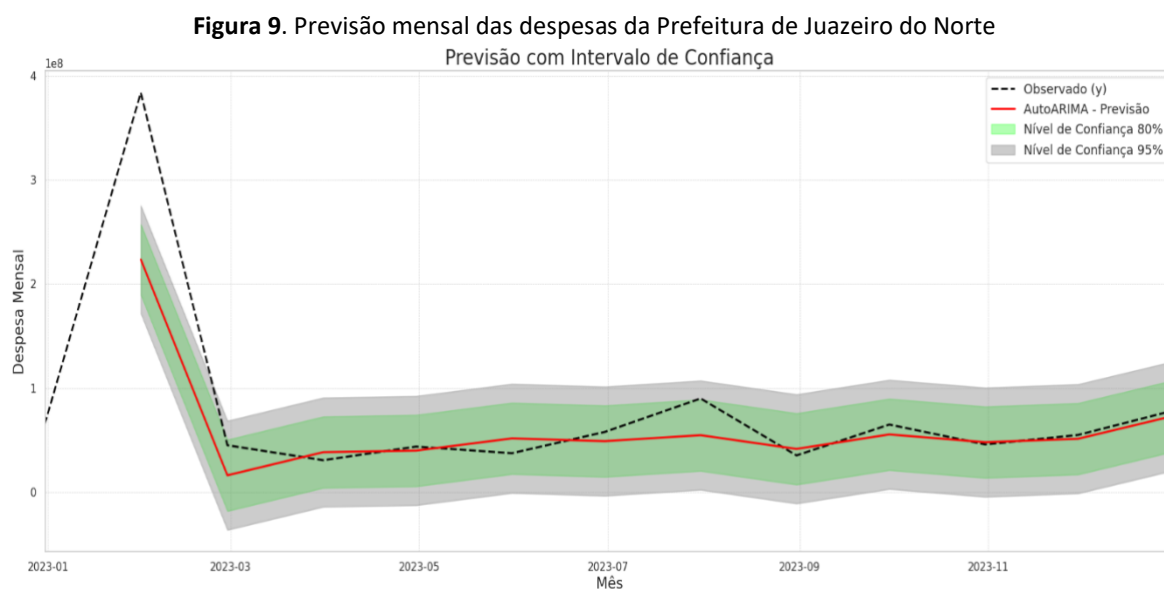
4. Interpretação dos Níveis de Confiança:

- 80%: Existe uma probabilidade de 80% de que as despesas reais se situem dentro do intervalo de confiança mais interno (menos abrangente).

- 95%: Esse intervalo mais amplo indica maior incerteza, mas inclui uma probabilidade maior (95%) de que as despesas reais estejam dentro desse intervalo. Isso é útil para um planejamento mais conservador, onde se espera cobrir cenários com maior probabilidade.

5. Análise da Precisão do Modelo: O Auto ARIMA captura a tendência média da série, mas a incerteza aumenta ao prever meses com picos fortes, como janeiro. A presença dos intervalos de confiança reflete que, embora a previsão forneça um valor central esperado, a variabilidade dos gastos torna necessário considerar uma faixa mais ampla de valores, especialmente ao lidar com períodos de pico.

Abaixo na Figura 9, temos o gráfico da previsão com níveis de confiança.



Neste gráfico acima, o foco está exclusivamente no período de previsão para o ano de 2023. A visualização permite observar com mais clareza o comportamento esperado das despesas nos meses restantes, conforme projetado pelo modelo Auto ARIMA.

### Complemento da Análise

1. Comportamento Pós-Pico de janeiro: Após o pico em janeiro, o modelo prevê uma estabilização gradual das despesas ao longo dos meses subsequentes, sem indicar picos adicionais até o final do ano. Isso sugere que, com base nos dados históricos, o modelo não espera variações sazonais ou flutuações significativas.

2. Convergência da Previsão com o Observado: A linha de previsão (vermelha) se aproxima dos valores observados (linha preta tracejada) após o mês de janeiro, indicando uma aderência maior do modelo ao comportamento recente das despesas. Isso pode ser um sinal de que o modelo se ajusta bem a períodos de estabilidade, mas enfrenta dificuldades com picos repentinos.

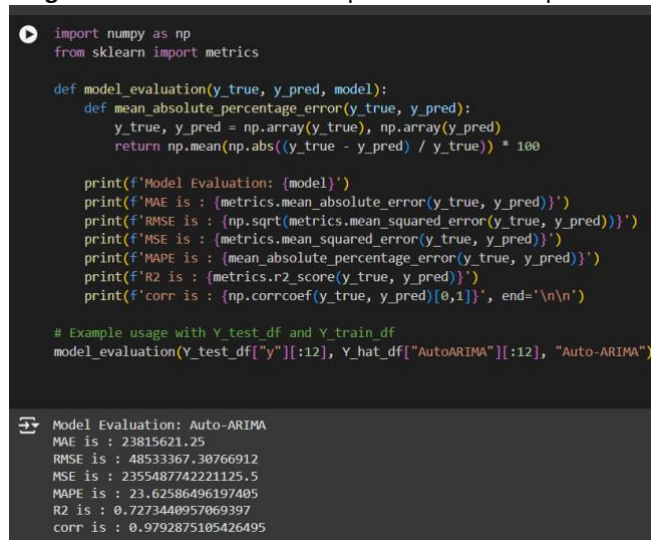
3. Intervalo de Confiança em Detalhe:

- No período entre fevereiro e dezembro de 2023, a faixa de confiança de 80% (verde) mantém-se relativamente estreita, o que indica uma menor variabilidade esperada. Esse comportamento sugere que o modelo prevê com razoável confiança que as despesas mensais permanecerão em um intervalo estreito, refletindo uma expectativa de estabilidade.
- A faixa de 95% (cinza), embora mais ampla, ainda sugere uma previsão sem grandes incertezas, reforçando a ideia de uma projeção estável.

Essa previsão indica que, para os meses restantes do ano, a prefeitura pode esperar uma estabilidade nas despesas, com pouca probabilidade de ocorrências de valores extremos. Esse dado pode ser útil para um planejamento financeiro mais seguro e com menor margem de erro para o segundo semestre dos anos.

A partir dessas análises, podemos utilizar as métricas de desempenho para averiguar a capacidade preditiva da série temporal (Figura 10).

**Figura 10.** Métricas de desempenho do modelo preditivo



```
import numpy as np
from sklearn import metrics

def model_evaluation(y_true, y_pred, model):
    def mean_absolute_percentage_error(y_true, y_pred):
        y_true, y_pred = np.array(y_true), np.array(y_pred)
        return np.mean(np.abs((y_true - y_pred) / y_true)) * 100

    print(f'Model Evaluation: {model}')
    print(f'MAE is : {metrics.mean_absolute_error(y_true, y_pred)}')
    print(f'RMSE is : {np.sqrt(metrics.mean_squared_error(y_true, y_pred))}')
    print(f'MSE is : {metrics.mean_squared_error(y_true, y_pred)}')
    print(f'MAPE is : {mean_absolute_percentage_error(y_true, y_pred)}')
    print(f'R2 is : {metrics.r2_score(y_true, y_pred)}')
    print(f'corr is : {np.corrcoef(y_true, y_pred)[0,1]}', end='\n\n')

# Example usage with Y_test_df and Y_train_df
model_evaluation(Y_test_df["y"][:12], Y_hat_df["AutoARIMA"][:12], "Auto-ARIMA")
```

Model Evaluation: Auto-ARIMA  
MAE is : 23815621.25  
RMSE is : 48533367.30766912  
MSE is : 2355487742221125.5  
MAPE is : 23.62586496197405  
R2 is : 0.7273440957069397  
corr is : 0.9792875105426495

Fonte: Autores (2024).

O código avalia o desempenho do modelo Auto ARIMA em uma série temporal, comparando os valores reais ( $y_{true}$ ) com os valores previstos ( $y_{pred}$ ). O modelo Auto ARIMA apresenta desempenho razoável, com um erro percentual médio (MAPE) de 23,63% e uma correlação forte (0.979) entre valores reais e previstos. No entanto, o  $R^2$  (72,7%) sugere que ele não explica tão bem a variabilidade dos dados em comparação a outros modelos. Apesar disso, o MAE e o RMSE indicam que o erro absoluto está dentro de uma faixa esperada, considerando a escala dos dados.

### Aplicação do Modelo ETS

Após apresentar os resultados do modelo ARIMA, seguimos para a análise do modelo ETS (Erro, Tendência e Sazonalidade). A implementação do ETS foi conduzida de forma semelhante ao ARIMA, garantindo a consistência na comparação entre os dois métodos.

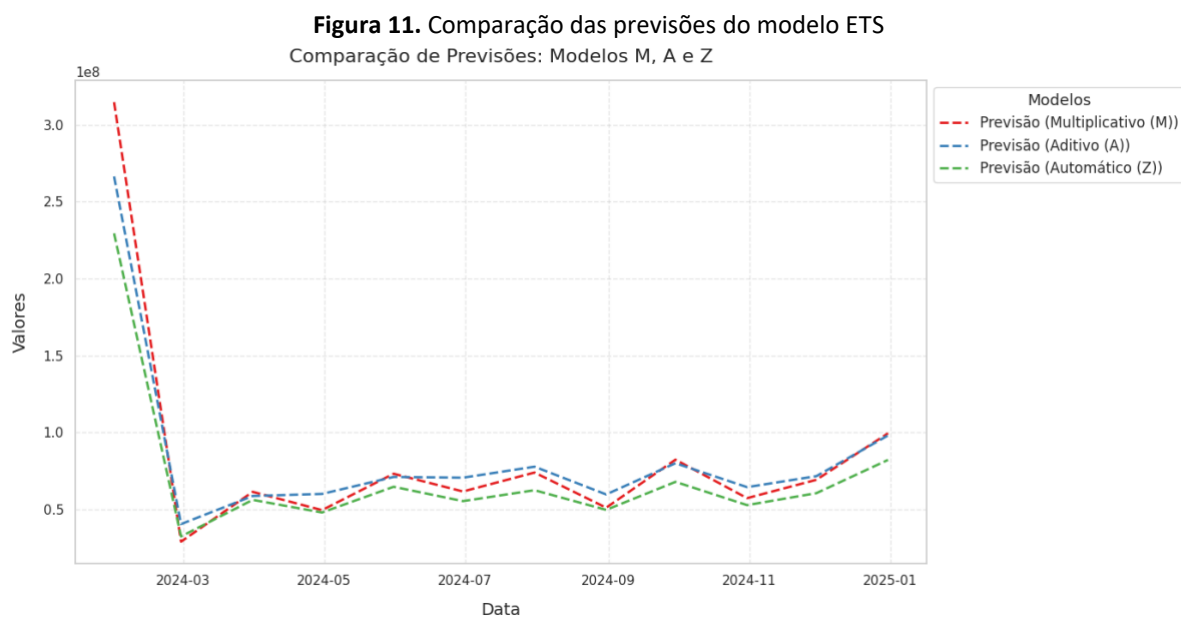
### Implementação do ETS

Divisão dos Dados:

Assim como no ARIMA, os dados foram divididos em treino e teste. Essa separação permite avaliar a capacidade preditiva do modelo ao aplicar previsões em dados não utilizados no treinamento. A separação dos dados se deu de forma idêntica ao modelo ARIMA, com o conjunto de treinamento contendo as observações até 1º de janeiro de 2023, enquanto o conjunto de teste abrange as observações posteriores a essa data.

### Configuração do Auto ETS:

Para explorar as diferentes combinações possíveis de ETS (erro, tendência e sazonalidade), utilizou-se o Auto ETS, uma ferramenta automatizada que testa múltiplas configurações e seleciona a melhor de acordo com os critérios de desempenho. A sazonalidade foi configurada conforme identificado na análise exploratória da série temporal, garantindo que o modelo capturasse os padrões periódicos adequados (Figura 11).



Fonte: Autores (2024).

Este gráfico apresenta a evolução das previsões realizadas por três modelos de séries temporais: Multiplicativo (M), Aditivo (A) e Automático (Z), ao longo do tempo. Apesar de todos os modelos capturarem tendências e padrões similares, os testes de desempenho mostraram que o modelo Multiplicativo (M) tem o melhor ajuste aos dados históricos.

#### Modelo Multiplicativo (M):

- Representado pela linha vermelha tracejada.

Com base nas métricas de avaliação (MSE, MAE, RMSE, MAPE,  $R^2$  e correlação), este modelo apresentou os melhores resultados, indicando maior precisão e menor erro nas previsões.

A proximidade dos valores previstos pelo modelo M aos valores reais sugere que ele é mais confiável para capturar padrões na série.

#### Modelo Aditivo (A):

- Representado pela linha azul tracejada.

Apresenta resultados muito próximos do modelo Multiplicativo (M), especialmente após a estabilização dos valores (abril de 2024 em diante). Apesar disso, o Aditivo (A) ficou em segundo lugar nas métricas de desempenho.

#### Modelo Automático (Z):

- Representado pela linha verde tracejada.

Embora também capture as tendências gerais da série, o modelo Z prevê valores ligeiramente mais baixos em relação aos outros dois modelos.

As métricas de desempenho sugerem que ele é menos preciso do que os modelos M e A, embora ainda apresente uma correlação aceitável com os dados reais.

#### Modelo Final:

Após o ajuste, o modelo selecionado apresentou a configuração [M, M, M] conforme a Figura 12, que utiliza componentes multiplicativos para erro, tendência e sazonalidade. Essa escolha indica que os padrões da série são proporcionais ao nível da série, comum em séries temporais que exibem crescimento ou sazonalidade com variações percentuais.

**Figura 12.** Modelo multiplicativo [M,M,M] com dados de treino e teste

Fonte: Autores (2024).

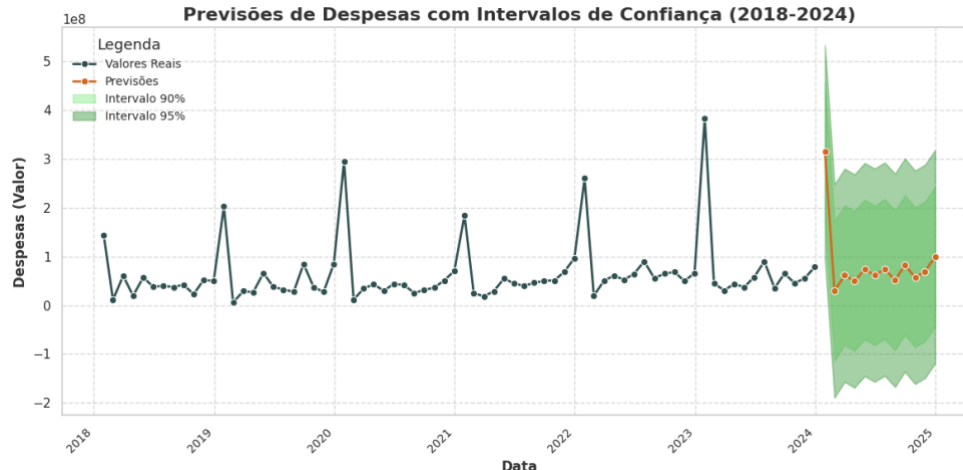
Este gráfico apresenta os dados históricos (treino e teste) e as previsões realizadas pelo Modelo M, que foi avaliado e se destacou em relação a outros modelos testados com base nas métricas de desempenho (MSE, MAE, RMSE, MAPE,  $R^2$  e correlação).

### Componentes do Gráfico

- Treino (linha azul): Dados históricos utilizados para treinar o modelo, representando as despesas de 2018 a 2023.
- Teste (linha vermelha): Dados reais reservados para validação (Teste), correspondendo a parte do ano de 2023.
- Previsão pelo Modelo M (linha amarela): Resultados gerados pelo modelo M para os anos de 2024 e 2025, com base no aprendizado obtido durante o treino.

O Modelo M apresentou um desempenho consistente ao longo de todo o período analisado. Durante o treino (2018-2023), foi capaz de capturar a alta volatilidade das despesas, incluindo os picos expressivos observados em anos como 2019 e 2022, mostrando-se bem ajustado às variações históricas. No teste (2023), o modelo apresentou uma forte aderência aos dados reais, demonstrando sua capacidade de reproduzir os padrões observados na série, com erros reduzidos nas previsões. Para o período de previsão (2024-2025), o modelo manteve a tendência e a sazonalidade da série histórica, projetando valores alinhados aos ciclos anteriores e suavizando picos extremos, o que contribui para a confiabilidade das estimativas futuras.

A Figura 13 apresenta a previsão do ETS das despesas com intervalos de confiança no período de 2018 a 2024.

**Figura 13.** Gráfico de previsão com intervalo de confiança  
**Previsões de Despesas com Intervalos de Confiança (2018-2024)**

Fonte: Autores (2024).

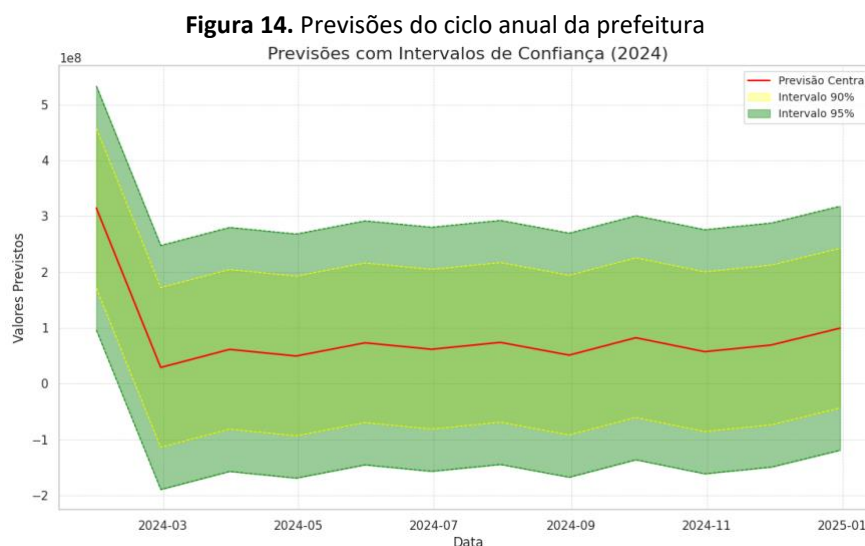
### Série temporal (2018-2024):

- Os pontos pretos representam os valores reais observados até o final de 2023.
- A linha laranja representa as previsões futuras, começando em 2024.

### Intervalos de Confiança:

- As áreas sombreadas indicam os intervalos de confiança para as previsões:
- 90% (verde claro): incerteza maior.
- 95% (verde escuro): maior incerteza.

O gráfico da Figura 14 abaixo apresenta as previsões com intervalos de confiança.



Fonte: Autores (2024).

### Previsão Central (Linha Vermelha):

- Representa os valores previstos de despesas, exibindo uma tendência de estabilização a partir de março de 2024.
- A previsão segue um padrão relativamente constante ao longo do ano, com pequenas variações mensais.

### Intervalos de Confiança:

- O intervalo de 90% (amarelo) é mais estreito, indicando maior precisão na estimativa central.
- O intervalo de 95% (verde) é mais amplo, refletindo uma incerteza maior, especialmente em março de 2024, onde há maior dispersão.
- Ambos os intervalos convergem para maior estabilidade a partir de abril de 2024.

Os gráficos apresentam a evolução das despesas municipais entre 2018 e 2024, evidenciando picos cíclicos que sugerem relação com períodos de festas e eventos sazonais do município, além de compromissos financeiros recorrentes. Os valores reais até 2023 mostram aumentos significativos em determinados meses, provavelmente associados a festividades locais ou eventos que demandam maiores investimentos públicos. As previsões para 2024 indicam uma tendência de estabilidade moderada, mas com um leve aumento, enquanto os intervalos de confiança refletem maior incerteza sobre a variabilidade desses gastos no futuro. Esse padrão reforça a influência de obrigações financeiras fixas e a sazonalidade dos eventos na gestão das despesas.

### Métricas de Precisão ETS

Após a análise preditiva do modelo ETS, utilizou-se métricas estatísticas para analisar a predição do modelo (Figura 15).



Figura 15. Métricas de Precisão ETS

```
[ ] from sklearn import metrics

def model_evaluation(y_true, y_pred, model):

    def mean_absolute_percentage_error(y_true, y_pred):
        y_true, y_pred = np.array(y_true), np.array(y_pred)
        return np.mean(np.abs((y_true - y_pred) / y_true)) * 100
    print(f'Model Evaluation: {model}')
    print(f'MSE is : {metrics.mean_squared_error(y_true, y_pred)}')
    print(f'MAE is : {metrics.mean_absolute_error(y_true, y_pred)}')
    print(f'RMSE is : {np.sqrt(metrics.mean_squared_error(y_true, y_pred))}')
    print(f'MAPE is : {mean_absolute_percentage_error(y_true, y_pred)}')
    print(f'R2 is : {metrics.r2_score(y_true, y_pred)}')
    print(f'corr is : {np.corrcoef(y_true, y_pred)[0,1]}',end='\n\n')

[ ] model_evaluation(test["y"][:12], forecast["hat"], "AutoETS")

Model Evaluation: AutoETS
MSE is : 742974609996429.4
MAE is : 21510499.880150557
RMSE is : 27257560.60245358
MAPE is : 36.67977583522682
R2 is : 0.9139981037362149
corr is : 0.9812674508454049
```

Fonte: Autores (2024).

O código avalia o desempenho do modelo de previsão comparando os valores reais ( $y_{true}$ ) com os valores previstos ( $y_{pred}$ ). O modelo "Auto ETS" apresenta um desempenho razoável, com boa explicação dos dados ( $R^2 = 0.913$ ) e uma forte correlação (0.981) entre os valores reais e previstos. No entanto, o MAPE de 36,7% sugere que há margem para melhoria, especialmente se a aplicação requer alta precisão. O MAE e o RMSE também indicam que os erros absolutos ainda podem ser reduzidos, dependendo da escala e da importância dos valores previstos.

### Comparação entre os modelos ETS e ARIMA

Tabela 3. Comparação entre os modelos – Autoarima vs AutoETS

| Métrica        | AutoARIMA            | AutoETS                | Melhor Modelo (Menor/Maior) |
|----------------|----------------------|------------------------|-----------------------------|
| MAE            | 23,815,621.25        | 21,510,499.88          | AutoETS (menor é melhor)    |
| RMSE           | 48,533,367.31        | 27,275,560.60          | AutoETS (menor é melhor)    |
| MSE            | 2,355,487,422,211.25 | 742,974,609,996,429.40 | AutoETS (menor é melhor)    |
| MAPE           | 23.63%               | 36.68%                 | AutoARIMA (menor é melhor)  |
| R <sup>2</sup> | 0.7273               | 0.9139                 | AutoETS (maior é melhor)    |
| Correlação     | 0.9793               | 0.9813                 | AutoETS (maior é melhor)    |

Fonte: Autores (2024).

#### 1) Precisão Absoluta:

a) O modelo **AutoETS** se destaca em métricas como **MAE**, **RMSE** e **MSE**, indicando melhor desempenho ao prever os valores absolutos dos dados.

b) Isso indica que, para aplicações onde a minimização dos erros absolutos é crucial (ex.: previsão de demanda de serviços, orçamento financeiro), o **AutoETS** seria a escolha mais apropriada.

#### 2) Precisão Relativa:

a) O **AutoARIMA** apresenta um **MAPE** menor (23,63%), o que indica maior precisão relativa, sendo mais adequado em contextos em que os erros proporcionais aos valores reais são mais relevantes.

b) Isso o torna mais adequado para contextos em que a magnitude relativa do erro é mais importante que o erro absoluto.

### 3) Explicação da Variância ( $R^2$ ):

a) O modelo **AutoETS** é superior ao **AutoARIMA**, explicando 91,3% da variação nos dados contra apenas 72,7% do ARIMA.

b) Isso reforça que o **AutoETS** possui maior capacidade de captar a estrutura dos dados, sendo mais eficaz para séries temporais com flutuações sazonais e padrões mais complexos.

### 4) Correlação:

a) Ambos os modelos apresentam alta correlação ( $> 0.97$ ) entre os valores reais e previstos, o que reflete um bom alinhamento geral.

b) No entanto, o **AutoETS** apresenta uma correlação ligeiramente superior (0.9813 vs. 0.9793), o que pode indicar um melhor ajuste fino às oscilações da série.

### Impacto Prático dos Resultados

- Se o objetivo for prever valores exatos (ex.: quanto será arrecadado no próximo mês em unidades monetárias), o **AutoETS** é a melhor escolha, pois apresenta menores erros absolutos e maior explicação da variância.
- Se o foco for proporcionalidade dos erros (ex.: prever a porcentagem de crescimento ou queda nas despesas), o **AutoARIMA** pode ser mais útil, pois seu erro percentual é menor.

### CONSIDERAÇÕES FINAIS

O modelo **Auto ETS (multiplicativo)** destacou-se como o mais eficiente na redução dos erros absolutos, apresentando os menores valores de **RMSE** e **MAE**. Esses indicadores reforçam a superioridade do modelo em fornecer previsões mais precisas quando a magnitude dos erros é o principal critério de avaliação. Além disso, o **Auto ETS** obteve um  $R^2$  mais elevado, evidenciando sua capacidade superior de capturar a variabilidade dos dados e de refletir as flutuações observadas na série temporal de maneira mais fiel. Essa característica torna o modelo especialmente útil na identificação de padrões sazonais e variações complexas, o que é essencial para o planejamento e a análise de séries temporais com sazonalidade significativa.

Por outro lado, o **Auto ARIMA** apresentou melhor desempenho em termos de **precisão relativa**, evidenciada por um **MAPE** mais baixo. Esse resultado sugere que o modelo é mais indicado em cenários onde a proporção do erro em relação aos valores reais é mais relevante. Contudo, o **Auto ARIMA** encontrou dificuldades em representar sazonalidades mais pronunciadas e em capturar adequadamente picos sazonais, especialmente no mês de janeiro. Apesar de fornecer previsões robustas para tendências e ciclos gerais, sua capacidade de lidar com padrões sazonais mais intensos mostrou-se limitada, mesmo ao considerar intervalos de confiança de 80% e 95%.

Essas limitações do **Auto ARIMA** reforçam a adequação do **Auto ETS (multiplicativo)** para séries temporais com variações sazonais acentuadas. A abordagem multiplicativa permitiu ao modelo ajustar-se de forma mais eficiente às alterações abruptas e aos picos sazonais, proporcionando previsões mais alinhadas à realidade observada. Em síntese, os resultados obtidos indicam que a escolha do modelo ideal depende do contexto da análise: enquanto o **Auto ARIMA** pode ser preferido para situações com menor ênfase em sazonalidades complexas e maior foco em tendências gerais, o **Auto ETS (multiplicativo)** é mais indicado para cenários com sazonalidades marcantes e padrões mais dinâmicos.

### Limitações e Considerações Adicionais

Este estudo focou exclusivamente na comparação entre os modelos **Auto ETS** e **Auto ARIMA**, sem explorar outros métodos comuns de previsão de séries temporais, como **Redes Neurais Recorrentes (RNNs)**, **Prophet** e **modelos híbridos**. A escolha desses dois modelos se justifica pela sua ampla utilização em cenários de previsão estatística e pela sua eficiência

computacional, permitindo análises rápidas e interpretáveis. No entanto, estudos futuros podem expandir essa análise, avaliando o desempenho de modelos baseados em aprendizado profundo ou abordagens híbridas que combinem técnicas estatísticas e de *machine learning*.

Além disso, é importante destacar que fatores **políticos e econômicos** desempenham um papel crucial na formação dos padrões de gastos ao longo do tempo. Mudanças em políticas fiscais, oscilações cambiais, crises econômicas e variações no mercado de trabalho podem influenciar significativamente as tendências identificadas pelos modelos. Embora os modelos analisados capturem padrões históricos e sazonais, sua capacidade de antecipar mudanças estruturais na economia é limitada. Dessa forma, a interpretação das previsões deve sempre ser contextualizada dentro do cenário macroeconômico vigente.

## REFERÊNCIAS

- Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. Brockwell, P. J. & Davis, R. A. (Eds.). (2002). *Introduction to time series and forecasting*. New York, NY: Springer New York.
- Vergara, S. H. C. & Almeida, C. V. L. de. (2003). *Propostas para uma gestão pública municipal efetiva*. FGV Editora.
- Chatfield, C. & Xing, H. (2019). *The analysis of time series: an introduction with R*. Chapman and hall/CRC.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approach* (4th ed.). SAGE Publications.
- Denhardt, R. B. & Denhardt, J. V. (2003). *The new public service: Serving, not steering* (2nd ed.). M.E. Sharpe.
- Ferreira, M. T. N., Silva, D. J. V. da, Santos, A. M. N. dos, Henrique, T. G., Tavares, V. H. N. D., & Santana, W. J. de. (2024). Gestão pública e desenvolvimento sustentável. *Brazilian Journal of Development*, 10(8), e71884-e71884.
- Draper, N. R. & Smith, H. (1998). *Applied regression analysis* (Vol. 326). John Wiley & Sons.
- Enders, W. (2008). *Applied econometric time series*. John Wiley & Sons.
- Field, A. (2024). *Discovering statistics using IBM SPSS statistics*. Sage publications limited.
- Hastie, T. (2009). The elements of statistical learning: data mining, inference, and prediction.
- Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. *International journal of forecasting*, 20(1), 5-10.
- Hyndman, R. J. & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3<sup>rd</sup> ed.). OTexts. <https://otexts.com/fpp3/>
- Hyndman, R. J. & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, 27(3). <https://doi.org/10.18637/jss.v027.i03>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning: With applications in R*. Springer.
- Makridakis, S., Spyropoulos, M., Assimakopoulos, V., & Nikolaidis, E. (2018). Statistical and machine learning forecasting methods: Concerns and ways forward. *PLOS ONE*, 13(3), 1-26. <https://doi.org/10.1371/journal.pone.0194889>
- Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (2003). *Forecasting methods and applications* (3rd ed.). Wiley.
- Morettin, P. A. & Toloi, C. M. C. (2006). *Análise de séries temporais* (2nd ed.). Blucher.
- Pereira, L. C. B. (1998). *Reforma do Estado para a cidadania: a reforma gerencial brasileira na perspectiva internacional*. Editora 34.
- Shumway, R. H. & Stoffer, D. S. (2017). *Time series analysis and its applications: With R examples* (4th ed.). Springer.
- Silva, M. C., Santos, P. R., & Almeida, J. F. (2020). Portal de Transparência como ferramenta de controle social: Uma análise do período de 2018 a 2020. *Revista Gestão Pública*, 12(4), 23-45.
- Willmott, C. J. & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30, 79-82.
- Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science*, 6(3), 324-342. <https://doi.org/10.1287/mnsc.6.3.324>