



**UNIVERSIDADE FEDERAL FLUMINENSE**  
**INSTITUTO DE CIÊNCIAS E TECNOLOGIA**  
**CURSO DE ENGENHARIA DE PRODUÇÃO**

**KATHERINE ZUNIGA**  
**MARCIO RODRIGUES**  
**MATHEUS PAIVA**

**ESTUDO DE CASO SOBRE A EMPRESA SEGROB NOTLAD**

**RIO DAS OSTRAS- RJ**

**2025**

**KATHERINE ZUNIGA**  
**MARCIO RODRIGUES**  
**MATHEUS PAIVA**

**ESTUDO DE CASO SOBRE A EMPRESA SEGROB NOTLAD**

Estudo de caso sobre a empresa Segrob Notlad no Curso de Graduação em Engenharia de Produção da Universidade Federal Fluminense – Rio das Ostras - RJ como método avaliativo da disciplina Análise Preditiva.

Docente: Dalton Borges.

**RIO DAS OSTRAS- RJ**

2025

## Resumo:

Este estudo de caso tem como foco o desenvolvimento de uma solução preditiva para o varejo de moda rápida, por meio da análise da demanda por camisetas básicas da empresa Segrob Notlad. A pesquisa parte da observação do crescimento expressivo e sustentado nas vendas do produto entre os anos de 2022 e 2024, com o objetivo de construir modelos de previsão capazes de auxiliar o planejamento operacional da marca. Para tanto, foram utilizados dados históricos de vendas diárias, tratados e analisados com base na metodologia CRISP-DM, aplicando técnicas de modelagem preditiva e métricas de avaliação como MAPE, RMSE e MAD. Os resultados apontam para a existência de padrões sazonais semanais e anuais, além da influência de campanhas promocionais no comportamento de consumo. A abordagem adotada permite não apenas prever com maior precisão a demanda futura, mas também gerar insumos estratégicos que fortalecem a tomada de decisão em um mercado altamente dinâmico como o *fast fashion*.

Palavras-chave: *Fast Fashion*. Demanda. Análise Preditiva. Séries Temporais

## SUMÁRIO

<b>1 - Introdução.....</b>	<b>5</b>
<b>2 - Objetivo.....</b>	<b>6</b>
<b>3 - Método.....</b>	<b>7</b>
<b>4 - Estudo de Caso.....</b>	<b>8</b>
<b>4.1 - Entendimento do Negócio.....</b>	<b>8</b>
<b>4.2 - Entendimento dos Dados.....</b>	<b>9</b>
4.2.1 - Gráfico de Nível.....	9
4.2.2 - Autocorrelação.....	11
4.2.3 - Média Quarter.....	12
4.2.4 - Registro Diário de Vendas.....	14
4.2.5 - Gráfico de Tendências.....	15
4.2.6 - Histograma Sazonalidade Temporal.....	16
4.2.7 - Modelos Preditivos.....	17
4.2.7.1 Cumulativo e Naive.....	17
4.2.7.2 Média Móvel.....	19
4.2.7.3 Suavização Exponencial Simples.....	20
4.2.7.4 Regressão Linear Simples.....	21
4.2.7.5 Análise dos Resíduos.....	22
4.2.7.6 Validação do Modelo.....	27
4.2.7.7 Regressão Linear Dinâmica.....	29
<b>4.3 - Entendimento Geral dos Dados de Vendas.....</b>	<b>29</b>
<b>4.4 - Previsão de vendas para Dezembro de 2024.....</b>	<b>30</b>
<b>4.5 Preparação de Dados.....</b>	<b>31</b>
4.5.1 Organização inicial.....	31
4.5.2 Criação de variáveis auxiliares.....	31
4.5.3 Verificação e tratamento de inconsistências.....	32
4.5.4 Classificação dos dados.....	32
4.5.5 Resultado final da preparação.....	32
<b>5. Fundamentação teórica.....</b>	<b>32</b>
<b>5.1 Validação Cruzada dos Modelos de Previsão.....</b>	<b>32</b>
<b>6. Referências.....</b>	<b>35</b>

## 1 - Introdução

Diante das rápidas transformações no comportamento do consumidor e da crescente competitividade no setor de moda, compreender a atuação de marcas consolidadas no mercado torna-se essencial para a elaboração de estratégias que aliam inovação, eficiência operacional e adaptação às dinâmicas do varejo. Inserida nesse cenário, a Segrob Notlad se destaca como uma marca brasileira de *fast fashion* cuja trajetória é marcada pela combinação entre ousadia, agilidade e espírito urbano.

Fundada no Rio de Janeiro, a empresa rapidamente consolidou sua presença nacional por meio de um portfólio focado em design versátil, preços acessíveis e campanhas publicitárias com forte apelo junto ao público jovem e conectado. Atualmente, conta com mais de 80 lojas no Brasil, além de presença internacional por meio de lojas conceito em pontos estratégicos da América do Sul e Europa.

A origem da marca está diretamente ligada à história de vida de seu fundador, Segrob Notlad, cuja trajetória entre o leste europeu e o Brasil influenciou diretamente a identidade da empresa, resultando em uma fusão entre estética minimalista e diversidade cultural. Ao longo de sua evolução, a marca passou a integrar de forma estratégica tecnologias de inteligência de mercado e automação de processos em sua cadeia de suprimentos, posicionando-se como referência em adaptação aos ciclos da moda e às exigências do consumidor contemporâneo.

Em sua fase mais recente, iniciada em 2025, a empresa reforça seu compromisso com a inovação ao incorporar soluções de inteligência artificial para previsão de demanda e apoio à tomada de decisões. Este movimento sinaliza uma transição para uma atuação ainda mais orientada por dados, destacando a análise preditiva como uma ferramenta central em sua busca por eficiência, assertividade e competitividade em um mercado em constante transformação.

## 2 - Objetivo

Este estudo de caso faz parte do movimento constante da Segrob Notlad em busca de inovação e melhorias no seu processo de decisão. O principal objetivo é entender e antecipar a demanda de um dos produtos mais importantes da marca: a camiseta básica. A partir da análise dos dados históricos de vendas, busca-se construir modelos de previsão que auxiliem o planejamento operacional e aumentem a agilidade da empresa para lidar com as mudanças e desafios do setor de *fast fashion*.

O objetivo geral deste trabalho é desenvolver uma previsão confiável da demanda diária de camisetas básicas, de modo a apoiar a tomada de decisão nas áreas de produção, logística e abastecimento. Para alcançar esse propósito, serão realizadas análises sobre o comportamento histórico de vendas, com a identificação de padrões sazonais, tendências de crescimento ou queda, e eventuais eventos atípicos que possam impactar o consumo.

A escolha desse enfoque justifica-se pela importância estratégica de prever corretamente a demanda. Antecipar as necessidades de produção permite à Segrob Notlad otimizar seus processos logísticos, reduzir estoques excessivos, minimizar perdas e melhorar o atendimento ao cliente final, fatores essenciais para manter a competitividade no mercado de *fast fashion*, que se caracteriza por alta volatilidade e forte pressão por agilidade e adaptação contínua.

### 3 - Método

A metodologia adotada para a condução deste estudo de caso baseia-se no modelo CRISP-DM (*Cross-Industry Standard Process for Data Mining*), amplamente reconhecido como uma referência na estruturação de projetos de análise de dados. O CRISP-DM organiza o desenvolvimento do projeto em etapas sequenciais e iterativas, proporcionando um fluxo lógico que facilita a compreensão dos objetivos de negócio, o tratamento dos dados e a construção de soluções analíticas eficazes.

O trabalho inicia-se com a fase de entendimento do negócio, na qual são analisadas as necessidades estratégicas da Segrob Notlad, especialmente no que se refere à antecipação da demanda por camisetas básicas. Nesta etapa, são definidos os objetivos de previsão e as métricas que serão utilizadas para a avaliação do desempenho dos modelos, alinhando as expectativas de análise com a realidade operacional da empresa.

Em seguida, é realizada a fase de entendimento dos dados, que consiste na coleta, organização e exploração preliminar do conjunto histórico de vendas. Nessa etapa, busca-se identificar padrões relevantes, possíveis inconsistências, comportamentos sazonais e tendências gerais, além de realizar o tratamento de dados ausentes ou incorretos.

A fase de preparação dos dados sucede o entendimento inicial, envolvendo a limpeza, transformação e formatação dos dados de forma adequada para a modelagem. Esse processo inclui a criação de variáveis derivadas, normalização de valores e a divisão dos dados em conjuntos de treinamento e teste, respeitando as características temporais da série.

A fase de modelagem contempla a aplicação de técnicas de previsão apropriadas, considerando a natureza da série temporal analisada. Serão testados diferentes modelos preditivos, com o objetivo de encontrar aquele que apresentar melhor desempenho nas previsões diárias de demanda. A avaliação dos modelos será baseada nas métricas previamente definidas: o erro percentual absoluto médio (MAPE), a raiz do erro quadrático médio (RMSE) e a média das diferenças absolutas (MAD).

Após a modelagem, é realizada a fase de avaliação, que busca validar a qualidade das previsões geradas e confirmar se os resultados atendem aos objetivos estabelecidos no início do projeto. Caso necessário, ajustes nos modelos ou no tratamento dos dados serão realizados de forma a melhorar a acurácia das previsões.

Por fim, será feita a fase de implantação, que, no contexto deste estudo, refere-se à apresentação dos resultados obtidos, à análise crítica das soluções propostas e às

recomendações de possíveis aplicações práticas dentro do planejamento da Segrob Notlad. Todo o processo é concebido de maneira flexível, permitindo revisões e aprimoramentos à medida que novos dados ou demandas estratégicas surgirem, em consonância com o dinamismo do ambiente corporativo da empresa.

## **4 - Estudo de Caso**

O estudo de caso é uma abordagem metodológica que permite investigar, com profundidade e contextualização, situações reais enfrentadas por empresas, organizações ou setores específicos. Ao focar em um único objeto de análise, como um produto, processo ou decisão empresarial, esse tipo de estudo busca compreender os fatores que influenciam determinada realidade, identificar padrões de comportamento e propor soluções baseadas em dados concretos. Diferente de análises puramente teóricas ou generalistas, o estudo de caso utiliza informações reais e circunstâncias específicas, o que o torna uma ferramenta valiosa para gerar diagnósticos detalhados e embasamento prático para a tomada de decisões estratégicas.

### **4.1 - Entendimento do Negócio**

A Segrob Notlad atua no mercado de *fast fashion*, um setor conhecido pela alta dinamicidade, pela exigência de respostas rápidas às mudanças de comportamento do consumidor e pela intensa pressão por inovação constante. Dentro desse cenário, a empresa construiu sua estratégia combinando inteligência de mercado, automação da cadeia de suprimentos e forte presença no ambiente digital e físico. O foco é atender de forma eficiente um público jovem e urbano, que valoriza tanto o acesso rápido a novidades quanto uma boa relação custo-benefício.

O produto central deste estudo é a camiseta básica, um dos itens mais relevantes dentro do portfólio desta empresa. Por seu perfil atemporal e sua alta rotatividade, a camiseta básica representa um componente fundamental para a geração de receita e para a manutenção do fluxo operacional de lojas e centros de distribuição. Ter previsibilidade de sua demanda é essencial para otimizar a produção, controlar estoques e assegurar a disponibilidade contínua desse item, minimizando perdas por excesso ou falta de produtos nas lojas.

A empresa, ciente da importância estratégica da camiseta básica, busca com este



estudo de caso aumentar sua capacidade de planejamento a partir da análise crítica dos dados de vendas históricos. A utilização de técnicas de análise preditiva visa não apenas antecipar o volume necessário para um período específico, mas também fortalecer sua estrutura de tomada de decisão frente a eventos sazonais, variações no comportamento do consumidor e mudanças de mercado.

O ambiente em que a Segrob Notlad está inserida é caracterizado por forte influência de datas comerciais, campanhas promocionais, variações climáticas e eventos sociais e culturais. Além disso, o próprio ciclo da moda, ainda que em um produto básico, pode impactar picos ou quedas de demanda em determinados períodos. Outro fator relevante é a operação multicanal da empresa, integrando lojas físicas e estratégias digitais, o que exige ainda mais precisão na previsão de demanda para garantir uma resposta eficiente em todos os pontos de venda.

O desafio proposto se insere diretamente na estratégia de fortalecer o uso de inteligência de dados como diferencial competitivo, permitindo à Segrob Notlad uma atuação ainda mais ágil e assertiva no mercado. A previsão de demanda para a camiseta básica, portanto, representa não apenas uma ação tática de curto prazo, mas também uma evolução no modo como a empresa utiliza dados para sustentar seu crescimento e eficiência operacional em um ambiente de alta volatilidade.

## **4.2 - Entendimento dos Dados**

### **4.2.1 - Gráfico de Nível**

Apresenta a evolução das vendas diárias de camisetas básicas masculinas da marca Segrob Notlad entre janeiro de 2022 e novembro de 2024. A linha azul representa o volume diário de vendas, enquanto a linha vermelha tracejada indica o nível médio histórico de vendas, calculado em 214,16 unidades por dia.

Ao longo do tempo, observa-se uma mudança progressiva no comportamento da demanda (Figura 1):

ANO	
2022	A maior parte das vendas diárias encontra-se <b>abaixo da média histórica</b> . Isso sugere um período inicial de maturação da operação, com menor volume de comercialização.
2023	Começa a haver uma <b>recuperação e estabilização</b> das vendas, com a curva oscilando em torno do nível médio. Isso pode estar associado a estratégias de expansão da marca ou maior aceitação do produto no mercado.
2024	Nota-se um <b>crescimento contínuo e sustentado nas vendas</b> , com a maioria dos pontos <b>acima da média histórica</b> . Esse padrão evidencia uma <b>mudança de nível</b> na série temporal, indicando que o comportamento da demanda está se consolidando em patamares mais altos.

FIGURA 1 - COMPORTAMENTO DE DEMANDA - FONTE: AUTORES,2025.

Além disso, o gráfico exibe picos regulares de venda, que podem estar relacionados a eventos promocionais, sazonalidade comercial ou datas comemorativas específicas (como *Black Friday* e Natal). Esses picos, visivelmente acima da média, reforçam a importância de ações estratégicas planejadas.

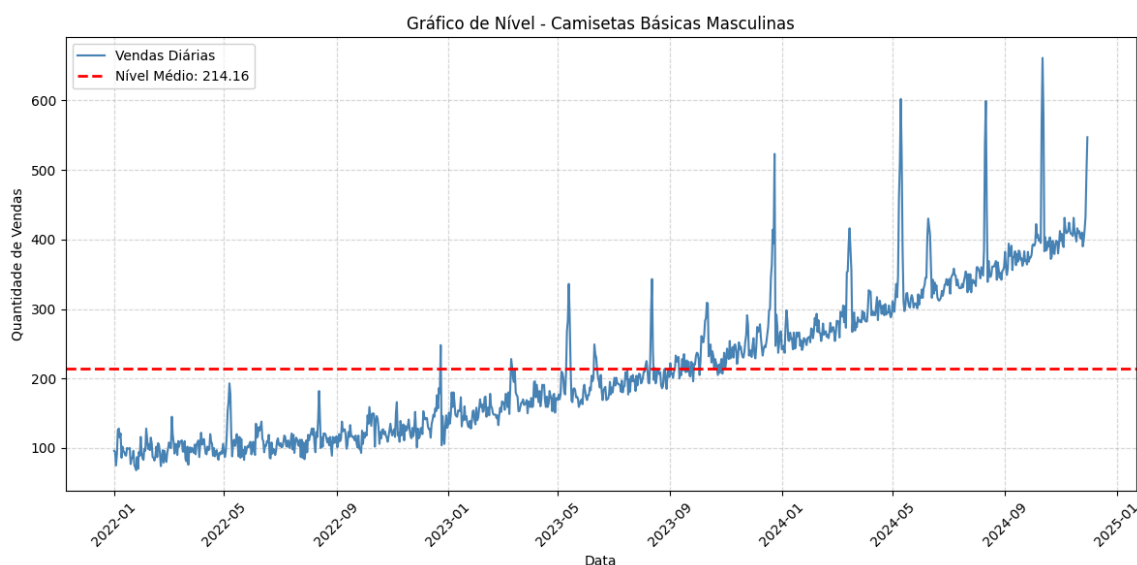


FIGURA 2 - GRÁFICO DE NÍVEL - FONTE: AUTORES,2025.

A curva de nível serve como uma ferramenta de referência para avaliação de desempenho. Sua aplicação permite identificar mudanças de comportamento na demanda ao longo do tempo, determinar períodos que exigem reforço logístico ou operacional (como os de pico), subsidiar a construção de modelos preditivos baseados em mudanças estruturais de patamar, fundamentais para o planejamento de estoques e produção.

Com isso, o gráfico de nível confirma que a Segrob Notlad vem apresentando

crescimento consistente nas vendas de camisetas básicas, com indicações claras de que a média histórica utilizada pode em breve tornar-se obsoleta diante do novo patamar de desempenho observado.

#### **4.2.2 - Autocorrelação**

Apresenta as médias mensais de vendas de camisetas básicas masculinas no período de Janeiro de 2022 a Novembro de 2024. Esse tipo de análise permite avaliar quanto os valores de uma série temporal estão relacionados com seus próprios valores defasados no tempo (lags), sendo uma ferramenta fundamental para identificar padrões sazonais e dependência temporal.

No gráfico, podemos observar que as primeiras defasagens (lags 1 a 5) apresentam valores de autocorrelação altos e positivos (acima de 0,6), indicando que as vendas de um mês estão fortemente correlacionadas com os meses imediatamente anteriores. A partir do lag 6, a autocorrelação vai diminuindo progressivamente, sugerindo uma perda gradual de memória da série ao longo do tempo. A curva cruza a linha de significância estatística (faixa azul claro) por volta do lag 10 a 11, momento em que os valores deixam de ser significativamente diferentes de zero. Isso indica que, após cerca de 10 meses, a influência das vendas passadas sobre as atuais se torna desprezível. A presença de autocorrelação significativa nos primeiros meses reforça a existência de padrões estruturais e possíveis tendências, que devem ser considerados na construção de modelos preditivos.

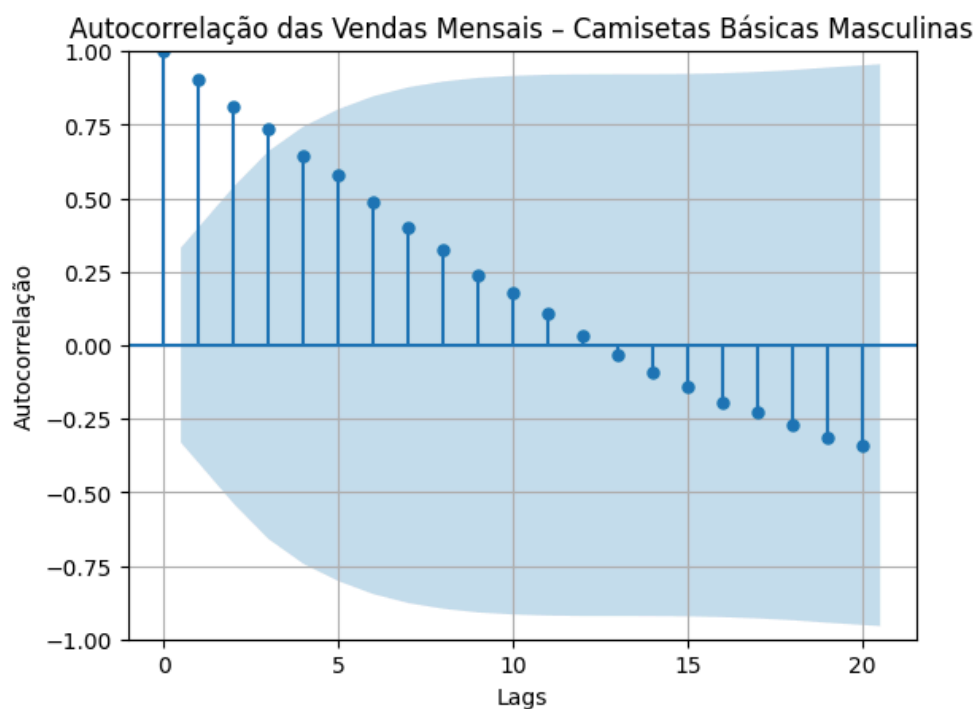


FIGURA 3 - AUTOCORRELAÇÃO - FONTE: AUTORES,2025.

A análise de autocorrelação confirma que as vendas mensais seguem uma estrutura temporal dependente, ou seja, os valores atuais são influenciados por períodos anteriores, especialmente nos primeiros 6 a 8 meses. Isso é consistente com comportamentos típicos de negócios de moda rápida, em que decisões de produção e marketing têm efeitos duradouros, mas não permanentes. A compreensão desses padrões é essencial para o desenvolvimento de modelos de previsão mais robustos, ou modelos baseados em janelas de tempo.

#### 4.2.3 - Média Quarter

Exibe a evolução da média trimestral de vendas de camisetas básicas masculinas da marca Segrob Notlad, com base nos dados compreendidos entre o primeiro trimestre de 2022 (Q1/2022) e o terceiro trimestre de 2024 (Q3/2024). A análise foi realizada a partir do agrupamento dos dados de vendas por trimestres civis e cálculo da média de unidades vendidas em cada período. A linha do gráfico mostra uma tendência de crescimento contínuo e acentuado da demanda ao longo do tempo:

- No início da série (Q1/2022), a média trimestral de vendas era inferior a 120 unidades por dia.

- A partir de Q3/2022, observa-se uma trajetória de alta constante, sem quedas ou estabilizações significativas.
- Em Q3/2023 há um salto mais expressivo, possivelmente relacionado a ações promocionais ou expansão da marca.
- O último ponto do gráfico, Q3/2024, apresenta uma média superior a 400 unidades diárias, representando um crescimento de mais de 250% em relação ao início da série.

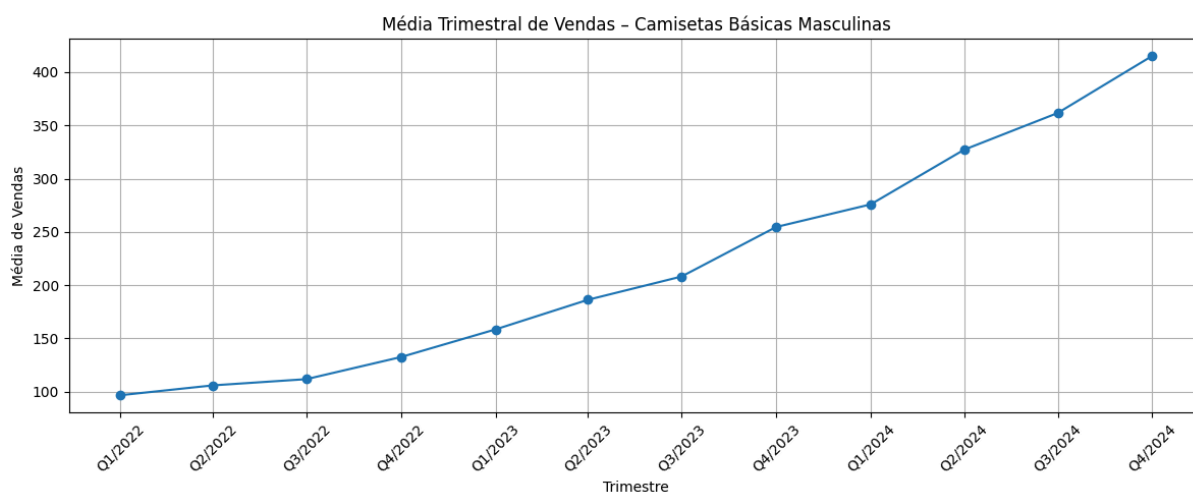


FIGURA 4 - GRÁFICO DE MÉDIA QUARTER - FONTE: AUTORES, 2025.

Este gráfico evidencia uma clara mudança de patamar na operação comercial da empresa. O comportamento crescente e estável da curva indica que o produto em análise apresenta forte aceitação no mercado e potencial de escalabilidade. Do ponto de vista analítico, a evolução trimestral:

- Confirma a validação do produto ao longo do tempo.
- Reforça a necessidade de ajustar modelos preditivos para acomodar uma tendência de alta acentuada.
- Serve como base para projeções futuras mais agressivas de produção e logística.

A análise trimestral suaviza variações diárias e mensais, permitindo visualizar com mais clareza a tendência estrutural de crescimento da demanda. Trata-se, portanto, de uma métrica importante para planejamento estratégico e previsão de longo prazo.

#### 4.2.4 - Registro Diário de Vendas

Apresenta o registro diário das vendas de camisetas básicas masculinas da marca Segrob Notlad, no período de janeiro de 2022 a novembro de 2024. O gráfico permite observar de forma clara o comportamento da demanda ao longo do tempo. No início da série, os volumes de venda diários são mais baixos, oscilando majoritariamente entre 80 e 150 unidades. Com o passar dos meses, especialmente a partir de 2023, nota-se uma tendência consistente de crescimento. Esse aumento se intensifica ao longo de 2024, quando os registros diários frequentemente ultrapassam 300 unidades, chegando, em alguns casos, a ultrapassar a marca de 600 unidades em um único dia.

Além da tendência geral de crescimento, o gráfico também evidencia a presença de picos regulares e espaçados, que podem estar relacionados a eventos promocionais, campanhas de liquidação, datas comemorativas ou ações de marketing da marca — o que é típico do setor de *fast fashion*. A progressão contínua da linha de vendas, com redução gradual dos dias de baixa performance, sugere um processo de consolidação de mercado, fidelização de clientes e possível ampliação da rede de distribuição.

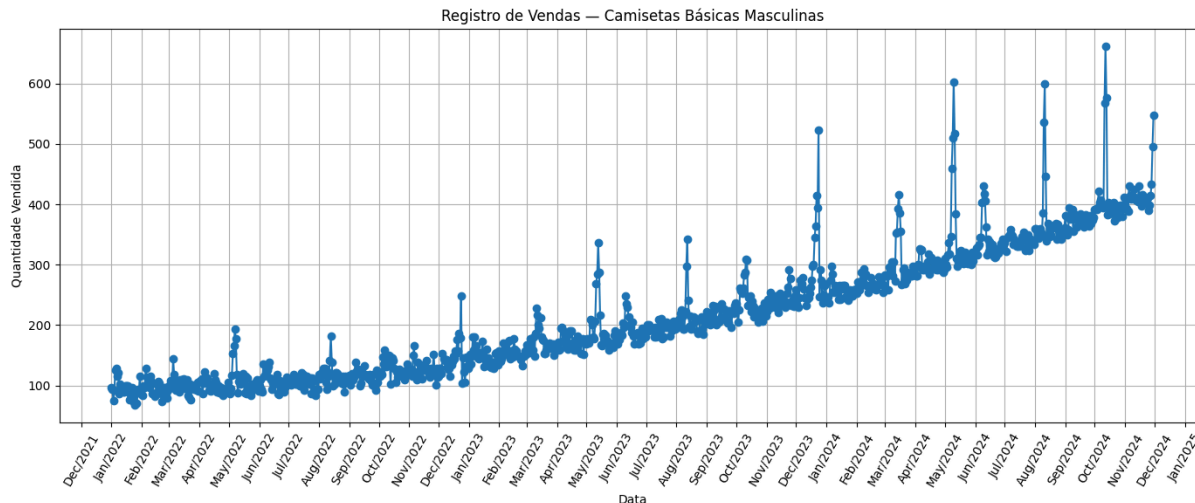


FIGURA 5 - GRÁFICO DE LINHA - FONTE: AUTORES,2025.

Em termos analíticos, o gráfico mostra que o padrão de vendas da empresa não é apenas aleatório ou sazonal, mas estruturado, com sinais claros de expansão. Essa leitura visual corrobora os dados anteriores de média trimestral e autocorrelação, reforçando que a empresa não só cresce, como cresce de forma consistente e sustentada ao longo do tempo.

#### 4.2.5 - Gráfico de Tendências

Apresenta a tendência mensal das vendas de camisetas básicas masculinas da marca Segrob Notlad, no período de janeiro de 2022 a novembro de 2024. O gráfico foi construído com base na soma das vendas realizadas a cada mês, o que permite suavizar variações diárias e destacar a evolução geral da demanda ao longo do tempo. A curva resultante mostra um padrão de crescimento nítido e contínuo, com poucos sinais de retração ou estagnação.

No início da série, os volumes mensais se concentram abaixo de 4.000 unidades, caracterizando um estágio inicial da operação. A partir do segundo semestre de 2022, nota-se um movimento ascendente consistente, que se intensifica ao longo de 2023 e se consolida em 2024 com volumes superiores a 12.000 unidades mensais. Esse padrão indica não apenas crescimento de vendas, mas também o fortalecimento da marca no mercado, possivelmente impulsionado por estratégias bem-sucedidas de marketing, expansão de canais de venda e fidelização do público-alvo.

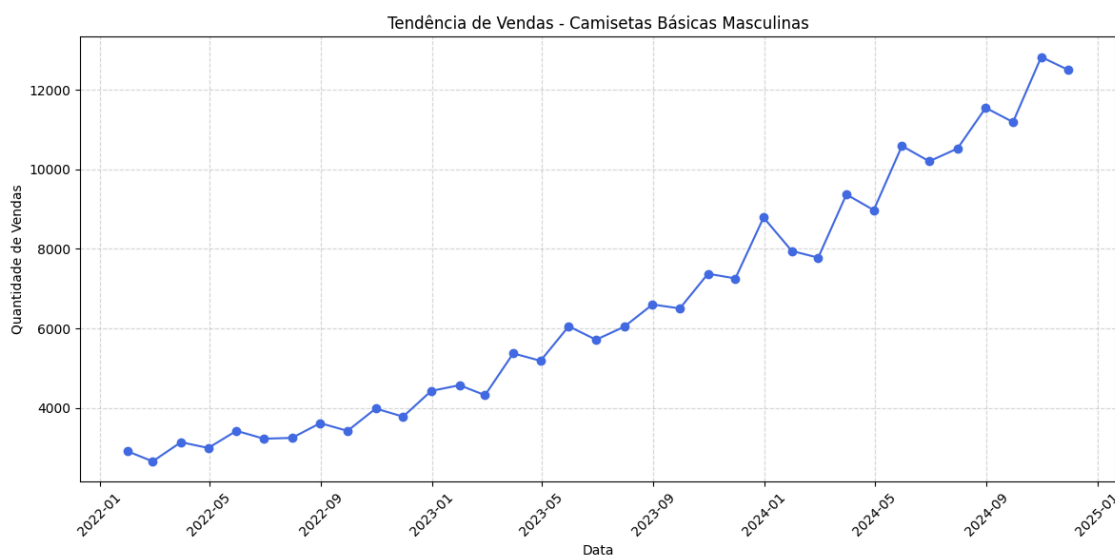


FIGURA 6 - GRÁFICO DE TENDÊNCIAS - FONTE: AUTORES,2025.

Além disso, a regularidade da curva ao longo dos trimestres reforça a hipótese de uma demanda crescente e sustentada, o que é um indicativo importante para o planejamento de estoques, logística e produção. A análise da tendência permite prever, com base em dados

históricos, a continuidade desse crescimento para períodos futuros, portanto, sendo uma ferramenta essencial para decisões estratégicas no contexto do *fast fashion*.

#### **4.2.6 - Histograma Sazonalidade Temporal**

Apresenta a distribuição total de vendas de camisetas básicas masculinas ao longo dos dias da semana, evidenciando a existência de um padrão sazonal semanal. A análise dos dados entre Janeiro de 2022 e Novembro de 2024 revela que as vendas tendem a aumentar progressivamente de segunda-feira até sábado, com pico máximo aos sábados, e leve queda no domingo. Já o outro histograma de sazonalidade temporal (Figura 8) apresenta a distribuição da média diária de vendas nos dias da semana, reforçando a tendência observada na Figura 7. Mesmo ao considerar a média, os sábados continuam se destacando como o dia com maior desempenho, o que confirma uma consistência na alta demanda nesse dia. Esse comportamento sugere que o consumidor da marca está mais ativo comercialmente nos fins de semana, o que pode estar associado a maior disponibilidade de tempo para compras, seja em lojas físicas ou online, e a campanhas promocionais pontuais.

Apesar da diferença entre os dias não ser abrupta, a constância do crescimento ao longo da semana e o destaque dos finais de semana apontam para uma oportunidade estratégica de reforço de ações de marketing, estoques e atendimento nessas datas. O padrão identificado é coerente com o perfil de consumo do público jovem e urbano da marca, e pode ser utilizado como insumo em modelos de previsão de demanda que levem em consideração a variável "dia da semana" como fator explicativo.



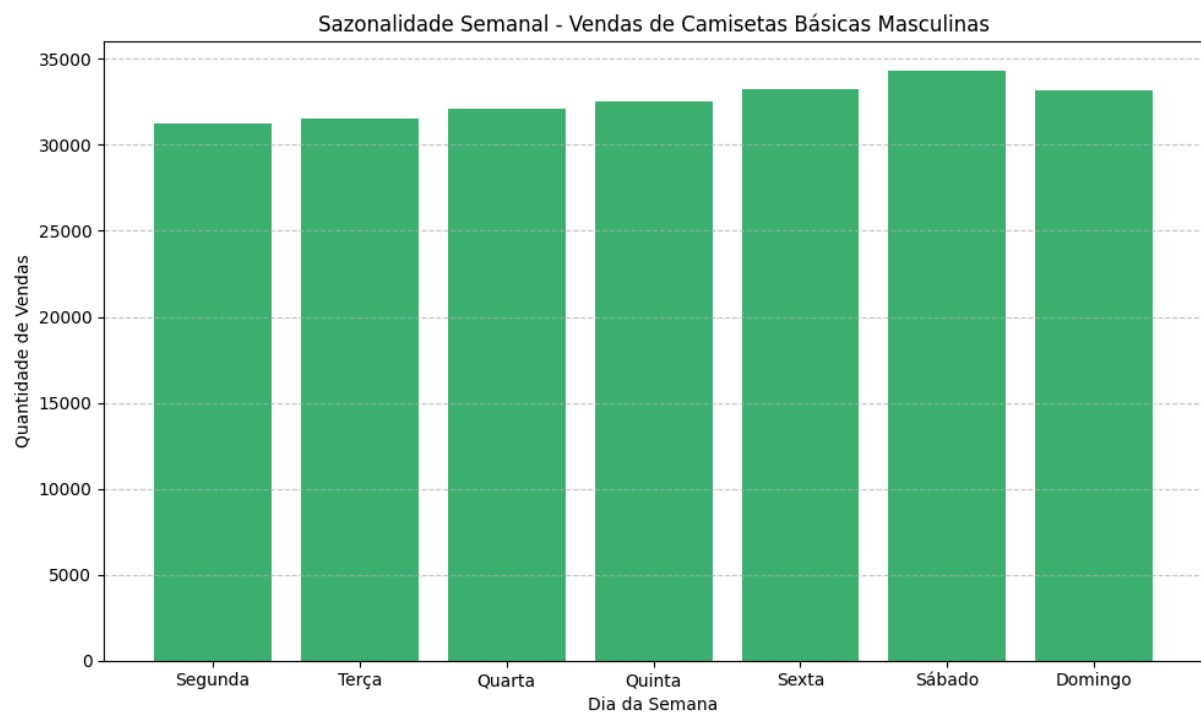


FIGURA 7 - SAZONALIDADE TEMPORAL TOTAL - FONTE: AUTORES,2025.

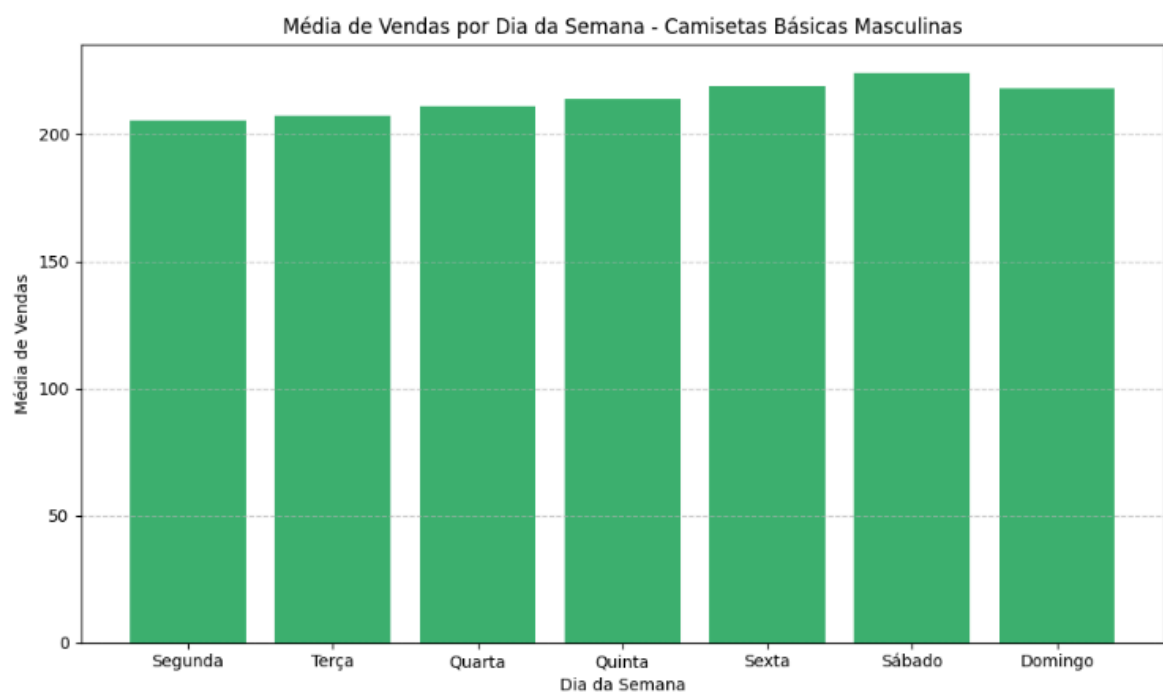


FIGURA 8 - SAZONALIDADE TEMPORAL EM MÉDIA - FONTE: AUTORES,2025.

## 4.2.7 - Modelos Preditivos

### 4.2.7.1 Cumulativo e Naive

No contexto da previsão de vendas diárias, é comum iniciar análises com modelos mais simples, que servem como benchmarks para métodos mais sofisticados. Entre os

modelos mais básicos estão o modelo Naive e o modelo Cumulativo, ambos aplicados neste estudo a uma série temporal de vendas com tendência crescente ao longo dos anos.

O modelo Naive assume que a melhor estimativa para o próximo período é o último valor observado. Essa abordagem, embora extremamente simples, é útil quando a série apresenta forte persistência. No gráfico correspondente, observa-se que a linha de previsão (vermelha tracejada) se mantém constante durante o período de teste, refletindo exatamente o último valor do conjunto de treino. Essa característica torna o modelo inadequado para séries com tendência, como é o caso das vendas analisadas, que exibem crescimento consistente ao longo do tempo. A comparação com os dados reais (linha azul) evidencia o descompasso entre a previsão e a realidade, especialmente nos meses mais recentes, em que há picos significativos de venda.

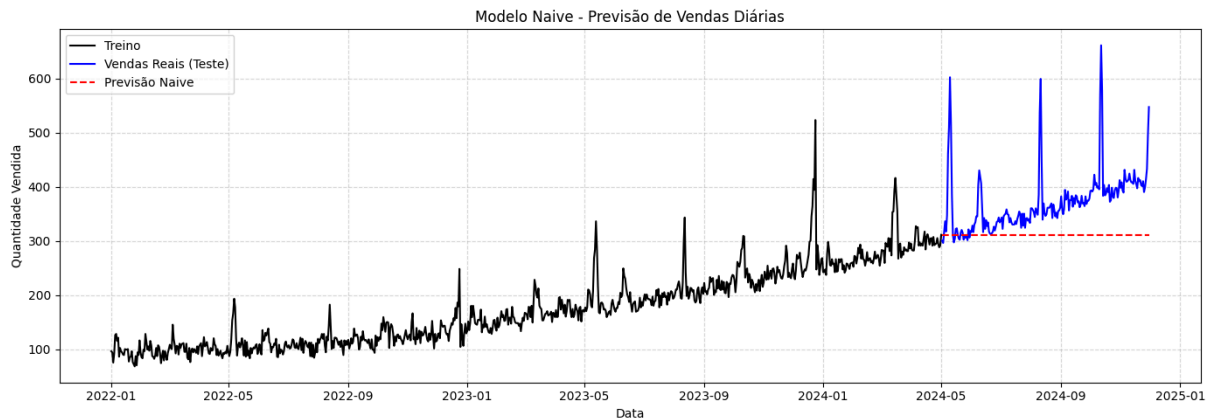


FIGURA 9 - MODELO NAIVE - FONTE: AUTORES,2025.

O modelo Cumulativo baseia-se na média de todos os valores observados até o ponto de corte. Ao suavizar as flutuações da série, ele oferece uma previsão mais conservadora e menos sensível a variações abruptas. No gráfico que representa esse modelo, a previsão (linha verde tracejada) também permanece constante no período de teste, porém em um patamar inferior ao da previsão Naive. Isso reflete a média histórica mais diluída da série, resultando em uma subestimação ainda maior das vendas reais.

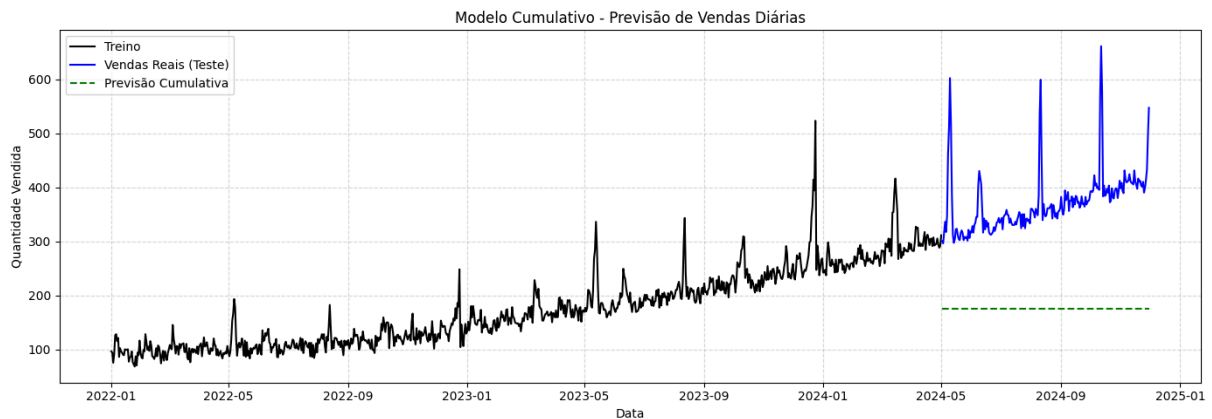


FIGURA 10 - MODELO CUMULATIVO - FONTE: AUTORES,2025.

Ambos os modelos demonstram desempenho limitado diante de uma série com tendência clara de crescimento e eventos pontuais que geram picos de venda. Embora úteis como referência inicial, suas limitações são evidentes: não capturam tendências, padrões sazonais nem reagem a variações estruturais ao longo do tempo. No entanto, essas abordagens são importantes para fins comparativos, permitindo avaliar o ganho de desempenho de modelos mais robustos, como ARIMA, Prophet ou métodos baseados em aprendizado de máquina.

#### 4.2.7.2 Média Móvel

A média móvel é uma ferramenta estatística fundamental na análise de séries temporais, utilizada para suavizar variações de curto prazo e evidenciar tendências de longo prazo em conjuntos de dados. Seu princípio consiste no cálculo da média de um número fixo de observações anteriores, proporcionando uma visão mais clara do comportamento da variável ao longo do tempo. Esse método é especialmente útil em contextos como o acompanhamento de vendas, onde oscilações diárias podem dificultar a interpretação direta dos dados brutos.

O gráfico apresentado ilustra a evolução das vendas diárias de camisetas entre 2022 e 2024, incorporando três médias móveis: de 7, 30 e 365 dias. A linha azul clara representa os dados diários, com variações pontuais que refletem picos sazonais, promoções ou eventos específicos. A média móvel de 7 dias, traçada em laranja, suaviza essas oscilações, revelando padrões semanais. Já a linha verde, correspondente à média móvel de 30 dias, fornece uma perspectiva mais estável, destacando a tendência mensal de crescimento nas vendas. Por fim, a média móvel de 365 dias, em vermelho, apresenta uma visão consolidada da trajetória de longo prazo, indicando um crescimento contínuo e consistente ao longo dos anos.

A leitura integrada dessas médias permite compreender tanto as flutuações sazonais quanto a tendência estrutural do negócio. Observa-se um aumento gradual e sustentado nas vendas, sinalizando expansão de mercado e possível consolidação da marca. A utilização de diferentes janelas de médias móveis, como exemplificado, é uma prática recomendada para apoiar decisões estratégicas em áreas como produção, logística e marketing, fornecendo uma base analítica sólida para o planejamento de ações futuras.

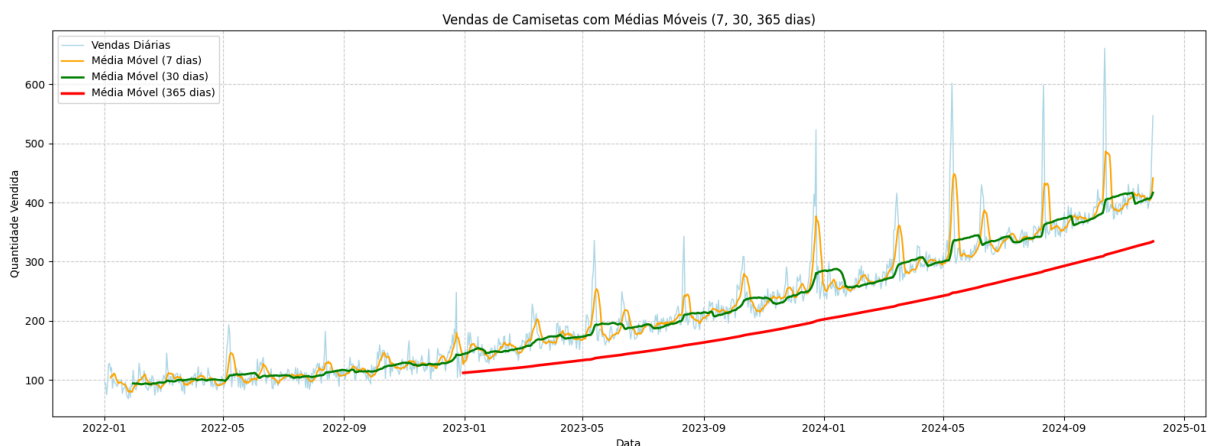


FIGURA 11 - MÉDIA MÓVEL - FONTE: AUTORES,2025.

#### 4.2.7.3 Suavização Exponencial Simples

A Suavização Exponencial Simples é uma técnica empregada na análise de séries temporais, especialmente quando o objetivo é reduzir a variabilidade dos dados e ressaltar comportamentos subjacentes. Fundamentada na atribuição de pesos exponencialmente decrescentes às observações passadas, a SES utiliza um parâmetro de suavização,  $\alpha$  (alfa), cujo valor determina o grau de sensibilidade da suavização em relação aos dados mais recentes. Valores baixos de  $\alpha$  resultam em uma série mais suavizada, menos sensível a flutuações de curto prazo, enquanto valores elevados aumentam a responsividade às variações mais recentes.

O gráfico apresentado demonstra a aplicação da SES a uma série de vendas diárias, utilizando três diferentes configurações de  $\alpha$  (0,1, 0,5 e 0,9). A linha original da série, em preto, é acompanhada por três curvas suavizadas que evidenciam comportamentos distintos. Com  $\alpha = 0,1$ , a suavização é mais acentuada, resultando em uma curva que reage lentamente a variações súbitas, ideal para mitigar ruídos em séries ruidosas. Com  $\alpha = 0,5$ , observa-se um equilíbrio entre suavização e reatividade, acompanhando mais de perto as oscilações da série sem perder a estabilidade. Já com  $\alpha = 0,9$ , a série suavizada torna-se altamente sensível,

refletindo praticamente todas as flutuações presentes nos dados originais, o que pode ser útil em contextos onde se deseja capturar alterações recentes com maior rapidez.

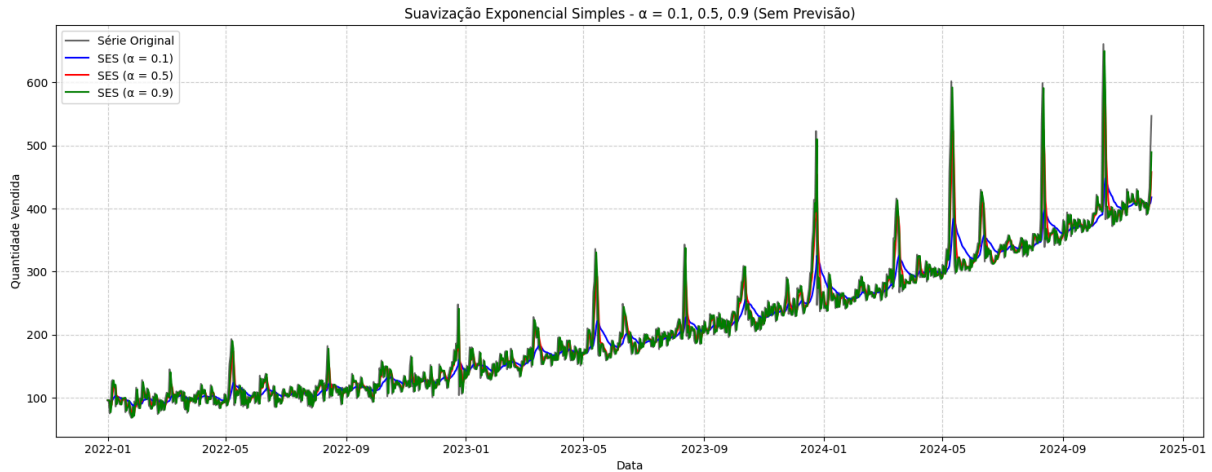


FIGURA 12 - SUAVIZAÇÃO EXPONENCIAL SIMPLES - FONTE: AUTORES,2025.

#### 4.2.7.4 Regressão Linear Simples

A regressão linear simples é uma técnica estatística fundamental utilizada para modelar a relação entre duas variáveis: uma variável dependente e uma variável independente. Trata-se da forma mais básica de regressão, em que se busca descrever, por meio de uma equação linear, como a variável explicativa (ou preditora) influencia a variável resposta ao longo do tempo ou de observações, como podemos ver na Equação 1.

Equação 1- Regressão Linear Simples

$$\hat{Y}_t = \beta_0 + \beta_1 X_t + \varepsilon_t$$

Onde:

- $\hat{Y}_t$  é o valor previsto da variável dependente no tempo  $t$ ;
- $X_t$  é a variável independente (ou preditora) no tempo  $t$ ;
- $\beta_0$  é o termo constante (intercepto), que representa o valor previsto de  $Y$  quando  $X = 0$ ;
- $\beta_1$  é o coeficiente angular da reta (inclinação), que representa a variação média de  $Y$  para cada unidade de aumento em  $X$ ;
- $\varepsilon_t$  é o erro aleatório, que contém a variação de  $Y$  não explicada por  $X$ .

Elaborado pelos autores,2025

O objetivo principal do modelo de regressão linear simples é encontrar a melhor reta

de ajuste que possa prever a variável dependente com o menor erro possível. Para isso, utiliza-se o método dos mínimos quadrados, que estima os coeficientes do modelo de forma a minimizar a soma dos quadrados dos resíduos, isto é, das diferenças entre os valores observados e os valores previstos. Embora simples, esse modelo assume que há uma relação linear entre as variáveis e que os erros são independentes, com média zero e variância constante.

O gráfico da figura 13 exibe uma análise de regressão linear simples entre as vendas reais (variável independente no eixo X) e as vendas previstas pelo modelo (variável dependente no eixo Y). Cada ponto do gráfico representa uma observação do conjunto de dados, e a linha laranja é a reta de regressão ajustada com a equação  $y=0,96x+8,45$ .

A interpretação dessa reta indica que, para cada unidade adicional nas vendas reais, o modelo prevê um aumento de aproximadamente 0,96 unidades nas vendas previstas, com um intercepto de 8,45. Esse coeficiente angular próximo de 1 sugere uma forte correlação linear entre os valores reais e os previstos, o que é desejável em modelos preditivos — idealmente, em um modelo perfeito, teríamos uma reta com inclinação 1 e intercepto 0, ou seja,  $y=x$ .

Visualmente, a concentração dos pontos ao redor da linha de regressão reforça a capacidade do modelo em replicar o comportamento da variável real. No entanto, observa-se certa dispersão crescente à medida que as vendas aumentam, o que pode indicar uma leve heterocedasticidade, um fenômeno onde os erros de previsão aumentam com o valor da variável independente. Além disso, a presença de outliers nas faixas mais altas de vendas reais sugere que o modelo pode ter dificuldade em prever com exatidão eventos extremos ou valores fora do padrão observado na maior parte da série.

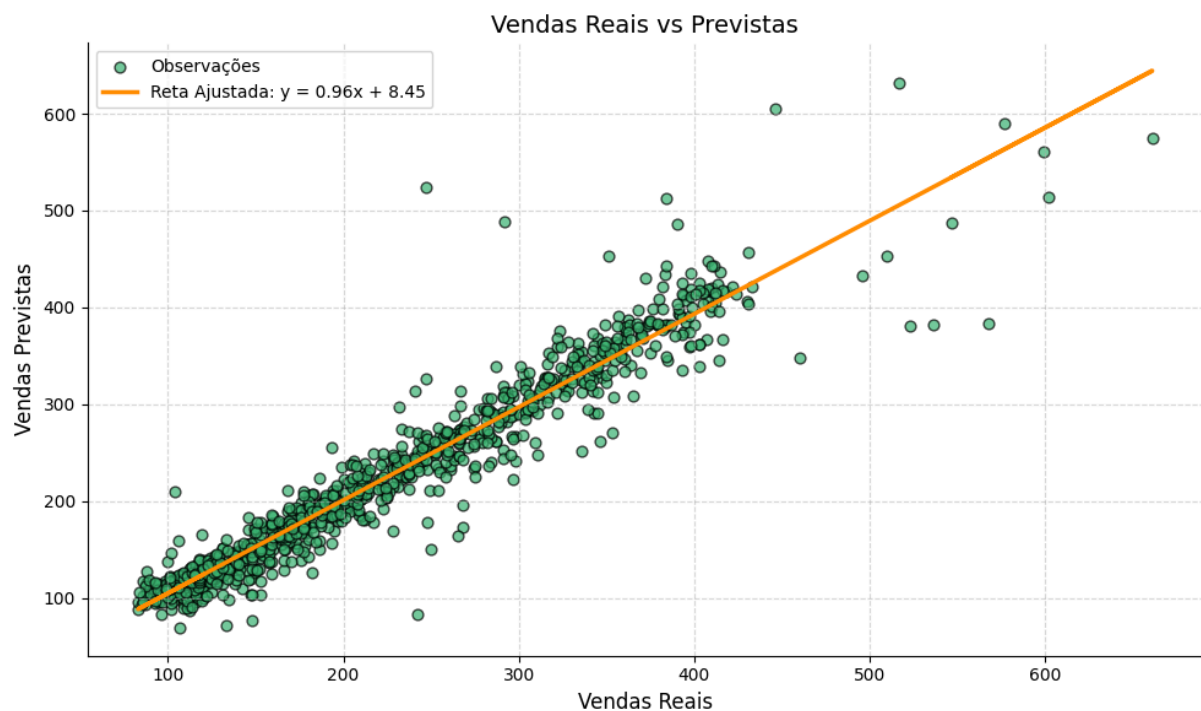


FIGURA 13 - REGRESSÃO LINEAR SIMPLES - FONTE: AUTORES,2025.

#### 4.2.7.5 Análise dos Resíduos

Os resíduos representam a diferença entre os valores reais observados e os valores previstos por um modelo. Sua análise é essencial para avaliar a precisão e a confiabilidade das previsões. Quando bem ajustado, um modelo apresenta resíduos distribuídos de forma simétrica em torno de zero, sem padrões visíveis, o que indica ausência de viés sistemático e bom desempenho preditivo. Essa distribuição ideal sugere que os erros são aleatórios e que o modelo está capturando adequadamente a estrutura dos dados.

Por outro lado, resíduos com distribuição assimétrica, presença de valores extremos ou padrões recorrentes podem indicar falhas no modelo, como sub ajuste, superajuste ou ausência de variáveis explicativas importantes. Nesses casos, a análise crítica dos resíduos permite identificar pontos de melhoria e orientar ajustes na modelagem. Assim, os resíduos não apenas quantificam o erro, mas também funcionam como uma ferramenta diagnóstica indispensável para o aperfeiçoamento de modelos preditivos.

As figuras abaixo apresentam três histogramas que descrevem, respectivamente, a distribuição das vendas reais, das vendas previstas por um modelo e dos resíduos (erros) provenientes da diferença entre os valores reais e os previstos. A análise desses gráficos

permite avaliar a precisão e o comportamento do modelo preditivo adotado, bem como a aderência das previsões ao comportamento observado nas vendas reais.

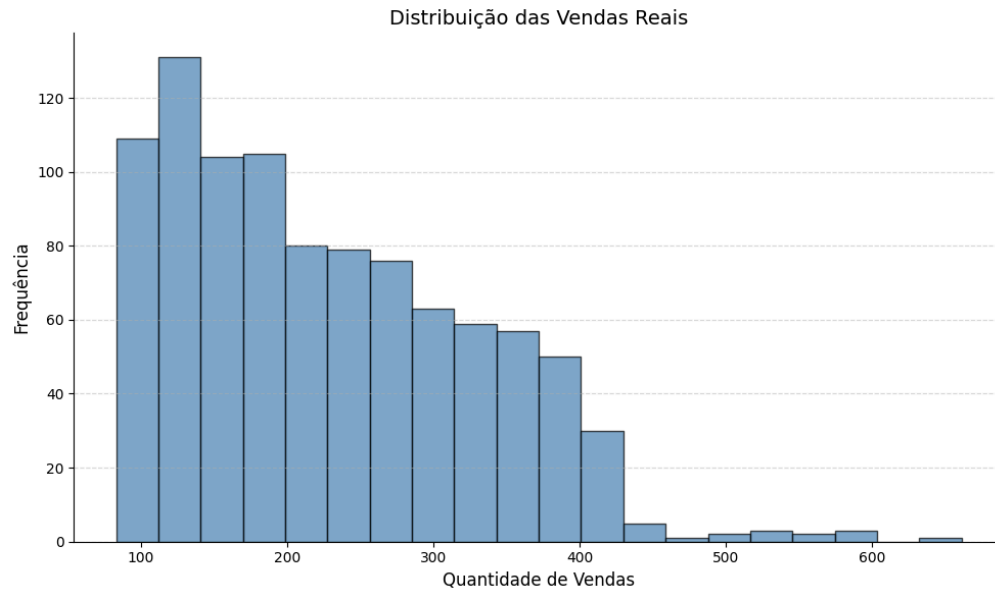


FIGURA 14 - DISTRIBUIÇÃO DE VENDAS REAIS - FONTE: AUTORES, 2025.

A Figura 14, "Distribuição das Vendas Reais", exibe uma distribuição assimétrica à direita (distribuição assimétrica positiva), indicando que a maioria dos valores de vendas reais se concentra nas faixas inferiores, entre aproximadamente 100 e 300 unidades. Há uma cauda longa que se estende até valores superiores a 600 unidades, sugerindo a presença de eventos menos frequentes com vendas excepcionalmente altas. Essa assimetria é característica comum em dados de vendas, onde picos de demanda podem ocorrer de maneira pontual e não representam o comportamento típico do mercado.



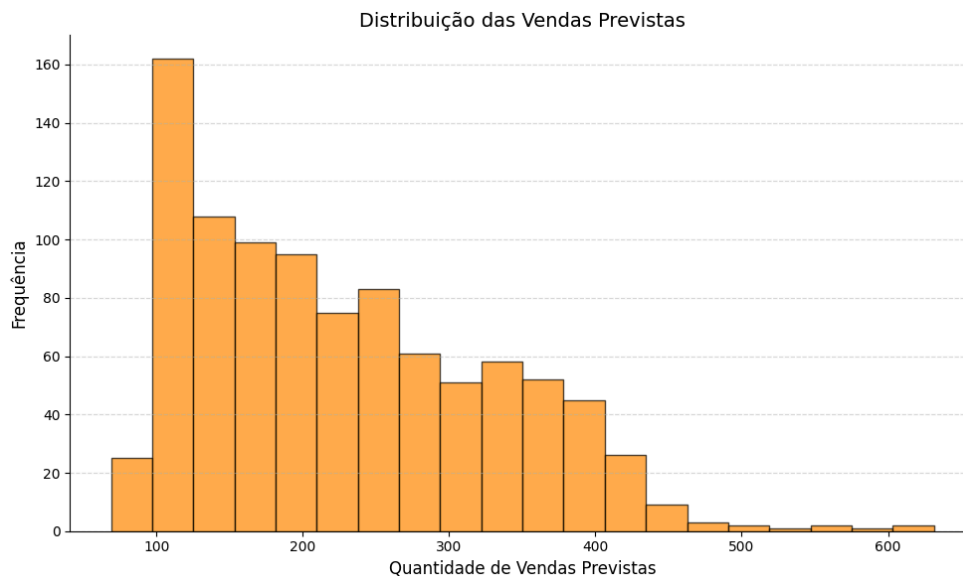


FIGURA 15 - DISTRIBUIÇÃO DE VENDAS PREVISTAS - FONTE: AUTORES,2025.

A Figura 15, "Distribuição das Vendas Previstas", observa-se uma distribuição com padrão semelhante à das vendas reais, também assimétrica à direita. No entanto, é perceptível que o modelo gera uma frequência levemente maior nas faixas mais baixas (entre 80 e 150 unidades), o que pode indicar uma tendência do modelo a subestimar ligeiramente os volumes de vendas. Ainda que as previsões acompanhem o padrão geral das vendas reais, essa leve discrepância nas frequências sugere que há espaço para ajustes nos parâmetros do modelo, a fim de melhorar a acurácia especialmente nos extremos da distribuição.

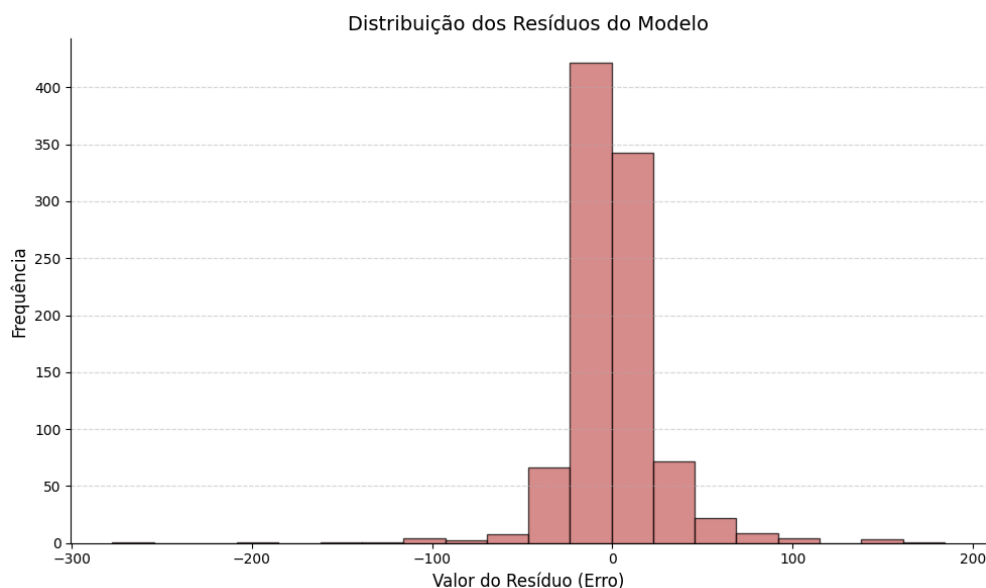


FIGURA 16 - DISTRIBUIÇÃO DOS RESÍDUOS DO MODELO - FONTE: AUTORES,2025.

Na Figura 16, "Distribuição dos Resíduos do Modelo", oferece uma visão crítica da qualidade do modelo preditivo. A distribuição dos resíduos está centrada em torno de zero e possui uma forma aproximadamente simétrica, o que é um indicativo positivo de que não há viés sistemático nas previsões — ou seja, o modelo não tende a superestimar ou subestimar de forma consistente. A maior parte dos erros está concentrada em um intervalo estreito, com a frequência mais alta próxima ao valor zero, indicando que, na maioria dos casos, os erros de previsão são pequenos. No entanto, também se observa a presença de valores extremos, tanto negativos quanto positivos, o que revela que o modelo ainda comete erros significativos em algumas instâncias, especialmente nos pontos de cauda da distribuição de vendas.

Os gráficos apresentados acima sugerem que o modelo preditivo utilizado possui um desempenho razoável, sendo capaz de capturar o padrão geral das vendas. Sua principal virtude está na simetria da distribuição dos resíduos e na centralização em torno de zero, o que denota ausência de viés. No entanto, a análise também evidencia oportunidades de melhoria, especialmente no tratamento de casos extremos e na correção de uma possível tendência à subestimação das vendas em determinados intervalos. Ajustes no modelo, como a consideração de variáveis adicionais ou a reformulação da função de perda, podem contribuir para reduzir os erros em cenários mais atípicos e refinar ainda mais a capacidade preditiva.

Os gráficos das Figuras 17, 18 e 19 apresentados revelam importantes aspectos sobre o comportamento dos resíduos do modelo de previsão de vendas. A Figura 17, mostra os resíduos em função das vendas previstas. A distribuição dos pontos em torno da linha de base (zero) é razoavelmente simétrica, com ligeira dispersão crescente à medida que as vendas previstas aumentam. Isso sugere a presença de heterocedasticidade, ou seja, a variabilidade dos erros tende a crescer com valores maiores de previsão. A reta de regressão ajustada, com inclinação próxima de zero, reforça que não há um viés evidente na média dos resíduos, embora o aumento da dispersão possa indicar que o modelo tem mais dificuldade em prever valores elevados de vendas com precisão.

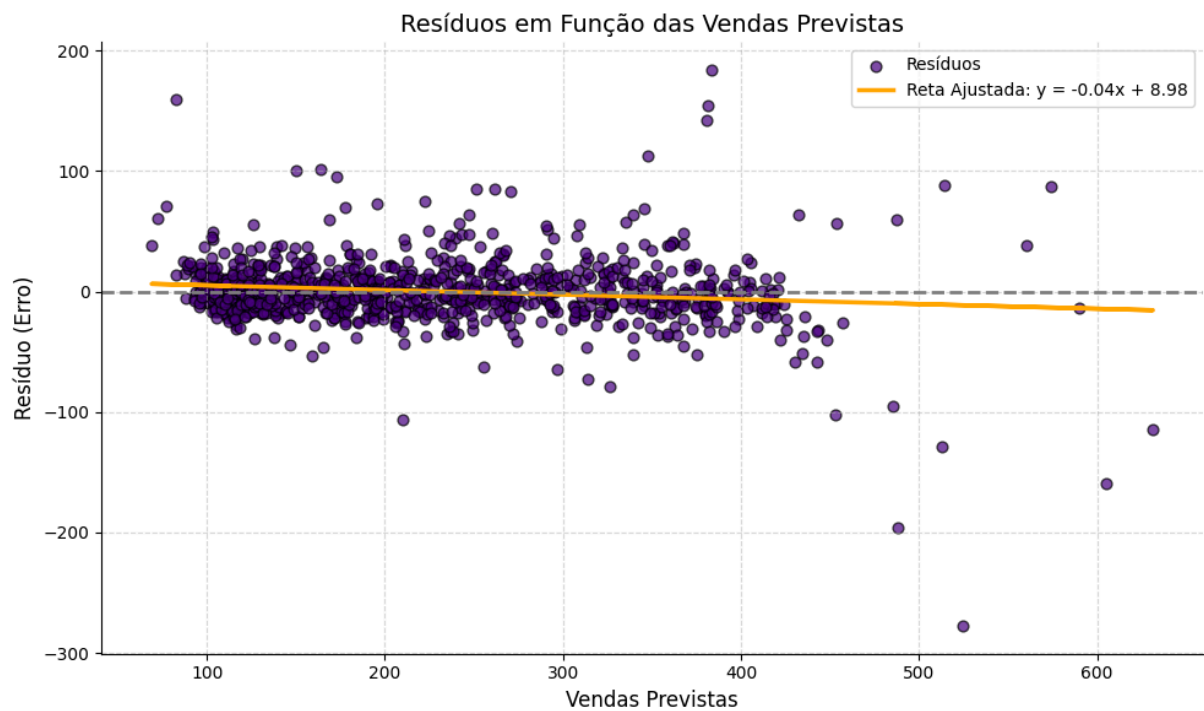


FIGURA 17 - RESÍDUOS EM FUNÇÃO DAS VENDAS PREVISTAS - FONTE: AUTORES, 2025.

No gráfico da Figura 18, observa-se o comportamento dos resíduos em relação aos valores reais de vendas. O padrão é semelhante ao observado anteriormente, com concentração de resíduos próximos a zero para a maioria dos pontos, mas com presença de valores extremos, tanto positivos quanto negativos. A inclinação da reta ajustada, também próxima de zero, aponta para ausência de tendência sistemática nos erros em relação aos valores reais, o que é um indicativo positivo de acurácia média. No entanto, os outliers e a dispersão irregular alertam para a necessidade de investigações mais profundas sobre eventos específicos que o modelo não está capturando adequadamente.

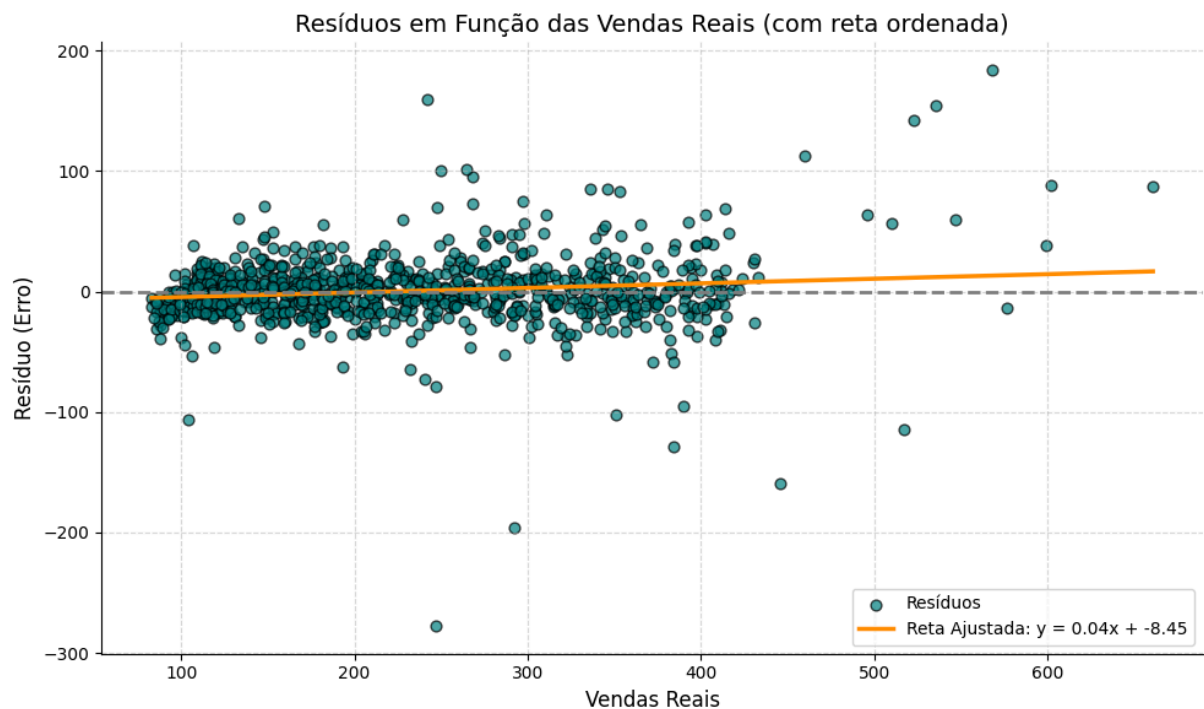


FIGURA 18 - RESÍDUOS EM FUNÇÃO DAS VENDAS REAIS - FONTE: AUTORES,2025.

A evolução temporal dos resíduos, apresentada na Figura 19, complementa a análise ao mostrar a variação dos erros ao longo do tempo. A maior parte dos resíduos oscila em torno de zero, como esperado, mas há picos relevantes em determinados períodos, evidenciando a ocorrência de erros mais expressivos em momentos específicos. Essa flutuação pode estar associada a fatores sazonais, alterações externas ao sistema modelado ou limitações temporais da base de dados utilizada para treinamento. Apesar disso, não se observa um padrão de tendência ou ciclo claro, o que sugere que o modelo não está sistematicamente errando em determinadas épocas, mas sim enfrentando dificuldades pontuais.

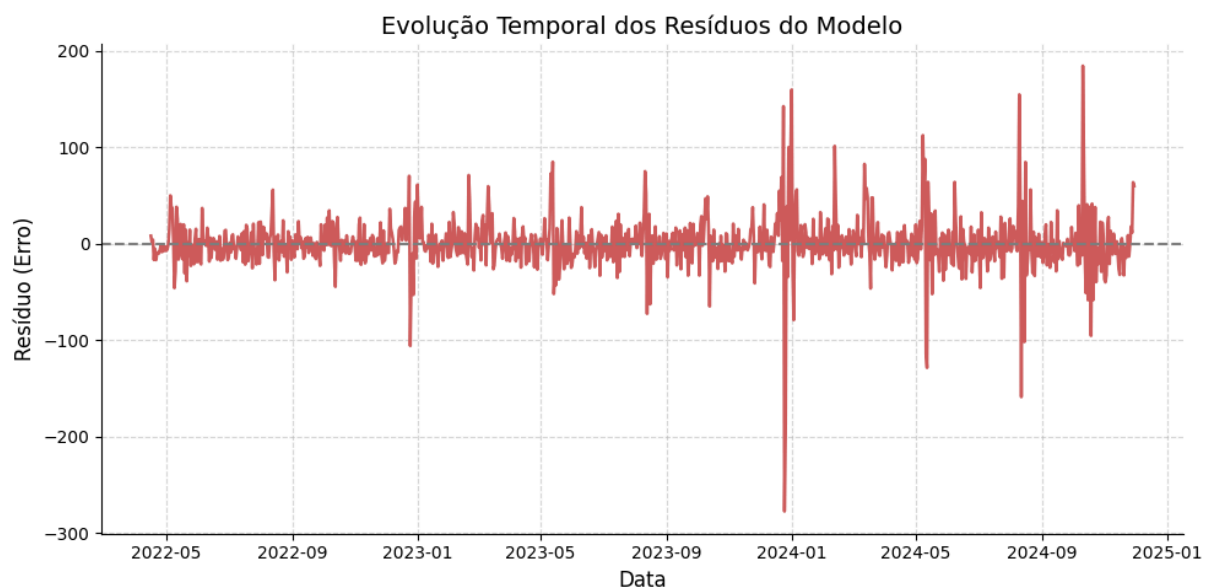


FIGURA 19 - EVOLUÇÃO TEMPORAL DOS RESÍDUOS DO MODELO - FONTE: AUTORES, 2025.

Em conjunto, os três gráficos demonstram que o modelo apresenta desempenho geral satisfatório, com ausência de viés médio, mas requer melhorias no tratamento de casos extremos e na estabilidade de previsões para altos volumes de vendas. A análise dos resíduos revela-se, portanto, uma ferramenta indispensável não apenas para validar o modelo, mas também para orientar estratégias de refinamento e aumentar sua robustez preditiva.

#### 4.2.7.6 Validação do Modelo

O gráfico de evolução do  $R^2$  ao longo do tempo (Figura 20) obtida por meio de regressões lineares móveis (Rolling OLS) mostra como a capacidade explicativa do modelo preditivo variou entre janeiro de 2022 e novembro de 2024. A linha azul representa o  $R^2$  por janela de tempo, enquanto a linha tracejada em 0,5 serve como referência crítica. Observa-se que, no início da série, o modelo apresentou baixos níveis de explicação (abaixo de 0,5), com alta flutuação, o que indica instabilidade e potencial ausência de variáveis explicativas relevantes nas primeiras fases do período analisado.

A partir de meados de 2023, o  $R^2$  se estabiliza acima da linha de 0,5, chegando a ultrapassar 0,75 em alguns momentos, especialmente entre o segundo trimestre de 2023 e meados de 2024. Isso sugere que, neste intervalo, os lags, dummies sazonais e demais variáveis explicativas conseguiram capturar de forma mais eficaz os padrões da demanda. A elevação e estabilidade do  $R^2$  reforçam a adequação do modelo em períodos mais recentes,

mas também alertam para a necessidade de atenção redobrada nas previsões realizadas com base em períodos instáveis, como o primeiro ano da série.

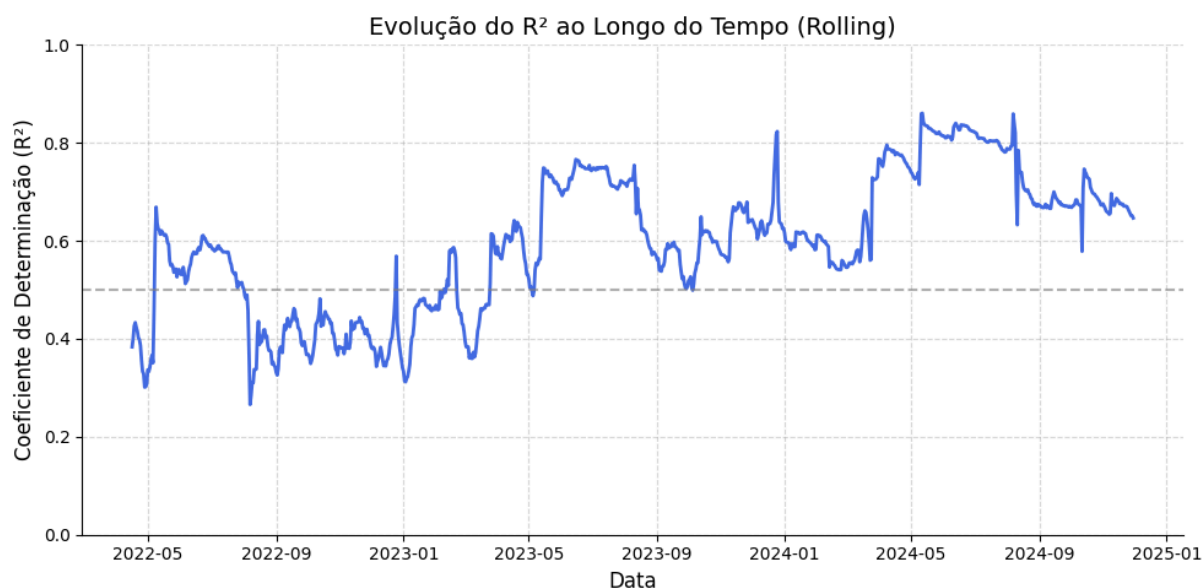


FIGURA 20 - EVOLUÇÃO DO  $R^2$  AO LONGO DO TEMPO - FONTE: AUTORES, 2025.

O gráfico comparativo de vendas (**Figura 21**) apresenta o desempenho do modelo de regressão linear com janelas móveis (Rolling OLS), utilizando 10 lags (defasagens) como variáveis explicativas para prever a demanda diária de camisetas básicas. A curva azul representa as vendas reais observadas entre janeiro de 2022 e novembro de 2024, enquanto a curva laranja exibe os valores previstos pelo modelo ao longo do mesmo período.

Observa-se que o modelo foi capaz de acompanhar razoavelmente bem a tendência geral de crescimento da série, especialmente a partir de meados de 2023. No entanto, verifica-se uma dificuldade recorrente na predição de picos de venda abruptos, nos quais o modelo tende a superestimar ou subestimar significativamente os valores reais. Essas discrepâncias indicam que os 10 lags utilizados capturam parte da estrutura temporal da série, mas não são suficientes para explicar eventos exógenos, como campanhas promocionais, feriados móveis ou variações climáticas. Ainda assim, o alinhamento estrutural entre as duas curvas reforça a utilidade do modelo como ferramenta de apoio ao planejamento operacional, desde que complementado por variáveis adicionais para maior precisão em períodos críticos.

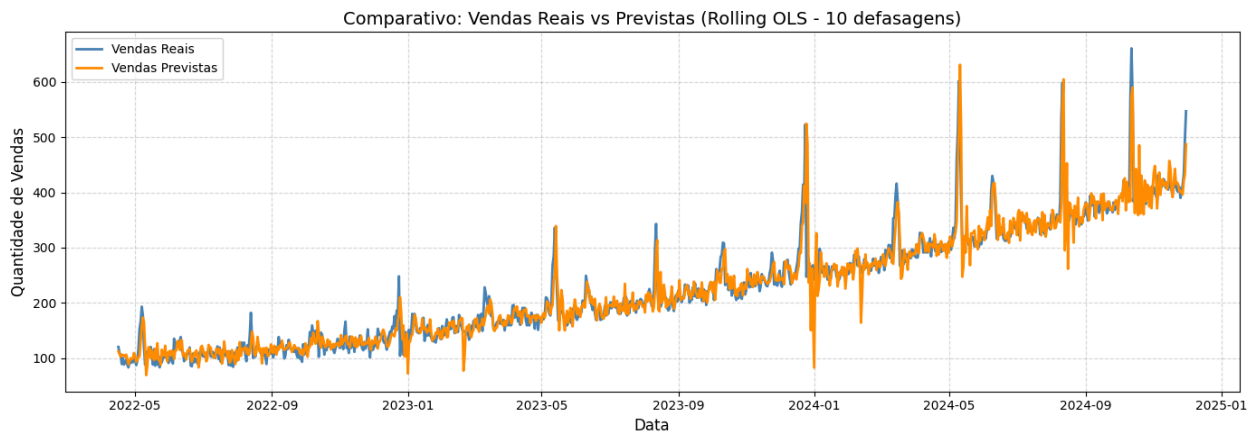


FIGURA 21 - COMPARATIVO: VENDAS REAIS x PREVISTAS - FONTE: AUTORES, 2025.

A análise do coeficiente de determinação  $R^2$  por meio da regressão linear com janelas móveis (Rolling OLS) permitiu avaliar a estabilidade temporal do modelo preditivo de demanda da Segrob Notlad. Observou-se que o  $R^2$  médio se manteve em torno de 0,596 ao longo do período de janeiro de 2022 a novembro de 2024, com flutuações relevantes. Houve momentos de boa aderência do modelo, principalmente entre meados de 2023 e meados de 2024, quando o  $R^2$  ultrapassou 0,70 em diversas janelas, indicando boa explicação da variabilidade nas vendas a partir das variáveis utilizadas (lags e sazonais).

Entretanto, períodos de baixa da explicabilidade do gráfico (Figura 22), como os observados no início de 2022 e nos últimos meses de 2024, sugerem que o modelo pode não capturar adequadamente eventos sazonais específicos ou variações de comportamento do consumidor em momentos críticos, como campanhas promocionais. Essa instabilidade reforça a importância de revisar variáveis externas no modelo e de aplicar validações adicionais antes de confiar plenamente nas previsões para dezembro de 2024, que é um mês de alta relevância para o planejamento logístico e de estoques da empresa.

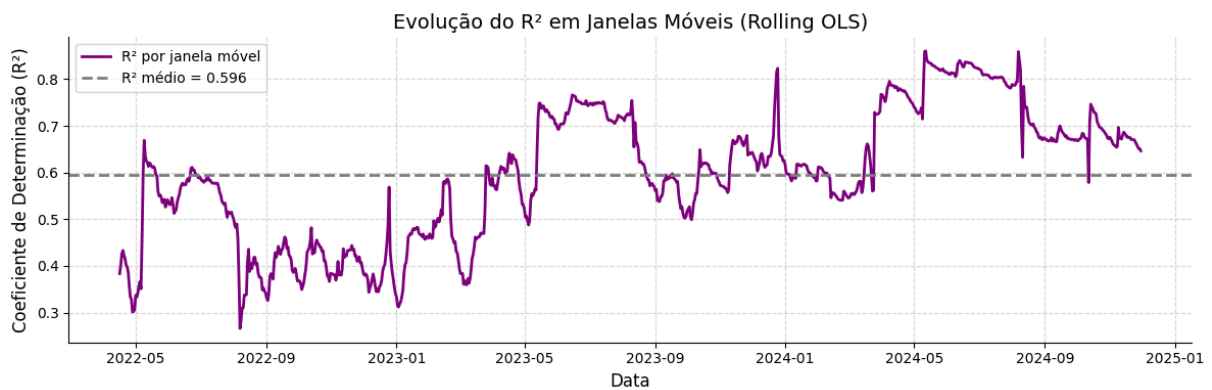


FIGURA 22 - EVOLUÇÃO DO R² EM JANELAS MÓVEIS - FONTE: AUTORES, 2025.

#### 4.2.7.7 Regressão Linear Dinâmica

O principal objetivo da regressão linear dinâmica é adaptar o modelo às mudanças graduais no padrão dos dados ao longo do tempo. Isso é feito ao permitir que o intercepto se ajuste em cada período com base em seu valor anterior, capturando, por exemplo, efeitos de crescimento de longo prazo, sazonalidades irregulares ou impactos acumulados de eventos de marketing. Embora mantenha a estrutura linear entre as variáveis, o modelo admite que o nível base da série (intercepto) não seja fixo, o que resulta em previsões mais realistas e responsivas para séries temporais em ambientes voláteis, como o da Segrob Notlad.

Equação 2 – Regressão Linear com Intercepto Dinâmico

$$Y_t = \alpha_t + \beta X_t + \varepsilon_t \quad \text{com} \quad \alpha_t = \alpha_{t-1} + \eta_t$$

Onde:

- $Y_t$ : variável dependente (resposta) no tempo  $t$ ;
- $X_t$ : variável independente (preditor);
- $\alpha_t$ : intercepto dinâmico no tempo  $t$ , que evolui ao longo do tempo;
- $\beta$ : coeficiente angular fixo (pode ser estendido para dinâmico em modelos mais complexos);
- $\varepsilon_t$ : erro aleatório da equação de observação, com média zero e variância constante;
- $\eta_t$ : erro da equação de estado, que controla a variação de  $\alpha_t$  no tempo.

Elaborado Pelos Autores

O gráfico (Figura 23) ilustra o desempenho de uma regressão linear dinâmica com intercepto adaptativo, comparando os valores observados de vendas diárias (linha azul) com os valores previstos pelo modelo (linha laranja), ao longo do período de janeiro de 2022 a



novembro de 2024. Este tipo de modelagem permite que o intercepto do modelo se ajuste ao longo do tempo, capturando variações estruturais e mudanças graduais no comportamento da série temporal, sem perder a linearidade dos coeficientes principais.

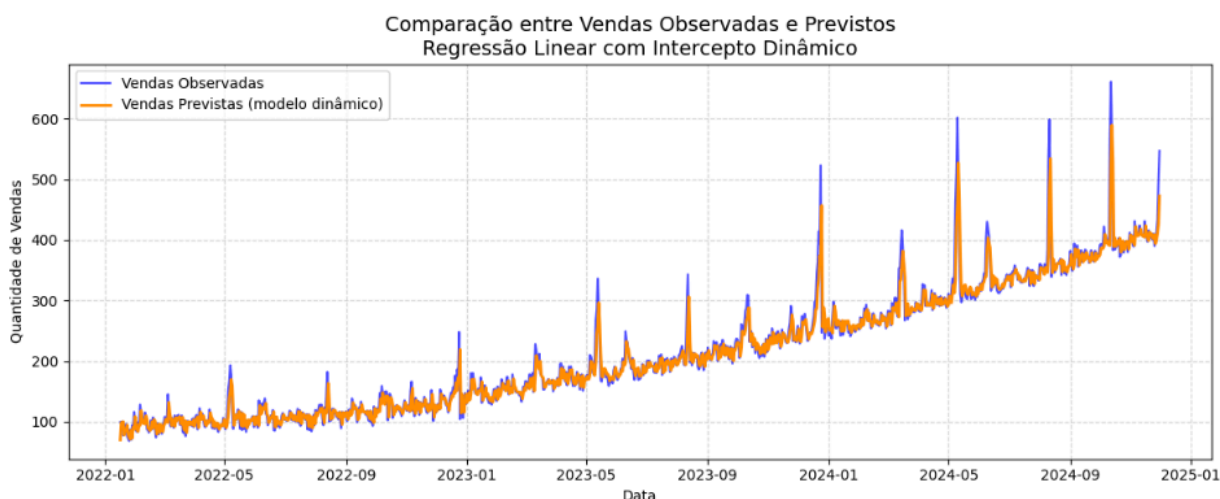


Figura 23- Comparação entre Vendas Observadas e Previstos - Fonte: Autores, 2025

Observa-se que o modelo apresenta excelente aderência à tendência de crescimento da demanda, com alta capacidade de acompanhar os valores reais inclusive nos períodos de aceleração mais acentuada a partir de 2023. Diferentemente de modelos mais rígidos, a regressão com intercepto dinâmico acompanha melhor os movimentos de subida gradual e os platôs sazonais, reduzindo erros sistemáticos de sub ou superestimação. No entanto, como é comum nesse tipo de abordagem, os picos extremos (provavelmente ligados a campanhas promocionais ou eventos sazonais) ainda apresentam pequenas discrepâncias, o que pode ser resolvido com a inclusão de dummies específicas ou variáveis externas.

### 4.3 - Entendimento Geral dos Dados de Vendas

A análise gráfica e de dados das vendas diárias de camisetas básicas masculinas da Segrob Notlad entre 2022 e 2024 revela uma forte tendência de crescimento, acompanhada por sazonalidades bem definidas e picos regulares de venda. Verificou-se que os meses de maio e dezembro se destacam por picos de demanda, possivelmente associados ao Dia das Mães, mudança brusca de tempo em várias regiões do país e festas de fim de ano. O comportamento semanal também é significativo: as vendas são mais altas aos sábados,

evidenciando maior atividade do consumidor no fim de semana, mostrando uma possível força da loja física nas vendas.

A autocorrelação mostra que há persistência no comportamento de consumo, com padrão sazonal próximo de 12 meses. Além disso, a presença de picos de venda recorrentes, com intervalos regulares, sugere que ações promocionais mensais podem ter impacto direto sobre as vendas, já que esses aumentos não se explicam por sazonalidades anuais nem por comportamento de dias da semana.

O aumento expressivo das vendas ao longo de 2024 pode ser atribuído a fatores como o fortalecimento da fidelização de clientes e o crescimento do interesse por peças básicas e versáteis, alinhadas ao conceito de praticidade do *fast fashion*, que se conecta bem com o público da marca.

Em resumo, a interpretação dos gráficos traz um entendimento do comportamento da demanda: trata-se de uma série com crescimento sustentado, forte influência de datas sazonais e resposta de possíveis campanhas promocionais recorrentes.

#### **4.4 - Previsão de vendas para Dezembro de 2024**

A Tabela 3 apresenta os resultados da previsão diária de vendas de camisetas básicas masculinas para o mês de dezembro de 2024, obtida por meio do modelo Prophet, treinado com os dados históricos entre Janeiro de 2022 e Novembro de 2024. Os valores indicam a estimativa do volume de vendas esperado para cada dia do mês, considerando a tendência e a sazonalidade semanal observadas no comportamento do consumidor.

Verifica-se uma variação cíclica nas previsões, com volumes mais baixos no início da semana e picos concentrados aos sábados. Esse padrão reforça o comportamento de compra identificado nos dados históricos, em que os fins de semana se destacam como os períodos de maior atividade comercial.

DATA	VENDAS (UN)
01/12/2024	429
02/12/2024	417
03/12/2024	420
04/12/2024	426
05/12/2024	430
06/12/2024	437
07/12/2024	445
08/12/2024	443
09/12/2024	432
10/12/2024	436
11/12/2024	443
12/12/2024	448
13/12/2024	455
14/12/2024	464
15/12/2024	461
16/12/2024	451
17/12/2024	455
18/12/2024	460
19/12/2024	464
20/12/2024	471
21/12/2024	477
22/12/2024	473
23/12/2024	461
24/12/2024	463
25/12/2024	466
26/12/2024	468
27/12/2024	472
28/12/2024	477
29/12/2024	471
30/12/2024	456
31/12/2024	456

TABELA 3 - PREVISÃO DE VENDAS PARA DEZEMBRO/2024 - FONTE: AUTORES,2025.

Além da tabela, a Figura 9 ilustra graficamente a distribuição diária dessas previsões, facilitando a visualização das oscilações e tendências ao longo do mês. Essa representação gráfica reforça a consistência do padrão semanal e evidencia os dias com maior potencial de vendas, fornecendo subsídios importantes para o planejamento de estoque, logística e ações promocionais específicas para o período.

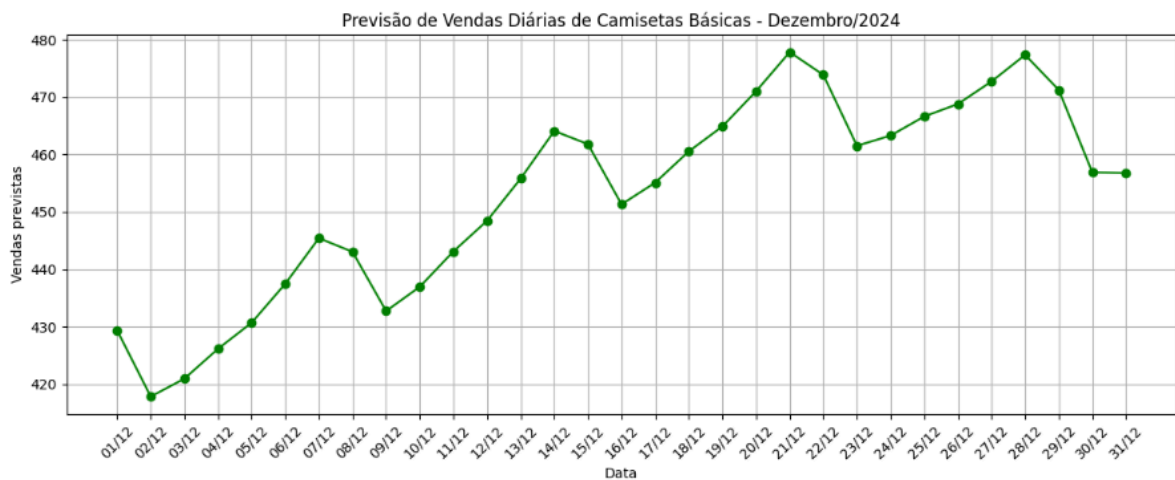


FIGURA 9 - MODELO DE PREVISÃO DE VENDAS DE DEZEMBRO/2024 - FONTE: AUTORES, 2025.

## 4.5 Preparação de Dados

Os dados utilizados neste estudo referem-se às vendas diárias de camisetas básicas masculinas da marca Segrob Notlad, no período de janeiro de 2022 a novembro de 2024. O arquivo original, no formato *xlsx*, apresentava duas colunas principais: a *Data* da venda e a quantidade de unidades *Vendas* registradas por dia.

### 4.5.1 Organização inicial

Inicialmente, foi realizada a padronização e renomeação das colunas para *Data* e *Vendas*, garantindo clareza e coerência com as etapas posteriores de análise. A variável *Data* foi convertida para o formato de data (*datetime*) e o conjunto de dados foi ordenado de forma crescente no tempo, respeitando a lógica da série temporal.

### 4.5.2 Criação de variáveis auxiliares

Com base na variável *Data*, foi criada uma nova variável categórica denominada *Dia\_da\_Semana*, contendo o nome do dia correspondente (*Segunda-feira*, *Terça-feira* etc.), uma separação do ano, mês e dia em colunas diferentes denominadas como *Ano*, *Mês* e *Dia* com o objetivo de permitir a análise de padrões sazonais ao longo da semana.

### 4.5.3 Verificação e tratamento de inconsistências

Durante a inspeção inicial, não foram identificados valores ausentes na série. No entanto, a estrutura do script desenvolvido prevê o tratamento automático de dados faltantes com base na imputação pela média, técnica comum em séries temporais, caso eventuais falhas surjam em atualizações futuras.

Além disso, verificou-se que todos os registros estão no formato esperado e não foram encontradas duplicidades ou inconsistências na codificação.

#### 4.5.4 Classificação dos dados

A variável de interesse (*Vendas*) é classificada como quantitativa discreta, uma vez que representa contagens inteiras de unidades vendidas. Já as variáveis *Dia\_da\_Semana*, *Ano*, *Mês* e *Dia* são variáveis qualitativas nominal, derivada para análise de padrões cíclicos. Assim, o conjunto pode ser classificado como bivariado, por envolver duas variáveis distintas.

#### 4.5.5 Resultado final da preparação

O conjunto de dados final apresenta estrutura limpa, com colunas bem definidas e formatadas, pronto para análises estatísticas e modelagem preditiva. A planilha tratada foi salva com o nome *Dados\_Preparados\_CRISP\_DM.xlsx*, estando disponível para consulta no *github*.

### 5. Fundamentação teórica

#### 5.1 Validação Cruzada dos Modelos de Previsão

Neste estudo, aplicamos e comparamos quatro modelos de previsão — Naive, Média Cumulativa, Média Móvel (MM7) e Suavização Exponencial com  $\alpha = 0,70$  — para estimar a demanda diária por camisetas básicas masculinas ao longo do mês de novembro de 2024.

Para avaliar o desempenho de cada modelo, utilizamos a validação cruzada do tipo *walk-forward*, onde, a cada dia, a previsão é comparada com o valor real observado, e os erros são acumulados para cálculo de métricas de desempenho. A métrica adotada foi o Erro Absoluto Médio (MAE), que mede a média dos valores absolutos das diferenças entre os valores reais e os valores previstos. Essa métrica foi escolhida por sua simplicidade e por não penalizar excessivamente grandes erros.

As Tabelas 1 e 2 apresentam, respectivamente, os valores diários previstos por cada modelo e os erros absolutos correspondentes em relação aos dados reais observados. Essas tabelas detalham o desempenho individual de cada abordagem ao longo do período de análise, permitindo uma comparação visual clara entre os métodos.

Timestamp	Camisetas Básicas Masculinas	Naive	Cumulativo	MM7	Suavização Exponencial $\alpha=0,7$
01/11/2024	398	412,00	413,81	393,14	403,61

02/11/2024	408	412,00	413,81	395,16	400,60
03/11/2024	393	412,00	413,81	395,47	401,50
04/11/2024	389	412,00	413,81	395,11	401,23
05/11/2024	431	412,00	413,81	394,84	401,31
06/11/2024	415	412,00	413,81	396,96	401,29
07/11/2024	409	412,00	413,81	397,53	401,29
08/11/2024	410	412,00	413,81	395,46	401,29
09/11/2024	413	412,00	413,81	395,79	401,29
10/11/2024	424	412,00	413,81	395,88	401,29
11/11/2024	412	412,00	413,81	395,94	401,29
12/11/2024	408	412,00	413,81	396,06	401,29
13/11/2024	408	412,00	413,81	396,23	401,29
14/11/2024	405	412,00	413,81	396,13	401,29
15/11/2024	431	412,00	413,81	395,93	401,29
16/11/2024	411	412,00	413,81	395,99	401,29
17/11/2024	405	412,00	413,81	396,02	401,29
18/11/2024	397	412,00	413,81	396,04	401,29
19/11/2024	416	412,00	413,81	396,06	401,29
20/11/2024	410	412,00	413,81	396,06	401,29
21/11/2024	412	412,00	413,81	396,03	401,29
22/11/2024	403	412,00	413,81	396,02	401,29
23/11/2024	401	412,00	413,81	396,03	401,29
24/11/2024	410	412,00	413,81	396,04	401,29
25/11/2024	390	412,00	413,81	396,04	401,29
26/11/2024	399	412,00	413,81	396,04	401,29
27/11/2024	414	412,00	413,81	396,04	401,29
28/11/2024	433	412,00	413,81	396,03	401,29
29/11/2024	496	412,00	413,81	396,03	401,29
30/11/2024	547	412,00	413,81	396,04	401,29

TABELA 1 - PREVISÃO DOS MODELOS PARA NOVEMBRO/2024 - FONTE: AUTORES,2025.

ERRO (NAIVE)	E (CUM)	E (MM7)	E.(SE)
14,00	15,81	4,86	5,61
4,00	5,81	12,84	7,40
19,00	20,81	2,47	8,50
23,00	24,81	6,11	12,23
19,00	17,19	36,16	29,69
3,00	1,19	18,04	13,71
3,00	4,81	11,47	7,71
2,00	3,81	14,54	8,71
1,00	0,81	17,21	11,71
12,00	10,19	28,12	22,71

0,00	1,81	16,06	10,71
4,00	5,81	11,94	6,71
4,00	5,81	11,77	6,71
7,00	8,81	8,87	3,71
19,00	17,19	35,07	29,71
1,00	2,81	15,01	9,71
7,00	8,81	8,98	3,71
15,00	16,81	0,96	4,29
4,00	2,19	19,94	14,71
2,00	3,81	13,94	8,71
0,00	1,81	15,97	10,71
9,00	10,81	6,98	1,71
11,00	12,81	4,97	0,29
2,00	3,81	13,96	8,71
22,00	23,81	6,04	11,29
13,00	14,81	2,96	2,29
2,00	0,19	17,96	12,71
21,00	19,19	36,97	31,71
84,00	82,19	99,97	94,71
135,00	133,19	150,96	145,71
MAE	15,40	16,06	21,70
			18,22

TABELA 2 - ERRO ABSOLUTO DOS MODELOS DE PREVISÃO - FONTE: AUTORES,2025.

Com base nos resultados da validação cruzada, o modelo Naive apresentou o menor Erro Absoluto Médio (MAE = 15,40), demonstrando ser o mais eficaz para esta série temporal específica. Isso indica que, em contextos onde os dados não apresentam tendência clara nem sazonalidade significativa, modelos simples como o Naive podem se mostrar mais robustos e eficazes do que métodos mais complexos, que podem ser sensíveis a ruídos ou variações abruptas.

## 6. Referências

MORETTIN, Pedro A.; TOLOI, Clélia M. Análise de séries temporais: modelos lineares univariados. 1a ed. São Paulo, Editora Blucher, 2021, 475 p.

Site Pessoal | Dalton Borges. Disponível em: <https://www.daltonborges.com.br/>. Acesso em 26 abr. 2025.

DATASCIENCE-PM. *CRISP-DM 2.0: Data Science Methodology*. Disponível em: <https://www.datascience-pm.com/crisp-dm-2/>. Acesso em: 13 maio 2025.

MORETTIN, P. A.; SINGER, J. M. Estatística e Ciência de Dados. Editora Blucher, 2021