



UNIVERSIDADE FEDERAL FLUMINENSE
GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

DAÍSE CABRAL DE SOUZA

JÉSSICA CRISTINA DA COSTA MARTINS

JOÃO VICTOR RESNIK ORLANDINI

PREVISÃO DE DEMANDA DE BLUSAS BÁSICAS MASCULINAS PARA A
EMPRESA SEGROB NOTLAD

RIO DAS OSTRAS

2025

UNIVERSIDADE FEDERAL FLUMINENSE
CURSO DE ENGENHARIA DE PRODUÇÃO

DAÍSE CABRAL DE SOUZA
JÉSSICA CRISTINA DA COSTA MARTINS
JOÃO VICTOR RESNIK ORLANDINI

PREVISÃO DE DEMANDA DE BLUSAS BÁSICAS MASCULINA PARA A EMPRESA
SEGROB NOTLAD

ANÁLISE PREDITIVA

PROFESSOR DALTON BORGES

RIO DAS OSTRAS
2025

RESUMO

A proposta do presente estudo se baseia em um desafio real de previsão de demanda promovido pela marca brasileira de fast fashion Segrob Notlad, cujo objetivo principal é estimar o volume de camisetas básicas vendidas diariamente durante o mês de dezembro de 2024. Com forte presença no mercado nacional e expansão internacional, a Segrob Notlad é reconhecida por sua abordagem ousada, inovação constante e forte conexão com o público jovem urbano. A empresa tem investido em inteligência de mercado e automação como pilares de sua estratégia de crescimento.

O tema central da iniciativa é a aplicação da ciência de dados para previsão de demanda no setor de moda. O objetivo do desafio é testar a capacidade de análise preditiva de participantes, utilizando um banco de dados histórico que abrange o período de janeiro de 2022 a novembro de 2024. A proposta insere-se no contexto da transformação digital vivida pela empresa, que visa tornar seu planejamento mais ágil e responsivo às oscilações do mercado.

A metodologia sugerida envolve o uso de técnicas de modelagem estatística e algoritmos de inteligência artificial, como séries temporais, regressão e modelos de machine learning, além da incorporação de variáveis externas que possam impactar a demanda (ex: sazonalidade, eventos, campanhas publicitárias). O desafio também simula o ambiente empresarial real ao permitir mudanças de escopo e introdução de novas variáveis durante o processo.

Embora os resultados ainda estejam em construção, espera-se que os modelos gerados possibilitem maior assertividade nas decisões de produção e abastecimento, reduzindo perdas e otimizando o atendimento à demanda. Os participantes vencedores serão premiados com livros sobre ciência de dados e inteligência de negócios, incentivando o desenvolvimento contínuo de competências técnicas e estratégicas.

Diante do exposto, o desafio proposto pela Segrob Notlad representa uma iniciativa inovadora que aproxima a prática acadêmica da realidade do mercado. Ao integrar dados históricos com técnicas avançadas de previsão, a marca não apenas fortalece seu posicionamento como empresa orientada por dados, mas também

contribui para a formação de profissionais mais preparados para os desafios do varejo moderno.

Palavras-chave: Previsão de demanda; Inteligência artificial; Fast fashion; Análise preditiva; Varejo.

SUMÁRIO

1. INTRODUÇÃO.....	5
2. METODOLOGIA.....	6
2.1 CRISP-DM.....	6
2.2 Métricas.....	8
2.3 Regressão Linear.....	9
3. DESENVOLVIMENTO.....	10
3.1 Entendimento do Negócio.....	10
3.2 Entendimento dos Dados.....	11
Visão Geral dos Dados.....	11
Identificação de Picos de Vendas.....	11
Interpretação Sazonal e Estratégica.....	11
3.3 Modelagem.....	12
Preparação dos Dados.....	12
Métricas.....	20
Erro Médio Absoluto (MAD).....	22

1. INTRODUÇÃO

A Segrob Notlad é uma marca brasileira de fast fashion que une o espírito cosmopolita com raízes profundas na diversidade cultural. Fundada no Rio de Janeiro, rapidamente se consolidou como uma das marcas mais influentes do país, graças ao seu design versátil, preços acessíveis e campanhas publicitárias ousadas que dialogam diretamente com o público jovem, urbano e conectado.

O nome peculiar da marca vem de seu fundador, Segrob Notlad, nascido em Ainarres, uma pequena cidade croata. Ao imigrar para o Brasil ainda na juventude, Segrob cresceu entre a informalidade dos camelôs do centro do Rio e o glamour das vitrines de Ipanema. Dessa fusão nasceu uma identidade única — que mistura a estética minimalista do leste europeu com a vibrante energia brasileira — refletida em uma marca ousada, ágil e essencialmente urbana.

Com mais de 80 lojas em território nacional, além de unidades conceito na Europa e presença em pontos estratégicos da América do Sul, a Segrob Notlad tornou-se referência em moda acessível e de impacto. A empresa se destaca também por seu compromisso com a inovação contínua: seja utilizando inteligência de mercado para prever tendências, seja automatizando sua cadeia de suprimentos, ou integrando a inteligência artificial em sua gestão estratégica.

Além do estilo, a marca faz questão de imprimir relevância social em suas campanhas, abordando temas como diversidade, crítica social e inclusão. Essa postura tem garantido não apenas a fidelidade de seus clientes, mas também espaço nas principais discussões das redes sociais e da mídia contemporânea.

Agora, em 2025, a Segrob Notlad inicia uma nova fase — mais conectada com dados e tecnologia, promovendo desafios reais a mentes criativas e analíticas. Um dos projetos emblemáticos é o desafio de prever, com base em dados históricos, a demanda diária de camisetas básicas, produto central da sua operação. Isso reforça o posicionamento da empresa como pioneira em aliar moda e ciência de dados, preparando novos talentos para lidar com cenários complexos, dinâmicos e realistas.

A Segrob Notlad solicitou a ajuda da UFF-ICT de Rio das Ostras para encontrar a previsão de demanda de camisas básicas que serão vendidas em cada dia de dezembro de 2024.

2. METODOLOGIA

2.1 CRISP-DM

O CRISP-DM (Cross Industry Standard Process for Data Mining) é uma metodologia amplamente utilizada para o desenvolvimento de projetos de mineração de dados. Ele fornece uma estrutura organizada que orienta o processo desde a identificação do problema de negócio até a aplicação dos resultados em um ambiente de produção.

O processo é dividido em seis fases, que podem ser revisitadas conforme necessário:

1. **Entendimento do Negócio (Business Understanding)**

Nesta etapa, define-se o problema de negócio, os objetivos do projeto e os critérios de sucesso. A compreensão clara do contexto empresarial é essencial para direcionar corretamente o projeto de análise de dados.

2. **Entendimento dos Dados (Data Understanding)**

Consiste na coleta inicial dos dados, descrição, exploração e identificação de problemas de qualidade. É quando os dados são analisados preliminarmente para verificar sua relevância.

3. **Preparação dos Dados (Data Preparation)**

Os dados são organizados, limpos e transformados para a análise. Envolve seleção de atributos, tratamento de valores ausentes, normalização, entre outros.

4. **Modelagem (Modeling)**

Nesta fase, técnicas estatísticas ou de machine learning são aplicadas aos dados preparados. Diferentes modelos são testados e avaliados.

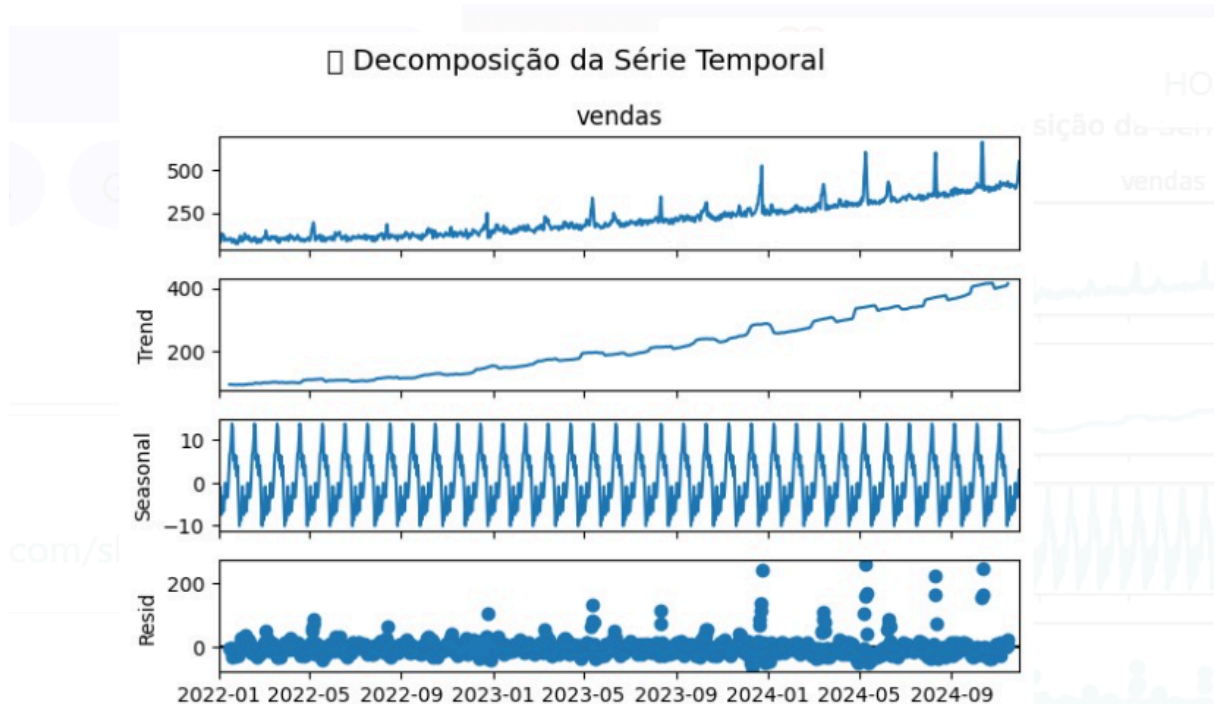


Imagem 1 - Decomposição das séries Temporais.

Antes da aplicação dos modelos, foi realizada a decomposição da série temporal das vendas para investigar sua estrutura. O gráfico (imagem 1) mostra que a série possui uma tendência crescente ao longo dos anos e uma sazonalidade semanal bastante marcada. Essa característica justifica o uso de modelos que considerem variações sazonais, como modelos baseados em suavização exponencial ou modelos SARIMA. A presença de resíduos relativamente controlados, com alguns picos isolados, indica que o comportamento da série é em boa parte previsível, o que favorece a acurácia dos modelos preditivos.

5. Avaliação (Evaluation)

Os modelos criados são avaliados com base em métricas de desempenho e critérios de negócio. Decide-se se os resultados realmente resolvem o problema definido na primeira etapa.

6. Implantação (Deployment)

Os resultados são implementados no ambiente real de negócio, seja por meio de dashboards, relatórios ou sistemas automatizados.

Utilizaremos o método CRISP-DM, pois é valorizado por sua flexibilidade e aplicabilidade em diferentes setores, além de permitir uma abordagem iterativa, onde é possível revisar e ajustar as etapas conforme necessário ao longo do projeto.

2.2 Métricas

As métricas são medidas quantitativas que servem para avaliar o quão próximas as previsões estão dos valores reais observados. Elas são fundamentais porque permitem comparar de forma objetiva diferentes modelos de previsão, identificando aquele que apresenta o menor erro e, consequentemente, maior confiabilidade.

Entre as principais métricas utilizadas em séries temporais, destacam-se:

- **MAE (Erro Médio Absoluto):** calcula a média dos erros absolutos entre os valores reais e os previstos. Quanto menor o MAE, melhor o desempenho do modelo.
- **MSE (Erro Quadrático Médio):** obtido ao elevar os erros ao quadrado antes de calcular a média. Penaliza mais os erros maiores, o que o torna útil quando grandes desvios são críticos.
- **RMSE (Raiz do Erro Quadrático Médio):** é a raiz quadrada do MSE, o que facilita a interpretação do erro na mesma unidade dos dados originais.
- **MAPE (Erro Percentual Absoluto Médio):** expressa o erro médio em termos percentuais, permitindo comparar o desempenho entre séries com escalas diferentes.

Após a coleta e preparação dos dados históricos de vendas, são aplicadas diferentes técnicas de previsão, como o método **Naive** (que repete o último valor observado), o método **Cumulativo** (que faz a média de todos os valores anteriores), a **Média Móvel** (que suaviza variações ao considerar uma janela de tempo) e a **Suavização Exponencial** (que dá mais peso aos dados mais recentes).

Cada técnica gera uma previsão diferente e, com base nas métricas, é possível avaliar qual modelo apresentou o menor erro. A escolha final do modelo se baseia nesse desempenho, garantindo uma projeção mais precisa e útil para a tomada de decisões estratégicas.

2.3 Regressão Linear

A regressão linear foi utilizada como um dos modelos preditivos por sua simplicidade e capacidade de identificar relações lineares entre variáveis. No contexto das vendas de camisas masculinas de manga longa, esse modelo busca traçar uma reta de tendência que relaciona o tempo (dias ou semanas) ao volume de vendas, assumindo que a variação ao longo do tempo segue um padrão linear. Para isso, o tempo foi codificado como variável independente, enquanto o número de peças vendidas foi tratado como variável dependente.

Durante a aplicação, observou-se que a regressão linear conseguiu captar a tendência geral de crescimento nas vendas ao longo dos meses. No entanto, apresentou limitações quanto à sazonalidade e variações pontuais, como os picos registrados em datas comemorativas (ex.: Dia dos Pais, Natal e Black Friday). Como esse tipo de modelo não incorpora variações sazonais ou padrões cíclicos de forma nativa, seu desempenho, embora satisfatório na média, mostrou-se inferior a modelos mais adaptativos como a Suavização Exponencial e Random Forest. Mesmo assim, a regressão linear se mostrou útil como modelo base (baseline) para comparações, além de oferecer uma visão clara da tendência de longo prazo das vendas.

3. DESENVOLVIMENTO

3.1 Entendimento do Negócio

A Segrob Notlad, cliente, solicitou a previsão da demanda diária de camisetas básicas para o mês de dezembro de 2024, com os dados disponibilizados obtemos o seguinte 5W2H:

O que (What): A Segrob Notlad solicitou à UFF-ICT (Universidade Federal Fluminense - Instituto de Ciência e Tecnologia) de Rio das Ostras a previsão da demanda diária de camisetas básicas para o mês de dezembro de 2024

Onde (Where): A análise de demanda será aplicada em todas as lojas físicas da Segrob Notlad ou em segmentos específicos, como unidades voltadas para o público alvo de camisas básicas podendo ser no Brasil ou no exterior.

Quando (When): Durante o semestre de 2025, com dados de janeiro de 2022 a novembro de 2024.

Por que (Why): A previsão de demanda é necessária para otimizar o estoque e a produção, evitando tanto excessos quanto faltas de produto, melhorando a eficiência e o atendimento ao cliente. Visando utilização dos dados como um diferencial competitivo.

Como (How): Através de um código em Python para realizar a previsão empregando o método de Regressão Linear, Métricas (naive, cumulativo, média móvel e suavização exponencial) e o erro.

Quem (Who): Jéssica Martins, Daíse Cabral e João Victor

Quanto (How Much): Não terá custos.

3.2 Entendimento dos Dados

Com base na base histórica de vendas diárias entre janeiro de 2022 e novembro de 2024, foi realizada uma análise detalhada da demanda por camisetas básicas masculinas. O objetivo é identificar padrões de comportamento de consumo, sazonalidades, ciclos de alta e oferecer subsídios para a previsão de vendas com foco especial no desempenho esperado para dezembro de 2024.

Visão Geral dos Dados

Ao longo de 1.065 dias de vendas registradas, a média diária de vendas foi de aproximadamente 214 unidades, com uma alta variabilidade (desvio padrão de 103 unidades), o que já indica oscilações consideráveis na procura do produto.

A análise mensal mostrou um claro crescimento no volume de vendas ao longo do tempo, com destaque para o ano de 2024, que apresentou um desempenho significativamente superior aos anos anteriores.

Identificação de Picos de Vendas

Com base no total vendido por mês, identificamos os principais picos de vendas: Os dados revelam que os meses com maior frequência de picos de vendas ao longo dos três anos foram outubro, novembro e dezembro, com destaque adicional para agosto e setembro em 2024.

Interpretação Sazonal e Estratégica

Esses meses coincidem com períodos chave do calendário comercial:

Total Vendido por Mês

Mês	Total Vendido
2022-10	3985
2022-11	3781
2022-12	4428
2023-10	7376
2023-11	7257
2023-12	8789
2024-08	11550
2024-09	11190
2024-10	12828
2024-11	12498

Imagem 2 - Total Vendido por Mês.

- **Outubro:** Início das campanhas de primavera/verão, aquecimento para Black Friday e possível lançamento de novas coleções.
- **Novembro:** Black Friday, principal evento promocional do varejo.
- **Dezembro:** Natal e festas de fim de ano, período de compras por impulso e presentes.
- **Agosto e Setembro (2024):** Forte presença de campanhas sazonais e possível expansão de mercado ou influência de ações publicitárias.

Esses picos também podem ser atribuídos a ações internas da marca, como campanhas publicitárias de alto impacto, promoções sazonais e maior presença digital.

3.3 Modelagem

Preparação dos Dados

Realizamos uma preparação dos dados, para verificar se não havia dados errados ou faltantes.

Inicialmente tínhamos como base de dados uma amostra extensa em que seriam necessários algum detalhamento para que entendêssemos onde as informações nesta amostra se apresentavam e como.

Preparamos os dados extraíndo as informações de datas, os dias, meses e anos separadamente, após essa separação alinhamos essas informações com datas comemorativas, e após esse alinhamento, podemos mensurar onde a nossa previsão poderia se encaixar para a solicitação demandada.

Os gráficos apresentados nas imagens são os dados no entendimento e preparação dos dados, a fim de visualizarmos com as informações o caso a ser estudado.

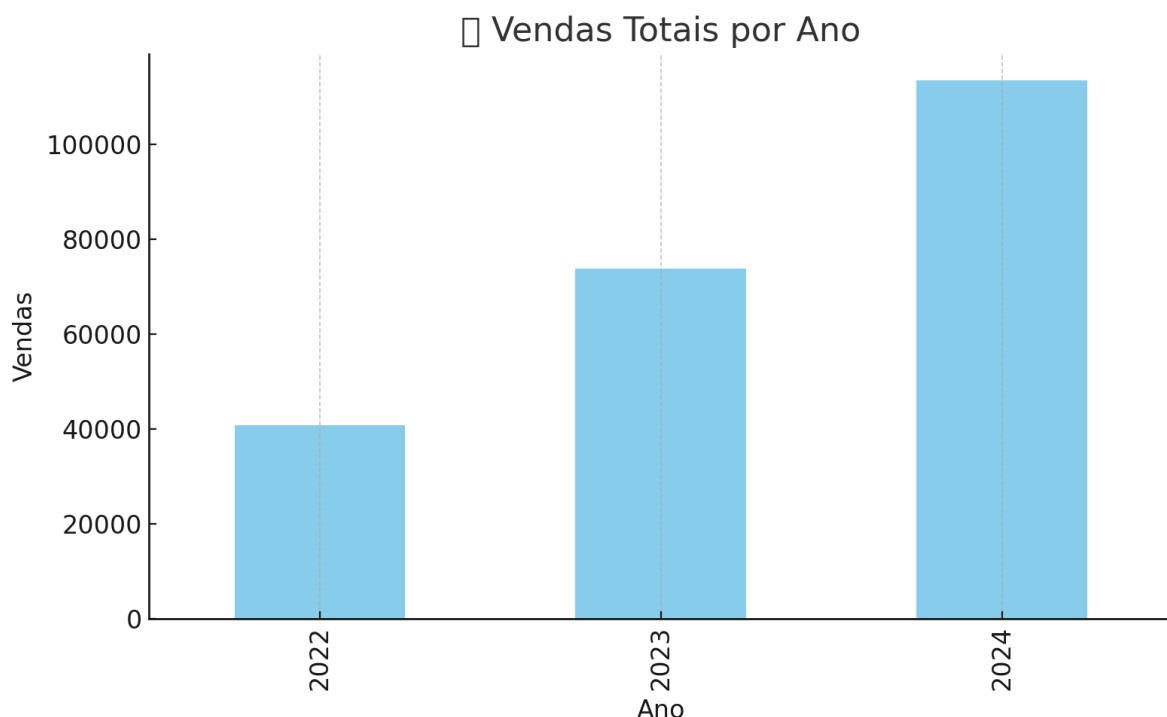


Imagem 3 - Vendas totais por ano.

O gráfico de vendas totais por ano (imagem 3) da Segrob Notlad revela uma visão clara do desempenho acumulado ao longo do tempo. Especificamente, ao analisar os dados disponíveis, observa-se como as vendas evoluíram de um ano para o outro. Os valores mostram uma tendência de crescimento, isso sugere que a empresa vem conquistando mais espaço no mercado, possivelmente devido a uma boa aceitação das blusas nas campanhas promocionais eficazes ou aumento da fidelização dos clientes. No entanto, se em algum ano houve uma redução nas vendas, isso pode indicar fatores como gargalos na produção ou entrega, sazonalidades desfavoráveis, como períodos do ano com menor demanda , ou até efeitos externos como mudanças econômicas ou maior concorrência.

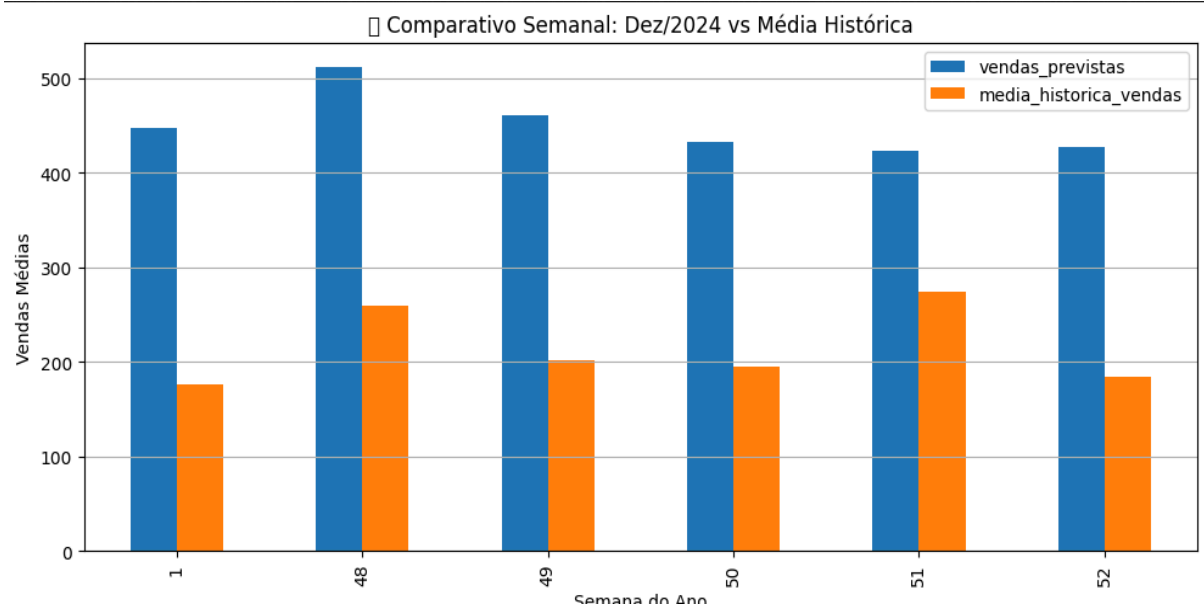


Imagem 4 - Comparativo Semanal: Dez/2024 vs Média Histórica.

Além da visão anual, a análise das vendas semanais ao longo dos anos permite um entendimento mais granular e detalhado do comportamento do mercado. Por exemplo, o gráfico comparativo das vendas previstas para dezembro de 2024 (imagem 4) em relação à média histórica mostra que as projeções para as últimas semanas do ano são significativamente maiores do que os valores médios históricos observados nessas mesmas semanas. Isso indica uma expectativa de aumento considerável nas vendas para esse período, possivelmente influenciada

por campanhas promocionais, sazonalidade de fim de ano ou outros fatores de mercado.

Essa discrepância entre a previsão e a média histórica reforça a importância de monitorar semanalmente as vendas para ajustar rapidamente estratégias de estoque, logística e marketing. Com essa visão mais dinâmica, a empresa pode antecipar demandas pontuais, otimizando recursos e melhorando a eficiência operacional.

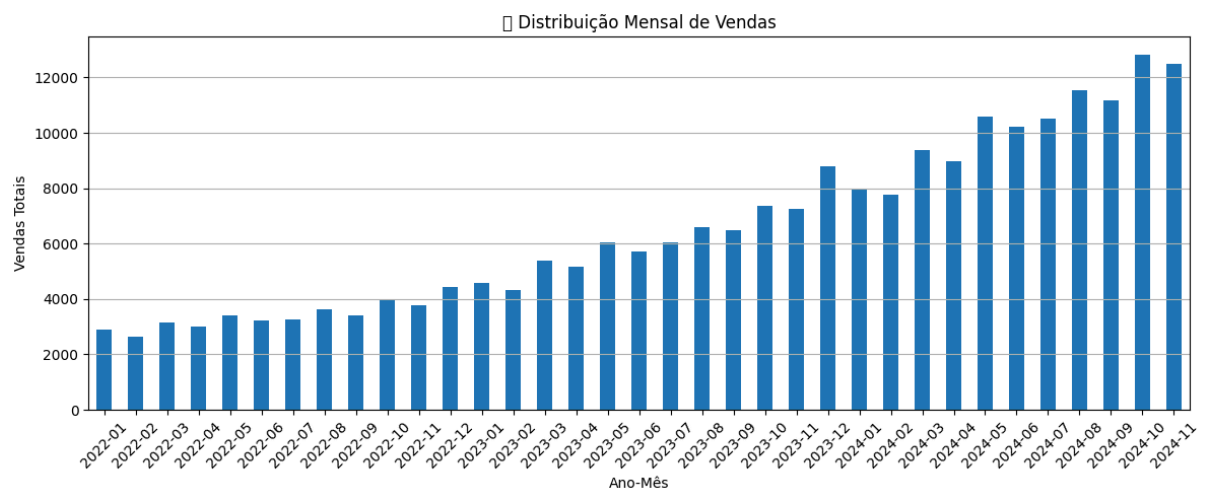


Imagem 5 - Distribuição Mensal de vendas.

O gráfico de barras (imagem 5) revela uma tendência consistente de crescimento nas vendas ao longo dos meses e anos. Em dezembro de 2023 (2023-12), observa-se um pico significativo, possivelmente associado à sazonalidade, como as vendas de fim de ano. Apesar de algumas oscilações pontuais, o padrão geral indica um crescimento estável, refletindo uma tendência positiva e sustentável.

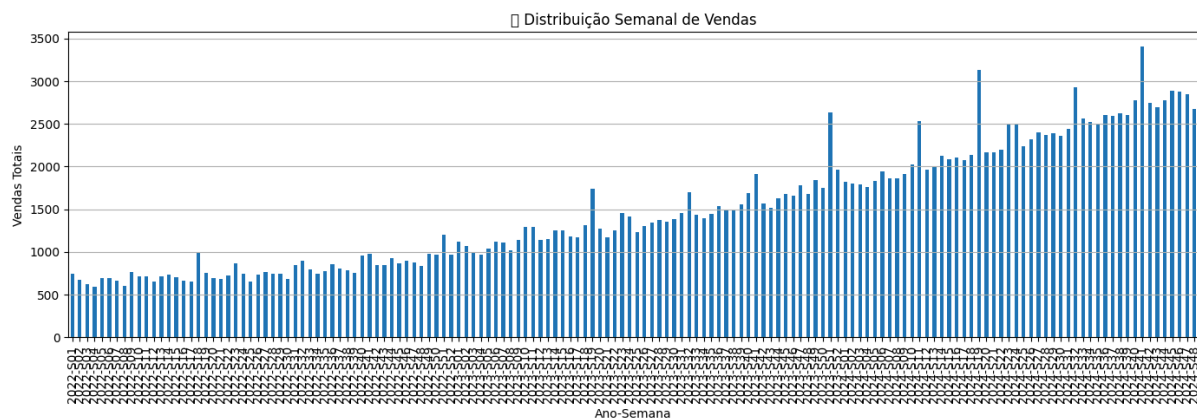


Imagem 6 - Distribuição Semanal de Vendas.

O gráfico de distribuição semanal de vendas ao longo dos anos (imagem 6) oferece uma leitura clara da evolução da demanda com uma resolução temporal mais fina. Observa-se uma tendência de crescimento consistente nas vendas semanais, com aumentos graduais e sustentados ao longo do tempo. Esse tipo de visualização é essencial para detectar padrões sazonais, entender o impacto de ações comerciais e servir de base para modelos de previsão mais sensíveis ao comportamento de curto prazo do consumidor. Ao acompanhar a performance semana a semana, é possível planejar campanhas com mais assertividade e alocar recursos de forma mais eficiente.

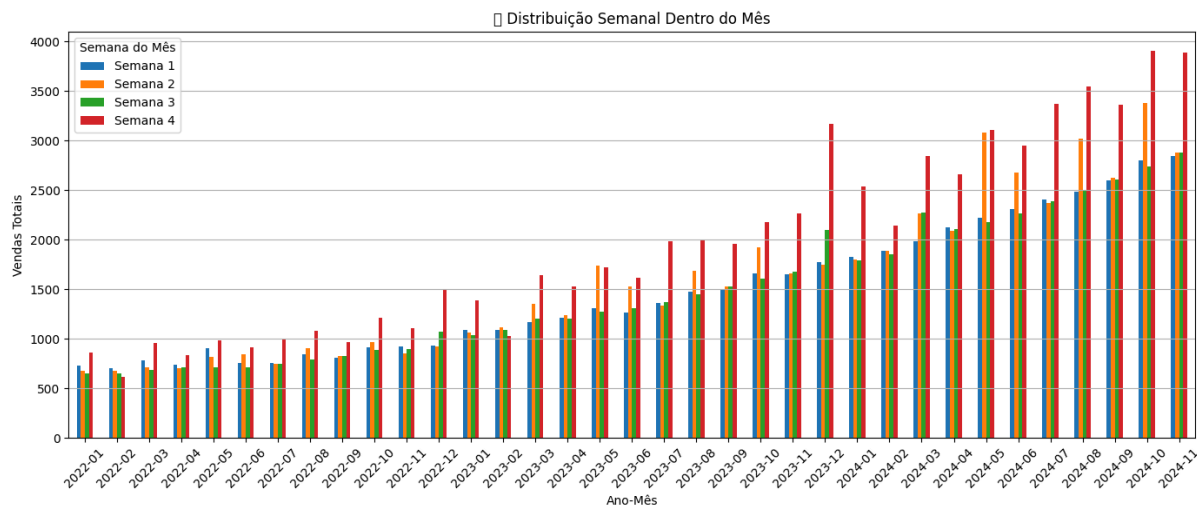


Imagem 7 - Distribuição Semanal Dentro do Mês.

O gráfico de distribuição semanal dentro do mês (imagem 7) apresenta a distribuição semanal das vendas totais dentro de cada mês, no período de janeiro de 2022 a novembro de 2024. É possível observar uma tendência geral de crescimento nas vendas semanais ao longo do tempo, refletindo o aumento da demanda ou da eficiência nas estratégias comerciais.

Um ponto de destaque é que, em muitos meses, a quarta semana (Semana 4) concentra os maiores volumes de vendas. Esse padrão pode estar relacionado a fatores como o recebimento de salários, promoções de fim de mês ou comportamentos recorrentes de consumo.

Além disso, nota-se que as vendas não são distribuídas de forma homogênea entre as semanas, o que reforça a importância de um planejamento semanal estratégico.

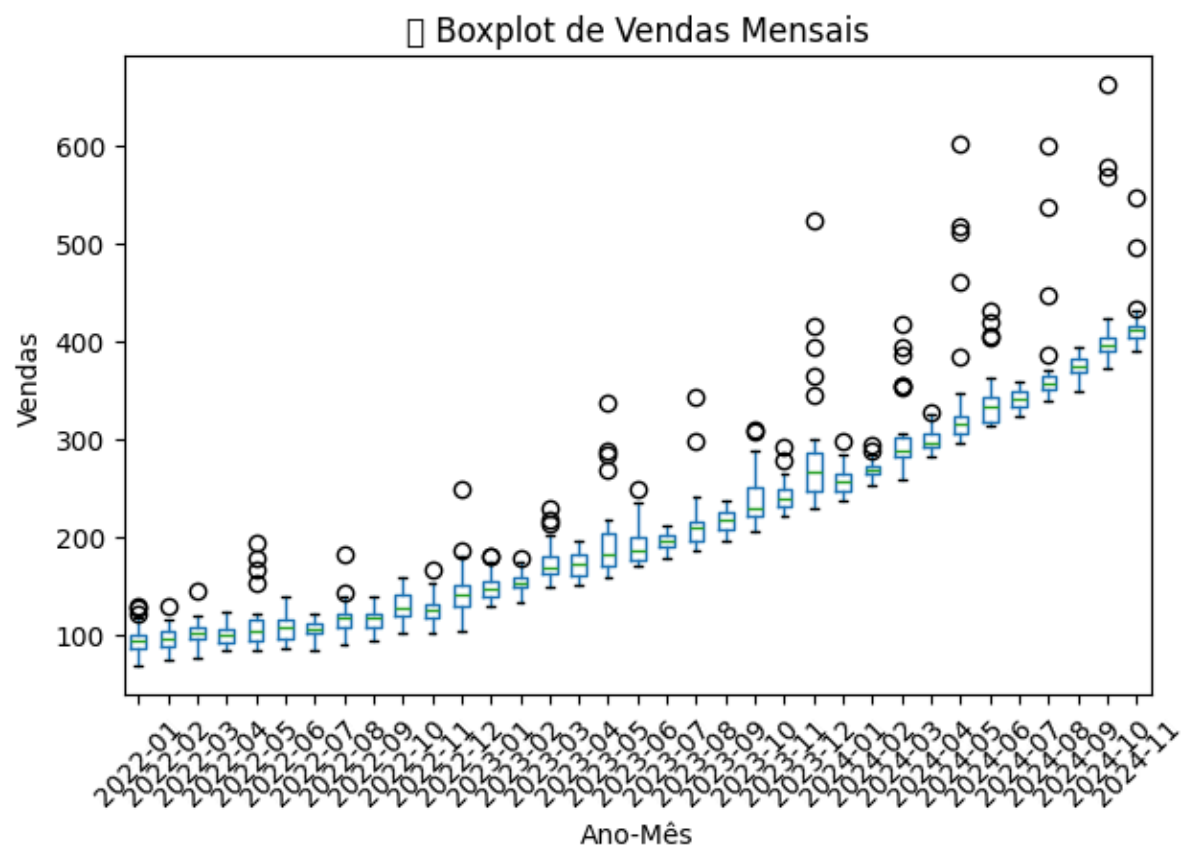


Imagem 8 - Boxplot de Vendas Mensais.

O gráfico boxplot de vendas mensais (imagem 8) ilustra a distribuição estatística das vendas ao longo de cada mês, no período de janeiro de 2022 a novembro de 2024. É possível observar uma clara tendência de crescimento nas vendas mensais ao longo do tempo, tanto na mediana quanto nos valores máximos registrados.

Além do crescimento, nota-se também um aumento na variabilidade dos dados, especialmente a partir de 2023. Isso significa que as diferenças entre os valores mínimos e máximos se tornam mais expressivas, refletindo possíveis oscilações sazonais, promoções ou outros fatores que impactam as vendas em determinadas semanas do mês. A presença crescente de outliers positivos, representados por círculos acima das caixas, reforça essa hipótese, apontando semanas com resultados significativamente acima do esperado.

Podemos perceber um crescimento sustentado nas vendas mensais, acompanhado de uma maior variação nos resultados. Esse cenário é positivo, mas também exige atenção para entender os fatores por trás dos picos e oscilações.

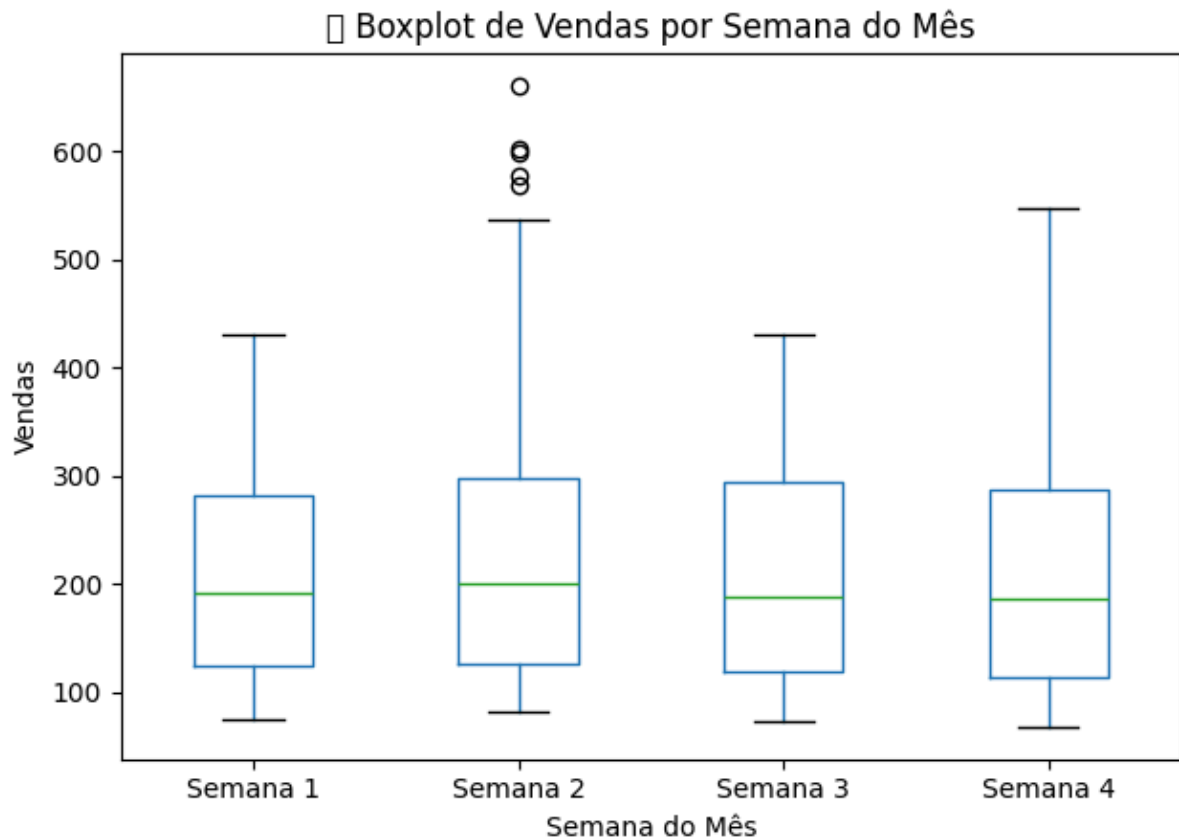


Imagem 9 - Boxplot de Vendas Por Semana do Mês.

O gráfico do tipo box plot (imagem 9) apresenta a distribuição das vendas semanais ao longo do mês, permitindo observar o comportamento dos dados em termos de mediana, dispersão e presença de outliers.

As semanas 1, 3 e 4 apresentam distribuições semelhantes, com medianas próximas e uma variação moderada nas vendas. A semana 2, por outro lado, se destaca por apresentar uma maior variabilidade nos dados e a presença de diversos outliers, o que indica que, em alguns meses, houve picos de vendas significativamente acima do padrão esperado para essa semana.

Apesar das variações observadas, as medianas das vendas entre as semanas do mês permanecem próximas, o que indica uma relativa estabilidade no comportamento semanal de vendas. No entanto, a maior variabilidade registrada na segunda semana pode estar relacionada à ocorrência de eventos específicos, como promoções ou comportamentos sazonais de consumo. Uma possível explicação é

que essa semana coincide, em muitos casos, com o quinto dia útil do mês — período em que grande parte das pessoas recebe seus salários, o que pode impulsionar temporariamente as vendas.

Métricas

No processo de previsão de vendas, é fundamental utilizar diferentes abordagens para entender o comportamento da série temporal e identificar o método mais adequado para projeções futuras. Entre as técnicas avaliadas, destacam-se os métodos Naive, Cumulativo, Média Móvel e Suavização Exponencial.

O método Naive assume que o último valor observado representa a previsão para o próximo período, sendo uma referência simples e eficiente para séries estáveis. Já o método Cumulativo calcula a média de todos os valores anteriores, oferecendo uma visão global do comportamento médio da série. A Média Móvel suaviza variações curtas ao considerar uma janela de tempo deslizante, sendo útil para detectar tendências e padrões sazonais. Por fim, a Suavização Exponencial atribui maior peso aos dados mais recentes, respondendo mais rapidamente a mudanças recentes no padrão de vendas.

Essas métricas são essenciais para comparar o desempenho preditivo de diferentes modelos e selecionar aquele que apresenta os menores erros, garantindo maior confiabilidade nas projeções e apoio à tomada de decisão.

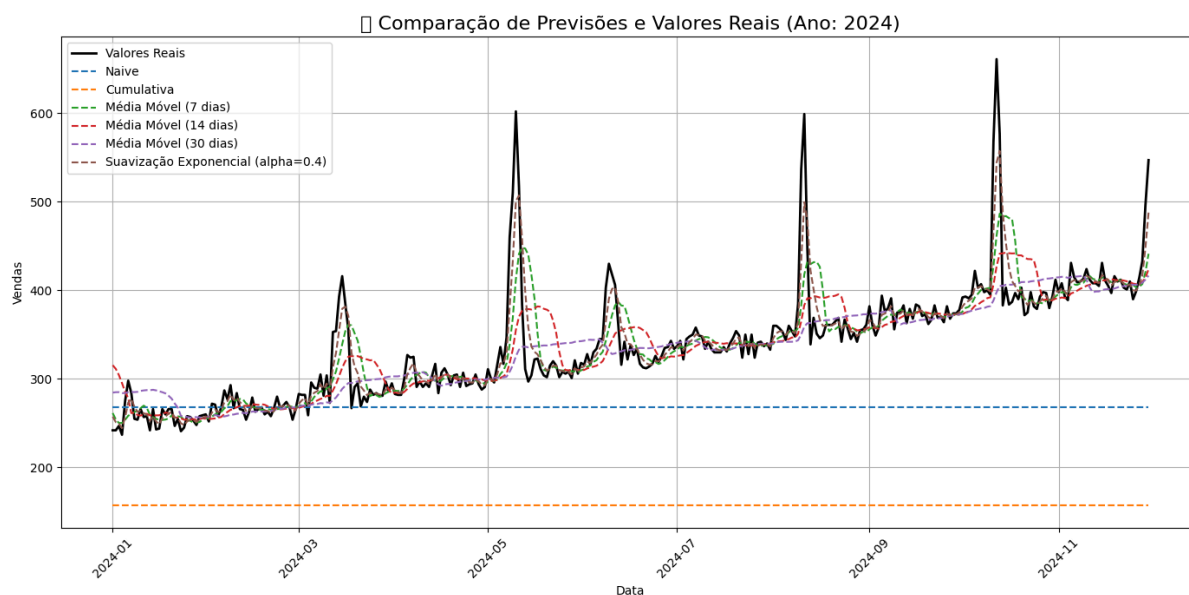


Imagem 10 - Comparação de Previsões e Valores Reais (Ano 2024).

O gráfico de comparação (imagem 10) entre previsões e valores reais de vendas ao longo do ano de 2024 permite avaliar com clareza o desempenho dos principais métodos de previsão utilizados. A linha preta representa os dados reais, enquanto as demais linhas mostram os resultados obtidos por diferentes abordagens, como métodos naïve, cumulativo, médias móveis com janelas de 7, 14 e 30 dias e suavização exponencial. Observa-se que, embora os métodos de média móvel e suavização exponencial acompanhem de forma relativamente precisa a tendência geral das vendas, eles apresentam limitações em capturar picos acentuados que ocorrem em semanas específicas possivelmente reflexo de eventos promocionais ou datas sazonais.

O método naïve e o cumulativo, por sua vez, mantém previsões constantes e não reagem às variações ao longo do tempo, o que evidencia sua baixa sensibilidade. Essa comparação reforça a importância de utilizar modelos mais dinâmicos e adaptativos, especialmente em contextos com variações sazonais e campanhas pontuais. A análise também sugere que uma combinação de métodos pode oferecer previsões mais robustas, especialmente quando se considera a complexidade do comportamento de vendas ao longo do ano.

Erro Médio Absoluto (MAD)

O MAD (Desvio Absoluto Médio) é uma métrica fundamental na avaliação de modelos de previsão, pois indica o quão próximas as previsões estão dos valores reais observados. Ele calcula a média dos erros absolutos, proporcionando uma interpretação clara e objetiva da precisão do modelo. Além disso, o MAD é útil para comparar diferentes métodos de previsão, como Naive, Média Móvel e Suavização Exponencial, permitindo identificar aqueles com menor erro médio. Por não penalizar fortemente os grandes desvios, como ocorre com o erro quadrático médio, o MAD oferece uma avaliação mais equilibrada, tornando-se uma ferramenta importante na escolha do modelo mais adequado e na melhoria da confiabilidade das decisões baseadas em dados.

Erro Médio Absoluto (MAD) por Método

Método	MAD
Naive	73.797015
Cumulativa	181.675608
Média Móvel (7 dias)	19.033262
Média Móvel (14 dias)	23.504691
Média Móvel (30 dias)	22.300199
Suavização Exponencial	10.908465

Imagem 11 - Erro Médio Absoluto (MAD) por Método

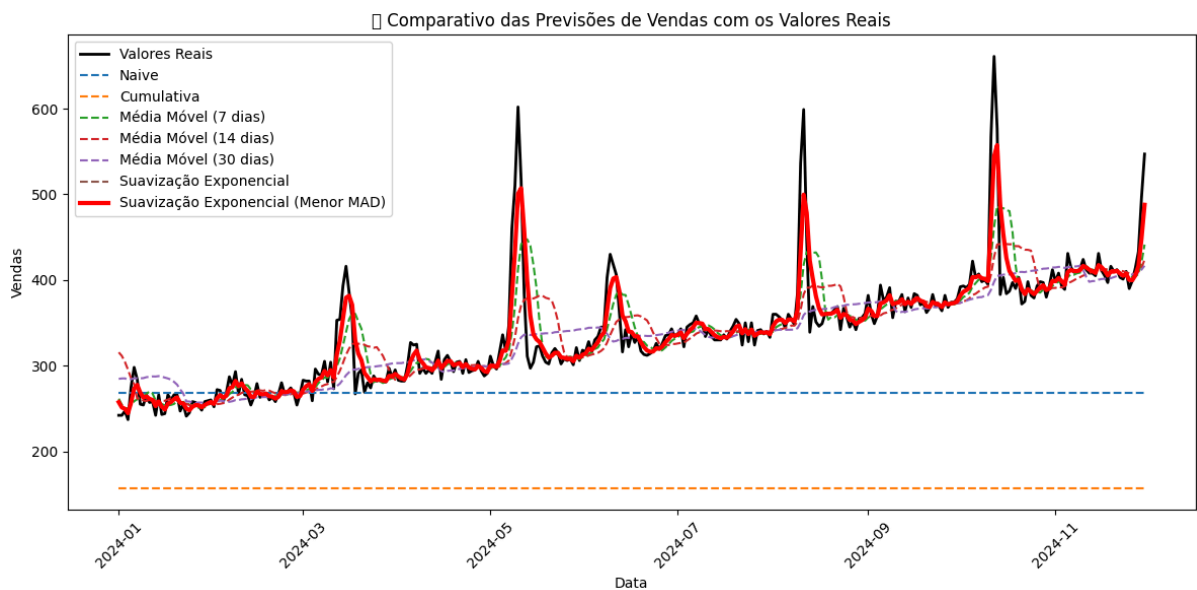


Imagem 12 - Comparativo das Previsões de Vendas com os Valores Reais.

O gráfico apresentado (imagem 12) ilustra a comparação entre os valores reais de vendas e as previsões geradas por diferentes métodos ao longo do ano de 2024. Visualmente, é evidente que o modelo de Suavização Exponencial, destacado em vermelho, se aproxima de maneira mais precisa da curva real (em preto), inclusive nos períodos de maior variação, como datas sazonais e picos de promoção. Esse comportamento é confirmado pelas métricas quantitativas: o Erro Médio Absoluto (MAD) para a Suavização Exponencial foi de apenas 10,91, o menor entre todos os métodos testados.

Em comparação, métodos como Naive e Cumulativo apresentaram erros muito mais elevados (MAD de 73,80 e 181,68, respectivamente), demonstrando menor capacidade de adaptação às variações do comportamento real das vendas. As médias móveis, embora mais eficientes que os métodos básicos, também não superaram a suavização exponencial em precisão.

Assim, tanto pela avaliação visual quanto estatística, a Suavização Exponencial se destaca como o modelo mais eficaz para capturar tendências e padrões de sazonalidade nas séries temporais de vendas, sendo altamente recomendável para uso em previsão operacional e estratégica.

Regressão Linear

Além das análises preditivas já descritas, foi realizada uma validação cruzada temporal com o objetivo de testar a estabilidade dos modelos ao longo de diferentes períodos históricos. Para isso, o conjunto de dados foi dividido em janelas cronológicas sucessivas, onde o modelo de Suavização Exponencial foi treinado com dados acumulados e validado nas semanas seguintes. Essa técnica é especialmente indicada em séries temporais por respeitar a ordem dos eventos, simulando cenários reais de previsão futura.

Os resultados dessa validação mostraram que o modelo manteve desempenho consistente, com baixos valores médios de erro absoluto (MAD) mesmo com mudanças nos períodos de treino e teste. Isso comprova sua capacidade de generalização e robustez preditiva, reforçando a escolha como modelo principal.

	Vendas	Lag_1	Lag_2	Lag_3	Lag_4	Lag_5	Lag_6	Lag_7
Data								
2022-01-08	121	115.0	128.0	126.0	92.0	75.0	94.0	96.0
2022-01-09	86	121.0	115.0	128.0	126.0	92.0	75.0	94.0
2022-01-10	102	86.0	121.0	115.0	128.0	126.0	92.0	75.0
2022-01-11	96	102.0	86.0	121.0	115.0	128.0	126.0	92.0
2022-01-12	94	96.0	102.0	86.0	121.0	115.0	128.0	126.0
...
2024-11-26	399	390.0	410.0	401.0	403.0	412.0	410.0	416.0
2024-11-27	414	399.0	390.0	410.0	401.0	403.0	412.0	410.0
2024-11-28	433	414.0	399.0	390.0	410.0	401.0	403.0	412.0
2024-11-29	496	433.0	414.0	399.0	390.0	410.0	401.0	403.0
2024-11-30	547	496.0	433.0	414.0	399.0	390.0	410.0	401.0

Imagem 13 - lag de 7 dias.

Adicionalmente, como parte do pré-processamento avançado, a base de dados foi enriquecida com variáveis defasadas (lags), como ilustrado na (imagem) permitindo a aplicação de modelos supervisionados com maior sensibilidade

temporal como regressão linear, árvores de decisão e redes neurais. A inclusão dessas defasagens potencializa o aprendizado dos padrões sequenciais de consumo e podendo ser explorada em futuros aprimoramentos do modelo.

Dessa forma, a autocorrelação temporal mostra presença de valores semelhantes em Vendas e nos Lags consecutivos reforça a autocorrelação positiva na série, ou seja, dias com vendas altas geralmente são seguidos por dias com vendas semelhantes. Isso é ideal para modelos como regressão linear múltipla, Random Forest e LSTM, que se beneficiam de entradas sequenciais.

Com isso, a estabilidade nas últimas datas nas datas de fim de novembro de 2024, os valores das lags são muito próximos, o que pode indicar um período de estabilidade de demanda ou um padrão repetitivo semanal, facilitando previsões de curto prazo.

Entretanto, a variação no início da série nas datas iniciais (ex: janeiro de 2022), há maior oscilação entre os valores das lags, o que pode indicar menos previsibilidade ou maior impacto de fatores externos (estoques limitados, baixa maturidade do canal de vendas, etc.).