



Previsão de Demanda em Frigorífico para Apoio à Gestão de Estoque

**Artigo relativo ao Trabalho de Conclusão de Curso
do Bacharelado em Sistemas de Informação**

Aluno

Fellipe de Brito Lira Batista

Orientador

Cícero Garrozi

Departamento de Estatística e Informática

Recife – PE
Junho de 2026

Fellipe de Brito Lira Batista

Previsão de Demanda em Frigorífico para Apoio à Gestão de Estoque

Artigo apresentado ao Curso de Bacharelado em Sistemas de Informação da Universidade Federal Rural de Pernambuco, como requisito parcial para obtenção do título de Bacharel em Sistemas de Informação.

Universidade Federal Rural de Pernambuco – UFRPE
Departamento de Estatística e Informática
Curso de Bacharelado em Sistemas de Informação

Orientador: Cícero Garrozi

Recife
Junho de 2026

Previsão de Demanda em Frigorífico para Apoio à Gestão de Estoque

Fellipe de Brito Lira Batista ¹, Cícero Garrozi ¹

¹Departamento de Estatística e Informática
Universidade Federal Rural de Pernambuco
Rua Dom Manuel de Medeiros, s/n – CEP: 52171-900
Recife – PE – Brasil

fellipe.brito@ufrpe.br, cicero.garrozi@ufrpe.br

Resumo. *Este trabalho apresenta o desenvolvimento e a avaliação de uma solução de previsão de vendas por produto aplicada ao contexto do Frigorífico Candeias. A proposta utiliza dados históricos reais de vendas com o objetivo de apoiar decisões relacionadas a compras, estoque e planejamento operacional. Para isso, foi estruturada uma pipeline de processamento de dados, foi adotado o modelo Prophet como ponto de partida e foi realizada a comparação com outros métodos de previsão e estratégias complementares de avaliação. Os experimentos foram conduzidos por meio de backtesting, buscando verificar o comportamento dos modelos em condições próximas ao uso prático. Como parte da solução, também foi desenvolvido um aplicativo de apoio à decisão para facilitar a visualização e a interpretação das previsões pelo dono da empresa e por gestores. Como contribuição, o trabalho integra tratamento de dados, modelagem preditiva, avaliação experimental e disponibilização dos resultados em uma solução aplicada ao ambiente empresarial.*

Abstract. *This work presents the development and evaluation of a product-level sales forecasting solution applied to the context of Frigorífico Candeias. The proposal uses real historical sales data in order to support decisions related to purchasing, inventory, and operational planning. For this purpose, a data processing pipeline was structured, the Prophet model was adopted as a starting point, and comparisons were made with other forecasting methods and complementary evaluation strategies. The experiments were conducted through backtesting, aiming to verify model behavior under conditions close to practical use. As part of the solution, a decision-support application was also developed to facilitate the visualization and interpretation of forecasts by company owners and managers. As a contribution, the work integrates data processing, predictive modeling, experimental evaluation, and result delivery in a solution applied to the business environment.*

1. Introdução

A previsão de vendas representa uma atividade essencial para empresas que buscam aprimorar seus processos de planejamento, controle de estoque, definição de metas comerciais e tomada de decisão [15, 4]. Em ambientes comerciais cada vez mais competitivos,

a capacidade de estimar a demanda futura permite que gestores organizem melhor suas operações, reduzam desperdícios e antecipem possíveis variações no comportamento de consumo [15, 5]. Dessa forma, a utilização de métodos de previsão torna-se relevante para apoiar decisões relacionadas à compra de produtos, reposição de estoque, estratégias de venda e planejamento financeiro.

Apesar de sua importância, a previsão de vendas ainda é um desafio em muitos contextos empresariais. Variações sazonais, promoções, mudanças no comportamento dos consumidores, oscilações econômicas e limitações na qualidade dos dados históricos podem comprometer a precisão das estimativas [6, 5]. Em empresas que comercializam grande variedade de produtos, esse desafio torna-se ainda mais evidente, pois cada item pode apresentar um padrão próprio de demanda, incluindo produtos de venda recorrente, itens com picos ocasionais e séries com muitos dias sem venda [11, 15]. Como consequência, previsões inadequadas podem gerar excesso de estoque, falta de produtos, perda de vendas, aumento de custos operacionais e dificuldade na definição de estratégias comerciais [5, 4, 15].

O estudo foi realizado no contexto do Frigorífico Candeias, uma empresa de pequeno porte do setor alimentício, composta por aproximadamente 12 funcionários e voltada principalmente à comercialização de carnes. A empresa realiza vendas em loja física e por *delivery*, atendendo tanto consumidores finais quanto clientes empresariais. Atualmente, parte das decisões de compra e reposição de estoque é realizada de forma empírica pelo gestor, com base na experiência prática e na observação do movimento de vendas. Como o planejamento de compras ocorre em frequência semanal e mensal, problemas como compras emergenciais e variações de demanda em feriados e datas comemorativas tornam a previsão de vendas uma ferramenta relevante para apoiar decisões operacionais.

Nesse cenário, técnicas computacionais baseadas em séries temporais podem contribuir para uma análise mais estruturada do comportamento das vendas [7, 13]. Ao utilizar registros anteriores, modelos de previsão permitem identificar padrões, tendências e possíveis variações ao longo do tempo. Entretanto, a escolha do método de previsão deve considerar as características dos dados, uma vez que diferentes abordagens de previsão, incluindo modelos estatísticos e redes neurais, nem sempre apresentam melhor desempenho em séries com baixo volume, alta variabilidade ou demanda intermitente [7, 4, 11]. Por esse motivo, a comparação entre modelos de previsão e métodos de referência torna-se uma etapa importante para avaliar a qualidade das estimativas geradas [7, 9].

O problema investigado neste trabalho consiste em verificar se registros históricos de venda do Frigorífico Candeias podem ser transformados em previsões úteis ao planejamento comercial e operacional. Em termos práticos, busca-se estimar a quantidade futura vendida por produto de forma suficientemente compreensível para apoiar decisões relacionadas a compras, estoque, metas e análise de desempenho.

Com base nesse problema, o objetivo do trabalho é desenvolver e avaliar uma solução de previsão de vendas por produto que combine tratamento dos dados, aplicação do modelo Prophet, comparação com métodos de referência e visualização dos resultados em uma aplicação mobile de apoio à decisão. Para isso, o estudo contempla a organização da base histórica, a preparação das séries temporais, a realização de experimentos de *back-testing*, a comparação entre previsões diárias e semanais e a análise da utilidade prática

dos resultados gerados.

A abordagem adotada parte da utilização de dados históricos reais como base para geração das previsões. Após a preparação da base, aplica-se o modelo Prophet [13] e realiza-se a comparação com métodos de referência, como previsões baseadas no último valor observado, médias móveis e médias por dia da semana. A avaliação é conduzida por meio de experimentos de *backtesting*, nos quais os modelos são treinados com parte do histórico e testados em períodos posteriores [7]. Também são avaliadas estratégias de previsão semanal e de seleção de modelos por produto, uma vez que decisões de compra, estoque e produção podem ser mais bem representadas pelo volume acumulado em períodos operacionais.

Como contribuição, o trabalho apresenta uma análise aplicada da previsão de vendas em contexto empresarial, integrando dados históricos reais, *pipeline* de processamento, modelagem preditiva, comparação experimental e uma ferramenta de apoio à visualização dos resultados. Além disso, busca evidenciar que a escolha do modelo deve considerar o perfil de demanda de cada produto e que a granularidade temporal da previsão interfere diretamente em sua aderência às decisões reais de compra, estoque e produção.

1.1. Organização do Trabalho

O restante deste artigo está organizado da seguinte forma. A Seção 2 apresenta os trabalhos relacionados, destacando estudos e abordagens existentes para previsão de vendas, previsão de demanda e apoio à decisão. A Seção 3 apresenta a fundamentação teórica, com os conceitos de séries temporais, métricas de avaliação, modelos de previsão e estratégias complementares utilizadas no trabalho. A Seção 4 descreve a metodologia utilizada, incluindo a base de dados, o processamento das informações e o procedimento experimental. A Seção 5 apresenta a arquitetura de integração e o aplicativo de apoio à decisão. A Seção 6 apresenta os experimentos e a análise dos resultados, com foco na comparação entre estratégias de previsão e na justificativa da escolha dos modelos. Por fim, a Seção 7 apresenta as conclusões, limitações do trabalho e sugestões para trabalhos futuros.

2. Trabalhos Relacionados

Diversos trabalhos têm investigado o problema de previsão de vendas e demanda no varejo, considerando tanto métodos estatísticos tradicionais quanto abordagens baseadas em aprendizado de máquina e aprendizado profundo. Em geral, esses estudos mostram que a acurácia das previsões depende fortemente das características da série analisada, como sazonalidade, influência de feriados, intermitência e variabilidade entre produtos.

Huber e Stuckenschmidt [6] analisaram a previsão diária de demanda no varejo com ênfase em dias especiais de calendário, como feriados e datas com comportamento atípico. Os autores trataram o problema como uma tarefa supervisionada de aprendizado de máquina e compararam diferentes métodos, incluindo redes neurais e árvores impulsionadas por gradiente. Os resultados indicaram superioridade dos métodos de aprendizado de máquina em relação às abordagens tradicionais, além de mostrarem que os efeitos de calendário precisam ser tratados de forma explícita quando se deseja melhorar a previsão em nível diário.

No trabalho de Guo et al. [5], foi proposta uma abordagem híbrida que combina Prophet e *Support Vector Regression* (SVR) para previsão de demanda com sazonalidade. Nesse modelo, o Prophet é utilizado para capturar tendência, sazonalidade e feriados, enquanto o SVR é empregado para modelar padrões não lineares. Os autores observaram que a abordagem híbrida apresentou desempenho superior ao Prophet isolado e a outros métodos de comparação, indicando que a combinação de modelos pode ser vantajosa em cenários com comportamento mais complexo.

Žunić et al. [15] apresentaram um *framework* de previsão de vendas baseado no Prophet e em estratégias de *backtesting*, utilizando dados reais de uma grande empresa varejista. Além da previsão em si, os autores também propuseram uma classificação do portfólio de produtos de acordo com o nível esperado de confiabilidade das previsões. O trabalho evidencia que, em bases reais, produtos distintos podem apresentar diferentes níveis de previsibilidade, o que torna relevante a avaliação retroativa do desempenho antes da adoção prática do modelo.

Nasseri et al. [11] compararam modelos baseados em árvores com redes *Long Short-Term Memory* (LSTM) em dados reais de supermercado, incluindo variáveis externas como preço, clima e dados relacionados à COVID-19. Os autores concluíram que os modelos baseados em árvores apresentaram desempenho superior ao LSTM nos cenários avaliados. Esse resultado reforça que modelos mais complexos nem sempre produzem as melhores previsões e que a comparação entre diferentes abordagens continua sendo necessária em aplicações reais de varejo.

De forma semelhante, Ecevit et al. [4] compararam LSTM, Prophet e *Seasonal Autoregressive Integrated Moving Average* (SARIMA) em um problema de previsão de curto prazo para vendas em *e-commerce*. Os resultados mostraram que o modelo LSTM obteve melhor desempenho no conjunto analisado, superando o Prophet e o SARIMA. Embora o contexto seja diferente do varejo físico, o estudo mostra que o desempenho do modelo depende do tipo de dado, da granularidade da série e do horizonte de previsão adotado.

Com base nesses trabalhos, observa-se que não existe um único modelo universalmente superior para todos os problemas de previsão de vendas. A literatura indica que fatores como sazonalidade, feriados, intermitência, granularidade temporal e perfil dos produtos influenciam diretamente os resultados. Embora os estudos relacionados apresentem diferentes abordagens baseadas em aprendizado de máquina e redes neurais, este trabalho concentrou-se principalmente na avaliação de modelos estatísticos e de séries temporais. Essa escolha está relacionada à maior interpretabilidade dos resultados, ao menor custo computacional de execução e à aderência ao contexto operacional da empresa estudada, no qual a previsão deve ser compreensível e viável para uso recorrente em apoio à decisão. Nesse sentido, o presente trabalho diferencia-se por investigar a previsão de vendas em uma base real, inicialmente com o uso do Prophet, mas também por comparar diferentes estratégias, avaliar o desempenho em nível semanal e considerar a seleção de modelos por produto como forma de apoiar a tomada de decisão gerencial.

3. Fundamentação Teórica

Esta seção apresenta os principais conceitos utilizados no desenvolvimento e na avaliação da solução proposta. O objetivo é separar a base conceitual do procedimento metodológico, deixando a metodologia concentrada nas etapas executadas no projeto.

3.1. Previsão de Séries Temporais

A previsão de séries temporais busca estimar valores futuros a partir de observações históricas ordenadas no tempo. No contexto deste trabalho, cada série temporal representa a quantidade vendida de um produto ao longo dos dias. Esse tipo de abordagem permite observar padrões como tendência, sazonalidade, recorrência de vendas, períodos sem movimentação e variações atípicas de demanda [7].

Em problemas de previsão de vendas, a granularidade temporal influencia diretamente a utilidade da previsão. Previsões diárias podem ser úteis para acompanhamento detalhado, mas também são mais sensíveis a deslocamentos de venda entre dias próximos. Já previsões semanais tendem a representar melhor decisões de compra, produção e estoque, pois essas decisões normalmente consideram o volume acumulado em um período operacional.

3.2. Avaliação de Modelos de Previsão

A avaliação de modelos de previsão exige a comparação entre valores previstos e observados. Neste trabalho, foram utilizadas três métricas quantitativas: *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE) e *Weighted Absolute Percentage Error* (WAPE) [9].

O MAE representa o erro absoluto médio entre previsão e valor observado, sendo interpretado na mesma unidade da variável prevista. O RMSE também mede erro em unidade da variável analisada, mas penaliza desvios maiores de forma mais intensa. O WAPE relaciona o erro absoluto total ao volume observado no período avaliado, sendo útil para comparar produtos com diferentes volumes de venda.

Os experimentos utilizam *holdout* e *backtesting*. O *holdout* corresponde a um período final da série histórica separado para teste, sem ser utilizado no treinamento. O *backtesting* consiste em simular previsões em períodos passados, treinando os modelos com dados anteriores e comparando as previsões com os valores observados. Quando esse processo é repetido em diferentes janelas temporais, tem-se o *backtesting* multijanelas, que reduz a dependência de uma única data de corte.

3.3. Modelos Estatísticos e Métodos de Referência

O Prophet é um modelo de previsão de séries temporais desenvolvido para lidar com tendência, sazonalidade e efeitos de calendário, como feriados e datas especiais [13]. Por permitir configuração relativamente simples e boa interpretabilidade, o Prophet é frequentemente utilizado como ponto de partida em problemas aplicados de previsão.

Além de modelos mais estruturados, é comum utilizar métodos de referência, ou *baselines*, para verificar se um modelo apresenta ganho real em relação a estratégias simples. Entre esses métodos estão a repetição do último valor observado, médias móveis e

médias históricas por dia da semana. Em séries de baixo volume, esses métodos podem apresentar desempenho competitivo, especialmente quando o comportamento da demanda é irregular.

Também existem modelos estatísticos automáticos, como AutoARIMA, AutoETS e AutoTheta. Esses métodos buscam selecionar automaticamente estruturas e parâmetros adequados para a série analisada, reduzindo a necessidade de ajuste manual e permitindo comparação com modelos de referência e com o Prophet [8, 10, 2].

3.4. Demanda Intermitente e Tratamento de Zeros

Parte dos produtos avaliados apresenta demanda intermitente, isto é, muitos dias sem venda e poucos dias com registro de venda. Esse comportamento dificulta a previsão diária, pois pequenas diferenças entre o dia previsto e o dia em que a venda ocorre podem gerar erro elevado, mesmo quando o volume semanal previsto é próximo ao observado.

Modelos como Croston, *Syntetos-Boylan Approximation* (SBA) e *Teunter-Syntetos-Babai* (TSB) foram propostos para lidar com séries de demanda intermitente, separando ou ajustando componentes relacionados ao tamanho da demanda e ao intervalo entre ocorrências de venda [3, 12, 14].

Outra forma de lidar com muitos zeros é aplicar estratégias de pós-processamento, chamadas neste trabalho de *zero handling*. No *prediction floor*, previsões muito pequenas são zeradas com base em um limiar associado ao histórico dos dias com venda. No *occurrence gate*, um modelo auxiliar estima a probabilidade de ocorrer venda em determinado dia, mantendo a previsão apenas quando essa probabilidade ultrapassa um limiar. A estratégia *combined* combina as duas regras.

3.5. Estratégias de Combinação e Agregação Temporal

Um *ensemble* combina previsões geradas por mais de um modelo, buscando reduzir a dependência de uma única estratégia. Essa combinação pode ser feita por médias simples ou ponderadas, como nos casos de pesos 50/50 e 70/30 avaliados neste trabalho.

Outra estratégia utilizada foi o modelo semanal híbrido. Nessa abordagem, a previsão é realizada inicialmente em nível semanal, estimando o volume total vendido por produto em cada semana. Em seguida, esse total é distribuído entre os dias da semana com base no perfil histórico de participação de cada dia nas vendas do produto. Assim, a soma das previsões diárias preserva o total semanal previsto, combinando uma previsão mais alinhada à decisão operacional com uma visualização diária quando necessário.

3.6. Computação em Nuvem e AWS

A computação em nuvem permite o uso de recursos computacionais, como armazenamento, processamento e serviços de aplicação, por meio da internet, sem que a organização precise manter toda a infraestrutura localmente¹. No contexto deste trabalho, esse conceito é relevante porque a solução proposta precisa receber novos arquivos,

¹Amazon Web Services. *What is Cloud Computing?*. Disponível em: <https://aws.amazon.com/what-is-cloud-computing/>. Acesso em: 08 jul. 2026.

armazenar dados, executar etapas de processamento e disponibilizar previsões para consulta pelo aplicativo.

A Amazon Web Services (AWS) é uma plataforma de computação em nuvem da Amazon que oferece serviços para armazenamento, processamento, orquestração de fluxos e disponibilização de aplicações. Neste projeto, a AWS foi considerada como camada de suporte à operacionalização da solução, permitindo que a *pipeline* de previsão seja integrada ao envio de documentos e ao consumo dos resultados pelo aplicativo.

4. Metodologia

Esta seção descreve os procedimentos adotados para o desenvolvimento da solução de previsão de vendas. A metodologia contempla a caracterização da pesquisa, a descrição da base de dados, o processamento das informações, a preparação das séries temporais, a seleção dos produtos piloto, os modelos avaliados e o procedimento experimental.

4.1. Caracterização da Pesquisa

O trabalho caracteriza-se como uma pesquisa aplicada, pois busca resolver um problema prático relacionado ao planejamento de vendas e estoque em um contexto empresarial. A proposta parte da necessidade de estimar a quantidade futura vendida por produto, de forma a apoiar o dono da empresa e gestores em decisões relacionadas a compras, reposição de estoque, acompanhamento de demanda e análise de desempenho comercial.

A pesquisa possui abordagem quantitativa e caráter experimental. A abordagem quantitativa ocorre pelo uso de dados históricos de vendas, métricas de erro e comparação numérica entre modelos de previsão. O caráter experimental está relacionado à execução de diferentes métodos de previsão sobre séries temporais reais, com avaliação por *back-testing* e comparação entre valores previstos e observados.

A solução desenvolvida foi organizada em três camadas: modelagem preditiva, infraestrutura de processamento e publicação, e aplicativo de apoio à decisão. A primeira camada prepara os dados, treina os modelos e avalia os resultados. A segunda organiza a ingestão de novos arquivos, o armazenamento, a execução da *pipeline* e a publicação das previsões. A terceira apresenta os resultados ao usuário final.

4.2. Base de Dados

A base de dados utilizada é composta por registros históricos de vendas do Frigorífico Candeias. Os dados têm início em 01/09/2023 e são atualizados semanalmente com novos registros, formando uma base incremental de vendas. Para fins de avaliação experimental, foi considerada uma fotografia da base consolidada em maio de 2026. Nessa versão, a base contém 226.503 registros de venda, abrangendo 1.685 produtos distintos e 24 grupos comerciais.

Os dados brutos foram fornecidos pelo sistema utilizado pela empresa em dois formatos principais: planilhas eletrônicas e documentos em PDF. Como o próprio sistema disponibiliza informações nesses formatos, foi criada uma etapa de ingestão capaz de converter registros com estruturas diferentes para um formato comum. Para isso, foi definido um esquema canônico de vendas, contendo campos como data da venda, nome

do produto, grupo do produto, quantidade vendida, preço unitário, valor total do item, origem do registro e identificador do lote de ingestão.

Na versão da base utilizada nesta etapa, o arquivo canônico consolidado de vendas contém 226.503 registros, provenientes de lotes de ingestão associados aos arquivos de vendas disponíveis. Após limpeza, enriquecimento com o cadastro de produtos, agregação por produto e data, tratamento de *outliers* e preenchimento dos dias sem venda com valor zero, a série temporal diária final passou a conter 472.660 observações produto-data, abrangendo 1.680 produtos distintos e 25 grupos comerciais. Assim, a diferença entre os 226.503 registros e as 472.660 observações ocorre porque a base de modelagem passa a representar uma grade diária contínua por produto. Cada observação representa a quantidade vendida de um produto em uma determinada data. Portanto, neste trabalho, a previsão de vendas refere-se à previsão da quantidade vendida por produto, e não diretamente à previsão de faturamento.

Os principais campos utilizados na base de modelagem estão descritos na Tabela 1. Esses campos representam a estrutura mínima utilizada para organizar as vendas como séries temporais por produto.

Tabela 1. Dicionário dos principais campos utilizados na modelagem

Campo	Tipo	Descrição	Uso na modelagem
ds	Data	Data da observação	Ordenação temporal da série e separação dos períodos de treino e teste
Produto	Texto	Nome do produto vendido	Identificação da série temporal de cada produto
Grupo	Texto	Agrupamento comercial do produto	Apoio à análise por categoria e interpretação dos resultados
y	Numérico	Quantidade vendida do produto na data correspondente	Variável-alvo prevista pelos modelos

Fonte: autoria própria (2026).

Embora a base canônica possua outros campos auxiliares, como valor total, preço unitário, origem do arquivo e identificador do lote de ingestão, os campos apresentados na Tabela 1 correspondem aos atributos diretamente utilizados na construção das séries temporais e na avaliação dos modelos.

Nos testes finais de modelagem, foi utilizada a versão mais recente da série diária consolidada, cobrindo aproximadamente o período de setembro de 2023 a maio de 2026. Durante a preparação da base, o intervalo de 09/06/2025 a 29/08/2025 exigiu tratamento específico em razão da troca de sistema e da perda ou incompletude inicial de registros. Após a reconstrução da série, esse período passou a compor a base final utilizada nos testes comparativos.

4.3. Pipeline de Processamento dos Dados

O processamento dos dados foi estruturado em uma *pipeline* com etapas separadas para entrada, tratamento, modelagem e geração de resultados. Essa organização permite que novos arquivos de venda sejam incorporados ao fluxo sem a necessidade de refazer manualmente todas as etapas de preparação da base.

O fluxo geral da *pipeline* é apresentado na Figura 1. O processo parte dos relatórios de venda disponibilizados pelo sistema da empresa, realiza a ingestão e a padronização dos registros, consolida a base histórica, gera a série temporal diária por produto, executa o tratamento dos dados, treina e avalia os modelos de previsão e publica os resultados para consumo pelo aplicativo desenvolvido.



Figura 1. Fluxo de dados da solução proposta

Fonte: autoria própria (2026).

Essa estrutura separa os dados brutos dos dados tratados e dos resultados finais. Os arquivos originais permanecem preservados, enquanto as bases intermediárias e processadas são utilizadas nas etapas de análise e modelagem. O aplicativo não realiza o treinamento dos modelos diretamente; sua função é permitir o envio de novos documentos e consumir os resultados publicados pela camada de processamento.

4.4. Pré-processamento e Preparação das Séries Temporais

Antes da aplicação dos modelos, os registros extraídos de planilhas e documentos PDF foram convertidos para o esquema canônico de vendas. Nessa etapa, foram padronizados campos como data da venda, nome do produto, grupo comercial, quantidade, preço unitário, valor total, origem do arquivo e identificador do lote de ingestão. Também foram removidas linhas-resumo, registros agregados, produtos nulos, datas ausentes, quantidades inválidas e itens não diretamente relacionados à venda de produtos.

Após a limpeza inicial, os registros foram enriquecidos com o cadastro de produtos, permitindo associar cada item ao seu grupo comercial. Também foram tratadas

inconsistências observadas em parte das exportações do sistema, incluindo correções associadas ao período afetado pela troca de sistema. Em seguida, as vendas foram agregadas por produto e por data, resultando em uma série temporal diária para cada produto.

Para permitir a comparação entre produtos e a aplicação dos modelos em uma malha temporal contínua, os dias sem venda foram preenchidos com valor zero. Assim, cada produto passou a possuir uma sequência diária completa dentro do período analisado. Em seguida, foi aplicado o tratamento de *outliers* por produto, com identificação de picos a partir do limite superior definido pelo intervalo interquartil. O intervalo interquartil, representado por IQR, corresponde à diferença entre o terceiro quartil ($Q3$) e o primeiro quartil ($Q1$). Neste trabalho, valores acima de $Q3 + 1,5 \times IQR$ foram considerados candidatos a valores atípicos. Conforme o perfil da série, os valores extremos foram limitados por *capping* ou removidos da base de treino. No *capping*, o valor extremo é substituído por um limite máximo aceitável, preservando a observação na série. Na remoção, o registro é excluído da base de treino quando o pico é considerado pouco representativo do comportamento usual do produto. Essas regras reduzem o impacto de eventos isolados sobre o ajuste dos modelos.

As séries também foram classificadas segundo indicadores calculados a partir do histórico de cada produto. O indicador *zero ratio* representa a proporção de dias sem venda; *sum_y* representa o volume total vendido na série; *cv_positive* representa o coeficiente de variação calculado apenas sobre os dias em que houve venda; e *spike_ratio* representa a razão entre a maior quantidade vendida em um dia e a mediana dos dias com venda.

A literatura sobre previsão de demanda reconhece que séries com baixa frequência de venda, muitos zeros e comportamento intermitente exigem atenção específica na avaliação e modelagem [3, 12, 14]. Neste trabalho, entretanto, os limiares apresentados na Tabela 2 não foram adotados como valores universais da literatura, mas como critérios heurísticos definidos a partir da análise exploratória da base estudada.

O limiar de *zero ratio* maior ou igual a 0,75, combinado com *sum_y* menor ou igual a 200, foi utilizado para identificar séries com baixa frequência de venda e baixo volume acumulado. Já o limiar de *zero ratio* maior ou igual a 0,55 foi adotado para classificar produtos em que os dias sem venda representam mais da metade da série, indicando que os zeros deixam de ser exceção e passam a ser uma característica dominante do comportamento observado. O critério de *spike_ratio* maior ou igual a 3,5 foi utilizado para identificar séries em que o maior valor diário é pelo menos 3,5 vezes superior à mediana dos dias com venda, sinalizando a presença de picos fortes. Por fim, o perfil de séries mais estáveis foi associado a *cv_positive* menor ou igual a 0,80 e *zero ratio* menor ou igual a 0,50, indicando menor variabilidade relativa nos dias com venda e menor predominância de dias zerados. A Tabela 2 apresenta os critérios utilizados.

Os limiares foram definidos como critérios práticos de análise, com o objetivo de apoiar o tratamento de valores atípicos e a interpretação dos resultados. Eles não têm a finalidade de estabelecer uma regra universal para séries de vendas, mas sim de criar uma segmentação coerente com as características da base analisada.

Tabela 2. Critérios utilizados para classificação do perfil das séries

Perfil	Critério utilizado
<i>Intermitente de baixo volume</i>	$zero_ratio \geq 0,75$ e $sum_y \leq 200$
<i>Muitos zeros</i>	$zero_ratio \geq 0,55$
<i>Picos fortes</i>	$spike_ratio \geq 3,5$
<i>Mais estáveis</i>	$cv_positive \leq 0,80$ e $zero_ratio \leq 0,50$
<i>Misto</i>	Séries que não atendem aos critérios anteriores

Fonte: autoria própria (2026).

4.5. Seleção dos Produtos Piloto

A avaliação inicial foi realizada em um conjunto amplo de produtos, permitindo observar o comportamento geral dos modelos em diferentes perfis de demanda. Para a análise experimental principal, entretanto, foi utilizado um recorte de produtos piloto com maior interesse prático para a empresa e melhor aderência ao contexto de varejo.

Inicialmente, os testes do piloto consideraram oito produtos. Durante a análise, o produto MOIDA DE PATINHO foi identificado como um item com perfil distinto, associado a vendas para outros estabelecimentos e comportamento mais próximo de *Business-to-Business* (B2B) ou carne *delivery*. Por esse motivo, ele foi retirado do piloto de varejo e deve ser tratado como um caso separado. Além disso, o produto PAO DE ALHO TRADICIONAL SANTA MASSA 400G foi removido da comparação principal por apresentar comportamento atípico e possível mudança, substituição ou instabilidade de *Stock Keeping Unit* (SKU), o que poderia distorcer a média geral dos experimentos.

Assim, a comparação final foi conduzida com seis produtos do piloto limpo:

- FAROFA DA CASA;
- GALETO ASSADO;
- FILE MIGNON;
- KIT 10 HAMBURGUESES;
- FILE DE PEITO DE FRANGO;
- CHA DE DENTRO (COXAO MOLE).

A definição do piloto limpo não significa que os demais produtos da base tenham sido desconsiderados no desenvolvimento da solução. A base completa foi utilizada nas etapas de consolidação, tratamento e análise exploratória, permitindo observar diferentes perfis de venda. Entretanto, para a comparação experimental principal, optou-se por um recorte menor e mais controlado, de modo a reduzir distorções provocadas por produtos com comportamento muito distinto do varejo analisado.

A escolha dos seis produtos do piloto final teve como objetivo melhorar a qualidade do recorte analítico e a comparabilidade entre os modelos, e não apenas selecionar produtos com melhor desempenho de previsão. Foram considerados quatro critérios principais: relevância operacional para decisões de compra e estoque, disponibilidade de histórico suficiente para treinamento e teste, aderência ao comportamento típico de venda ao consumidor final e presença de diferentes perfis de demanda dentro do recorte. Assim, o piloto inclui produtos com maior recorrência de venda, produtos de menor volume e produtos com maior irregularidade, permitindo avaliar os modelos em cenários distintos sem misturar, na média principal, itens com perfil B2B ou possível instabilidade cadastral.

Dessa forma, os seis produtos mantidos no piloto limpo representam um conjunto controlado para avaliação comparativa das estratégias de previsão. Os produtos removidos, como MOIDA DE PATINHO e PAO DE ALHO TRADICIONAL SANTA MASSA 400G, permanecem relevantes para estudos futuros, mas foram separados da comparação principal por apresentarem características que poderiam dificultar a interpretação dos resultados agregados.

4.6. Modelos e Estratégias Avaliados

A escolha dos modelos avaliados foi organizada de forma incremental. O Prophet foi adotado como modelo inicial de referência por permitir uma primeira modelagem estruturada das séries temporais, considerando tendência, sazonalidade e efeitos de calendário, além de apresentar configuração compatível com a reexecução periódica da *pipeline*. Essa escolha também se justifica pela interpretabilidade dos resultados e pela possibilidade de comparação com estudos relacionados que aplicam o Prophet em problemas de previsão de vendas.

A partir desse modelo inicial, foram avaliadas diferentes configurações do Prophet: treinamento diário, avaliação diária agregada por semana, treinamento semanal puro e inclusão de calendário de feriados. O calendário considerou feriados nacionais, feriados móveis, datas estaduais e municipais relevantes para Pernambuco e Jaboatão dos Guararapes, além de datas especiais de varejo, como Dia das Mães, Dia dos Pais, vésperas de Natal e Ano Novo.

Para verificar se o Prophet apresentava ganho real em relação a estratégias mais simples, foram incluídos métodos de referência, como último valor observado, média móvel de sete dias, média por dia da semana, repetição da última semana e médias móveis semanais. Esses métodos serviram como base comparativa para identificar situações em que um modelo mais estruturado de fato acrescentava desempenho.

Também foram avaliados modelos voltados à demanda intermitente, como Croston, *Syntetos-Boylan Approximation* (SBA) e *Teunter-Syntetos-Babai* (TSB), devido à presença de produtos com muitos dias sem venda. Além disso, foram testados modelos estatísticos automáticos por meio da biblioteca *StatsForecast*², estratégias de *zero handling*, combinações por *ensemble* e o modelo semanal híbrido. Esses grupos foram incluídos para comparar alternativas com diferentes níveis de complexidade e verificar quais estratégias eram mais adequadas aos perfis de demanda observados no piloto.

Experimentos exploratórios com Chronos/Bolt também foram realizados [1], mas essa abordagem não foi selecionada como estratégia principal. A comparação final priorizou modelos com menor custo de execução, maior interpretabilidade e maior aderência ao processo de revisão periódica da solução em ambiente operacional.

4.7. Procedimento Experimental

A avaliação dos modelos foi realizada por meio de *backtesting*, conforme definido na fundamentação teórica. O desenho experimental contemplou três níveis principais de avaliação. Primeiro, o Prophet diário foi avaliado com janelas de *holdout* de 28 e 80 dias,

²Nixtla. *StatsForecast*. Disponível em: <https://github.com/Nixtla/statsforecast>. Acesso em: 08 jul. 2026.

com o objetivo de verificar se o erro observado estava associado ao tamanho do período de teste ou à própria granularidade diária. Em seguida, foram comparadas estratégias diárias e semanais, incluindo a avaliação diária agregada em blocos semanais e o treinamento diretamente sobre séries semanais.

A escolha das janelas de *holdout* considerou diferentes objetivos de avaliação. A janela de 28 dias foi adotada como principal por representar aproximadamente quatro semanas de operação, permitindo avaliar o desempenho dos modelos em um período curto, mas ainda capaz de capturar variações semanais relevantes para decisões de compra, produção e estoque. A janela de 80 dias foi utilizada como avaliação complementar, com o objetivo de verificar se os erros observados no Prophet diário estavam associados apenas ao tamanho reduzido do período de teste ou se permaneciam em um horizonte mais longo. Por fim, a janela de 7 dias foi usada como diagnóstico operacional de curto prazo, aproximando a avaliação do cenário de previsão da próxima semana. Entretanto, por ser mais sensível a oscilações pontuais de demanda, essa janela não foi adotada como critério principal de escolha dos modelos.

Além disso, a avaliação incluiu testes com feriados, tratamentos de zeros, modelos intermitentes, modelos estatísticos automáticos, *baselines* e *ensembles*. Para reduzir a dependência de uma única data de corte, a comparação principal adotou *backtesting* multijanelas, utilizando janelas de 28 dias como critério principal de seleção de modelos.

A comparação principal foi conduzida no piloto limpo com seis produtos de varejo, considerando a base reconstruída, a métrica principal semanal e condições equivalentes entre os modelos avaliados. O WAPE semanal foi utilizado como métrica principal, enquanto WAPE diário, MAE e RMSE foram mantidos como métricas complementares.

5. Arquitetura de Integração

Além da modelagem preditiva, a solução contempla uma arquitetura de integração e publicação dos resultados em ambiente de nuvem. Essa arquitetura foi planejada para transformar a *pipeline* local e experimental em um fluxo operacional, capaz de receber novos documentos de venda, armazenar arquivos, executar etapas de processamento e disponibilizar previsões para consumo pelo aplicativo desenvolvido.

5.1. Arquitetura de Integração e Publicação dos Resultados

O Amazon S3 foi definido como repositório central dos dados e artefatos do projeto. A organização proposta utiliza separação lógica entre dados brutos, dados intermediários, bases processadas, resultados, modelos e arquivos publicados para consumo externo. Assim, os prefixos `raw/`, `interim/`, `processed/`, `reports/`, `models/` e `public/` representam, respectivamente, a entrada original de arquivos, os dados extraídos ou parcialmente tratados, as bases prontas para modelagem, os relatórios e métricas, os artefatos de modelo e os resultados finais publicados.

A integração entre o aplicativo e a camada de armazenamento foi estruturada por meio de uma fachada de *Application Programming Interface* (API) composta por Amazon API Gateway e funções AWS Lambda. Para o envio de documentos, foi desenvolvido um handler compatível com Lambda, responsável por receber a solicitação do aplicativo e

gerar uma URL pré-assinada do S3³. Com isso, o arquivo é enviado diretamente para o S3, sem trafegar pelo *backend* da aplicação.

Após o armazenamento do arquivo bruto, a *pipeline* executa as etapas de ingestão, processamento canônico, geração da série diária por produto, tratamento de valores atípicos, treinamento, avaliação e publicação. Para a orquestração desse fluxo, o AWS Step Functions é previsto como componente responsável por controlar a sequência de execução das etapas, registrar o andamento do processamento e tratar falhas⁴. Nos casos em que o treinamento em lote ultrapassar o perfil adequado para execução em Lambda, o Amazon ECS com AWS Fargate é considerado como alternativa para executar a modelagem em contêiner.

Com base nos experimentos finais, a saída principal recomendada para a *pipeline* passa a ser a previsão semanal, por estar mais alinhada ao planejamento de compra, estoque e produção. Quando houver necessidade de visualização diária, a distribuição do total semanal pode ser realizada como etapa posterior, inicialmente com base no perfil histórico de venda por dia da semana.

5.2. Aplicativo de Apoio à Decisão

Como parte da solução, foi desenvolvido um aplicativo mobile voltado à visualização dos resultados de previsão e ao apoio à tomada de decisão. O público principal da ferramenta é o dono da empresa ou gestor responsável por acompanhar vendas, produtos e possíveis necessidades de reposição.

O aplicativo apresenta uma visão resumida das previsões, uma lista de produtos monitorados, detalhes por produto, histórico recente, métricas de erro e indicadores de atenção. Essas informações são organizadas com o objetivo de facilitar a interpretação dos resultados técnicos por usuários que precisam tomar decisões práticas no dia a dia da empresa.

Na tela de resumo, o usuário visualiza a base mais recente publicada, a quantidade de produtos monitorados, a previsão acumulada para o período e informações de destaque, como grupo líder e produtos em atenção. Essa tela tem função gerencial, pois resume o estado atual das previsões sem exigir que o gestor consulte cada produto individualmente. Já a tela de produtos em destaque apresenta uma lista priorizada dos itens com maior previsão, permitindo identificar rapidamente quais produtos tendem a concentrar maior demanda no horizonte avaliado.

A Figura 2 apresenta telas do resumo executivo e dos produtos em destaque, utilizadas para sintetizar as principais informações da previsão para o gestor.

³Amazon Web Services. *Download and upload objects with presigned URLs*. Disponível em: <https://docs.aws.amazon.com/AmazonS3/latest/userguide/using-presigned-url.html>. Acesso em: 08 jul. 2026.

⁴Amazon Web Services. *What is AWS Step Functions?*. Disponível em: <https://docs.aws.amazon.com/step-functions/latest/dg/welcome.html>. Acesso em: 08 jul. 2026.

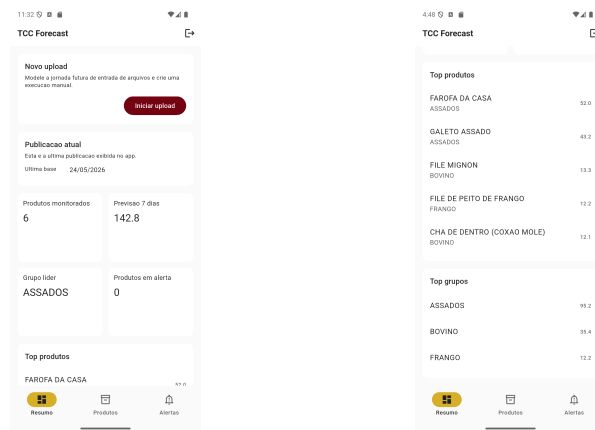


Figura 2. Telas de resumo executivo e produtos em destaque no aplicativo desenvolvido

Fonte: autoria própria (2026).

Além da visualização dos resultados, o aplicativo também participa do fluxo de atualização da base. Para isso, ele solicita ao backend uma URL de *upload*, envia o documento diretamente ao S3 e registra ou acompanha a execução da *pipeline*. Após o processamento, o aplicativo consome os dados publicados pela camada de previsão, incluindo resumos gerais, listagem de produtos, detalhes por produto, previsões futuras e status de execução.

Na listagem de produtos, o aplicativo permite consultar os itens monitorados, aplicar filtros e ordenar os resultados conforme critérios relevantes para o gestor, como grupo, confiança e maior previsão. Essa organização busca transformar a saída dos modelos em uma interface de consulta operacional, em que o usuário pode localizar rapidamente um produto e acessar sua previsão detalhada.

A Figura 3 mostra a tela de consulta dos produtos monitorados, enquanto a Figura 4 apresenta exemplos de visualização detalhada por produto.



Figura 3. Tela de listagem e consulta dos produtos monitorados

Fonte: autoria própria (2026).

Nas telas de detalhe, o aplicativo apresenta a previsão futura e o histórico recente

do produto selecionado. Com isso, o gestor pode comparar o comportamento passado com a estimativa gerada, avaliar a coerência da previsão e utilizar essa informação como apoio ao planejamento de compra, reposição ou produção.

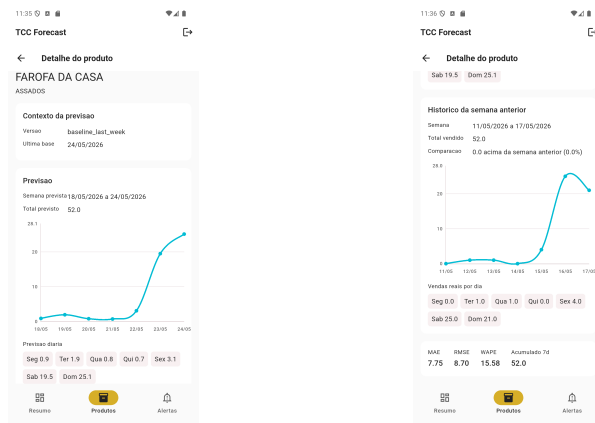


Figura 4. Telas de previsão e histórico por produto no aplicativo desenvolvido

Fonte: autoria própria (2026).

No escopo deste trabalho, o aplicativo atua como camada de apoio à decisão e de interação com a *pipeline*, mas não realiza o treinamento dos modelos localmente. O processamento analítico permanece no backend, enquanto o aplicativo consome dados reais publicados e apresenta informações consolidadas ao dono da empresa ou gestor.

6. Experimentos e Análise dos Resultados

Esta seção apresenta os experimentos realizados para avaliar as estratégias de previsão de vendas aplicadas aos produtos piloto, bem como a análise dos resultados obtidos. Para tornar a leitura mais clara, os testes foram organizados em quatro grupos: avaliação inicial do Prophet, aprimoramentos do Prophet, comparação com modelos alternativos e consolidação da estratégia final. Essa organização permite discutir não apenas o desempenho isolado de cada teste, mas também o papel de cada etapa na escolha da abordagem adotada para a solução.

6.1. Configuração Geral dos Experimentos

Os experimentos finais utilizaram a base diária consolidada a partir dos arquivos de vendas disponíveis. Essa base contém séries temporais por produto, com dias sem venda preenchidos com zero. Nos testes finais, foi utilizada a versão reconstruída da base, incluindo o período de 09/06/2025 a 29/08/2025, após o tratamento das inconsistências associadas à troca de sistema e à incompletude inicial dos registros.

A comparação principal foi realizada com seis produtos do piloto de varejo: FAROFA DA CASA, GALETO ASSADO, FILE MIGNON, KIT 10 HAMBURGUES, FILE DE PEITO DE FRANGO e CHA DE DENTRO (COXAO MOLE). Os produtos MOIDA DE PATINHO e PAO DE ALHO TRADICIONAL SANTA MASSA 400G foram retirados da média principal, respectivamente, por apresentarem perfil B2B associado a carne para *delivery* e comportamento atípico possivelmente relacionado a mudança ou instabilidade de cadastro do produto.

A métrica principal adotada foi o WAPE semanal, pois as decisões de compra, produção e estoque são mais alinhadas ao volume acumulado por semana do que ao erro em cada dia isolado. O WAPE diário, o MAE e o RMSE foram utilizados como métricas complementares. Em alguns testes, também foram observadas a mediana do WAPE e a pior janela, especialmente para avaliar a estabilidade operacional em produtos de baixo volume.

A Tabela 3 resume a organização dos experimentos e o objetivo principal de cada grupo de testes.

Tabela 3. Organização dos experimentos realizados

Grupo	Testes incluídos	Objetivo principal
Avaliação inicial	Prophet diário e avaliação semanal	Verificar granularidade da previsão
Aprimoramentos	Feriados e tratamento de zeros	Avaliar ajustes sobre o Prophet
Modelos alternativos	Seleção por produto e multijanelas	Comparar famílias de modelos
Estratégia final	Híbrido semanal e <i>holdout</i> de 7 dias	Definir uso operacional

Fonte: autoria própria (2026).

6.2. Experimento 1: Avaliação Inicial do Prophet

6.2.1. Prophet Diário

O primeiro teste avaliou o Prophet treinado em frequência diária. Nesse formato, cada produto possui uma observação por dia, incluindo dias sem venda com valor zero. A avaliação inicial utilizou *holdout* de 28 dias, correspondente aos 28 dias mais recentes disponíveis na base utilizada no experimento. Assim, os modelos foram treinados com os registros anteriores a esse intervalo e as previsões geradas para esse período final foram comparadas com os valores observados. Os resultados estão dispostos na Tabela 4.

O Prophet diário apresentou WAPE médio de 67,52, MAE médio de 1,65 e RMSE médio de 2,64. Apesar de os erros absolutos não serem elevados em todos os casos, o WAPE diário permaneceu alto devido à presença de muitos dias sem venda e à dificuldade de acertar a distribuição exata da demanda entre os dias.

Para complementar a análise numérica, a Figura 5 apresenta a comparação entre os valores realizados e previstos para o produto GALETO ASSADO, que apresentou o maior volume realizado entre os produtos do piloto limpo no período de *holdout*. Observa-se que o Prophet conseguiu representar alguns períodos de maior venda, mas também apresentou deslocamentos entre os dias previstos e os dias em que as vendas efetivamente ocorreram. Esse comportamento ajuda a explicar o WAPE diário elevado e reforça a necessidade de avaliar também estratégias agregadas semanalmente.

Tabela 4. Métricas por produto no Prophet diário com *holdout* de 28 dias

Produto	WAPE (%)	MAE	RMSE
FAROFA DA CASA	34,71	2,32	3,76
GALETO ASSADO	25,86	1,75	3,95
FILE MIGNON	75,64	1,59	2,22
KIT 10 HAMBURGUES	90,04	1,22	1,42
FILE DE PEITO DE FRANGO	87,58	1,21	1,59
CHA DE DENTRO	91,27	1,80	2,89

Fonte: autoria própria (2026).

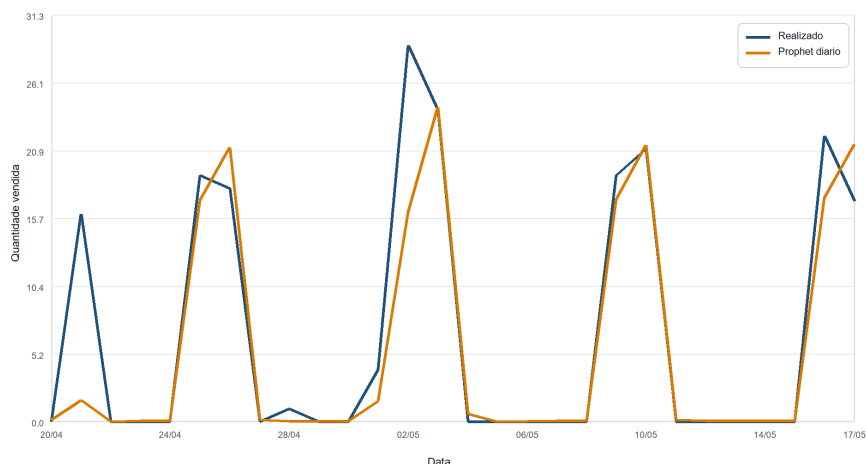


Figura 5. Comparação entre quantidade realizada e prevista pelo Prophet diário para o produto GALETO ASSADO no *holdout* de 28 dias.

Fonte: autoria própria (2026).

Os resultados por produto mostram que o erro diário continuou heterogêneo. O menor WAPE foi observado em GALETO ASSADO, com 25,86, enquanto CHA DE DENTRO e KIT 10 HAMBURGUESES permaneceram entre os casos mais difíceis. Isso reforça que o problema estava associado à granularidade diária e à intermitência das vendas, e não somente a um produto isolado.

Também foi realizado um teste com *holdout* de 80 dias para verificar se o erro elevado era consequência de uma janela de teste curta. Nessa rodada, o WAPE médio passou de 67,52, no *holdout* de 28 dias, para 70,74, no *holdout* de 80 dias, enquanto o MAE médio passou de 1,65 para 1,87 e o RMSE médio de 2,64 para 3,12. Esse resultado indicou que o principal problema não estava apenas no tamanho da janela de teste, mas na granularidade diária e na irregularidade das vendas.

6.2.2. Avaliação Diária Agregada e Semanal

Nesta etapa, foi avaliada a diferença entre prever e medir diariamente ou semanalmente. Para isso, foram comparados três cenários: Prophet diário avaliado dia a dia, Prophet diário com previsões agregadas em quatro blocos semanais e Prophet treinado diretamente sobre dados semanais. A comparação entre esses cenários está apresentada na Tabela 5.

Tabela 5. Comparação entre avaliação diária e semanal

Cenário	WAPE médio (%)
Prophet diário 28 dias	67,52
Prophet diário agregado em 4 semanas	37,34
Prophet semanal puro em 4 semanas	44,32

Fonte: autoria própria (2026).

Os resultados indicaram que a agregação semanal reduziu consideravelmente o erro em relação à avaliação diária. Isso mostra que o modelo pode errar a distribuição

da venda entre os dias, mas ainda assim estimar melhor o volume acumulado da semana. Como o planejamento de compra, estoque e produção tende a depender mais do volume semanal do que do acerto diário exato, a avaliação semanal mostrou-se mais adequada ao contexto do projeto.

Na base final avaliada, o Prophet semanal puro não superou o Prophet diário agregado em quatro semanas, o que sugere que, para este piloto, a agregação da saída diária permaneceu mais aderente ao comportamento observado. A Figura 6 ilustra esse comportamento para GALETO ASSADO, evidenciando que a comparação semanal é suficientemente estável para apoiar decisões operacionais, ainda que a dinâmica diária permaneça mais irregular.

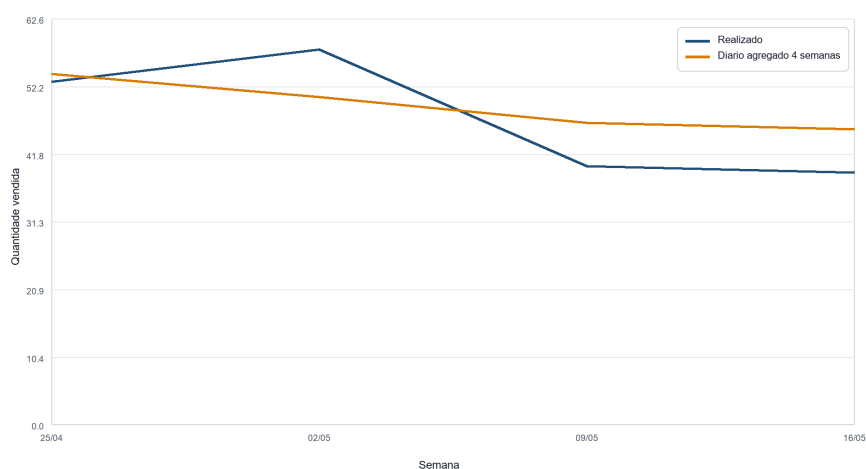


Figura 6. Comparação entre quantidade realizada e prevista pelo Prophet diário agregado em quatro semanas para o produto GALETO ASSADO.

Fonte: autoria própria (2026).

6.3. Experimento 2: Aprimoramentos do Prophet

6.3.1. Inclusão de Feriados

Também foi avaliado o efeito da inclusão de um calendário de feriados no Prophet. O calendário considerou feriados nacionais fixos, feriados móveis, datas estaduais e municipais relevantes para Pernambuco e Jaboatão dos Guararapes, além de datas especiais de varejo. A comparação geral entre os cenários sem e com feriados está apresentada na Tabela 6.

Tabela 6. Efeito da inclusão de feriados no Prophet agregado semanalmente

Cenário	WAPE semanal médio (%)
Sem feriados	37,34
Com feriados	39,33

Fonte: autoria própria (2026).

A inclusão de feriados elevou o WAPE semanal médio de 37,34 para 39,33. Na comparação realizada com a base final avaliada, o calendário de feriados não produziu

ganho agregado. Ainda assim, essa avaliação foi mantida por se tratar de uma variável de domínio relevante, especialmente em produtos sensíveis a datas especiais. Para complementar a análise, a Tabela 7 apresenta o WAPE semanal por produto nos cenários sem e com feriados.

Tabela 7. WAPE semanal por produto sem e com feriados

Produto	Sem feriados	Com feriados	Delta
GALETO ASSADO	15,45	21,96	6,51
FILE DE PEITO DE FRANGO	29,60	28,53	-1,06
FAROFA DA CASA	14,69	20,49	5,80
FILE MIGNON	51,20	50,25	-0,95
KIT 10 HAMBURGUES	36,26	35,14	-1,12
CHA DE DENTRO	76,84	79,58	2,74

Fonte: autoria própria (2026).

O efeito foi heterogêneo entre os produtos. Houve pequena melhora em FILE DE PEITO DE FRANGO, FILE MIGNON e KIT 10 HAMBURGUES. Em contrapartida, observou-se piora em FAROFA DA CASA, GALETO ASSADO e CHA DE DENTRO (COXAO MOLE).

Assim, os feriados permanecem como variável possivelmente útil em casos específicos, mas não como solução geral para o piloto analisado.

6.3.2. Estratégias de Tratamento de Zeros

Como muitos produtos apresentam grande quantidade de dias sem venda, foram testadas estratégias de tratamento de zeros. A primeira estratégia, chamada *prediction floor*, zera previsões muito pequenas com base em um limite definido a partir da mediana histórica dos dias com venda. A segunda, chamada *occurrence gate*, estima a probabilidade de ocorrer venda em determinado dia e zera a previsão quando essa probabilidade fica abaixo de um limiar. A estratégia *combined* combina as duas anteriores.

No *prediction floor*, o limite mínimo não foi definido por um valor fixo igual para todos os produtos. Para cada série, foi calculada a mediana dos dias com venda e foram testados multiplicadores dessa mediana. Assim, um produto com vendas típicas maiores teria um limite absoluto maior, enquanto um produto de menor volume teria um limite menor. A intenção era remover previsões residuais muito pequenas, comuns em modelos contínuos, sem impor o mesmo corte para produtos com escalas diferentes.

No *occurrence gate*, a estimativa de quantidade foi separada da decisão sobre haver ou não venda no dia. Primeiro, um modelo auxiliar estimou essa probabilidade.

Essa probabilidade pode ser representada por $P(venda > 0)$.

Depois, a previsão do Prophet era mantida apenas quando essa probabilidade ultrapassava um limiar. Quando a probabilidade ficava abaixo desse valor, a previsão era zerada. Essa abordagem buscou reduzir falsos positivos em dias sem venda, mas introduziu o risco de zerar previsões em datas com venda registrada.

Os valores testados foram definidos como uma grade simples de pós-processamento. Para o *prediction floor*, foram avaliados percentuais da mediana histórica

dos dias com venda. Para o *occurrence gate*, foram avaliados limiares de probabilidade.

Na versão conservadora, a escolha do cenário também passou a considerar o risco operacional de zerar volume observado. Para isso, foi limitada a parcela da venda que poderia ser indevidamente transformada em zero.

Foram consideradas duas configurações principais para o *zero handling*: uma orientada à redução direta do erro e outra conservadora, com maior restrição ao risco de zerar previsões em datas com venda registrada. Na configuração inicial, o Prophet base apresentou WAPE semanal de 36,69, enquanto a melhor configuração de *zero handling* reduziu esse valor para 30,11. Como a proposta da solução exige maior cautela operacional, a comparação final desta subseção prioriza a versão conservadora. Os resultados por produto estão apresentados na Tabela 8.

Tabela 8. Resultado do WAPE para o tratamento conservador de zeros por produto

Produto	Cenário	Base	Melhor	Delta	Mult.	Limiar	zerado(%)
FAROFA DA CASA	Combined 10	30,18	26,77	-3,41	0,00	0,12	0,02
GALETO ASSADO	Base	20,63	20,63	0,00	0,00	0,00	0,00
FILE DE PEITO DE FRANGO	Combined 05	40,61	40,42	-0,19	0,03	0,24	0,03
FILE MIGNON	Combined 05	41,62	41,30	-0,33	0,00	0,21	0,03
KIT 10 HAMBURGUERES	Gate 05	58,69	58,69	0,00	0,00	0,05	0,00
CHA DE DENTRO	Combined 05	69,86	69,02	-0,84	0,00	0,29	0,01

Fonte: autoria própria (2026).

Na tabela, *Floor* indica *prediction floor*, *Gate* indica *occurrence gate*, e *Combined* indica a combinação das duas regras. *Mult.* representa o multiplicador médio aplicado sobre a mediana histórica dos dias com venda, *Limiar* representa o limiar médio de probabilidade do *occurrence gate*, e *zerado* representa a parcela média do volume observado que foi indevidamente transformada em zero pelo pós-processamento nas janelas avaliadas.

Na rodada final, o *zero handling* conservador melhorou alguns produtos, mas pouco alterou a média geral, passando de 43,60 no Prophet base multijanelas para 42,80. Por esse motivo, a estratégia foi considerada útil como opção por produto, mas não como padrão global da solução.

A Figura 7 ilustra o comportamento do *zero handling* em GALETO ASSADO. Nesse caso, a estratégia não produziu redução relevante do erro em relação ao cenário base, o que está de acordo com os resultados apresentados na Tabela 8, na qual o produto permaneceu sem ganho efetivo. Esse resultado reforça que o tratamento de zeros não deve ser adotado como política global, mas apenas como alternativa seletiva para séries em que a calibração produza benefício consistente.

6.4. Experimento 3: Comparação com Modelos Alternativos

6.4.1. Seleção de Modelos por Produto

Após os testes com Prophet e tratamento de zeros, foi realizada uma seleção de modelos por produto. Foram comparados Prophet com diferentes configurações, *baselines* simples, modelos intermitentes e modelos estatísticos automáticos. Em uma avaliação com uma única janela de 28 dias, o Prophet base apresentou WAPE semanal de 36,69, enquanto o melhor modelo por produto reduziu o WAPE semanal para 22,59.

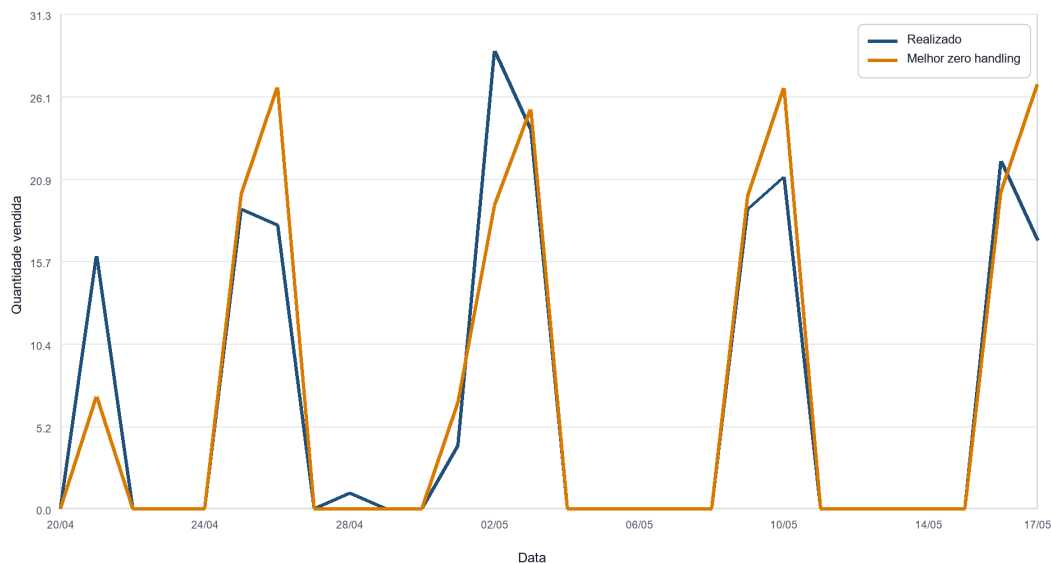


Figura 7. Comparação entre quantidade realizada e melhor cenário de zero handling para o produto GALETO ASSADO.

Fonte: autoria própria (2026).

Essa etapa tratou cada produto como uma série com comportamento próprio. Além do Prophet com e sem pós-processamento, foram incluídos *baselines* baseados em médias por dia da semana e médias móveis, modelos intermitentes e modelos automáticos da família StatsForecast. O critério de escolha foi o menor WAPE semanal no período de validação, mantendo o WAPE diário, o MAE e o RMSE como métricas complementares. Os resultados dessa seleção por produto em uma janela de 28 dias estão apresentados na Tabela 9.

Tabela 9. Seleção de modelos por produto em uma janela de 28 dias

Produto	Melhor cenário	Família	Base	Melhor	Delta
GALETO ASSADO	Combined	Prophet	13,23	11,52	-1,71
FILE DE PEITO DE FRANGO	Prophet com feriados	Prophet	23,19	23,19	0,00
FAROFA DA CASA	Occurrence gate	Prophet	36,51	16,05	-20,47
FILE MIGNON	AutoARIMA	StatsForecast	27,61	16,24	-11,38
KIT 10 HAMBURGUES	Moving avg. 7	Baseline	65,99	31,58	-34,41
CHA DE DENTRO	SBA	Intermitente	53,63	36,94	-16,69

Fonte: autoria própria (2026).

Os resultados da janela única indicaram que não havia uma família dominante para todos os produtos. Dois produtos foram melhor atendidos por configurações do Prophet, três por *baselines* de dia da semana ou média móvel, e um por StatsForecast. Essa diversidade justificou a adoção de seleção por produto como hipótese central para os testes seguintes.

Apesar do bom resultado, essa avaliação com uma única janela poderia estar otimista por depender de apenas um período de teste. Por isso, foi realizado um *backtesting* multijanelas, que será apresentado na próxima subseção. A seleção por produto continuou sendo importante, pois os resultados indicaram que diferentes produtos foram me-

lhor representados por diferentes famílias de modelos, como Prophet, *baselines*, modelos intermitentes e StatsForecast.

6.4.2. *backtesting* Multijanelas

O *backtesting* multijanelas foi utilizado para tornar a avaliação mais robusta e reduzir a dependência de uma única data de corte. Em uma configuração exploratória anterior, com seis janelas de 28 dias, 13 cenários por produto e oito produtos avaliados, o Prophet base apresentou WAPE semanal médio de 55,30, enquanto o melhor modelo por produto apresentou WAPE semanal médio de 39,81. A comparação principal desta subseção, entretanto, considera a base reconstruída e o piloto limpo de seis produtos. Nessa avaliação, foram comparadas quatro famílias de resultados: Prophet base multijanelas, seleção de modelos multijanelas, modelo semanal híbrido e *zero handling* conservador.

O desenho multijanelas reduziu a dependência de uma única data de corte. Em cada janela, os modelos eram treinados com o histórico disponível antes do *holdout* e avaliados no período seguinte. Em seguida, as métricas eram consolidadas por produto, usando média, mediana, desvio e pior janela. Essa estrutura permitiu observar não apenas o desempenho médio, mas também a estabilidade dos modelos ao longo do tempo. A comparação final entre as principais estratégias avaliadas no piloto limpo está apresentada na Tabela 10.

Tabela 10. Comparação final dos modelos no piloto limpo

Teste	WAPE semanal médio
Prophet base multijanelas	43,60
Seleção de modelos	33,53
Modelo semanal híbrido	31,27
<i>Zero handling</i> conservador	42,80
Melhor por produto entre os multijanelas	31,22

Fonte: autoria própria (2026).

Os resultados mostram que o Prophet base não foi a melhor escolha como modelo único para o piloto. O ganho principal ocorreu ao permitir seleção de modelo por produto. O modelo semanal híbrido apresentou o melhor desempenho médio entre as estratégias globais avaliadas, com WAPE semanal médio de 31,27, enquanto a seleção de modelos apresentou resultado próximo, com WAPE semanal médio de 33,53.

A Figura 8 reforça esse resultado em visão agregada, mostrando que o ganho principal não veio de um único modelo universal, mas do uso de estratégias alternativas avaliadas ao longo de múltiplas janelas.

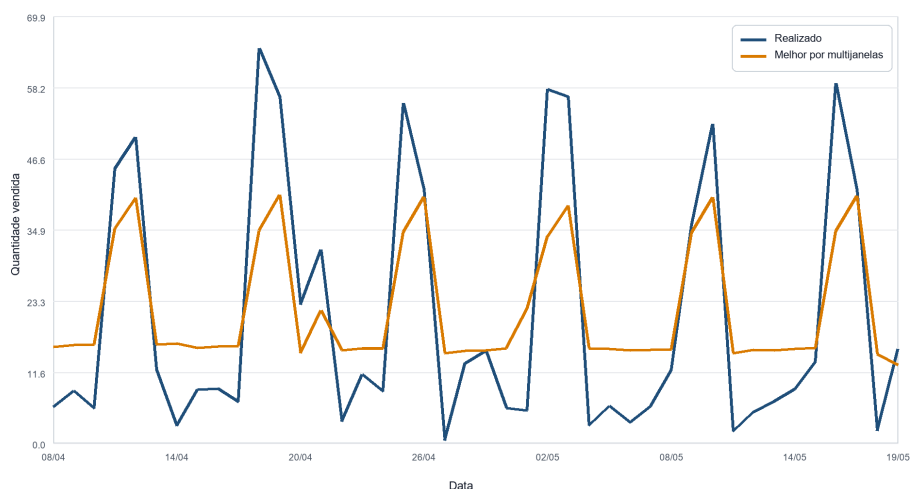


Figura 8. Comparação agregada entre quantidade realizada e melhor estratégia multijanelas.

Fonte: autoria própria (2026).

Por produto, o modelo semanal híbrido foi o melhor em quatro dos seis itens do piloto limpo: FAROFA DA CASA, FILE DE PEITO DE FRANGO, KIT 10 HAMBURGUERES e CHA DE DENTRO. A seleção multijanelas foi superior para FILE MIGNON, enquanto GALETO ASSADO apresentou desempenho muito próximo entre as principais alternativas avaliadas. A análise detalhada por produto é apresentada na Tabela 11.

Tabela 11. Comparação final por produto no *backtesting* multijanelas

Produto	Prophet	Seleção	Híbrido	Zeros	Melhor
FAROFA DA CASA	30,18	20,48	14,20	26,77	Híbrido
GALETO ASSADO	20,63	20,63	20,97	20,63	Empate
FILE DE PEITO DE FRANGO	40,61	39,77	36,96	40,42	Híbrido
KIT 10 HAMBURGUERES	58,69	40,72	38,15	58,69	Híbrido
FILE MIGNON	41,62	36,34	39,88	41,30	Seleção
CHA DE DENTRO	69,86	43,26	37,47	69,02	Híbrido

Fonte: autoria própria (2026).

A Figura 9 exemplifica esse comportamento, mostrando um caso em que o ganho entre estratégias foi pequeno, mesmo com a avaliação em múltiplas janelas. Esses resultados mostram que a comparação por produto foi importante para compreender o comportamento das séries e identificar alternativas competitivas ao Prophet base. Entretanto, para a operacionalização da solução, adotou-se o modelo semanal híbrido como estratégia padrão, por apresentar o melhor desempenho médio global entre as abordagens avaliadas, maior alinhamento ao horizonte semanal de decisão e menor custo operacional do que uma política de reavaliação completa a cada atualização da base.

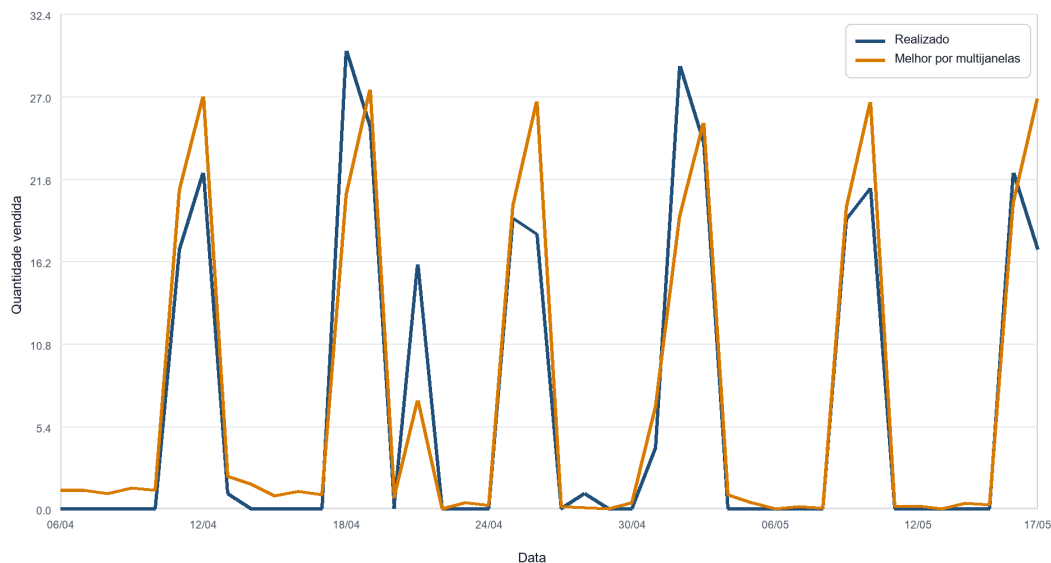


Figura 9. Comparação entre quantidade realizada e melhor estratégia multijanelas para o produto GALETO ASSADO.

Fonte: autoria própria (2026).

6.5. Experimento 4: Consolidação da Estratégia Final

6.5.1. Modelo Semanal Híbrido

O modelo semanal híbrido foi avaliado para separar duas decisões: estimar o volume total da semana e distribuir esse total entre os dias. Nessa abordagem, primeiro é prevista a demanda semanal do produto. Em seguida, a previsão semanal é distribuída entre segunda-feira, terça-feira, quarta-feira e os demais dias com base em um perfil histórico de venda por dia da semana.

Essa abordagem foi motivada pelos resultados anteriores, que indicaram que a previsão semanal é mais estável do que a previsão diária. Além disso, a estratégia preserva a possibilidade de exibir uma estimativa diária no aplicativo, sem obrigar o modelo principal a acertar exatamente o dia de ocorrência da venda.

Na prática, o híbrido usa o modelo semanal para controlar o volume total e uma regra de distribuição para produzir saídas diárias compatíveis com o aplicativo. A regra de distribuição utiliza o perfil histórico do produto por dia da semana e normaliza as parcelas para que a soma das previsões diárias seja igual ao total previsto para a semana. Assim, o modelo evita que pequenos erros diários alterem o total semanal planejado.

Na comparação final com seis produtos, o modelo semanal híbrido obteve WAPE semanal médio de 31,27, sendo a melhor estratégia geral agregada entre as avaliadas. Os resultados por produto estão apresentados na Tabela 12.

O delta da tabela compara o WAPE semanal do híbrido com o melhor modelo diário salvo para o mesmo produto. Valores negativos indicam melhora do híbrido em relação à alternativa diária. O híbrido melhorou quatro produtos, mas foi ligeiramente pior em GALETO ASSADO e FILE MIGNON. Mesmo assim, sua média geral foi a melhor

Tabela 12. Resultados por produto do modelo semanal híbrido

Produto	Cenário híbrido	Família	WAPE sem.	WAPE dia.	Delta sem.
FAROFA DA CASA	Last week	Baseline	14,20	40,95	-6,27
GALETO ASSADO	Prophet weekly	Prophet	20,97	34,83	0,34
FILE MIGNON	Moving avg. 8	Baseline	39,88	88,26	3,55
CHA DE DENTRO	Moving avg. 8	Baseline	37,47	111,44	-5,79
KIT 10 HAMBURGUES	Last week	Baseline	38,15	79,17	-2,57
FILE DE PEITO DE FRANGO	AutoARIMA	StatsForecast	36,96	97,11	-2,81

Fonte: autoria própria (2026).

entre as estratégias globais avaliadas.

A Figura 10 exemplifica esse comportamento em FAROFA DA CASA. Embora a distribuição diária não reproduza perfeitamente todos os picos observados, a estratégia híbrida mantém um padrão semanal mais consistente e contribui para reduzir o erro acumulado no horizonte de decisão utilizado no projeto.

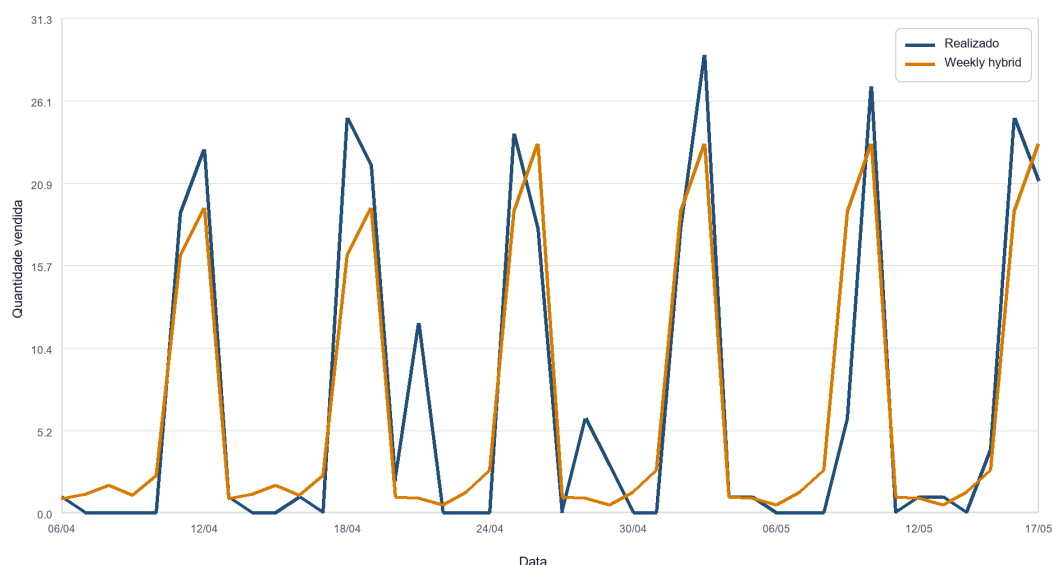


Figura 10. Comparação entre quantidade realizada e prevista pelo modelo semanal híbrido para o produto FAROFA DA CASA.

Fonte: autoria própria (2026).

6.5.2. holdout de 7 Dias

Além do *backtesting* com *holdout* de 28 dias, foi realizado um teste de curto prazo com *holdout* de 7 dias. O objetivo foi avaliar a seguinte questão operacional: se a *pipeline* em AWS rodar semanalmente, quão bem a solução consegue prever a próxima semana?

Esse teste utilizou 12 janelas, passo de 7 dias, janelas completas de segunda a domingo, período reconstruído incluído e os mesmos seis produtos do piloto limpo. Foram avaliadas as mesmas famílias de modelos do *backtesting* principal, incluindo Prophet base com feriados, Prophet com *zero handling* conservador, *baselines*, Croston, SBA, TSB, AutoARIMA, AutoETS e AutoTheta.

O *holdout* de 7 dias foi desenhado para aproximar o cenário de uso operacional semanal. Diferentemente do *holdout* de 28 dias, que avalia um horizonte mais longo e mais estável, esse teste mede a capacidade de prever apenas a próxima semana completa. Por isso, além da média, foram observadas a mediana e a pior janela, pois uma única semana atípica pode ter impacto grande no WAPE de produtos com baixo volume. O resumo dos resultados obtidos com *holdout* de 7 dias está apresentado na Tabela 13.

Tabela 13. Resumo do *backtesting* com *holdout* de 7 dias

Métrica	Valor
Prophet base 7 dias	50,48
Melhor modelo 7 dias	39,14
Mediana do melhor modelo 7 dias	32,49
Pior janela média do melhor modelo	136,04
WAPE diário médio do melhor modelo	76,76

Fonte: autoria própria (2026).

O teste de 7 dias reduziu o WAPE médio em relação ao Prophet base, de 50,48 para 39,14. Entretanto, esse resultado ficou ligeiramente pior que a melhor estratégia de 28 dias, que obteve WAPE semanal médio de 31,22. A mediana menor que a média indica que muitas semanas tiveram bom desempenho, mas algumas semanas atípicas elevaram o erro médio, especialmente em produtos de baixo volume. A análise por produto desse teste de curto prazo está apresentada na Tabela 14.

Tabela 14. *backtesting* de 7 dias por produto

Produto	Melhor cenário	Família	Base	Média	Mediana	Pior	WAPE dia.
FAROFA DA CASA	Weekday mean	<i>Baseline</i>	24,13	16,00	17,21	39,99	37,79
GALETO ASSADO	Floor 05	Zeros	20,66	19,53	17,92	46,59	28,36
FILE DE PEITO DE FRANGO	Combined 10	Zeros	80,32	80,08	50,92	423,57	133,20
KIT 10 HAMBURGUES	Weekday mean	<i>Baseline</i>	61,78	44,97	51,97	65,66	83,62
FILE MIGNON	SBA	Intermitente	60,62	41,01	41,24	100,00	76,00
CHA DE DENTRO	Weekday mean	<i>Baseline</i>	55,36	33,23	15,70	140,44	101,57

Fonte: autoria própria (2026).

Os resultados por produto mostram que o teste semanal curto favoreceu soluções simples em parte dos casos. *Baselines* venceram em FAROFA DA CASA, KIT 10 HAMBURGUES e CHA DE DENTRO (COXAO MOLE), enquanto estratégias com tratamento de zeros venceram em GALETO ASSADO e FILE DE PEITO DE FRANGO. Já FILE MIGNON apresentou melhor resultado com um modelo da família intermitente. A pior janela de CHA DE DENTRO (COXAO MOLE) foi muito alta, indicando que a avaliação operacional semanal ainda é sensível a semanas com baixo volume ou comportamento atípico.

Dessa forma, o *holdout* de 7 dias foi considerado útil como diagnóstico operacional da próxima semana, mas não substitui o *backtesting* de 28 dias como critério principal de seleção de modelos. A Figura 11 mostra que, nesse horizonte curto, o melhor modelo pode acompanhar melhor a estrutura semanal do que o Prophet base, mas ainda permanece sensível a oscilações mais intensas do volume realizado.

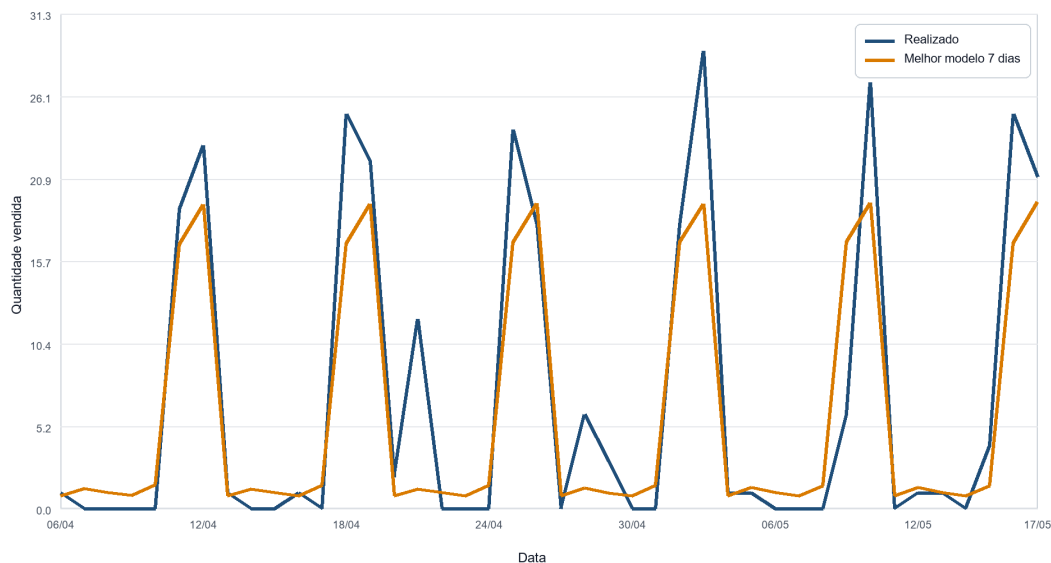


Figura 11. Comparação entre quantidade realizada e melhor modelo no *holdout* de 7 dias para o produto FAROFA DA CASA.

Fonte: autoria própria (2026).

6.6. Síntese dos Resultados

Os experimentos indicaram que a previsão diária pura com Prophet apresentou erro elevado, principalmente devido à presença de muitos zeros e ao deslocamento de vendas entre dias. Ao agregar a avaliação por semana, o erro foi reduzido de forma significativa, indicando que o volume semanal é mais adequado ao contexto operacional do projeto.

A inclusão de feriados não trouxe ganho agregado na comparação final, enquanto o tratamento de zeros mostrou potencial em alguns produtos, mas não deve ser adotado como padrão global devido ao risco de gerar falsos zeros. A comparação com modelos alternativos mostrou que diferentes produtos podem ser melhor representados por famílias distintas de modelos. Na comparação principal com seis produtos, o Prophet base obteve WAPE semanal médio de 43,60%, enquanto a seleção de modelos multijanelas reduziu esse valor para 33,53% e o modelo semanal híbrido para 31,27%.

6.7. Política Operacional de Uso do Modelo

Os resultados permitem distinguir dois níveis de decisão. Do ponto de vista analítico, o melhor modelo por produto entre os multijanelas apresentou o menor WAPE semanal médio, com 31,22%, mostrando que uma seleção específica por série pode gerar ganhos adicionais. Entretanto, essa alternativa exige maior custo de execução e manutenção, pois demanda reavaliar diferentes famílias de modelos para cada produto sempre que a base for atualizada.

Para a operacionalização da solução, foi adotado o modelo semanal híbrido como estratégia padrão. Essa decisão se justifica porque o híbrido apresentou o melhor desempenho médio entre as estratégias globais avaliadas, com WAPE semanal médio de 31,27%, além de manter aderência ao horizonte semanal de decisão e menor comple-

xidade operacional. A abordagem também preserva a possibilidade de exibir previsões diárias no aplicativo, por meio da distribuição do total semanal conforme o perfil histórico de venda por dia da semana.

Assim, a seleção por produto permanece como achado analítico e como referência para revisões futuras, mas não como rotina obrigatória a cada nova carga de dados. A política proposta é reavaliar a estratégia em frequência trimestral ou semestral, permitindo verificar se o modelo padrão continua adequado diante de mudanças no comportamento de consumo, sem elevar desnecessariamente o custo operacional da solução em ambiente AWS.

7. Conclusão

Este trabalho teve como objetivo desenvolver e avaliar uma solução de previsão de vendas por produto para apoiar decisões de compra, estoque e planejamento operacional no contexto do Frigorífico Candeias. Para isso, foi estruturado um processo de tratamento dos dados históricos de venda, geração de séries temporais por produto, avaliação comparativa de modelos de previsão e disponibilização dos resultados por meio de uma arquitetura de integração e de um aplicativo de apoio à decisão.

Os experimentos mostraram que a previsão diária pura com Prophet apresentou limitações relevantes, principalmente em produtos com muitos dias sem venda, baixo volume ou vendas deslocadas entre dias próximos. Ao avaliar os resultados em frequência semanal, observou-se maior aderência ao contexto operacional da empresa, uma vez que decisões de compra e reposição tendem a considerar o volume acumulado da semana, e não apenas o acerto exato de cada dia.

A comparação entre diferentes estratégias indicou que não há um único modelo superior para todos os produtos em todas as situações. Modelos simples, métodos estatísticos automáticos, estratégias de tratamento de zeros e configurações do Prophet apresentaram desempenhos distintos conforme o perfil de demanda. Entretanto, para a operacionalização da solução, o modelo semanal híbrido foi adotado como estratégia padrão, pois apresentou o melhor desempenho médio entre as estratégias globais avaliadas, manteve alinhamento com o horizonte semanal de decisão e reduziu a complexidade operacional em relação a uma seleção completa de modelos a cada atualização da base.

Dessa forma, a principal contribuição do trabalho está na integração entre tratamento de dados reais, avaliação experimental de modelos, definição de uma política operacional de previsão e disponibilização dos resultados em uma solução aplicada ao ambiente empresarial. O trabalho não se limitou à escolha de um algoritmo, mas estruturou um processo de apoio à decisão que pode ser atualizado periodicamente conforme novos dados sejam incorporados.

Apesar dos resultados obtidos, o estudo possui limitações. A comparação principal foi realizada com um recorte piloto de seis produtos, selecionados para permitir uma análise mais controlada e comparável. Produtos com comportamento atípico, possível instabilidade cadastral ou perfil diferente do varejo tradicional foram separados da avaliação principal e devem ser investigados em estudos futuros. Além disso, a qualidade da base histórica permanece um fator crítico, especialmente devido ao período reconstruído após a troca de sistema e à possibilidade de mudanças futuras no comportamento de consumo.

Como trabalhos futuros, recomenda-se ampliar a avaliação para um conjunto maior de produtos, incorporar informações comerciais adicionais, como preço e promoções, considerar variáveis externas, como calendário comercial e eventos locais, e aprimorar a distribuição diária das previsões semanais. Também se recomenda avaliar modelos de aprendizado de máquina e aprendizado profundo, com o objetivo de verificar se essas abordagens são capazes de capturar padrões mais complexos da demanda em comparação aos modelos estatísticos utilizados neste trabalho. Essa investigação deve considerar a disponibilidade de maior volume de dados, maior qualidade cadastral e variáveis explicativas adicionais, pois esses fatores tendem a influenciar diretamente o desempenho de modelos mais complexos. No aspecto operacional, recomenda-se evoluir a integração em AWS e o aplicativo desenvolvido, permitindo maior automação do envio de arquivos, acompanhamento das execuções e consulta das previsões pelo gestor.

Conclui-se que a solução proposta contribui para aproximar técnicas de previsão de vendas de um cenário real de gestão. A adoção do modelo semanal híbrido como política padrão, associada à revisão periódica trimestral ou semestral, oferece um equilíbrio entre desempenho preditivo, custo operacional e utilidade prática para o apoio à tomada de decisão.

Referências

- [1] Abdul Fatir Ansari, Lorenzo Stella, Caner Turkmen, Xiyuan Zhang, Pedro Mercado, Hui-bin Shen, Oleksandr Shchur, Syama Sundar Rangapuram, Sebastian Pineda Arango, Shubham Kapoor, Jasper Zschiegner, Danielle Maddix Robinson, Hao Wang, Michael W. Mahoney, Kari Torkkola, Andrew G. Wilson, Michael Bohlke-Schneider, and Yuyang Wang. Chronos: Learning the language of time series. *Transactions on Machine Learning Research*, 2024. Acesso em: 08 jul. 2026.
- [2] Vassilis Assimakopoulos and Konstantinos Nikolopoulos. The theta model: A decomposition approach to forecasting. *International Journal of Forecasting*, 16(4):521–530, 2000.
- [3] J. D. Croston. Forecasting and stock control for intermittent demands. *Operational Research Quarterly*, 23(3):289–303, 1972.
- [4] Alp Ecevit, Irem Ozturk, Mustafa Dag, and Tuncay Ozcan. Short-term sales forecasting using LSTM and prophet based models in e-commerce. *Acta Infologica*, 7(1):59–70, 2023.
- [5] Liang Guo, Weiguo Fang, Qiuhong Zhao, and Xu Wang. The hybrid PROPHET-SVR approach for forecasting product time series demand with seasonality. *Computers & Industrial Engineering*, 161:107598, 2021.
- [6] Jakob Huber and Heiner Stuckenschmidt. Daily retail demand forecasting using machine learning with emphasis on calendric special days. *International Journal of Forecasting*, 36(4):1420–1438, 2020. Artigo citado no trabalho relacionado.
- [7] Rob J. Hyndman and George Athanasopoulos. *Forecasting: Principles and Practice*. OTexts, Melbourne, Australia, 3 edition, 2021. Acesso em: 08 jul. 2026.
- [8] Rob J. Hyndman and Yeasmin Khandakar. Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, 27(3):1–22, 2008.

- [9] Rob J. Hyndman and Anne B. Koehler. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4):679–688, 2006.
- [10] Rob J. Hyndman, Anne B. Koehler, Ralph D. Snyder, and Simone Grose. A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting*, 18(3):439–454, 2002.
- [11] Mehran Nasseri, Taha Falatouri, Patrick Brandtner, and Farzaneh Darbanian. Applying machine learning in retail demand prediction—a comparison of tree-based ensembles and long short-term memory-based deep learning. *Applied Sciences*, 13(19):11112, 2023.
- [12] Aris A. Syntetos and John E. Boylan. The accuracy of intermittent demand estimates. *International Journal of Forecasting*, 21(2):303–314, 2005.
- [13] Sean J. Taylor and Benjamin Letham. Forecasting at scale. *The American Statistician*, 72(1):37–45, 2018.
- [14] Ruud H. Teunter, Aris A. Syntetos, and M. Zied Babai. Intermittent demand: Linking forecasting to inventory obsolescence. *European Journal of Operational Research*, 214(3):606–615, 2011.
- [15] Emir Žunić, Kemal Korjenić, Kerim Hodžić, and Dženana Đonko. Application of facebook’s prophet algorithm for successful sales forecasting based on real-world data. *International Journal of Computer Science & Information Technology*, 12(2):23–43, 2020.