

# Case: Previsão de demanda

**Letícia Rodrigues Chaparro Corrêa**

# Contexto e objetivo

---

Desenvolver uma solução de previsão de vendas (sell-out)

## Contexto do Negócio

Planejamento de vendas e operações depende de previsões confiáveis.

Erros de previsão impactam diretamente:

- excesso de estoque;
- perda de produtos;
- capital de giro e nível de serviço.

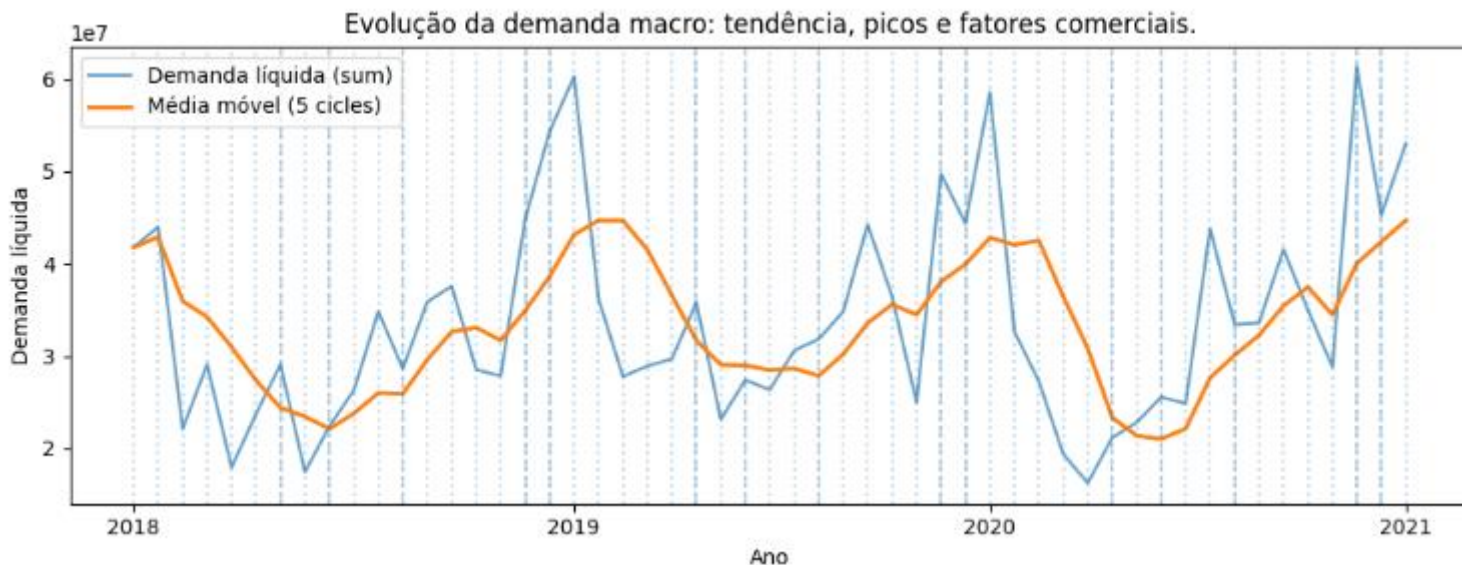
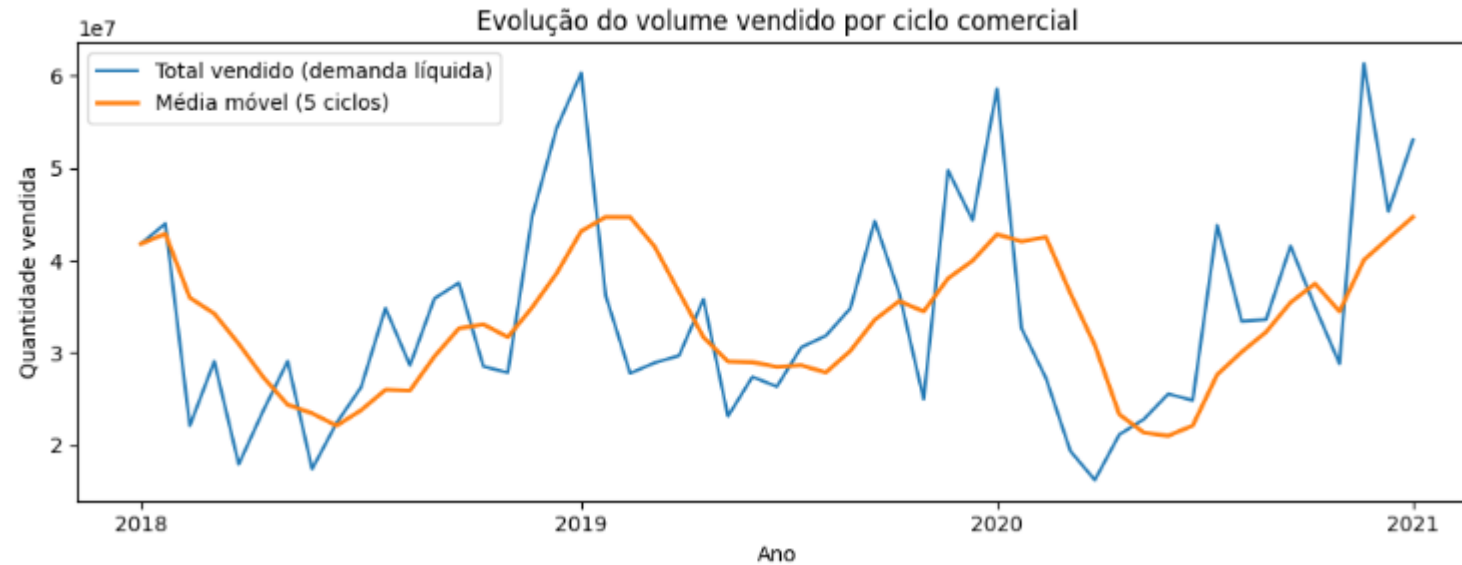
O desafio aumenta com:

- múltiplos SKUs, canais e regiões;
- efeitos de preço, desconto e campanhas;
- volatilidade temporal da demanda.

**Necessidade:** uma solução escalável e precisa para curto e médio prazo.

# Análise exploratória dos dados

## Análise dos dados históricos de venda – distribuição temporal



### Comportamento histórico da demanda

- Crescimento gradual da demanda entre 2018 e 2019;
- Sazonalidade recorrente, com ciclos de alta e queda bem definidos;
- Há picos e vales pontuais, associados a fatores comerciais e operacionais - não a ruído;
- A média móvel suaviza a série e evidencia o comportamento macro - previsibilidade da demanda.

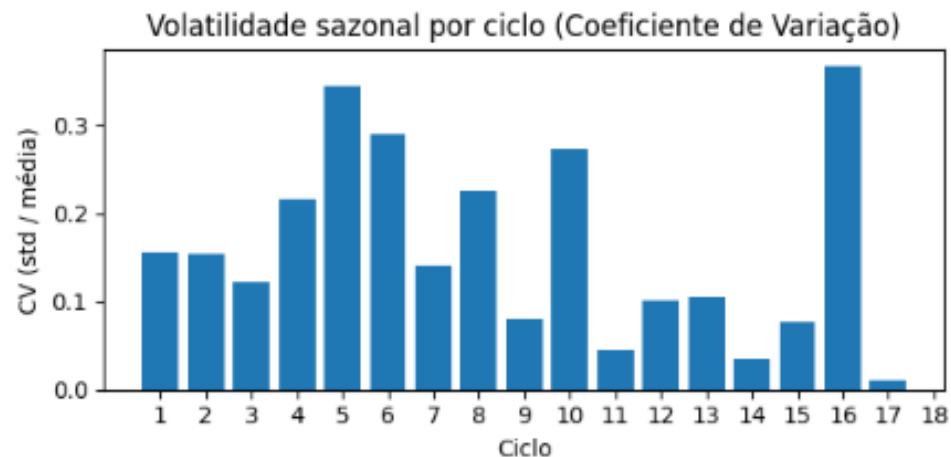
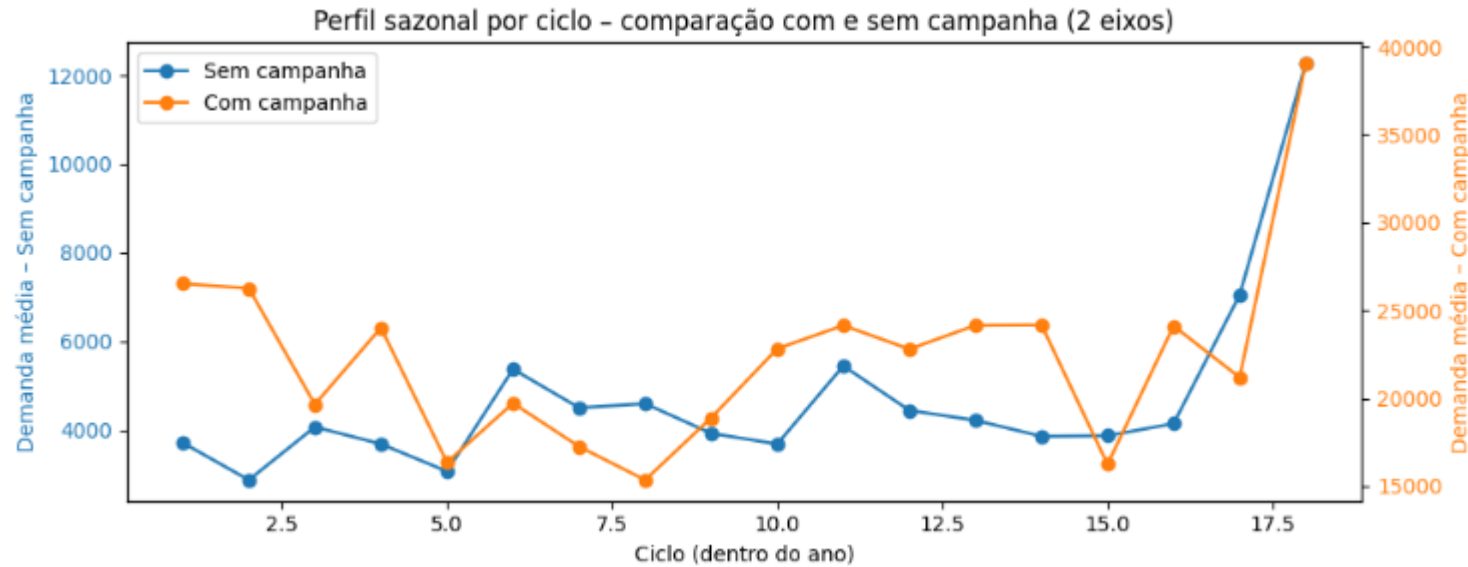
A série contém sinal consistente (tendência + sazonalidade), justificando o uso de modelos preditivos e análises de risco.

### Ações:

- Uso de médias móveis (rolling statistics) da demanda histórica;
- Criação de features de rolling mean e rolling std.

# Análise exploratória dos dados

## Análise dos dados históricos de venda – sazonalidade



### Sazonalidade

- **Padrão sazonal recorrente** ao longo dos ciclos do ano;
- Sazonalidade existe **mesmo sem campanhas**;
- **Campanhas elevam o nível da demanda - não alteram o padrão sazonal**;
- Alguns ciclos apresentam **maior volatilidade** (maior risco operacional);
- A demanda é previsível, não aleatória;
- Há períodos estruturalmente mais arriscados.

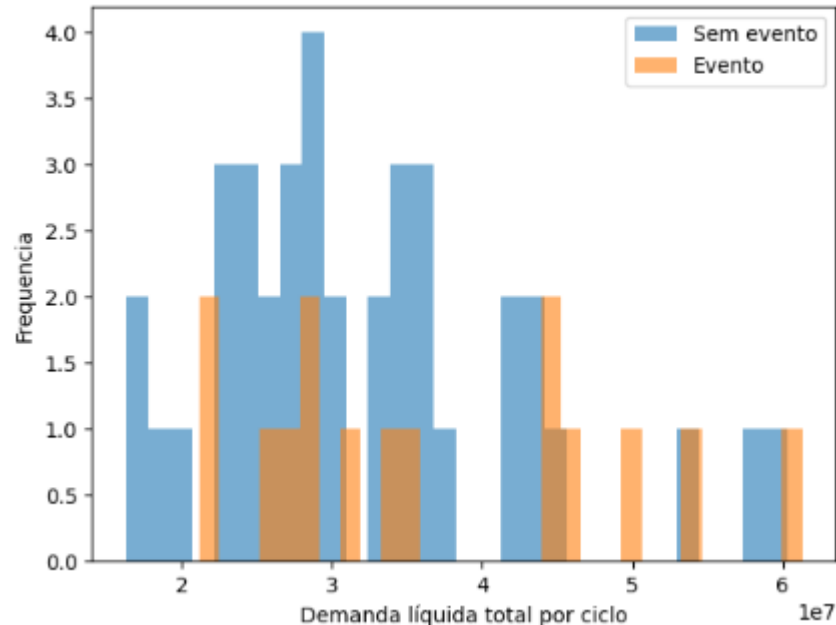
### Ações:

- Criação de features sazonais por ciclo;
- Inclusão de lags e estatísticas móveis;
- Modelagem explícita de campanhas;
- Uso da volatilidade para orientar análise de risco e estoque.

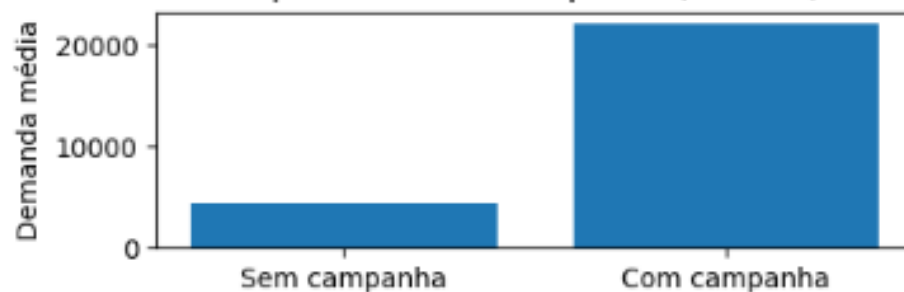
# Análise exploratória dos dados

## Análise dos dados históricos de venda – ciclos comemorativos

Distribuição da demanda total por indicador de evento (nível de ciclo)



Uplift médio de campanha (398.2%)



### Ciclos comemorativos

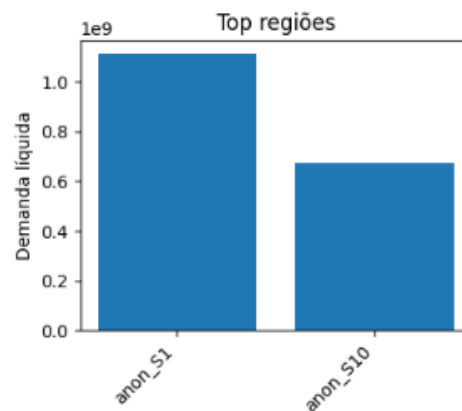
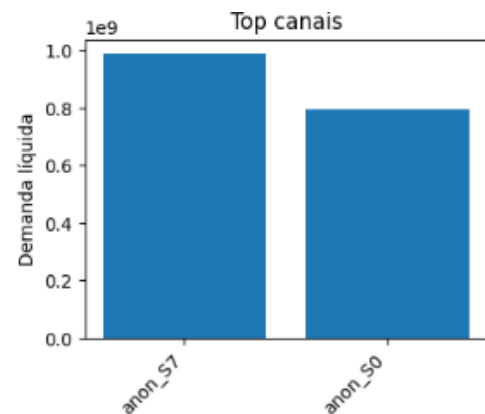
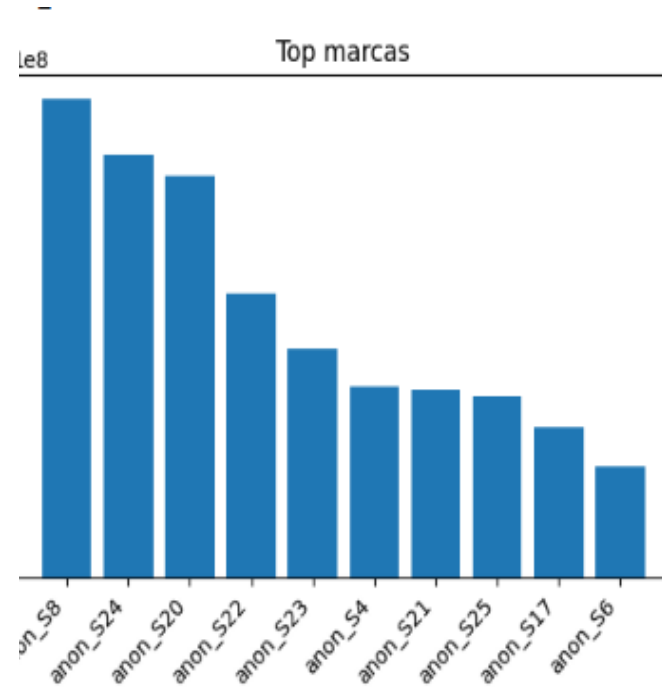
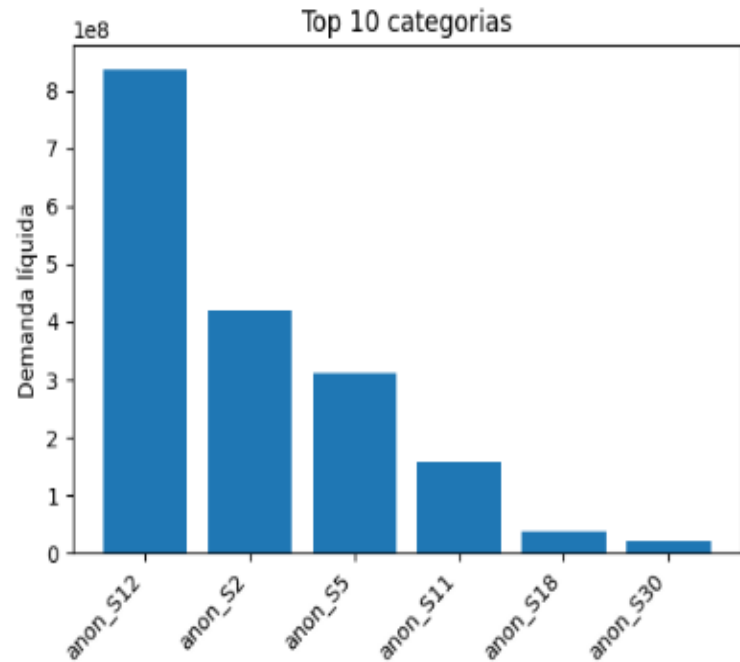
- Distribuição da demanda **com e sem eventos** apresenta padrões distintos;
- Períodos com campanha concentram **volumes significativamente mais altos**;
- A variabilidade aumenta em ciclos comemorativos;
- Uplift médio de campanha de aproximadamente **+400%**
- Campanhas são drivers relevantes de volume, não ruído;
- Planejamento sem considerar eventos subestima demanda;
- Ciclos comemorativos exigem: maior capacidade logística e maior estoque de segurança.

#### Ações:

- Criação de flag resumo de campanha;
- Separação explícita entre períodos com e sem campanha;
- Uso de campanhas como variável;
- Avaliação do impacto de eventos em múltiplos horizontes

# Análise exploratória dos dados

## Análise dos dados históricos de venda – segmentação por categoria



### Ciclos comemorativos

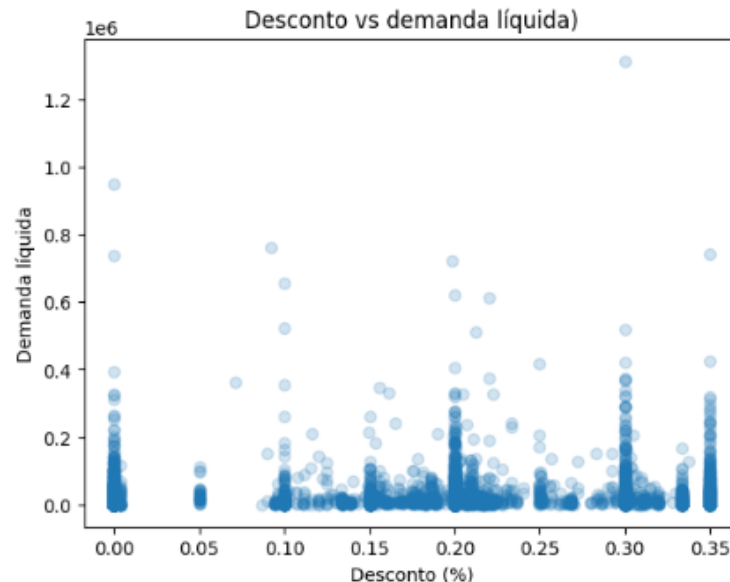
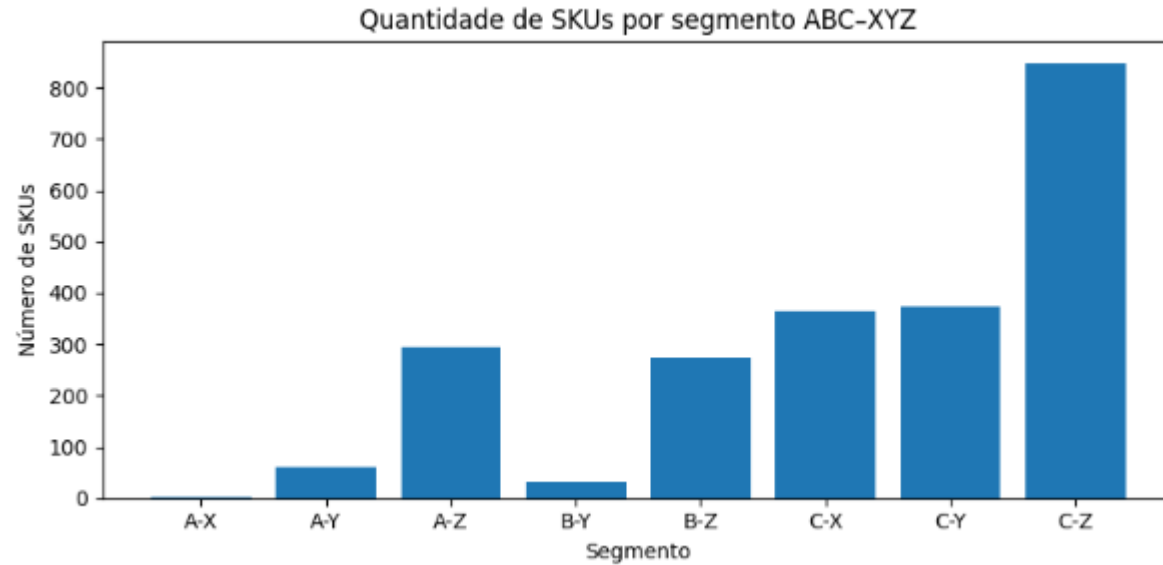
- Forte concentração de demanda em poucas categorias;
- Regiões e canais apresentam distribuição assimétrica – região mais forte;
- Marcas líderes concentram parcela relevante da demanda total – maior impacto exigem maior precisão;
- Segmentos de menor volume coexistem com alta variabilidade;
- Estratégias únicas para todos os segmentos não são eficientes
- Segmentos de cauda longa: maior risco relativo demandam políticas diferenciadas de estoque.

### Ações:

- Modelagem no nível granular (SKU × canal × região)
- Uso de modelo global, aprendendo padrões comuns entre séries
- Avaliação de métricas por segmento
- Base para reconciliação hierárquica (categoria/região)

# Análise exploratória dos dados

Análise dos dados históricos de venda – variedade de SKUs e preço



## Variedade de SKUs e preço

- O portfólio apresenta alta concentração de SKUs em segmentos C e Z - cauda longa
- Poucos SKUs concentram grande parte do volume (segmentos A e B);
- A relação entre desconto e demanda é não linear, com grande dispersão e outliers.
- Abordagens simples ou médias globais não capturam a realidade do portfólio;
- Estratégias de preço e estoque devem ser segmentadas.

## Ações:

- Uso de modelo global não linear para múltiplas séries
- Inclusão explícita de preço e desconto como variáveis;
- Criação de features de histórico, volatilidade e campanha;
- Avaliação por segmentos de risco e volume.

# Modelo de previsão da demanda

Metodologia utilizada



## Modelo de previsão

**HistGradientBoostingRegressor**

### Por quê?

- Captura relações não lineares reais do negócio (preço, desconto, campanhas);
- Robusto a ruído, outliers e heterogeneidade entre séries;
- Escalável para múltiplos SKUs, canais e regiões;
- Boa performance sem necessidade de escalonamento
- Previsões para múltiplos horizontes:
  - curto prazo ( $h=1$ )
  - médio prazo ( $h=2, h=3$ )



## Métricas de avaliação

- **WAPE** – erro relativo alinhado ao volume de vendas;
- **MAE** – erro absoluto médio (impacto operacional);
- **Bias** – controle de super / subestimação.



## Validação do modelo

- Avaliação com walk-forward validation;
- Gap treino × teste controlado – overfitting;
- Bias próximo de zero no curto prazo;
- Importância de variáveis.



# Modelo de previsão da demanda

## Análise dos parâmetros avaliados



### Métricas do modelo

	horizon	n	mean_mae	mean_wape	mean_bias
0	1	10.0	7243.469269	0.680188	0.006752
1	2	10.0	7665.414138	0.759274	0.079105
2	3	10.0	7905.413358	0.849449	0.207316

- WAPE < 1% até o horizonte 3;
- No horizonte **h = 1**, o erro relativo é **inferior a 0,7%**
- Em benchmarks de mercado, erros aceitáveis costumam ser **10–20%**;
- Erro cresce de forma gradual e esperada com o horizonte;
- Alta confiabilidade no curto prazo (h = 1–2)



### Robustez e Generalização

- Gap controlado entre treino e teste – **0,19-0,31 p.p.**;
- Ausência de colapso fora da amostra;
- Comportamento consistente ao longo dos splits.



**O modelo generaliza bem e é seguro para produção.**



### O que explica a demanda?

- **Comportamento recente da demanda** (rolling stats);
- **Histórico direto** (lags);
- **Preço e desconto**;
- **Campanhas** como fator de ajuste.

# Do forecast à decisão

Como o modelo gera valor real para o negócio



## Gestão de Risco Logístico

- Erro de previsão: incerteza de demanda no lead time;
- Estoque de segurança calculado com base em:
  - Variabilidade do erro;
  - Nível de serviço (95%).
- Estoque deixa de ser “regra empírica” e passa a ser quantitativo.



## Consistência Hierárquica

### Garantir que:

- Previsões por SKU agregam corretamente;
- Números fecham em categoria e região;
- Planejamento, orçamento e operação em sintonia.

### Proposta:

- Reconciliação hierárquica proporcional
- Base para métodos avançados (MinT)



## Aplicações

- Planejamento de vendas e compras;
- Definição de estoque por SKU × região;
- Simulações de campanha (what-if);
- Suporte ao S&OP e orçamento.



## Impacto no negócio

- Redução de rupturas;
- Menor capital imobilizado;
- Planejamento mais confiável;
- Decisões defendidas por dados.

# Próximos passos

Ideias de melhorias e evolução da solução proposta

---



## Evolução do modelo

- Integração de dados operacionais de estoque:
  - estoque atual
  - estoque de segurança
  - estoque máximo / mínimo
  - dias de estoque (DOH)
- Uso explícito de restrições físicas e financeiras no modelo;
- Diferenciação entre: erro de previsão e decisão de abastecimento.



## Modelo complementar: Sell-in vs Sell-out

Modelo complementar de sell-in reconciliando:

- demanda prevista;
- política de estoque;
- capacidade logística.

Reduz efeito chicote (bullwhip) e mantém um planejamento mais alinhado com compras e produção



## Integração Forecast + Política de Estoque



## Estoque como Variável do Modelo



## Modelos distintos para áreas de impacto/risco



## Operacionalização e MLOps