

Previsão de demanda: uma aplicação no setor de varejo farmacêutico utilizando análises estatísticas multivariadas e séries temporais

Autor: Isabela Schossler

isabela.schossler@edu.pucrs.br, PUCRS, Brasil

Orientador: Helio Radke Bittencourt

heliorb@pucrs.br, PUCRS, Brasil

Resumo: Diante do crescimento contínuo do mercado farmacêutico e da concorrência acirrada entre as empresas, busca-se constantemente novas oportunidades de negócio para melhorar a qualidade do serviço, com o objetivo de manter a participação no mercado. Nesse contexto, a análise e previsão de demanda desempenham um papel crucial como diferencial competitivo. Este estudo teve como objetivo geral a aplicação de uma ferramenta de clusterização para melhorar a acurácia das previsões estatísticas. As séries temporais de um conjunto de dezesseis produtos da empresa foram analisadas. Todos os produtos obtiveram resultados melhores com a aplicação do método proposto neste estudo em relação ao método atual da empresa. A aplicação do método estudado obteve uma previsão de demanda 64% mais acurada que a previsão gerada pela empresa em um dos produtos analisados. A aplicação da metodologia proposta contribuiu para o aprimoramento do nível de serviço da empresa, que passa a ter como diferencial competitivo uma maior disponibilidade do produto para o consumidor final.

Palavras-chave: *Feature Extraction*; *Cluster*; Previsão de Demanda, Séries Temporais; Varejo Farmacêutico.

1. Introdução

O mercado farmacêutico no Brasil vem evoluindo de forma constante nos últimos anos. No ano de 2021 foi ultrapassada a marca de R\$ 145 bilhões em venda dentre todos os laboratórios instalados no país, representando um crescimento de 13,6% em relação ao ano anterior. Já o varejo farmacêutico apresentou um crescimento de 55,1% nos últimos quatro anos, correspondendo a aproximadamente R\$ 88,3 bilhões em vendas (INTERFARMA, 2022). Há tendência de continuidade no crescimento para os próximos anos, impulsionada pelo aumento da expectativa de vida da população e a oferta de tratamentos disponibilizados. Segundo cálculos da consultoria IQVIA, para o ano de 2023 está prevista uma evolução de vendas de 9,8% em comparação ao ano de 2022 (SINDUSFARMA, 2022).

Diante deste crescimento consolidado do mercado farmacêutico, a competitividade tem feito com que as empresas busquem por novas alternativas de negócio visando o aumento da qualidade do serviço e a produtividade, para garantir a participação no mercado em que atuam. Desta forma, a análise e previsão de demanda têm se destacado cada vez mais no que se refere ao diferencial competitivo. É necessário que ocorra a identificação das oscilações do ambiente de negócio e a previsão correta da demanda para que se obtenha sucesso na formulação do planejamento e na orientação do plano estratégico da organização, com o intuito de que seja possível obter uma maior eficiência na resolução rápida de problemas (MONTGOMERY *et al.*, 1990; BALLOU, 2001).

Segundo Slack *et al.* (2009), técnicas de previsão de demanda são rotineiramente empregadas como ferramentas de suporte à tomada de decisões pelas organizações. Como visto em Bastos e Werner (2010), a gestão adequada dos estoques físicos no setor farmacêutico varejista pode gerar benefícios significativos para as empresas, especialmente considerando a ampla variedade de itens comercializados. A previsão de demanda desempenha um papel fundamental nesse contexto, pois fornece informações relevantes sobre o que, quanto e quando comprar, conferindo uma vantagem competitiva significativa. Por meio dessa previsão, é possível obter um maior controle sobre os produtos, contribuindo para que os níveis de estoque se aproximem do ideal, proporcionando à empresa maior agilidade operacional e redução de custos (BASTOS; WERNER, 2010).

A acurácia do modelo estatístico para previsão de demanda é um dos maiores desafios enfrentados dentro do planejamento de compras e estoques. A literatura é vasta no que se refere à previsão de demanda e a melhor forma de se utilizar os modelos qualitativos e quantitativos, ou a combinação de ambos. Werner (2004), por exemplo, destaca que a combinação de previsões de demanda pode contribuir para no fornecimento de previsões mais robustas e estruturadas. Contudo, ainda há espaço para o desenvolvimento de novas abordagens acerca do tratamento prévio de dados – como a ferramenta de clusterização – e seus impactos na acurácia dos modelos gerados.

A ferramenta de clusterização pode ser utilizada para aumentar a acurácia da previsão de demanda em relação aos dados analisados individualmente e facilitar a modelagem e o agrupamento das séries históricas de dados. Carazza (2020), realizou

um estudo sobre a aderência da utilização desta ferramenta no setor alimentício. A partir deste, foi constatado que o erro percentual absoluto médio (*mean absolute percentage error* - MAPE) foi superior nas previsões isoladas quando comparadas as previsões de demanda agregada, evidenciando que o agrupamento dos itens possibilitou um aumento na aderência da previsão estatística. Outro exemplo é a aplicação do método de clusterização no setor de varejo especializado no comércio de materiais voltados a área da saúde e bem-estar, onde foi possível verificar uma significativa redução no custo de processamento das previsões e um grande potencial para aplicação da metodologia (RODRIGUES, 2021). Além de aprimorar a precisão das previsões e reduzir o custo de processamento de dados, a abordagem proposta por Amalnick *et al.* (2020) por meio da utilização da ferramenta de clusterização também proporcionou benefícios aos gerentes de *marketing* no que se refere ao planejamento das unidades farmacêuticas.

Desta forma, este trabalho busca contribuir acerca da aplicação da clusterização para melhoria a acurácia em previsões estatísticas. Portanto, a questão de pesquisa que surge é: como incorporar uma etapa de clusterização em um modelo de previsão de demanda? Para responder à questão de pesquisa, este estudo tem como objetivo geral propor e aplicar uma sistemática baseada em técnicas de previsão de demanda integradas a modelos de agrupamento de dados históricos de itens. Dentre os objetivos específicos podem ser citados: (i) avaliação dos potenciais ganhos do agrupamento das categorias de produtos na previsão de demanda em relação a previsão individual do item e (ii) aumento da acurácia nas previsões de demanda dos diferentes tipos de itens.

Como delimitações de pesquisa, considera-se que: (i) serão utilizados apenas os modelos de previsão disponíveis no *software* da empresa analisada, sendo eles: média móvel, suavização exponencial simples, Holt e Holt-Winters, e (ii) a pesquisa não terá como foco a análise financeira dos resultados obtidos com a previsão de demanda, mas, sim, a análise do impacto da ferramenta de clusterização na acurácia das previsões estatísticas.

O artigo está estruturado em quatro seções distintas. Na introdução são apresentados o contexto em que o estudo foi desenvolvido, sua justificativa, objetivos e delimitações. Na segunda seção são abordadas a caracterização da pesquisa e a metodologia utilizada para o desenvolvimento do artigo. Os resultados obtidos

seguindo a metodologia escolhida são discriminados na terceira seção. Por fim, na quarta seção, são apontadas as considerações finais a respeito deste estudo.

2. Método

Nesta seção são apresentados os métodos de pesquisa e trabalho utilizados no desenvolvimento deste estudo. Sendo assim, este se encontra dividido em duas subseções: (i) método de pesquisa, no qual são apresentadas as características do estudo desenvolvido; e (ii) método de trabalho, onde são descritas as etapas de desenvolvimento, assim como as técnicas utilizadas e referencial teórico.

2.1. Método de pesquisa

Segundo Silva e Menezes (2005), a pesquisa é definida por meio de ações concebidas para solucionar um problema, empregando métodos racionais e sistemáticos. Sendo assim, este estudo é classificado como uma pesquisa aplicada, uma vez que tem como objetivo gerar conhecimento para aplicação prática na resolução de problemas específicos, através do aprimoramento da precisão dos métodos estatísticos utilizados por uma empresa do ramo farmacêutico (MARCONI; LAKATOS, 2002). A abordagem adotada para tratar o problema consiste em quantificar as informações obtidas, caracterizando, assim, a pesquisa como de natureza quantitativa (SILVA; MENEZES, 2005).

No que se refere ao tempo, o estudo tem caráter longitudinal, uma vez que foram analisados dados históricos de um longo período (SILVA; MENEZES, 2005). O objetivo da pesquisa é classificado como exploratório, visto que finalidade da pesquisa consiste em desenvolver e esclarecer conceitos e ideias visando à formulação de problemas ou hipóteses que possam ser investigados em estudos subsequentes (GIL, 1995). Por fim, pode-se caracterizar o procedimento como um estudo de caso, haja vista que se trata de uma investigação que propicia uma compreensão abrangente e minuciosa do fenômeno em questão (GIL, 1995).

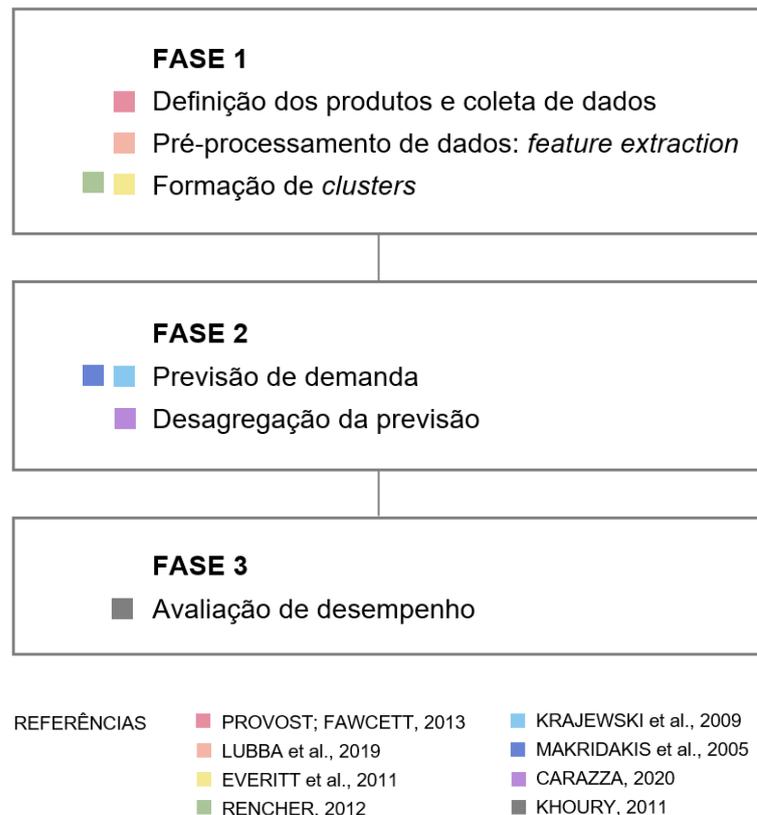
2.2. Método de trabalho

Neste tópico são detalhadas as etapas operacionais em que foram apoiados o método proposto, o qual será constituído de três fases, conforme apresentado na

Figura 1. A primeira fase consiste na definição dos produtos que serão utilizados com base na metodologia apresentada por Provost e Fawcett (2013), posteriormente é realizado o pré-processamento dos dados coletados por meio da metodologia *feature extraction*, na linguagem de programação *R* com aplicação do pacote *22 Canonical Time-series Characteristics* (LUBBA *et al.*, 2019), e a formação dos *clusters* por meio do método de *K-means*, descrito por Everitt *et al.* (2011) e Rencher (2012), e com auxílio do *software* RStudio.

Na segunda fase são aplicados os métodos estatísticos de previsão de demanda sobre os *clusters* com base na bibliografia de Krajewski *et al.* (2009) e Makridakis *et al.* (2005), utilizando o *software* Minitab. Posteriormente é realizada a desagregação de demanda dos *clusters* para cada um dos itens que o contemplam (CARAZZA, 2020). Por fim, na última fase é realizada a avaliação de desempenho como apresentado por Khoury (2011).

Figura 1 – Método de trabalho



Fonte: Elaborado pelo Autor

2.2.1. Definição dos produtos e coleta de dados

De acordo com Provost e Fawcett (2013), a escolha dos dados é um processo complexo e crucial na análise de dados, que pode ser estruturado em quatro etapas distintas. A primeira etapa da fase 1 envolve a compreensão do contexto no qual o problema está inserido e a definição clara dos objetivos do projeto, considerando os impactos esperados da análise. Posteriormente, devem ser formuladas as perguntas analíticas específicas que serão respondidas por meio da análise de dados. Na terceira etapa, realiza-se a identificação dos dados relevantes que serão utilizados para responder às perguntas analíticas formuladas, isso implica na seleção de conjuntos de dados que contenham as variáveis relevantes e que sejam capazes de influenciar o fenômeno em estudo. Por fim, a última etapa do processo envolve a limpeza, a transformação e a integração dos dados selecionados.

A seleção dos dados a serem analisados pode ser realizada através do emprego de um procedimento de priorização de itens. Um método viável para essa priorização é a utilização da Curva ABC, que permite a classificação dos itens de acordo com sua relevância para o negócio, visto que a maior parte do investimento de materiais está concentrada em um número reduzido de itens (BALLOU, 2001; SLACK, 2009). Segundo Guimarães (2020), a classificação é estabelecida da seguinte maneira: (i) curva A - corresponde a 20% dos produtos que representam 80% das vendas; (ii) curva B - engloba 30% dos produtos que correspondem a 15% das vendas; e (iii) curva C - envolve 50% dos produtos que correspondem a 5% das vendas.

Visto que os modelos tradicionais de previsão de demanda são negativamente impactados por valores espúrios nas séries temporais, todos os *outliers* foram retirados dos dados históricos obtidos através do *software* utilizado pela empresa, conforme procedimentos descritos por Pellegrini e Fogliatto (2001). Quando o *outlier* encontra-se no final da série temporal e for possível gerar uma previsão de demanda utilizando os dados que antecedem o valor espúrio, este deverá ser substituído pela previsão relativa correspondente ao mesmo período. Contudo, quando não houver dados suficientes para que o *outlier* seja expurgado pela previsão relativa, sugere-se que a substituição ocorra pelo valor médio das observações imediatamente adjacentes (PELLEGRINI; FOGLIATTO, 2001).

2.2.2. Pré-processamento de dados

O pré-processamento de dados é uma etapa crucial na análise de dados que envolve a transformação e preparação dos dados antes de aplicar uma técnica de análise específica. O objetivo do pré-processamento de dados é garantir que os dados estejam em um formato adequado e prontos para serem analisados, evitando problemas e melhorando a qualidade dos resultados (PROVOST; FAWCETT, 2013).

Devido a grande quantidade de dados gerados pelas séries temporais analisadas, encontrar padrões de dados pode tornar a análise mais clara e precisa. Desta forma a aplicação de técnicas de *feature extraction* se faz útil nesta etapa de pré-processamento de dados (BISHOP, 2006). De acordo com Kelleher e Tierney (2018), a *feature extraction* consiste em um processo que envolve a identificação e extração de características úteis de um conjunto de dados, a fim de melhorar a precisão e eficiência da análise, reduzindo a dimensionalidade dos dados e fornecendo um conjunto mais compacto e informativo de características que podem ser usadas para modelagem e análise.

Diante disto, há um destaque crescente para *softwares* que fornecem suporte para análises exploratórias de dados apoiadas por processos automatizados de extração de características em séries temporais multidimensionadas. Estes *softwares* permitem capturar uma variedade de aspectos e padrões presentes nos dados, a fim de facilitar seu posterior agrupamento e análise (LUBBA *et al.*, 2019).

2.2.3. Formação de *clusters*

Segundo Rencher (2012) e Everitt *et al.* (2011), o propósito da clusterização consiste em agrupar objetos em conjuntos homogêneos, com base em suas similaridades. Esse processo viabiliza a organização dos dados em grupos menores e mais gerenciáveis, facilitando a identificação de padrões e sua compreensão.

Existem duas abordagens para a clusterização, os métodos hierárquicos e os métodos não hierárquicos. A escolha entre a clusterização hierárquica e a não-hierárquica depende das características dos dados e dos objetivos da análise. A clusterização hierárquica pode ser mais adequada quando a estrutura dos dados é desconhecida ou quando é necessário explorar diferentes níveis de agrupamento. Já a clusterização não-hierárquica pode ser mais adequada quando for necessário obter

um número fixo de grupos ou quando os dados possuem uma estrutura clara e bem definida (EVERITT *et al.*, 2011; RENCHER, 2012).

Devido ao pré-processamento realizado nos dados na etapa anterior, onde são obtidas características para a clusterização, os métodos não-hierárquicos para a formação dos *clusters* são indicados (HAN *et al.*, 2011). Uma abordagem bastante conhecida dentro dos métodos não-hierárquicos, e que será utilizada nesta pesquisa, é o método *K-Means*. Inicialmente, este método seleciona aleatoriamente *k* objetos como centroides iniciais. Em seguida, cada objeto é alocado ao centroide mais próximo e os centroides são atualizados pela média dos objetos em cada *cluster*. Esse processo é repetido até que os centroides não mudem mais ou que um critério de convergência seja atingido (EVERITT *et al.*, 2011; RENCHER, 2012).

2.2.4. Previsão de demanda

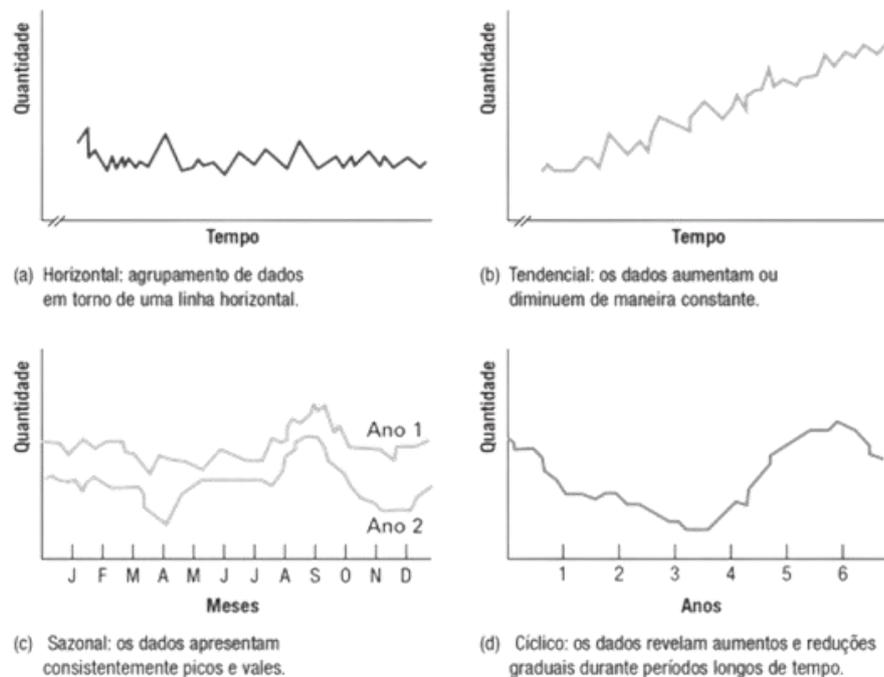
A previsão de demanda de produtos e serviços é uma atividade de extrema importância para empresas de todos os segmentos, uma vez que permite que decisões estratégicas sejam tomadas de forma mais precisa e assertiva. Ao prever a demanda futura é possível realizar ajustes na produção, nos níveis de estoque e na capacidade de atendimento de forma adequada, evitando desperdícios e aumentando a eficiência operacional. Além disso, a previsão de demanda permite que a empresa antecipe possíveis oscilações do mercado, identificando tendências e oportunidades de crescimento (MAKRIDAKIS *et al.*, 2005).

Os modelos de previsão de demanda podem ser apresentados como qualitativos e quantitativos. Modelos qualitativos são de natureza subjetiva, originados através da opinião de especialistas com base na experiência e conhecimento adquirido, correlacionando circunstâncias que possam gerar uma predição. Visto que diferentes especialistas podem apresentar preferências distintas, este método apresenta-se vulnerável a tendências, podendo comprometer a confiabilidade de seus resultados (PELLEGRINI, 2000). Neste trabalho são utilizados os modelos de previsão de demanda baseados em séries temporais, classificados como quantitativos. Estes modelos apresentam um processo de análise de dados bem definido, possibilitando a replicação dos resultados por diferentes especialistas. Este modelo utiliza as séries

temporais para determinar a tendência de comportamento futuro por meio de um modelo matemático (PELLEGRINI; FOGLIATTO, 2001).

Os modelos são aplicados com base na análise do comportamento histórico dos produtos e serviços, também conhecidas como séries temporais. Conforme Krajewski *et al.* (2009), há cinco padrões a serem verificados em séries temporais, vistos na Figura 2: (i) padrão horizontal, onde ocorrem flutuações dos dados em torno de uma média constante; (ii) padrão de tendência, onde há um aumento ou redução sistemática na média da série ao longo de um período; (iii) padrão sazonal, que consiste em um padrão de aumentos ou reduções na demanda de modo repetido dependendo do período de curto e médio prazo; (iv) padrão cíclico, onde ocorrem aumentos ou reduções graduais menos previsíveis na demanda por períodos mais longos de tempo; e (v) padrão aleatório, onde a variação de demanda ocorre de forma imprevisível.

Figura 2 – Padrões de Demanda



Fonte: KRAJEWSKI *et al.*, 2009

Os padrões aleatórios são resultados de causas eventuais, por este motivo sua previsão não pode ser realizada. Já os padrões horizontais, de tendência, de sazonalidade e de ciclo, podem ter combinações em graus variados para que seja definido um padrão de demanda para análise de um determinado produto ou serviço (KRAJEWSKI *et al.*, 2009).

2.2.4.1. Média móvel simples

O método de média móvel simples se mostra apropriado para previsão de demandas de curto prazo, não tendo necessidade de longos períodos históricos para que seja possível a realização de seu cálculo (MAKRIDAKIS *et al.*, 2005). Este método também se mostra mais eficiente quando aplicado a séries temporais que não apresentam tendências expressivas e influências de sazonalidade (KRAJEWSKI *et al.*, 2009).

Segundo Krajewski *et al.* (2009), a aplicação do método de média móvel simples consiste em calcular a média para os n períodos históricos mais recentes e utilizá-la como previsão para o período seguinte, conforme equação (1). Desta forma para o próximo período a ser analisado, onde a demanda calculada já foi realizada, o período mais antigo da média anterior deve ser substituído pela demanda mais recente, e o método recalculado (KRAJEWSKI *et al.*, 2009).

$$F_{t+1} = \frac{Y_t + Y_{t-1} + Y_{t-2} + \dots + Y_{t-n+1}}{n} \quad (1)$$

onde F_{t+1} é a previsão de demanda para o período $t+1$, Y_t é a demanda realizada no período t e n é o número total de períodos na média.

Para que a suavização das variações de demanda seja maior e a sensibilidade do modelo a mudanças na demanda seja reduzida, o número de períodos analisados na equação deve ser maximizado. Ao maximizar o número de períodos, o modelo de média móvel torna-se mais estável e menos sensível a mudanças abruptas na demanda, uma vez que as mudanças de curto prazo têm menos impacto na média calculada. (CORREA *et al.*, 2019).

2.2.4.2. Suavização exponencial simples

O método de suavização exponencial simples é semelhante ao modelo de média móvel ponderada, onde a ideia principal é atribuir pesos aos valores passados da série temporal, de modo que os valores mais recentes tenham mais peso em relação os valores mais antigos (KRAJEWSKI *et al.*, 2009). A fórmula do método de suavização exponencial simples é representada na equação (2), onde é possível observar que para o valor de Y_t é designado um parâmetro de suavização exponencial α , além de um peso para os valores previstos mais recentes dado por $1-\alpha$ (MAKRIDAKIS *et al.*, 2005).

$$F_{t+1} = \alpha Y_t + (1 - \alpha)F_t \quad (2)$$

onde F_{t+1} é a previsão de demanda para o período $t+1$, F_t é a previsão de demanda para o período t , Y_t é a demanda realizada no período t , α é a constante de suavização e n será o tamanho da série temporal.

Os pesos são controlados pelo parâmetro da constante de suavização, limitada a valores entre zero e um, que determinam o quão rapidamente estes pesos decaem. Quanto mais os valores de α se aproximam de um, maior será o ajuste do erro da previsão anterior, mostrando que o modelo ressalta demandas mais recentes e é sensível a mudanças. Já quando α se apresenta mais próximo a zero, menor será o ajuste, e o modelo tratará a demanda de forma mais uniforme e com previsões mais estáveis (MAKRIDAKIS *et al.*, 2005).

A suavização exponencial simples é amplamente utilizada em previsão de séries temporais, especialmente quando a série apresenta uma tendência suave e sazonalidade moderada. No entanto, existem algumas limitações, como a incapacidade de lidar com tendências abruptas e mudanças repentinas na série temporal (MAKRIDAKIS *et al.*, 2005).

2.2.4.3. Suavização linear de Holt

O método de suavização linear de Holt, também conhecido como método da suavização exponencial dupla, expande o método de suavização exponencial simples para contemplar previsões com dados que apresentem tendência, mas não sazonalidade. O método de suavização linear de Holt realiza a previsão de demanda

com duas constantes de suavização, α e β – que não estão correlacionadas e variam seus valores entre zero e um – além das equações de previsão (3), de nível (4) e de tendência (5) (MAKRIDAKIS *et al.*, 2005).

$$F_{t+m} = L_t + b_t m \quad (3)$$

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + b_{t-1}) \quad (4)$$

$$b_t = \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1} \quad (5)$$

onde F_{t+m} é a previsão de demanda para o período $t+m$, m é o horizonte de previsão, b_t é a estimativa de tendência da série temporal, L_t é a estimativa de nível da série temporal no período t e α e β são as constantes de suavização.

A equação (3) é utilizada para prever os valores futuros, onde o valor base L_t é somado a tendência b_t multiplicada pelo número de períodos do horizonte de planejamento m . A equação (4) de nível ajusta a previsão para a tendência do período anterior, por meio da adição da estimativa de tendência b_{t-1} ao último valor suavizado L_{t-1} . Esta técnica auxilia na eliminação da demora em verificar as inclinações da demanda, conduzindo o valor de previsão para um nível próximo ao valor da demanda atual. Já a equação (5) de tendência é atualizada a partir da diferença entre os dois últimos valores de nível suavizados, L_t e L_{t-1} . A tendência destes valores será modificada pela suavização β e adicionada a estimativa de tendência anterior, b_{t-1} , multiplicada por $1-\beta$ (MAKRIDAKIS *et al.*, 2005).

2.2.4.4. Método de Holt-Winters multiplicativo

O método de Holt-Winters é utilizado quando a série temporal apresenta um padrão na tendência e sazonalidade acentuada. Séries sazonais possuem padrões cíclicos de variação, repetidos por um mesmo período (PELLEGRINI; FOGLIATTO, 2001). A forma multiplicativa é aplicada a séries temporais onde a amplitude da sazonalidade varia de acordo com o nível da demanda. O método é composto pelas equações de previsão (6), de nível (7), de tendência (8) e de sazonalidade (9). (MAKRIDAKIS *et al.*, 2005).

$$F_{t+m} = (L_t + b_t m)S_{t-s+m} \quad (6)$$

$$L_t = \alpha \frac{Y_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + b_{t-1}) \quad (7)$$

$$b_t = \beta(L_t + l_{t-1}) + (1 - \beta)b_{t-1} \quad (8)$$

$$S_t = \gamma \frac{Y_t}{S_{t-s}} + (1 - \gamma)S_{t-s} \quad (9)$$

onde F_{t+m} é a previsão de demanda para o período $t+m$, m é o horizonte de previsão, b_t é a estimativa de tendência da série temporal, S_t é a estimativa do componente sazonal da série temporal no período t , s é o número de períodos por ciclo sazonal, L_t é a estimativa de do nível da série temporal no período t e α , β e γ são as constantes de suavização.

A equação (6) da previsão é resultado da equação de Holt (3), apresentada no tópico anterior, multiplicada pela estimativa de um componente sazonal S_{t-s+m} . A equação (7) de nível ajusta o valor base L_t a tendência realizada no período anterior, a partir da adição da estimativa de tendência da série temporal do período antecedente, b_{t-1} , ao último valor suavizado de nível, L_{t-1} . Para que seja possível eliminar as flutuações sazonais do cálculo do valor base na equação de nível (7), a demanda realizada no período t deve ser dividida pelo termo sazonal, S_{t-s} . A tendência neste método será calculada pela diferença entre os dois últimos valores suavizados de nível, L_t e L_{t-1} . Por fim, a equação de sazonalidade (9) estima o componente de sazonalidade através da multiplicação da constante de suavização γ pela razão da sazonalidade do período t , representada por Y_t , e a sazonalidade do período anterior, S_{t-s} (MAKRIDAKIS *et al.*, 2005).

2.2.5. Desagregação da previsão

Após a realização da previsão de demanda do *cluster* por meio do melhor modelo estatístico, é necessário que seja realizada a desagregação de tal demanda para cada um dos itens que compõem o *cluster*. A desagregação será realizada pelo índice de desagregação, que tem como base de cálculo os seguintes procedimentos: (i) para cada *cluster* avalia-se a participação percentual do *sell out* realizado de cada item ao longo do período analisado; e (ii) este percentual será então utilizado para desdobrar a demanda agregada para cada um dos produtos (CARAZZA, 2020).

2.2.6. Avaliação do desempenho

O principal objetivo desta análise é verificar se houve perda ou ganho na acurácia em decorrência da desagregação da demanda de cada um dos *clusters* para o nível de item e se a clusterização tem impactos positivos na previsão de demanda.

A medida de acurácia de um modelo de previsão é calculada por meio da soma dos erros produzidos por cada método. Deste modo, o modelo satisfatório para uma série de dados é aquele que apresenta menor erro associado. A avaliação da acurácia pode ser realizada utilizando diversas medidas, que permitem mensurar o quão próximos os valores previstos estão dos valores reais da série (PELLEGRINI, 2000). De acordo com Krajewski *et al.* (2009), o erro percentual absoluto médio (*mean absolute percentage error* - MAPE), é a melhor medida de erro para ser utilizada quando há comparações entre séries temporais. Khoury (2011) afirma que, quanto menor forem os valores obtidos através da equação (10), melhor a precisão dos dados previstos.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - F_i}{y_i} \right| * 100 \quad (10)$$

onde, n é o número total de períodos, y_i é o valor de demanda real e F_i será o valor de demanda previsto para o período i .

3. Resultados

No presente momento, a organização em estudo faz uso dos quatro métodos estatísticos disponível no *software* utilizado na empresa, sendo o método escolhido aquele que apresentar o menor MAPE histórico. Atualmente, alguns produtos demonstram uma acurácia estatística insatisfatória em suas previsões de demanda, sendo possível observar que o erro percentual absoluto médio pode atingir valores acima de 70% em determinados itens.

Desta forma, se objetiva alcançar uma acurácia maior na previsão estatística realizada. Nesta seção, são expostos os resultados provenientes da aplicação prática

do método de trabalho, com os resultados subdivididos de acordo com a estrutura proposta.

3.1. Definição dos produtos e coleta de dados

Para realização deste estudo, foi disponibilizada a base de dados utilizada pelo *software* da organização. Esta base de dados empregada para fins de análise é composta por dois principais grupos de produtos, medicamento e perfumaria. Dentre essa divisão, optou-se por focalizar no grupo de medicamento devido à sua significativa contribuição financeira para a estratégia da empresa, conforme identificado por meio da aplicação da Curva ABC. Esse grupo enquadra-se na categoria A da curva. No entanto, devido ao grande número de produtos presentes neste grupo, selecionou-se aqueles que detinham a maior representatividade em termos de volume de vendas e menor acurácia nas previsões estatísticas, resultando em uma amostra dos 16 principais produtos, doravante denominados produtos A até P.

Com base nessa abordagem, foram coletados dados históricos de *sell out* referente aos últimos 36 meses para cada um dos itens. O período histórico vai de junho de 2020 a maio de 2023, onde os valores realizados nos últimos três meses foram retirados da amostra a fim de serem utilizados no cálculo da acurácia dos valores preditos.

3.2. Pré-processamento de dados

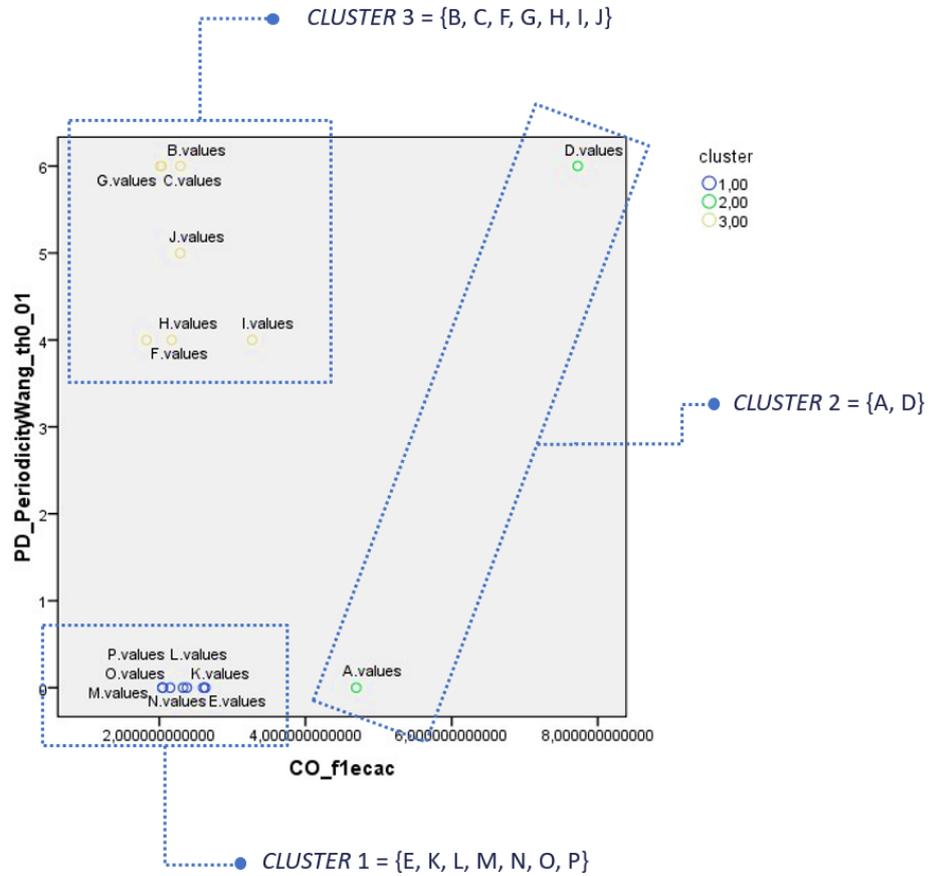
Nesta etapa da pesquisa, foi aplicada a função *catch22_all*, disponível no pacote "*catch22*" da linguagem de programação *R*, por meio do ambiente de desenvolvimento RStudio. Essa função foi utilizada com o objetivo de extrair um conjunto de características das séries temporais analisadas, descritas no Apêndice A, visando capturar uma variedade de aspectos e padrões presentes nos dados, a fim de facilitar seu posterior agrupamento e análise. Desta forma, a função forneceu um total de 22 características (*features*) para cada uma das 16 séries temporais, abrangendo medidas estatísticas, transformadas e propriedades não binárias.

3.3. Formação de *clusters*

Para a formação dos *clusters*, foi utilizado o algoritmo de aprendizado não supervisionado *K-means*, disponível no RStudio. Contudo, o referido algoritmo demanda que seja parametrizado o valor de *k*, que corresponde ao número de *clusters* desejados. Com o intuito de determinar o valor mais adequado para esta variável, foram consultados especialistas na área, que realizaram uma análise criteriosa dos dados. Após essa avaliação, que constatou que a amostra continha itens com alta sazonalidade, promocionais e itens regulares, foi recomendaram o uso de três *clusters* como o número mais apropriado para o comportamento das séries temporais utilizadas.

A partir da determinação do valor de *k*, foram utilizadas as características extraídas na etapa anterior como base para o agrupamento dos itens. Como resultados foram obtidas as seguintes formações de *clusters*: (a) *cluster* 1, formado pelos itens E, K, L, M, N, O e P; (b) *cluster* 2, formado pelos itens A e D; e (c) *cluster* 3, formado pelos itens B, C, F, G, H, I e J. Após a definição dos *cluster* utilizamos um algoritmo do tipo *STEPWISE* para selecionar as variáveis que melhor os discriminam. Desta forma, para uma melhor visualização dos *clusters*, conforme Figura 3, foram utilizadas as características *PD_PeriodicityWang_th0_01*, que descreve a ocorrência de padrões de periodicidade na série temporal, e a *CO_f1ecac* que está relacionada com o nível de incerteza ou aleatoriedade de uma variável.

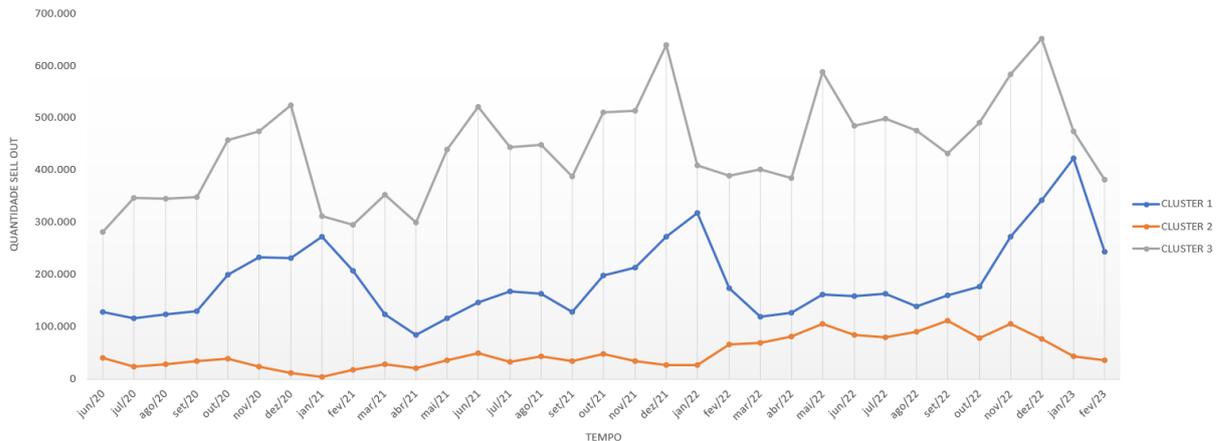
Figura 3 – Gráfico de dispersão entre as duas principais *features* e identificação dos *clusters*



Fonte: Elaborado pelo Autor

3.4. Previsão de demanda

Nesta etapa do processo, foi realizada a previsão de demanda para cada um dos *clusters* previamente estabelecidos na etapa anterior. Para isso, primeiramente, procedeu-se à soma da demanda dos itens individuais que compõem cada *cluster*, com o intuito de obter as séries temporais correspondentes, conforme a Figura 4.

Figura 4 – Séries Temporais por *cluster*

Fonte: Elaborado pelo Autor

Observa-se que o comportamento das séries temporais dos *clusters* 1 e 3 apresentam tendências de sazonalidade, demonstrando um padrão ao longo do período analisado. Em contraste, a série temporal do *cluster* 2 apresenta um comportamento relativamente horizontal, com flutuações ocorrendo em torno de um valor médio constante.

Após a consolidação dos agrupamentos de demanda, foram aplicados métodos de previsão específicos a cada *cluster* identificado. Sendo assim, utilizaram-se os métodos de Média Móvel, Suavização Exponencial Simples, Suavização Exponencial Dupla e Holt-Winters, para cada um dos *clusters* por meio do *software* estatístico Minitab, os resultados são apresentados na Tabela 1.

Os parâmetros empregados nos métodos estatísticos, dentre eles o número de períodos históricos para o cálculo da média móvel e as constantes de suavização α , β e γ para cálculo da suavização exponencial simples, Holt e Holt-Winters, foram estabelecidos com base nos mesmos valores empregados pelo *software* utilizado pela empresa. Para fins de cálculo da medida de avaliação de desempenho, foram utilizados os períodos retirados da amostra de séries temporais analisada, compreendendo o período de março de 2023 a maio de 2023.

Tabela 1 – Resultados das previsões de demanda agregada nos *clusters* para os três períodos considerados, separadamente por técnica

MÉDIA MÓVEL

TEMPO	CLUSTER 1	CLUSTER 2	CLUSTER 3
03/2023	269.604	75.488	502.079
04/2023	269.604	75.488	502.079
05/2023	269.604	75.488	502.079
MAPE	43%	29%	9%

SUAVIZAÇÃO EXPONENCIAL SIMPLES

TEMPO	CLUSTER 1	CLUSTER 2	CLUSTER 3
03/2023	205.841	38.712	487.757
04/2023	205.841	38.712	487.757
05/2023	205.841	38.712	487.757
MAPE	9%	63%	12%

SUAVIZAÇÃO EXPONENCIAL DUPLA

TEMPO	CLUSTER 1	CLUSTER 2	CLUSTER 3
03/2023	203.022	77.687	397.744
04/2023	206.506	115.649	399.262
05/2023	209.989	146.542	400.781
MAPE	10%	4%	28%

HOLT-WINTERS

TEMPO	CLUSTER 1	CLUSTER 2	CLUSTER 3
03/2023	186.824	75.990	496.635
04/2023	184.157	103.961	453.228
05/2023	200.271	131.933	575.714
MAPE	3%	7%	6%

Fonte: Elaborado pelo Autor

A escolha do método de previsão para cada *cluster* foi baseada na análise comparativa do MAPE resultante para os períodos. As previsões para os *clusters* 1 e 3 foram mais acuradas na aplicação do método de Holt-Winters, enquanto as previsões para o *cluster* 2 foram mais precisas utilizando o método de Suavização Exponencial Dupla (Tabela 2).

Tabela 2 – Previsões Escolhidas por *cluster*

TEMPO	CLUSTER 1	CLUSTER 2	CLUSTER 3
03/2023	186.824	77.687	496.635
04/2023	184.157	115.649	453.228
05/2023	200.271	146.542	575.714
MAPE	3%	4%	6%

Fonte: Elaborado pelo Autor

3.5. Desagregação da previsão

Uma vez determinado o método adequado para cada *cluster*, procedeu-se à desagregação da demanda de cada item componente. Tal desagregação foi realizada por meio do cálculo de um índice, resultante da representatividade da soma dos históricos de cada item ao somatório do *cluster* correspondente, conforme detalhado na Tabela 3.

Tabela 3 – Índice de desagregação por item dentro de cada *cluster*

ÍNDICE DE DESAGREGAÇÃO - CLUSTER 1							
TEMPO	E	K	L	M	N	O	P
∑ REALIZADO ITEM	3.015.115	1.954.844	281.929	248.489	8.938	317.527	406.783
∑ REALIZADO CLUSTER	6.233.625	6.233.625	6.233.625	6.233.625	6.233.625	6.233.625	6.233.625
ÍNDICE	48,37%	31,36%	4,52%	3,99%	0,14%	5,09%	6,53%

ÍNDICE DE DESAGREGAÇÃO - CLUSTER 2		
TEMPO	A	D
∑ REALIZADO ITEM	655.643	973.807
∑ REALIZADO CLUSTER	1.629.450	1.629.450
ÍNDICE	40,24%	59,76%

ÍNDICE DE DESAGREGAÇÃO - CLUSTER 3							
TEMPO	B	C	F	G	H	I	J
∑ REALIZADO ITEM	225.087	1.080.456	9.172.322	1.877.558	413.731	735.550	1.077.261
∑ REALIZADO CLUSTER	14.581.965	14.581.965	14.581.965	14.581.965	14.581.965	14.581.965	14.581.965
ÍNDICE	1,54%	7,41%	62,90%	12,88%	2,84%	5,04%	7,39%

Fonte: Elaborado pelo Autor

Com base nos resultados obtidos do índice de desagregação, foi realizada a distribuição das previsões do *cluster* para cada item componente. Com isso, foi possível calcular o MAPE de cada item em relação aos meses de março a maio de 2023, utilizando os dados retirados da amostra das séries temporais especificamente para essa finalidade, conforme Tabela 4.

Tabela 4 – Previsão desagregada por item

PREVISÃO CLUSTER 1 DESAGREGADA							
TEMPO	E	K	L	M	N	O	P
03/2023	90.364	58.587	8.450	7.447	268	9.516	12.191
04/2023	89.074	57.751	8.329	7.341	264	9.381	12.017
05/2023	96.868	62.804	9.058	7.983	287	10.201	13.069
MAPE	3%	10%	19%	19%	27%	11%	9%

PREVISÃO CLUSTER 2 DESAGREGADA

TEMPO	A	D
03/2023	31.259	46.428
04/2023	46.534	69.115
05/2023	58.964	87.578
MAPE	18%	9%

PREVISÃO CLUSTER 3 DESAGREGADA

TEMPO	B	C	F	G	H	I	J
03/2023	7.666	36.798	312.392	63.946	14.091	25.051	36.690
04/2023	6.996	33.582	285.089	58.357	12.859	22.862	33.483
05/2023	8.887	42.658	362.135	74.128	16.335	29.040	42.532
MAPE	9%	21%	8%	9%	10%	28%	20%

Fonte: Elaborado pelo Autor

3.6. Avaliação de desempenho

Por fim, foram comparados os valores do MAPE obtidos através da desagregação da previsão de cada *cluster* com a previsão realizada para cada item individualmente. A previsão individual foi realizada utilizando os quatro métodos estatísticos empregados neste estudo, sendo selecionado o método com maior acurácia, conforme resultados descritos no Apêndice B. Os resultados deste comparativo estão apresentados na Tabela 5.

Tabela 5 – Comparativo do MAPE entre os métodos de previsão agregada e individual separadamente por *cluster*

CLUSTER 1	E	K	L	M	N	O	P
MAPE - DESAGREGAÇÃO DO CLUSTER	3%	10%	19%	19%	27%	11%	9%
MAPE - PREVISÃO INDIVIDUAL DO ITEM	9%	13%	38%	55%	33%	31%	73%
CLUSTER 2	A	D					
MAPE - DESAGREGAÇÃO DO CLUSTER	18%	9%					
MAPE - PREVISÃO INDIVIDUAL DO ITEM	23%	19%					
CLUSTER 3	B	C	F	G	H	I	J
MAPE - DESAGREGAÇÃO DO CLUSTER	9%	21%	8%	9%	10%	28%	9%
MAPE - PREVISÃO INDIVIDUAL DO ITEM	13%	62%	9%	14%	15%	32%	10%

Fonte: Elaborado pelo Autor

Conclui-se que houve uma redução MAPE em todos os itens ao realizar a desagregação da demanda dos *clusters* em comparação com as previsões individuais de cada item. Notavelmente, os itens pertencentes ao *cluster 1* apresentaram uma

redução significativa no erro, destacando-se os itens M, O e P, que aumentaram consideravelmente a acurácia dos valores preditos. Os *clusters* 2 e 3 também obtiveram resultados positivos, embora as reduções no erro tenham sido mais modestas.

4. Considerações Finais

Este estudo teve como objetivo principal apresentar e aplicar uma sistemática baseada em técnicas de previsão de demanda integradas a modelos de agrupamento de dados históricos de categoria de itens para responder à questão de pesquisa definida. A questão de pesquisa buscava compreender como incorporar uma etapa de clusterização em um modelo de previsão de demanda que assegurasse uma maior acurácia nas previsões de demanda da empresa analisada. Sendo assim, a sistemática apresentada propôs a utilização de uma ferramenta antes do início do processo de previsão de demanda, visando garantir resultados mais acurados na compra realizada pela empresa, impactando diretamente o nível de satisfação dos clientes.

A previsão de demanda desempenha um papel fundamental na gestão dos principais processos internos da empresa, incluindo processos de importação, dimensionamento de estoque e atividades comerciais. Uma previsão de demanda mais acurada possibilita uma utilização mais eficiente dos recursos disponíveis e um planejamento mais eficaz dos recursos financeiros da empresa, resultando em um aumento no nível de atendimento aos clientes.

Os objetivos específicos descritos para este trabalho foram: (i) avaliação dos potenciais ganhos do agrupamento das categorias de produtos na previsão de demanda em relação a previsão individual do item e (ii) aumento da acurácia nas previsões de demanda dos diferentes tipos de itens. A partir da aplicação dos métodos estudados, foi possível verificar que houve um ganho significativo no emprego da ferramenta de clusterização sobre os itens analisados, havendo um aumento de acurácia em todos os itens da amostra.

Como sugestão para trabalhos futuros, recomenda-se considerar a utilização de períodos alternativos de desagregação de demanda caso os resultados obtidos por meio da aplicação de um índice de desagregação simples, que utiliza todo o período

histórico analisado, não atendam a tendência da série histórica do item em períodos imediatamente anteriores a previsão. Indica-se também a abrangência da aplicação da metodologia a amostras maiores, contemplando diversos grupos de itens.

Referências

- AMALNICK, M.S., HABIBIFAR, N., HAMID, M. *et al.* **An intelligent algorithm for final product demand forecasting in pharmaceutical units.** *Int J Syst Assur Eng Manag* 11, 481–493 (2020). <https://doi.org/10.1007/s13198-019-00879-6>
- BALLOU, R. H. **Gerenciamento da Cadeia de Suprimentos: Planejamento, Organização e Logística Empresarial.** 4. ed., Porto Alegre: Bookman, 2001.
- BARANDAS, M. *et al.* **TSFEL: Time Series Feature Extraction Library.** *SoftwareX*, v. 11, p. 100456, jan. 2020.
- CONSUL, F. B.; WERNER, L. **Avaliação de técnicas de previsão de demanda utilizadas por um software de gerenciamento de estoques no setor farmacêutico.** In: CONGRESSO BRASILEIRO DE PESQUISA OPERACIONAL, 40, 2010, São Paulo. *Anais do 40º Congresso Brasileiro de Pesquisa Operacional.* São Paulo: ABEPRO, 2010. p. 1-10.
- BISHOP, C. M. **Pattern Recognition and Machine Learning.** 1st ed. New York: Springer, 2006.
- CARAZZA, H. A. V. **Integração de ferramentas de clusterização e modelos quantitativos de previsão de demanda: uma aplicação no setor alimentício.** Dissertação (Pós-Graduação em Engenharia de Produção) - Departamento de Engenharia de Produção e Transportes. Porto Alegre – Universidade Federal do Rio Grande do Sul, 2020.
- CORREA, Henrique Luiz; GIANESI, Irineu Gustavo Nogueira; CAON, Mauro. **Planejamento, programação e controle da produção.** 6. ed. São Paulo: Atlas, 2019.
- EVERITT, Brian S.; LANDAU, Sabine; LEESE, Morven; STAHL, Daniel. **Cluster Analysis.** 5ed. New Jersey: John Wiley & Sons, Ltd., 2011.
- GIL, A. **Métodos e Técnicas de Pesquisa Social.** São Paulo: Atlas, 1995.
- GUIMARÃES, V. **“Curva ABC: o que é e por que ela importa?”.** 2017. Disponível em: <<https://www.escoladeecommerce.com/artigos/curva-abc-o-que-e-e-por-que-ela-importa/>> Acesso em: 25 de jun. 2023.
- HAN, J.; KAMBER, M.; PEI, J. **Data Mining: Concepts and Techniques.** 3rd ed. Burlington: Morgan Kaufmann, 2011.

HOLT, C. C. **Forecasting Seasonals and Trends by Exponentially Weighted Moving Averages**. International Journal of Forecasting. v. 20, n. 1, p. 5-10, 2004.

INTERFARMA. Associação da Indústria Farmacêutica de Pesquisa. **Guia 2022 Interfarma**. Disponível em: <<https://www.interfarma.org.br/library/guia-2022-interfarma/>>. Acesso em: 20 set. 2022.

KELLEHER, J. D.; TIERNEY, B. **Data science**. Cambridge, Massachusetts; London, England: The Mit Press, 2018.

KHOURY, F. K. C. B. **Minimização de custos de produção via programação inteira mista: estudo de caso de planejamento de produção de luminárias**. Orientador: Profa. Fernanda Maria Pereira Raupp. 2011. Dissertação de Mestrado (Mestre em Engenharia de Produção) - Pontifca Universidade Católica do Rio de Janeiro, 2011.

KRAJEWSKI, L. J.; RITZMAN, L. P.; MALHOTRA, M. K. **Administração de produção e operações**. São Paulo: Prentice-Hall, 2009.

LUBBA, C.H., SETHI, S.S., KNAUTE, P. et al. **catch22: CAnonical Time-series CHaracteristics**. Data Min Knowl Disc 33, 1821–1852 (2019). <https://doi.org/10.1007/s10618-019-00647-x>

MAKRIDAKIS, S. G.; HYNDMAN, R. J.; WHEELWRIGHT, S. C. **Forecasting: methods and applications**. New Delhi: Wiley, 2005.

MARCONI, M.A.; LAKATOS, E.M. **Técnicas de pesquisa: Planejamento e execução de pesquisas, amostragens e técnicas de pesquisas, elaboração, análise e interpretação de dados**. 5 ed. São Paulo, 2002.

MONTGOMERY, D.; JOHNSON, L.; GARDINER, J. **Forecasting and Time Series Analysis**. New York: McGraw-Hill, 1990.

PELLEGRINI, F. R. **Metodologia para implementação de sistemas de previsão de demanda**. Dissertação (Mestrado em Engenharia de Produção) - Departamento de Engenharia de Produção e Transportes. Porto Alegre – Universidade Federal do Rio Grande do Sul, 2000.

PELLEGRINI, F. R.; FOGLIATTO, F. S. **Passos para implantação de sistemas de previsão de demanda: técnicas e estudo de caso**. Production, v. 11, n. 1, p. 43–64, jun. 2001.

PROVOST, F.; FAWCETT, T. **Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking**. O'Reilly Media, 2013.

RENCHE, A. C. **Methods of Multivariate Analysis**. 3. ed., New York: John Wiley & Sons, 2012.

RODRIGUES, V. **Um modelo de previsão de demanda no varejo do setor de saúde e bem-estar**. Disponível em: <<https://repositorio.ufsc.br/handle/123456789/223370>>. Acesso em: 26 set. 2022.

SILVA, E. L.; MENEZES, E. M. **Metodologia da Pesquisa e Elaboração da Dissertação**. Florianópolis, Santa Catarina: UFSC, 4 ed. rev. atual, 2005.

SINDUSFARMA. **Mercado farmacêutico deve crescer 12% este ano e 10% em 2023**. Disponível em: <<https://pfarma.com.br/noticia-setor-farmaceutico/mercado/7695-mercado-farmaceutico-deve-crescer-12-este-ano-e-10-em-2023.html>>. Acesso em: 20 set. 2022.

SLACK, N.; CHAMBERS, S.; JOHNSTON, R.; **Administração da Produção**. São Paulo: Atlas, 2009.

VERGARA, S. C. **Projetos e relatórios de pesquisa em administração**. 3. ed. São Paulo: Atlas, 2000.

WERNER, L. **Um modelo composto para realizar previsão de demanda através da integração da combinação de previsões e do ajuste baseado na opinião**. Tese (Doutorado em Engenharia de Produção) - Departamento de Engenharia de Produção e Transportes. Porto Alegre – Universidade Federal do Rio Grande do Sul, 2004.

Apêndice A – Descrição das *Features*

Feature	Descrição
DN_HistogramMode_5	Moda do histograma com 5 intervalos (<i>bins</i>) da série temporal.
DN_HistogramMode_10	Moda do histograma com 10 intervalos (<i>bins</i>) da série temporal.
CO_f1ecac	Autoentropia do primeiro coeficiente cepstral da série temporal
CO_FirstMin_ac	Primeiro mínimo do autocorrelograma da série temporal.
CO_HistogramAMI_even_2_5	Informação mútua média entre duas séries temporais, calculada usando histograma com número par de intervalos (<i>bins</i>) e largura de banda de 0,5.
CO_trev_1_num	Número de vezes em que a série temporal reverte sua tendência ao longo do tempo.
MD_hrv_classic_pnn40	Porcentagem de intervalos RR consecutivos com diferenças superiores a 40 ms em relação ao intervalo RR anterior.
SB_BinaryStats_mean_longstretch1	Média dos estiramentos longos de valor um na série binária.
SB_TransitionMatrix_3ac_sumdiagcov	Soma das covariâncias na diagonal principal da matriz de transição de três símbolos em uma série binária.
PD_PeriodicityWang_th0_01	Porcentagem de ocorrência de padrões de periodicidade na série temporal, de acordo com o algoritmo de Wang <i>et al.</i> com limiar de 0,01.
CO_Embed2_Dist_tau_d_expfit_mean_diff	Diferença média entre as distâncias dos vetores incorporados em duas dimensões com diferentes valores de atraso (τ) na série temporal.
IN_AutoMutualInfoStats_40_gaussian_fmfi	Valor da função de informação mútua para uma série temporal gaussiana estimada a partir de 40 pontos.
FC_LocalSimple_mean1_taresrat	Taxa de resíduos entre média e média móvel para uma janela de um ponto na série temporal.
DN_OutlierInclude_p_001_mdrmd	Média dos resíduos médios absolutos da diferença média rolante (MDRMD) para uma janela de 0,01% dos pontos da série temporal.
DN_OutlierInclude_n_001_mdrmd	Número de pontos anômalos detectados pela diferença média rolante (MDRMD) para uma janela de 0,01% dos pontos da série temporal.
SP_Summaries_welch_rect_area_5_1	Área de um retângulo específico no espectro Welch da série temporal.
SB_BinaryStats_diff_longstretch0	Diferença entre a quantidade de estiramentos longos de zeros e uns na série binária.
SB_MotifThree_quantile_hh	Quantil alto-alto do histograma de padrões de três símbolos em uma série binária.
SC_FluctAnal_2_rsrangeftit_50_1_logi_prop_r1	Proporção de pontos em uma série temporal que se ajustam a uma função logística para uma determinada faixa de resíduos.

SC_FluctAnal_2_dfa_50_1_2_logi_pro p_r1	Proporção de pontos em uma série temporal que se ajustam a uma função logística para uma determinada faixa de resíduos, calculada usando análise de flutuação fractal (DFA).
SP_Summaries_welch_rect_centroid	Coordenada do centroide de um retângulo específico no espectro Welch da série temporal.
FC_LocalSimple_mean3_stderr	Erro padrão da média móvel para uma janela de três pontos na série temporal.

Fonte: LUBBA, 2019.

Apêndice B – Previsão individual do item

TEMPO	A	B	C	D	E	F	G	H
03/2023	28.531	9.504	17.703	59.159	114.500	315.071	33.348	26.313
04/2023	28.531	10.733	17.703	62.925	118.568	306.275	33.504	28.093
05/2023	28.531	10.225	17.703	69.627	131.642	332.290	33.661	29.873
TEMPO	I	J	K	L	M	N	O	P
03/2023	43.769	24.790	54.647	7.074	1.558	312	11.420	2.550
04/2023	43.769	24.790	54.647	3.454	911	312	5.656	1.490
05/2023	43.769	24.790	54.647	8.615	10.073	312	13.568	16.488

Fonte: Elaborado pelo autor