

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ADMINISTRAÇÃO
DEPARTAMENTO DE CIÊNCIAS ADMINISTRATIVAS**

BERNARDO CARDOZO BIANCHESSI

**ANÁLISE DE SÉRIES TEMPORAIS: COMPARAÇÃO ENTRE MODELOS
PREDITIVOS EM UMA MICROCERVEJARIA DO RIO GRANDE DO SUL**

Porto Alegre
2023

BERNARDO CARDOZO BIANCHESSI

**ANÁLISE DE SÉRIES TEMPORAIS: COMPARAÇÃO ENTRE MODELOS
PREDITIVOS EM UMA MICROCERVEJARIA DO RIO GRANDE DO SUL**

Trabalho de Conclusão de Curso a ser apresentado ao Departamento de Ciências Administrativas da Universidade Federal do Rio Grande do Sul, como requisito parcial para a obtenção do grau de Bacharel em Administração.

Orientador: Prof. Antônio Carlos Gastaud
Maçada

Porto Alegre
2023

SUMÁRIO

1 INTRODUÇÃO	8
2 JUSTIFICATIVA	11
3 OBJETIVOS	13
3.1 OBJETIVO GERAL	13
3.2 OBJETIVOS ESPECÍFICOS	13
4 REVISÃO BIBLIOGRÁFICA	14
4.1 PREVISÃO DE DEMANDA	14
4.2 MÉTODOS QUANTITATIVOS DE PREVISÃO DE DEMANDA	15
4.2.1. Métodos causais	15
4.2.2 Séries temporais	16
4.2.3 Médias Móveis	19
4.2.4 Ponderação exponencial simples	20
4.2.5 ARIMA	21
4.3 MÉTODOS DE INTELIGÊNCIA ARTIFICIAL DE PREVISÃO DE DEMANDA	21
4.4 MEDIDAS DE ACURÁCIA	22
4.4.1 Desvio padrão	23
4.4.2 Média do quadrado de erros (MQE)	23
4.4.3 Média absoluta de erros (MAE)	24
4.4.4 Percentual de erro absoluto médio (MAPE)	24
4.5 TOMADA DE DECISÃO BASEADA EM DADOS	25
5 PROCEDIMENTOS METODOLÓGICOS	27
5.1 ENTENDIMENTO DO NEGÓCIO	29
6 RESULTADOS	31
6.1 ENTENDIMENTO E PREPARAÇÃO DOS DADOS	31
6.2 MODELAGEM E AVALIAÇÃO	35
6.3 ANÁLISE DOS RESULTADOS	38
7 CONCLUSÃO	41
REFERÊNCIAS BIBLIOGRÁFICAS	43
ANEXO A - ROTEIRO DE ENTREVISTA SEMIESTRUTURADA PRÉ-MODELAGEM	48
Anexo B - ROTEIRO DE ENTREVISTA PÓS-MODELAGEM	51

LISTA DE FIGURAS

Figura 1 – Etapas da Metodologia	28
Figura 2 – Série temporal de litros vendidos por mês	34
Figura 3 – Decomposição da série temporal	35
Figura 4 – Valores das métricas resultantes da predição da série temporal	38
Figura 5 – Comparação entre os valores previstos pelo modelo ARIMA e	39
Figura 6 – Valores reais e previsão do modelo ARIMA	40

LISTA DE QUADROS

Quadro 1 – Atributos do Dataset	33
Quadro 2 – Estatísticas do conjunto de dados	36

AGRADECIMENTOS

Com profunda gratidão, estendo meus sinceros agradecimentos à minha família, cujo suporte inabalável e incentivo contínuo foram pilares fundamentais em minha jornada acadêmica.

À Paula, cujo apoio emocional foi uma luz nos momentos mais desafiadores. Ao Prof. Maçada, cujas orientações e esclarecimentos valiosos enriqueceram imensamente este trabalho.

Por fim, expresso minha gratidão à empresa que abriu suas portas para a realização deste estudo, permitindo-me a oportunidade de desenvolver e aprofundar minha pesquisa.

RESUMO

A previsão de demanda é essencial para empresas, sobretudo em mercados com intensa concorrência e sazonalidade marcante, como o das microcervejarias. Em ambientes como este, antecipar-se às tendências e compreender os padrões de consumo torna-se uma vantagem competitiva inestimável. Este estudo aborda tais desafios, objetivando a identificação do método mais eficiente para realizar previsões de demanda. Em modelos quantitativos de séries temporais, foi utilizado o ARIMA, assim como foram utilizadas técnicas avançadas de inteligência artificial, destacando-se os modelos de Redes Neurais Recorrentes (RNNs) com Memória de Longo Prazo (LSTM) e XG Boost. A partir de dados históricos fornecidos por uma microcervejaria, foi criada a série temporal de demanda da empresa que, posteriormente, foi analisada e utilizada para desenvolver um modelo de previsão, utilizando a biblioteca Pycaret. As métricas MAPE, MASE e MAE avaliaram a precisão das previsões. O modelo ARIMA, tradicionalmente conhecido por sua eficácia em análises de séries temporais, obteve os melhores resultados em todas as métricas de acurácia, alcançando um MAPE de 10,72%. Conclui-se que o ARIMA é o modelo com os melhores resultados de acurácia, representando uma opção para as microcervejarias realizarem a previsão de demanda. Ao adotar essa abordagem, as empresas podem obter vantagens competitivas significativas em seu planejamento.

Palavras-chave: Previsão de demanda. Predição de séries temporais. ARIMA. Microcervejarias. Aprendizado de Máquina.

ABSTRACT

Demand forecasting is essential for companies, especially in markets with intense competition and pronounced seasonality, such as the craft brewery sector. In environments like this, anticipating trends and understanding consumption patterns becomes an invaluable competitive advantage. This study addresses such challenges, aiming to identify the most efficient method for demand forecasting. In quantitative time series models, ARIMA was used, as well as advanced artificial intelligence techniques, notably the Recurrent Neural Network (RNN) models with Long Short-Term Memory (LSTM) and XG Boost. Based on historical data provided by a craft brewery, a company's demand time series was established, which was later analyzed and used to develop a forecasting model using the Pycaret library. The MAPE, MASE, and MAE metrics evaluated the accuracy of the forecasts. The ARIMA model, traditionally known for its efficacy in time series analyses, achieved the best results in all accuracy metrics, reaching a MAPE of 10.72%. It is concluded that ARIMA is the model with the best accuracy results, representing an option for craft breweries to forecast demand. By adopting this approach, companies can gain significant competitive advantages in their planning.

Keywords: Demand forecasting. Time series prediction. ARIMA. Microbreweries. Machine Learning.

1 INTRODUÇÃO

O setor de cerveja no Brasil tem uma longa história, tendo sido introduzido pelos colonizadores europeus, principalmente os alemães, no século XVI (Limberger, 2013). Atualmente, a produção e consumo de cerveja no Brasil é extremamente concentrada, considerando que 68% do mercado industrial de cerveja é detido pela AB InBev, 17,4% pela Heineken e 14,1% pela cervejaria nacional Petrópolis, é possível afirmar que essas três empresas conjuntamente possuem 98,5% do mercado brasileiro (Limberger, 2019). Porém, nos últimos dez anos, tem-se assistido a um crescimento exponencial do número de cervejarias em território nacional (MAPA, 2021). Além disso, a cerveja é hoje um dos principais produtos de consumo no país, e sua produção e comercialização tem se mantido em alta constante.

Segundo estudo realizado pela BarthHaas (2021), o Brasil encontra-se como o terceiro maior produtor de cerveja do mundo, produzindo 14,3 bilhões de litros no ano de 2021, estando atrás dos Estados Unidos com 20,3 bilhões e China com 35,9 bilhões. Por possuir um importante efeito multiplicador na economia, a produção de cerveja impacta uma extensa cadeia produtiva, a qual é responsável por 1,6% do PIB e 14% da indústria de transformação nacional, além de ser um dos maiores empregadores do Brasil, com 2,2 milhões de pessoas empregadas (Cervbrasil, 2016). Desta forma, o setor é um dos mais relevantes da economia brasileira.

De acordo com dados do Anuário da Cerveja (2021), o número de cervejarias no período entre 2011 e 2021 cresceu aproximadamente 1200%, sendo que o Rio Grande do Sul, com 285 cervejarias, ocupa o segundo lugar no ranking entre os estados brasileiros, estando atrás apenas de São Paulo em relação ao número de cervejarias, quem tem 340. Com esse aumento significativo de empresas atuando no mercado, a concorrência vem crescendo muito no setor, o que ocasiona uma maior competitividade e direciona cada vez mais as empresas a terem maior assertividade quanto a sua demanda.

A previsão de demanda é um tema amplamente discutido na literatura empresarial, e sua importância para o sucesso e a sustentabilidade econômica das microempresas no período pós-pandemia é amplamente reconhecida (Dos Reis, 2022). Pesquisas recentes indicam que a acuracidade na projeção da demanda

desempenha um papel essencial na rentabilidade das empresas (Veiga, Veiga, Vieira & Tortato, 2012). Esta acurácia tem impactos diretos sobre diversos aspectos, incluindo o nível de serviço ao consumidor, o estabelecimento de estoques de segurança, a minimização de custos totais e o desempenho global da cadeia de suprimentos. Uma previsão precisa permite uma melhor antecipação da demanda do cliente por parte da produção, conforme salientado por Meijden, Nunen & Ramondt (1994). Em contrapartida, projeções imprecisas podem afetar adversamente a eficácia da empresa, conduzindo a cenários como insuficiência no atendimento da demanda, excesso de estoque, resultando em produtos obsoletos, gerando prejuízos à empresa.

Isso se aplica também às microcervejarias, cuja capacidade de prever a demanda para seus produtos permite que as empresas façam planejamentos eficientes e tomem decisões assertivas sobre planejamento da produção e venda de seus produtos (Diniz, 2019). Além disso, a diferenciação e a sobrevivência de empresas atuantes em mercados competitivos dependem destes processos de planejamento. Sendo assim, o processo de previsão é de extrema relevância para atingir um efetivo e eficiente planejamento (Wheelwright *et al.*, 1980).

Atualmente os métodos de previsão podem ser classificados em três categorias: qualitativos, quantitativos e inteligência artificial. Os métodos qualitativos são recomendados quando não há dados quantitativos históricos disponíveis ou quando esses dados não forem confiáveis. Eles utilizam a expertise de especialistas, permitindo alta flexibilidade e análises mais amplas sobre aspectos estratégicos. São simples na implantação e utilizados para previsões de médio e longo prazos. Já os métodos quantitativos são recomendados quando existem dados históricos disponíveis. Eles se baseiam em análises estatísticas e matemáticas para prever tendências e padrões futuros. Esses métodos são úteis para prever pontos de inflexão na demanda e fornecem subsídios para previsões mais sofisticadas. Por fim, os métodos de inteligência artificial são capazes de modelar situações simples e complexas, possibilitando identificar relações não lineares e interativas entre os dados. Eles são recomendados para situações que requerem análise de realidades complexas e demandam recursos computacionais avançados (Ackermann, 2022).

Em relação às pesquisas recentes sobre previsão de demanda em empresas, segundo o estudo de caso realizado por Capitanio (2019), a utilização de técnicas de *Machine Learning* e Rede Neural Artificial (RNA), que são métodos de inteligência artificial, tem se mostrado uma opção promissora para a previsão de demanda, pois permite ajustes precisos em função de mudanças no mercado. Além disso, de acordo com Calôba (2002), a cooperação das redes neurais com métodos clássicos lineares mostrou-se um método de previsão mais assertivo do que o uso dessas técnicas separadas. Em complemento a estes pontos, a previsão de demanda também pode ser integrada pela análise de dados qualitativos, como pesquisas de mercado e entrevistas com especialistas (Lemos, 2006), o que permite à empresa entender as tendências e as necessidades dos consumidores de maneira mais precisa.

Diante da relevância dos métodos de previsão de demanda para cervejarias, essa monografia busca responder a seguinte questão de pesquisa: como prever a demanda de uma microcervejaria utilizando métodos quantitativos e redes neurais?

Apresentado o problema de pesquisa, o capítulo seguinte apresenta a justificativa para escolha do tema proposto.

2 JUSTIFICATIVA

A crescente concorrência no mercado de cervejas brasileiro, principalmente no Rio Grande do Sul, torna o processo de previsão de vendas cada vez mais complexo para pequenas e médias empresas que atuam no setor. Conforme exposto no capítulo anterior, existem diversas técnicas e modelos que podem ser utilizados para realizar a previsão de demanda de forma precisa. Levando isso em consideração, o presente trabalho propõe explorar a vasta base de dados históricos, fornecida pelo sistema ERP da empresa a ser estudada, a qual contém informações sobre a demanda dos últimos seis anos.

Sendo assim, o presente estudo tem como propósito utilizar abordagens quantitativas e de inteligência artificial para previsão de demanda. Nesse sentido, a análise das vantagens e desvantagens de ambas metodologias revela-se de extrema importância, uma vez que permitirá que as microcervejarias identifiquem os modelos mais adequados para a realidade de seu mercado, tornando o mercado, em geral, mais competitivo.

A empresa em questão é uma microcervejaria localizada no Rio Grande do Sul, a qual teve uma crescente demanda de seus produtos. De fato, nos últimos dois anos, a empresa conseguiu dobrar sua produção anual para atender essa demanda. Apesar do recente aumento na demanda, a empresa enfrenta desafios na gestão dessa, principalmente devido à natureza sazonal do mercado de cervejas e à crescente concorrência no setor. Esses fatores tornam a previsão de demanda uma ferramenta essencial para a sustentabilidade econômica da empresa.

Além disso, a necessidade de uma melhor previsão de demanda origina-se a partir de uma série de desafios internos enfrentados pela empresa. Como mencionado, a falta de previsibilidade na demanda levou a problemas de gestão de estoque, com excesso de produtos em estoque em algumas ocasiões e escassez de produtos em outras. Além disso, a empresa busca constantemente investir em sua expansão produtiva, o que requer uma compreensão clara da demanda futura.

De acordo com os autores Martins e Maçada (2022), que identificaram uma lacuna na literatura existente a respeito do Big Data Analytics, ainda há uma compreensão limitada de como essa tecnologia contribui especificamente para aprimorar a qualidade das decisões tomadas. Isso sugere a necessidade de estudos

adicionais que aprofundem esses conceitos, com foco específico na tomada de decisões baseada em dados.

Dessa forma, o estudo proposto sobre a previsão de demanda nessa microcervejaria é de grande relevância, sabendo que o tema é essencial para o sucesso e a sustentabilidade econômica da empresa (Dos Reis, 2022). A utilização dos métodos propostos, combinada com a disponibilidade dos dados históricos de demanda, permitirá uma análise mais precisa e antecipada das tendências do mercado, auxiliando na minimização dos riscos operacionais.

Após a apresentação da justificativa que fundamenta a presente pesquisa, o próximo capítulo tem como propósito explicitar os objetivos que se pretende alcançar.

3 OBJETIVOS

Este capítulo tem como finalidade expor os objetivos da presente pesquisa, os quais foram divididos em objetivos gerais e específicos, conforme apresentado nos tópicos subsequentes.

3.1 OBJETIVO GERAL

Analisar e propor um modelo de previsão de demanda adequado ao processo de planejamento atual de uma microcervejaria situada na cidade de Sapucaia do Sul - RS.

3.2 OBJETIVOS ESPECÍFICOS

São objetivos específicos desta pesquisa:

- a) identificar na literatura os métodos e técnicas de previsão de demanda;
- b) elaborar simulações com dados reais, aplicando os modelos selecionados;
- c) analisar os resultados obtidos a partir das simulações;
- d) identificar o modelo mais adequado à série temporal de demanda da microcervejaria.

4 REVISÃO BIBLIOGRÁFICA

O presente capítulo tem por objetivo apresentar uma revisão bibliográfica referente ao tema da pesquisa. Para tanto, foram selecionados periódicos de referência, a fim de buscar informações de qualidade e relevância no âmbito acadêmico. Com o propósito de proporcionar clareza na exposição das informações, o capítulo foi dividido em quatro seções.

4.1 PREVISÃO DE DEMANDA

De acordo com a definição de Chase, Jacobs e Aquilano (2006, p. 382), previsão de demanda é "a projeção de vendas futuras com base em análises e comparações de dados históricos, tendências, variações sazonais e outras informações relevantes". A previsão de demanda é essencial para o sucesso das organizações de diversos portes e áreas de atuação. Planos de negócios frequentemente dependem de previsões de demanda para determinar a capacidade de produção, a quantidade de matéria-prima necessária, a força de trabalho e os recursos financeiros disponíveis. Trata-se de um processo crítico que permite que as empresas identifiquem tendências, antecipem mudanças no mercado e tomem decisões estratégicas, informadas para garantir a eficiência e a eficácia do seu negócio (Petropoulos *et al.*, 2018).

Segundo Ackermann e Sellitto (2022), existem três grupos de métodos de previsão de demanda atualmente: qualitativos, quantitativos e de inteligência artificial. Os métodos qualitativos se baseiam em informações subjetivas, como opiniões de especialistas, análise de tendências e pesquisas de opinião ou intenção. Os métodos quantitativos, por sua vez, são baseados em análises estatísticas de séries históricas de dados. Por fim, os métodos de inteligência artificial utilizam algoritmos para analisar grandes quantidades de dados e gerar previsões mais precisas. Como o presente estudo tem por objetivo a utilização de métodos quantitativos e de inteligência artificial, na sequência serão descritos esses dois grupos.

Com relação a pesquisas recentes sobre este tema no mercado de bebidas, Mircetic *et al.* (2016) identificou que modelos precisos de previsão de demanda são

cruciais para reduzir custos logísticos. No estudo conduzido, o modelo S-ARIMA (Autoregressive Integrated Moving Average sazonal) foi empregado como uma ferramenta precisa para antecipar os padrões de demanda de uma empresa de bebidas. O S-ARIMA mostrou-se eficaz como modelo para prever a demanda semanal de bebidas.

4.2 MÉTODOS QUANTITATIVOS DE PREVISÃO DE DEMANDA

Os métodos quantitativos de previsão de demanda, também chamados de "métodos matemáticos", empregam modelos matemáticos para estimar valores previstos (Ren *et al.*, 2016). Segundo Lemos (2006), estes métodos se caracterizam por apresentar processos bem definidos para análise dos dados, possibilitando a replicação da metodologia por diferentes especialistas e a obtenção de previsões idênticas. Estas previsões podem ser feitas através de métodos causais, ou séries temporais (Thomas, 1996), que são detalhados na sequência.

4.2.1. Métodos causais

No método causal, a previsão de demanda está altamente correlacionada com fatores específicos no ambiente, como a economia e a taxa de juros de um país. Em modelos causais os dados exibem um comportamento identificável ao longo do tempo e existem relações identificáveis entre as informações que se deseja prever e outros fatores (Arvan *et al.*, 2019). O método causal é conhecido como regressão linear, onde a relação entre uma variável dependente e uma ou mais variáveis independentes é frequentemente modelada por meio de uma equação linear que tem como objetivo prever a variação na variável dependente (Evans e Olson, 2000). A equação em questão, comumente utilizada para representar a regressão simples é expressa matematicamente como:

$$Y = b_0 + b_1X_1 + e \quad (1)$$

Onde:

Y = variável dependente

b_0 = intersecção da linha no eixo X

b_1 = inclinação da linha

X_1 = variável independente

e = variável de erro

Em modelos de regressão simples, a variável dependente é descrita como uma função de uma única variável independente, estabelecendo uma relação linear. Contudo, é crucial ressaltar que a regressão linear é primordialmente voltada para interpolação, isto é, prever dentro do alcance dos dados coletados, e não para extrapolação, que envolve previsões além desse alcance. Na regressão múltipla, a relação entre a variável dependente e as variáveis independentes é representada por uma equação linear com várias variáveis, o que possibilita avaliar o impacto de cada variável independente na variável dependente (Evans e Olson, 2000).

4.2.2 Séries temporais

As séries temporais são conjuntos de observações geradas sequencialmente ao longo do tempo, que analisam o padrão de comportamento passado de um fenômeno com o objetivo de prever seu comportamento futuro. Esses métodos são frequentemente empregados para fazer previsões de curto prazo quando as observações seguem um padrão de comportamento identificável ao longo do tempo (Ackermann, 2022). Segundo Ballou (2001), esse método exerce a capacidade de identificar pontos de inflexão e pode ser considerado uma ferramenta útil para prever comportamentos em períodos de até doze meses.

Deste modo, os métodos quantitativos por séries temporais baseiam-se na premissa de que o padrão de comportamento observado no passado continuará a se manifestar no futuro. Esses métodos utilizam o padrão conhecido para prever demandas desconhecidas, ou seja, demandas futuras (Ackermann, 2022). É possível decompor as séries temporais em quatro componentes (Deb *et al.* 2017), conforme descrito a seguir:

- a) horizontal ou variações irregulares: variações aleatórias e de curto prazo cujas causas não são identificadas. Não podem ser previstas por modelos estocásticos de previsão;

- b) efeito de tendência: aumento ou diminuição da demanda com o tempo. A demanda pode também permanecer estacionária, com valores próximos a uma média constante;
- c) efeito sazonal: se refere a variações previsíveis na demanda de determinados produtos ou serviços em épocas específicas do ano. Essas variações são geralmente causadas por fatores sazonais, como mudanças climáticas, feriados, festivais ou eventos específicos de determinada época do ano;
- d) ciclo de negócios: flutuação geral e de periodicidade variável resultante de várias causas diferentes. São difíceis de prever e são observadas em economias capitalistas modernas.

Através do método de decomposição do tempo, a previsão de demanda é calculada por meio da seguinte fórmula:

$$Y = (T) \times (S) \times (C) \times (i) \quad (2)$$

Onde:

Y = valor da série ou demanda prevista;

T = componente tendencial;

S = componente sazonal;

C = componente cíclico;

i = efeitos aleatórios ou erros causados por flutuação.

Os valores de 'T' são obtidos a partir da análise de regressão simples entre a demanda e o tempo, o que permite ajustar uma linha de tendência aos valores reais (Ackermann, 2022). Possibilita-se ser utilizada média móvel ou método dos mínimos quadrados conforme fórmula a seguir:

$$T = a + bt \quad (3)$$

Onde 'T' representa a tendência, 't' é tempo e 'a' e 'b' são coeficientes a serem determinados para uma série de tempo. Sendo assim, os coeficientes são calculados a partir dessa fórmula:

$$a = T - bt \quad (4)$$

e

$$b = \frac{\sum D_t(t) - N(\bar{D})(t)}{\sum \bar{t}^2 - Nt^2} \quad (5)$$

Onde:

N = número de observações usadas no desenvolvimento da linha de tendência

D_t = demanda real do período de tempo 't';

\bar{D} = demanda média para N períodos de tempo;

\bar{t} = média de 't' sobre N períodos de tempo;

O índice de sazonalidade em uma série temporal é uma medida que indica o grau de variação cíclica de uma série de dados ao longo do tempo, com base em uma frequência sazonal pré-determinada. Ele é calculado pela média aritmética dos valores da série em cada período sazonal, dividida pela média aritmética de todos os valores da série. O índice de sazonalidade é geralmente expresso em porcentagem e é usado para ajustar a série temporal para remover o efeito da sazonalidade e analisar as tendências e variações irregulares na série. Calcula-se esse índice conforme a fórmula (Ackermann, 2022):

$$S_t = (D_t / T_t) \quad (6)$$

Onde:

S_t = índice de sazonalidade para o período 't';

D_t = demanda real do período de tempo 't';

T_t = valor da tendência já calculada.

A fórmula abaixo pode ser utilizada para calcular a previsão de um período de tempo 't' no futuro, de forma simplificada, sem considerar variações cíclicas e residuais:

$$Y_t = (D_t) \times (S_{t-L}) \quad (7)$$

Onde:

Y_t = Demanda prevista no período 't';

L = número dos períodos do ciclo sazonal.

4.2.3 Médias Móveis

Sobre os métodos de previsão com médias móveis, segundo Ackermann (2022) o modelo de média móvel simples utiliza a média aritmética dos 'n' valores passados da demanda para prever o período futuro imediato. A técnica é apropriada quando a demanda é estacionária e varia em torno de um valor médio (LI *et al.*, 2017). A suavização pode ser feita usando diferentes períodos anteriores, no entanto, o método de média móvel simples é lento para reagir a mudanças na demanda e não é eficaz para capturar variações sazonais, podendo até mesmo ocultá-las, dependendo do valor escolhido para 'n' (Fortsch e Khapalova, 2016).

Já o modelo da média móvel ponderada é um modelo semelhante ao modelo de média móvel simples, pois ambos levam em conta 'n' valores reais anteriores da demanda para compor a média (Ackermann, 2022). No entanto, no modelo de média móvel ponderada, são atribuídos pesos distintos para cada um dos valores, a fim de dar maior importância aos dados mais recentes gerados da demanda (Van Der Laan *et al.*, 2016).

A utilização de pesos distintos possibilita que os valores mais recentes da demanda tenham maior influência na previsão, o que pode ajudar a revelar alguma tendência de variação na demanda. No entanto, é importante destacar que quanto maior for o valor de 'n', mais suavizada será a previsão e menor será a capacidade do modelo de responder rapidamente a variações na demanda (Evans e Olson, 2000).

4.2.4 Ponderação exponencial simples

No contexto de séries temporais, o modelo da média móvel exponencialmente ponderada de 1ª ordem é uma técnica utilizada para previsão de dados futuros. Esse modelo é capaz de suavizar os picos de dados e considerar as características de comportamento da série temporal ao atribuir pesos progressivamente menores aos dados passados. Além disso, o cálculo da previsão é simples, exigindo apenas a utilização dos dados mais recentes (Ackermann, 2022). Segundo Lucas e Zhang (2016), esse modelo permite estender a previsão para vários períodos à frente, com adaptações possíveis, e representa-se pela equação pela seguinte equação para o período 't':

$$D_t = D_{t-1} + \alpha(Y_{t-1} - D_{t-1}) \quad (8)$$

Onde:

D_t = previsão para o período 't';

D_{t-1} = previsão para o período (t-1);

α = constante de suavização;

Y_{t-1} = demanda real para o período (t-1).

Para a utilização dessa previsão para qualquer período futuro, necessita-se ter a informação da demanda do período imediatamente anterior. Outro aspecto relevante é a definição da constante de suavização, a qual determina o peso dos valores passados na previsão (Ackermann, 2022). Essa constante deve ser escolhida entre 0 e 1, sendo que valores mais próximos de 1 dão maior peso aos valores anteriores e tornam a previsão mais sensível a variações bruscas na demanda, enquanto valores mais próximos de 0 reduzem a influência dos dados passados e tornam a previsão mais suave. Portanto, a escolha da constante de suavização deve levar em conta o comportamento da série temporal e o objetivo da previsão.

4.2.5 ARIMA

O modelo ARIMA (Autoregressive Integrated Moving Average) é um dos modelos mais utilizados em séries temporais. É uma combinação de três modelos diferentes: Autoregressivo (AR), Integração (I) e Média Móvel (MA). O modelo ARIMA foi proposto por Box e Jenkins (1976) e tem sido amplamente utilizado em previsão de dados em diversas áreas de negócio.

Segundo Souza (2006), esse método funciona adequadamente para análise de séries temporais com comprimento médio a longo, sendo recomendado um mínimo de 50 observações e preferencialmente 100. Um dos principais processos da metodologia é a transformação de uma série não-estacionária em uma estacionária, o que simplifica sua análise e proporciona maior estabilidade aos parâmetros estimados.

Além disso, ele baseia-se na premissa de que os valores de uma série temporal estão altamente correlacionados entre si, de modo que cada valor pode ser explicado por valores anteriores da série (Brentan *et al.*, 2017). A abordagem Box-Jenkins assume que o padrão de variabilidade nos dados é constante e usa um procedimento iterativo para ajustar um modelo de previsão com base em padrões aleatórios e cíclicos de demanda, minimizando assim os erros de previsão (Fortsch e Khapalova, 2016).

4.3 MÉTODOS DE INTELIGÊNCIA ARTIFICIAL DE PREVISÃO DE DEMANDA

Segundo Ackermann (2022), os métodos baseados em computação são capazes de imitar algumas das capacidades de processamento do cérebro humano na modelagem de situações simples e complexas. Esses modelos, que fazem uso da inteligência artificial, possuem a habilidade de identificar relacionamentos não lineares e interativos. Além disso, segundo a autora, nos últimos dez anos, houve um aumento significativo na atenção dada pelos pesquisadores às redes neurais artificiais, que têm sido reconhecidas como uma ferramenta computacional poderosa na resolução de problemas complexos.

Dentre as técnicas de redes neurais, destacam-se as redes neurais recorrentes (RNNs) com memória de longo prazo (LSTM) e o Extreme Gradient

Boosting (XG Boost) como métodos eficientes para a previsão de demanda. As RNNs com LSTM foram introduzidas por Hochreiter e Schmidhuber (1997) e têm sido amplamente aplicadas em previsão de séries temporais. De acordo com Soares Júnior (2020), as RNNs com LSTM têm vantagens significativas em relação a outros métodos, uma vez que, de acordo com as análises realizadas em seu estudo, verificou-se que os modelos baseados em LSTM apresentaram resultados superiores aos modelos ARIMA. Dessa maneira, os modelos demonstraram um bom desempenho em séries que apresentaram comportamentos mais irregulares, podendo ser atribuído à capacidade dos modelos LSTM em lidar com tais irregularidades. Os presentes resultados indicam que o uso de modelos baseados em LSTM pode ser uma alternativa promissora na previsão de séries temporais complexas.

Por outro lado, verifica-se o XG Boost, proposto por Chen e Guestrin (2016), um método de aprendizagem de máquina baseado no algoritmo *gradient boosted tree* (GBT), o qual consiste em criar diversos modelos de árvore de decisão considerados mais simples, com o intuito de combiná-los e obter um modelo mais poderoso e robusto. Esse resultado é obtido através do treinamento iterativo de um conjunto de árvores de decisão que busca minimizar uma função de perda estabelecida. A respeito de estudos recentes sobre o tema, Bessa (2021), concluiu-se que o modelo XG Boost com tratamento sazonal com ajuste da curva e com seleção de variáveis por RFA apresentou o maior nível de acurácia para previsões de vendas no setor de varejo de moda em relação a outros métodos.

Além disso, a combinação dos métodos LSTM e XG Boost tem sido objeto de pesquisa recente. Em particular, Semmelmann *et al.* (2022) propuseram um modelo híbrido de LSTM-XG Boost para a previsão de demanda no setor de energia, obtendo resultados superiores em relação aos métodos separados.

4.4 MEDIDAS DE ACURÁCIA

As medidas de acurácia têm o propósito de avaliar a eficiência de um método de previsão e também podem ser utilizadas como critério para escolher entre diferentes métodos de previsão (Lemos, 2006). Dessa forma, é extremamente importante avaliar a precisão das previsões para determinar a confiabilidade dos

resultados. Neste capítulo, serão apresentadas as principais medidas de erro de previsão, incluindo a medida de desvio padrão, percentual de erro absoluto médio (MAPE), média do quadrado de erros (MQE) e a média absoluta de erros (MAE).

4.4.1 Desvio padrão

O desvio padrão é uma medida de dispersão em torno da média. Sendo assim, o desvio padrão pode ser usado para avaliar o grau de variação entre as previsões e os valores reais. Quanto maior o desvio padrão, maior a variação, o que significa que as previsões são menos precisas. O desvio padrão é calculado usando a seguinte fórmula (Ballou, 2001):

$$E_D = \left(\frac{\sqrt{\sum_t (Y_t - D_t)^2}}{N - 1} \right) \quad (9)$$

Onde:

E_D = erro padrão da previsão;

Y_t = demanda real no período 't';

D_t = previsão para o período 't';

N = número total de observações.

4.4.2 Média do quadrado de erros (MQE)

A média do quadrado de erros (MQE) é uma medida de erro quadrático médio que é frequentemente usada em previsão. Ele mede a média dos erros quadráticos e é calculado usando a seguinte fórmula (LEMOS, 2006):

$$MQE = \frac{1}{n} \sum_{i=1}^n (Y_t - D_t)^2 \quad (10)$$

Onde:

MQE = média do quadrado de erros;

Y_t = demanda real no período t;

D_t = previsão para o período t;

n = número total de observações.

O MQE é uma medida de erro positiva, o que significa que quanto maior o valor, pior é a precisão da previsão (Montgomery et al., 2012).

4.4.3 Média absoluta de erros (MAE)

A média absoluta de erros (MAE) é uma medida de erro absoluto médio que é frequentemente usada em previsão. Ele mede a média dos erros absolutos e é calculado usando a seguinte fórmula (Lemos, 2006):

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_t - D_t| \quad (11)$$

Onde:

MAE = média absoluta de erros;

Y_t = demanda real no período t ;

D_t = previsão para o período t ;

n = número total de observações.

4.4.4 Percentual de erro absoluto médio (MAPE)

O percentual de erro absoluto médio (MAPE) é uma medida de erro relativo amplamente utilizada em previsão. Ele mede a precisão das previsões em termos percentuais e é calculado usando a seguinte fórmula (Lemos, 2006):

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{(Y_t - D_t)}{Y_t} \right| \quad (12)$$

Onde:

MAPE = percentual de erro absoluto médio;

Y_t = demanda real no período t ;

D_t = previsão para o período t ;

n = número total de observações.

O percentual de erro absoluto médio é expresso em porcentagem e quanto menor o valor, melhor é a precisão da previsão (Hyndman; Koehler, 2006). De acordo com Lemos (2006), a medida mais utilizada em estudos empíricos que comparam métodos é o MAPE. Isso se deve, em grande parte, à sua menor sensibilidade a valores extremos em comparação às medidas quadráticas. Além disso, o MAPE utiliza percentuais de erro e é independente da unidade de dados, permitindo melhor controle de escala.

É relevante mencionar que, apesar do MAPE ser a métrica mais indicada, a avaliação de acurácia ocorre em conjunto com outras medidas de erro. Importante ressaltar que modelos mais eficazes exibirão valores de erro menores.

4.5 TOMADA DE DECISÃO BASEADA EM DADOS

A tomada de decisão baseada em dados destaca-se como de relevância crítica no ambiente organizacional contemporâneo. Este fenômeno é particularmente notável no contexto do advento de tecnologias emergentes como Big Data e Indústria 4.0, bem como a crescente adoção da Administração Baseada em Evidências (ABE). Essas tendências têm redefinido a maneira como as organizações tomam decisões, enfatizando a importância do uso de dados e evidências empíricas para orientar ações estratégicas e operacionais.

De acordo com Martins e Maçada (2022), o Big Data Analytics tem se tornado uma ferramenta essencial na melhoria da qualidade da tomada de decisões nas organizações. Sendo assim, as empresas estão cada vez mais recorrendo a essa tecnologia para auxiliar na formulação de decisões robustas, que possam proporcionar vantagens competitivas e otimizar processos.

No entanto, apesar da crescente adoção do Big Data Analytics, os autores identificaram uma lacuna na literatura existente. De acordo com esses autores, ainda há uma compreensão limitada de como essa tecnologia contribui especificamente para aprimorar a qualidade das decisões tomadas. Isso sugere a necessidade de estudos adicionais que aprofundem esses conceitos, com foco específico na tomada de decisões baseada em dados.

Além disso, trazendo essa revisão para o contexto das teorias aplicadas na Administração, Dias (2022) investiga a Administração Baseada em Evidências (ABE) como um procedimento que pode potencialmente elevar a qualidade do processo decisório. O autor questiona os tipos de evidências utilizados pelos gestores para embasar suas decisões e avalia se os resultados obtidos estão alinhados com as expectativas. A pesquisa conduzida por Dias (2022), conclui que a ABE, quando implementada de forma adequada, emerge como uma abordagem que aumenta a probabilidade de sucesso nas decisões tomadas pelos gestores, fornecendo embasamento sólido e confiável.

Esta perspectiva ressalta a importância de adotar uma abordagem baseada em evidências para a tomada de decisão, destacando a necessidade de utilizar informações sólidas e confiáveis como base para as escolhas organizacionais. Através da aplicação da ABE, os gestores são capazes de embasar suas decisões em dados concretos e evidências empíricas, reduzindo assim a probabilidade de erros e aumentando as chances de sucesso. Portanto, a implementação efetiva da ABE pode contribuir significativamente para o aprimoramento da qualidade e a eficácia das decisões tomadas pelas organizações.

Em um contexto mais específico de aplicação do tema na indústria, Bousdekis (2021), destaca-se a importância da tomada de decisões baseada em dados na Indústria 4.0, especialmente nas operações de fabricação e manutenção. A infraestrutura avançada de sensores na Indústria 4.0 permite a implementação de algoritmos que analisam dados em tempo real, desta forma, é possível identificar padrões, além de prever situações emergentes e recomendar ações de mitigação. Essa abordagem baseada em dados pode melhorar potencialmente a eficácia e a eficiência das decisões de indústria.

Dessa forma, a tomada de decisão baseada em dados, como foi exposto nesta revisão teórica, é um campo de estudo em constante e crescente evolução que oferece oportunidades significativas para melhorar a qualidade das decisões organizacionais. No entanto, como foi exposto, é necessário um maior entendimento e aplicação efetiva destas técnicas nos seus campos de pesquisa.

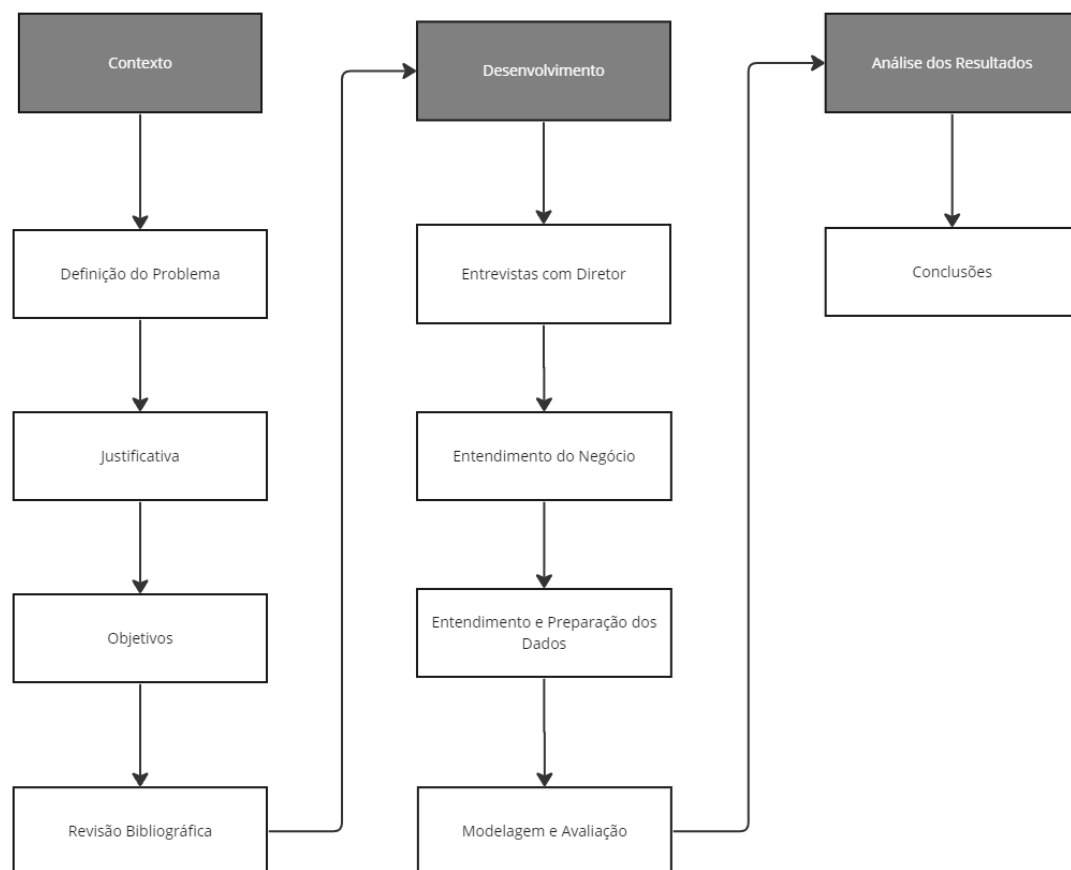
Após a revisão teórica concluída, a qual atua como base para a pesquisa, o próximo capítulo tem como objetivo definir a metodologia e apresentar o desenvolvimento da pesquisa.

5 PROCEDIMENTOS METODOLÓGICOS

Este capítulo objetiva apresentar a metodologia que foi utilizada para conduzir a pesquisa. Para garantir a validade e a precisão dos resultados, foram utilizados métodos e técnicas comprovados, amplamente reconhecidos no campo acadêmico.

O presente estudo adotou uma metodologia baseada no Cross Industry Standard Process for Data Mining (CRISP-DM), que consiste em um processo iterativo de seis etapas: Compreensão do Negócio, Entendimento dos Dados, Preparação dos Dados, Modelagem, Avaliação e Implantação. A metodologia CRISP-DM tem sido extensamente utilizada em projetos de Mineração de Dados, sendo considerada uma das melhores práticas para se desenvolver um projeto de Ciência de Dados (Gil *et al.*, 2021). A Figura 1 representa as etapas da metodologia adaptada neste estudo.

Figura 1 – Etapas da Metodologia



Fonte: elaborada pelo autor.

Esta metodologia permite uma abordagem sistemática e estruturada para lidar com problemas complexos, segundo Moro et al. (2011) e Wirth (2000), garantindo a replicação de projetos, melhoria na gestão, utilização de melhores práticas e maximização dos resultados.

As etapas do CRISP-DM foram adaptadas para atender às necessidades específicas do estudo. Inicialmente, foi realizada uma fase de compreensão do negócio, na qual foram identificados os objetivos da previsão de demanda na microcervejaria, bem como os requisitos e restrições do projeto. Durante esta fase, foi realizada uma entrevista semi-estruturada com os diretores da empresa. Segundo Bardin (1977), as entrevistas têm o objetivo de apurar informações, capturando o universo das percepções, emoções e interpretações dos informantes em seu contexto. O roteiro de entrevista, conforme anexo A, foi desenvolvido com base nos autores Makridakis et. al.(1998), Armstrong (2001), Chopra et al. (2016), Hyndman et al. (2018), Davydenko (2013), Svoronos (1997), Lee et al. (2004) e Syntetos et al. (2016), a entrevista visou compreender profundamente o contexto organizacional, a metodologia adotada pela empresa e os desafios que enfrenta no planejamento da demanda. Em seguida, a etapa de Entendimento dos Dados foi conduzida, envolvendo a coleta dos dados históricos da microcervejaria. Estes dados foram analisados para obter insights sobre as características e padrões presentes na série temporal (Wirth, 2000).

Após o entendimento dos dados, a etapa de Preparação dos Dados foi realizada, para a qual fez-se necessária fazer o tratamento dos dados e adequação das variáveis categóricas. A partir disto, foi possível garantir a qualidade e consistência dos dados utilizados na modelagem (Wirth, 2000).

A etapa de Modelagem é a principal do presente estudo, onde foram aplicados algoritmos de previsão de séries temporais utilizando a biblioteca Pycaret. Diferentes modelos foram explorados e testados com base na série temporal, a fim de encontrar o melhor desempenho preditivo para a microcervejaria (Wirth, 2000).

Em seguida, foi desenvolvida a Avaliação dos modelos, utilizando métricas de acurácia apropriadas para medir a qualidade das previsões. Essas métricas ajudaram a comparar e selecionar o modelo mais adequado para a microcervejaria, levando em consideração a etapa de Entendimento do Negócio (Wirth, 2000). Por

fim, a fase de Implantação foi adaptada para um contexto de pesquisa, onde se deu ênfase na extração de conhecimento a partir dos resultados obtidos. Nesta fase, que foi definida como Análise dos Resultados, foi feita a segunda entrevista com os diretores da empresa, que seguiu o roteiro de entrevistas semi-estruturada no Anexo B do presente estudo, e teve como objetivo avaliar com o corpo de diretores sob uma perspectiva mais prática a aplicabilidade do modelo. Os insights e conclusões derivados do estudo foram utilizados para melhorar o entendimento do comportamento da demanda da microcervejaria e selecionar o modelo mais eficiente para a empresa.

Para a realização das etapas do CRISP-DM foi necessário escolher uma ferramenta de trabalho, mais especificamente um ambiente de desenvolvimento integrado,(Integrated Development Environment, IDE) e a linguagem de programação que permita o desenvolvimento adequado do método. Desta forma, o IDE escolhido foi o Google Collab e a linguagem de programação Python, para os testes e predição dos dados, foi utilizado o módulo de Séries Temporais da biblioteca PyCaret, que já conta com modelos de predição disponíveis. Além disso, foi utilizada a biblioteca Pandas para fazer tratamento e manipulação dos dados. Esta abordagem em questão é inovadora ao aplicar estas técnicas ao contexto de microcervejarias.

5.1 ENTENDIMENTO DO NEGÓCIO

Como exposto no capítulo de objetivos, este estudo tem como objetivo compreender e analisar os modelos de previsão de séries temporais mais adequados para a realidade de uma microcervejaria brasileira localizada no Rio Grande do Sul. A empresa, que foi fundada em 1998, no último ano obteve uma produção de 1.254 mil litros de cerveja, resultando em um faturamento de aproximadamente R\$14,3 milhões. Sua operação está concentrada no âmbito estadual, mais especificamente na região metropolitana de Porto Alegre e no Vale do Rio dos Sinos. Quanto à sua estrutura organizacional, a empresa tem três sócios e conta com 19 colaboradores, sendo que seis atuam na área comercial, sete na operação, três na logística e três na área administrativa.

Inicialmente, foram realizadas entrevistas com o diretor de produção e o diretor comercial da fábrica. O objetivo da entrevista era compreender os objetivos que deveriam ser alcançados nos modelos que seriam desenvolvidos posteriormente, assim como entender o contexto da empresa e como atualmente a empresa desenvolve o processo de previsão de demanda. Nessas entrevistas foi possível entender a metodologia que atualmente emprega-se na empresa, além disso, foram explicadas todas as características da base de dados que seria utilizada no estudo.

Também foi relatado que devido à estrutura organizacional enxuta, a empresa enfrenta dificuldades em tomar decisões estratégicas de longo prazo com base nos dados internos disponíveis. Atualmente, o planejamento de produção é realizado por meio de uma análise histórica simplificada, utilizando o crescimento médio mensal das demandas em relação aos dois anos anteriores registradas no sistema, junto à projeção de crescimento que a empresa espera ter, após este cálculo, leva-se em consideração a sazonalidade da demanda por trimestre ao longo do ano. Portanto, a empresa está sujeita a enfrentar períodos em que a demanda por seus produtos excede a capacidade de produção, resultando em perda de vendas e receita, ou períodos em que a demanda é, significativamente, menor do que a produção, resultando em custos operacionais desnecessários.

Além disso, foi constatado que a empresa adota apenas um processo anual de previsão de demanda, no qual é planejado todo o ano subsequente, e ao longo do ano, conforme os diretores observam a necessidade, é feita uma revisão mensal. Esse processo envolve o cálculo da projeção das vendas, conforme explicado anteriormente, e com base nisso, são estabelecidas as metas para a área comercial. O planejamento da produção e a projeção de fluxo de caixa na área financeira, são, dessa maneira, elaborados. Contudo, é importante destacar que essa metodologia limita consideravelmente a flexibilidade da empresa para se adaptar às variações de demanda que ocorrem ao longo do ano. Além disso, o crescimento da empresa está fortemente vinculado ao comportamento registrado no passado, o que não necessariamente reflete as tendências atuais presentes nos dados.

Diante do exposto, o presente estudo concentrou-se em definir o modelo de previsão de demanda mais adequado para a realidade da empresa estudada, a fim

de melhorar a eficiência dos processos atuais. Os dados utilizados para o estudo foram extraídos do sistema de gestão atualmente empregado pela empresa, implantado em 2014.

6 RESULTADOS

Neste capítulo, serão apresentados os resultados obtidos pela pesquisa, expondo o desenvolvimento de cada etapa do estudo. A primeira etapa de desenvolvimento foi a de Entendimento e Preparação dos Dados, onde explana-se como os dados foram coletados, processados e preparados para análise, garantindo sua confiabilidade (Wirth, 2000). Em seguida, a etapa de Modelagem e Avaliação, demonstrando as metodologias empregadas, os critérios de avaliação adotados e os desafios enfrentados durante a modelagem (Wirth, 2000). Por fim, na seção de Análise dos Resultados, é feita a discussão das descobertas principais, interpretando seu significado e relevância no contexto da empresa estudada. Ao longo deste capítulo, o objetivo é oferecer uma visão abrangente e detalhada dos resultados da pesquisa.

6.1 ENTENDIMENTO E PREPARAÇÃO DOS DADOS

Nesta etapa do estudo, a exploração dos dados foi realizada para gerar insights e preparar os dados para a modelagem na próxima fase. Com base na entrevista realizada, o estudo concentrou-se nas vendas, em litros, dos três tipos de cerveja vendidos pela empresa - Pilsen, Cream Lager e Especiais - de Janeiro de 2015 a Junho de 2023, somando sete anos e meio, ou 102 períodos. O conjunto de dados analisado contém quatro atributos, que serão mostrados no quadro a seguir:

Quadro 1 – Atributos do Dataset

Atributo	Descrição
Data de Venda	Data que foi realizada a venda
Estilo	Estilo do produto
Produto	Descrição do produto
Litros Vendidos	Quantidade vendida em litros

Fonte: elaborado pelo autor.

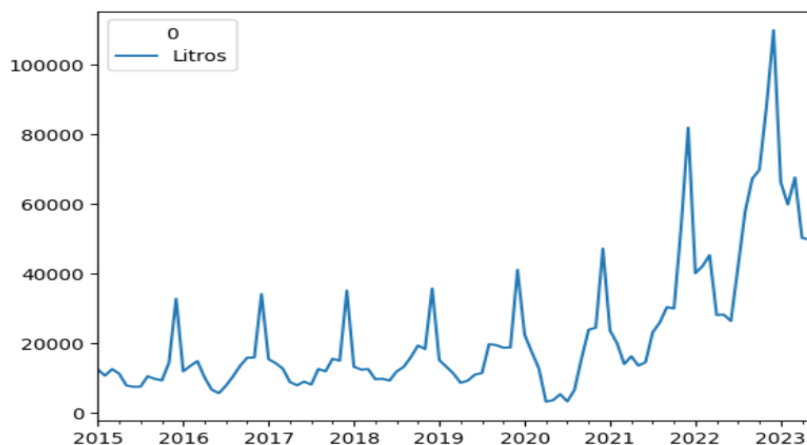
É importante ressaltar que os dados foram extraídos de forma bruta da base de dados, ou seja, cada venda que foi realizada na empresa estava discriminada. Nesse sentido, foi de extrema relevância para a correta análise dos dados realizar a preparação e transformação desses dados.

A preparação de dados é o processo de modificação dos dados brutos para que possam ser posteriormente processados e analisados, incluindo as tarefas de limpeza de dados, transformação de dados e redução de dados (Kochański, 2003). Sendo assim, a limpeza de dados consiste principalmente na substituição de valores ausentes e vazios, a eliminação de dados incorretos e a remoção de inconsistências. A transformação de dados implica em modificações que permitam que eles estejam em uma forma que possibilite sua exploração, como normalização ou agregação de dados.

A primeira etapa do processo de transformação envolveu a análise da evolução das vendas. Durante esta etapa, observou-se que alguns estilos vendidos anteriormente não são vendidos atualmente. Sendo assim, os dados referentes ao estilo de cerveja foram filtrados para evitar influências na análise da série temporal, uma vez que esse produto não é mais fabricado pela empresa.

Após este tratamento, foi feito o agrupamento dos dados por “Mês” e “Litros Vendidos”, uma técnica que envolve a organização dos dados em grupos homogêneos, fundamentada em características semelhantes (Jain, Murty, Flynn, 1999). A partir disso, obteve-se o gráfico de evolução dos “Litros Vendidos” por “Mês”. A figura a seguir mostrará o resultado deste gráfico:

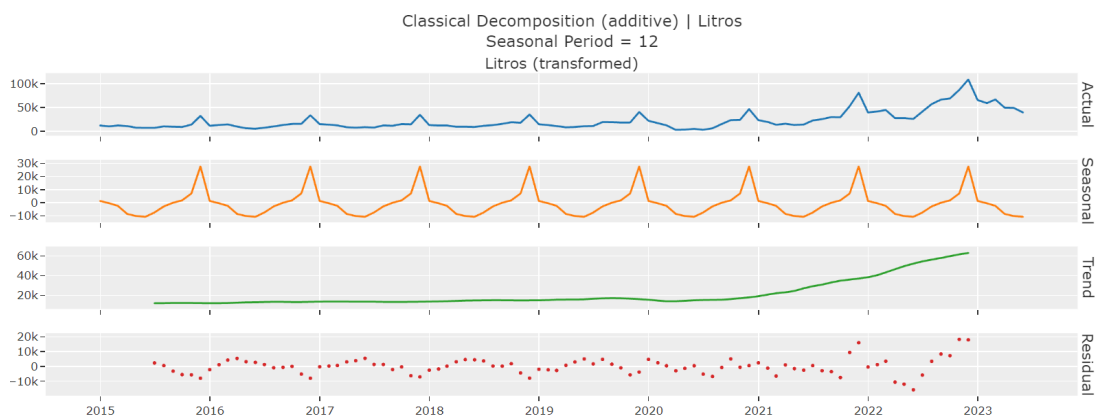
Figura 2 – Série temporal de litros vendidos por mês



Fonte: elaborada pelo autor.

Além disso, para analisar o comportamento das séries temporais, foi realizada sua decomposição em seus componentes. A decomposição de uma série temporal, como exposto no capítulo de revisão teórica, permite a análise dos elementos subjacentes como tendência, sazonalidade e resíduo. A tendência revela o padrão geral e a direção a longo prazo dos dados. A sazonalidade reflete as flutuações periódicas, e o resíduo representa as variações irregulares que não podem ser atribuídas à tendência ou sazonalidade. As imagens a seguir mostrarão o resultado dessas análises, ilustrando cada componente e proporcionando *insights* adicionais sobre a natureza e estrutura dos dados. Essa análise auxiliou na modelagem preditiva identificando a sazonalidade de 12 meses presente na série temporal assim como a tendência nos dados.

Figura 3 – Decomposição da série temporal



Fonte: elaborada pelo autor.

Ao examinar a decomposição da série temporal, permite-se observar uma mudança recente na tendência da demanda da empresa. Iniciando em 2021, ocorreu um aumento expressivo na demanda, modificando o padrão anterior da Série Temporal. No entanto, uma observação importante é que, apesar dessa alteração notável na tendência, o aspecto da sazonalidade se manteve estável durante todos os anos analisados. Dessa forma, mesmo com as mudanças decorrentes, a empresa registrou um pico de demanda no mês de dezembro, no período analisado. Essa constância, em meio à mudança de tendência, oferece informações valiosas que podem ser muito úteis para entender o comportamento do mercado e para planejar estratégias futuras no negócio. A respeito do ruído, é possível observar que ao mesmo passo que a tendência de demanda sofreu uma forte alteração em seu comportamento, os resíduos começaram a aumentar a sua amplitude, justamente pela alteração no comportamento da Série Temporal.

Além da análise da decomposição dos componentes da Série Temporal, foi utilizado a função *check_stats* da biblioteca Pycaret para obter informações estatísticas do conjunto de dados. A seguir será apresentada no quadro com o resultado dos testes estatísticos:

Quadro 2 – Estatísticas do conjunto de dados

Teste	Propriedade	Valor
Resumo	Comprimento	102,00
Resumo	# Valores Ausentes	0,0
Resumo	Média	23.163,45
Resumo	Mediana	14.908,00
Resumo	Desvio Padrão	20.096,00
Ruído Branco	Ljung-Box (K=24)	377.49
Ruído Branco	Ljung-Box (K=48)	402.68
Ruído Branco	p-value (K=24)	0,0
Ruído Branco	p-value (K=48)	0,0
Estacionaridade	p-value ADF	0,9986
Estacionaridade	p-value KPSS	0,01
Normalidade	p-value Shapiro	0,0

Fonte: elaborado pelo autor.

A partir da análise dos resultados, foi possível identificar *insights* fundamentais que guiaram o desenvolvimento das etapas subsequentes do método CRISP-DM.

Analisando o Quadro 2, o conjunto de dados possui um comprimento de 102 dados, o que indica que estamos lidando com uma quantidade moderada de dados, ficando dentro do ideal conforme apontado no capítulo de revisão teórica. É possível observar também que não há valores ausentes, sendo assim, o conjunto de dados é completo e não requer limpeza ou preenchimento de lacunas.

A média do conjunto de dados é de 23.163,45 e a mediana, que é o valor do meio quando os dados são ordenados, é de 14.908,00. A diferença significativa entre a média e a mediana sugere a presença de valores extremos que estão elevando a média.

Além disso, o desvio padrão é de 20.096,00. Um desvio padrão alto, como neste caso, indica que os valores estão demasiadamente espalhados em relação à média. Isto sugere uma grande volatilidade nos dados, com alguns valores muito

acima ou abaixo da média, o que pode ser causado pela natureza da sazonalidade existente na Série Temporal.

A ausência de normalidade nos dados, indica que os métodos estatísticos tradicionais que assumem uma distribuição normal podem não ser aplicáveis. Isso sugere a necessidade de utilizar técnicas alternativas que não façam essa suposição, ou de aplicar transformações específicas nos dados para torná-los mais adequados para a modelagem. Além disso, a rejeição da hipótese de ruído branco revela a existência de autocorrelação na Série Temporal, um sinal claro de que há padrões nos dados que podem ser explorados, através da modelagem. Essa descoberta é complementada pela observação da falta de estacionariedade, que aponta para a necessidade de modelos capazes de capturar tendências e sazonalidades. Considerando esses resultados em conjunto, é possível direcionar a seleção de técnicas de análise e modelagem de forma mais informada, permitindo uma abordagem mais adaptada e eficaz para a série histórica em questão.

6.2 MODELAGEM E AVALIAÇÃO

Esta seção tem como objetivo apresentar os algoritmos testados no conjunto de dados da série temporal, assim como avaliar os resultados obtidos a partir de medidas de acurácia. Como exposto no capítulo 4, existem diversos métodos para se desenvolver a previsão de uma série temporal, por este motivo a biblioteca PyCaret foi escolhida para executar a etapa de modelagem e avaliação. Esta escolha, se justifica pela capacidade da biblioteca de comparar diferentes métodos, permitindo uma análise com resultados bem fundamentados.

Para o desenvolvimento dos modelos e avaliação do método mais adequado, utilizou-se o módulo Séries Temporais presente na biblioteca Pycaret. Para criar o ambiente de treinamento, demandou-se utilizar a função *setup*, a qual requer quatro parâmetros essenciais. O primeiro parâmetro é a Série Temporal, que é o resultado da etapa anterior do estudo. O segundo parâmetro é o horizonte de previsão, que define a quantidade de dias escolhida para a realização dos testes, que serão comparados com os valores reais. O terceiro parâmetro é a variável a ser prevista, que no caso deste estudo é “Litros Vendidos”. Por fim, o quarto parâmetro é o de *folds*, que define a quantidade de subconjuntos de dados para se realizar os testes

cruzados, isso permite que o modelo seja treinado e testado em várias divisões diferentes dos dados, fornecendo uma avaliação mais robusta do seu desempenho.

A partir da definição de tais parâmetros foi utilizada a função *compare_models*, que treina e avalia o desempenho de todos os modelos disponíveis na biblioteca. A Figura 4, a seguir, demonstra os valores das medidas de acurácia obtidos para cada modelo testado:

Figura 4 – Valores das métricas resultantes da predição da série temporal

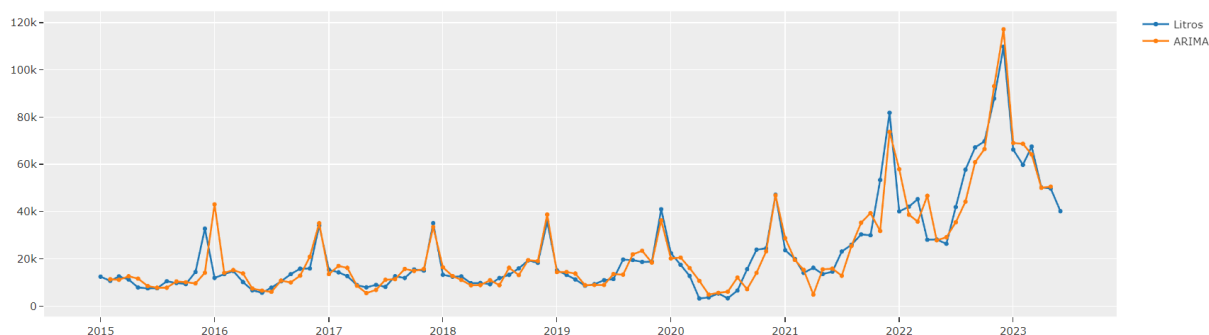
Modelo	MASE	MAE	MAPE
ARIMA	0,9887	6.202,8497	0,1072
Auto ARIMA	1,0433	6.613,5672	0,1251
STLF	1,2425	7.999,6326	0,1406
Orthogonal Matching Pursuit w/ Cond, Deseasonalize & Detrending	1,3314	8.778,9962	0,1469
Huber w/ Cond, Deseasonalize & Detrending	1,4825	9.641,8668	0,1631
Bayesian Ridge w/ Cond, Deseasonalize & Detrending	1,6165	10.456,0698	0,1784
ETS	1,6309	11.190,2787	0,1574
Theta Forecaster	1,6704	11.593,2990	0,1602
Extreme Gradient Boosting w/ Cond, Deseasonalize & Detrending	1,6795	10.927,5244	0,1779
Lasso Least Angular Regressor w/ Cond, Deseasonalize & Detrending	1,8268	11.743,6528	0,2051
Ridge w/ Cond, Deseasonalize & Detrending	1,8268	11.743,6552	0,2051
Lasso w/ Cond, Deseasonalize & Detrending	1,8268	11.743,6527	0,2051
Elastic Net w/ Cond, Deseasonalize & Detrending	1,8268	11.743,6534	0,2051
Linear w/ Cond, Deseasonalize & Detrending	1,8268	11.743,6552	0,2051
Exponential Smoothing	1,9021	13.079,8136	0,1795
Extra Trees w/ Cond, Deseasonalize & Detrending	1,9105	12.541,7081	0,2041
Naive Forecaster	1,9736	13.396,8333	0,2109
AdaBoost w/ Cond, Deseasonalize & Detrending	2,0205	13.050,2856	0,2045
CatBoost Regressor w/ Cond, Deseasonalize & Detrending	2,0535	13.290,6875	0,1982
Random Forest w/ Cond, Deseasonalize & Detrending	2,1612	14.104,6039	0,2235
Gradient Boosting w/ Cond, Deseasonalize & Detrending	2,1654	14.392,8774	0,2413
K Neighbors w/ Cond, Deseasonalize & Detrending	2,1743	13.844,8354	0,2050
Decision Tree w/ Cond, Deseasonalize & Detrending	2,1928	15.238,2961	0,2432
Light Gradient Boosting w/ Cond, Deseasonalize & Detrending	2,8539	18.761,9775	0,2657
Croston	3,2500	21.477,6592	0,3042
Seasonal Naive Forecaster	3,9022	25.943,3875	0,4257
Polynomial Trend Forecaster	4,0015	26.789,1970	0,3736
Grand Means Forecaster	6,3062	42.820,6526	0,6447

Fonte: elaborada pelo autor.

Com base na análise dos resultados obtidos, o modelo ARIMA se destaca como o melhor modelo entre todos os listados, com o menor MASE de 0,9887, o que indica que ele tem o menor erro médio em comparação com a escala dos dados. Além disso, o MAPE do ARIMA é de 0,1072, o que significa que, em média, suas previsões estão cerca de 10,72% distantes dos valores reais em termos

percentuais, o que foi considerado no estudo uma medida de erro aceitável. Dessa forma, definiu-se que ele seria escolhido como o mais adequado para o desenvolvimento do estudo. Além disso, entre os modelos de *machine learning*, o *Extreme Gradient Boosting* destacou-se, indicando que técnicas mais avançadas poderiam também ser aplicadas com possibilidade de sucesso. Na Figura 4 os dados previstos pelo modelo ARIMA são representados pela cor laranja, junto aos dados reais em azul. Possibilita-se evidenciar que, apesar do conjunto de dados não ser tão extenso, o modelo se adequou com eficiência ao comportamento existente na série histórica. Dessa maneira, em diversos pontos do gráfico, pode-se observar que o modelo conseguiu prever valores muito próximos da realidade da empresa.

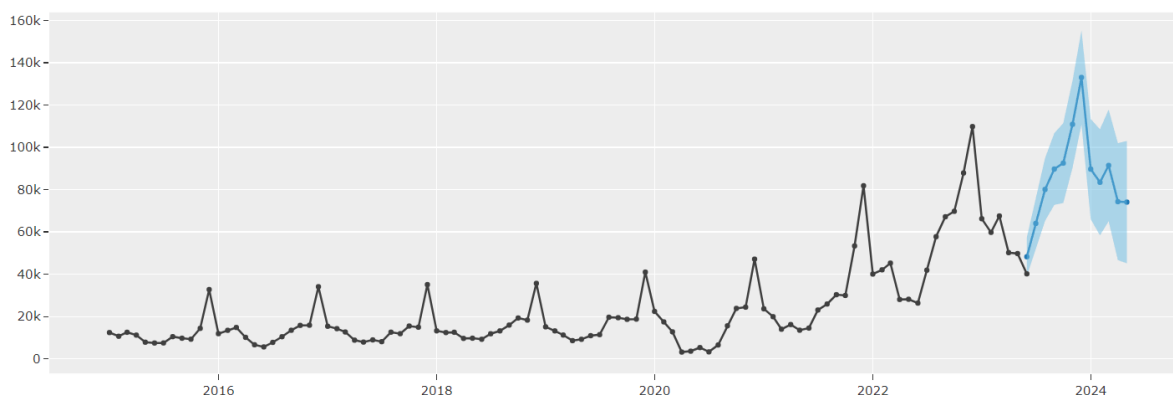
Figura 5 – Comparação entre os valores previstos pelo modelo ARIMA e os valores reais da Série Temporal



Fonte: elaborado pelo autor.

Sendo assim, após a análise dos testes, o modelo ARIMA foi criado e treinado com a função *create_model* e os próximos valores da série foram previstos. A Figura 6, a seguir, apresenta em preto os valores reais e em azul os resultados da previsão para um período de 12 meses.

Figura 6 – Valores reais e previsão do modelo ARIMA



Fonte: elaborada pelo autor.

Visualiza-se, nesse cenário, que as previsões obtidas pelo modelo seguem a tendência e a sazonalidade nos dados. Como retratado na literatura revisada, o modelo ARIMA é um dos mais utilizados para previsão de séries temporais com o comprimento de 50 a 100 observações. Além disso, um dos requisitos para a efetividade do modelo é a correlação existente entre os valores da série ao longo do tempo, o que foi confirmado pelo teste de ruído branco aplicado na fase de entendimento e preparação dos dados. Assim sendo, é possível destacar a relevância e acuracidade do modelo ARIMA no contexto desta pesquisa.

6.3 ANÁLISE DOS RESULTADOS

Os resultados obtidos a partir da modelagem e avaliação da série temporal, utilizando a biblioteca PyCaret, demonstram a eficácia do modelo ARIMA na previsão da demanda da microcervejaria. A análise dos valores das métricas de acurácia, como MASE e MAPE, revela que o modelo ARIMA não apenas se adequou ao comportamento da série histórica, mas também conseguiu prever valores muito próximos da realidade da empresa.

Em uma entrevista semiestruturada, que seguiu o roteiro presente no Anexo B deste estudo, com os diretores da empresa, foi possível obter uma perspectiva mais prática sobre os resultados. Os diretores expressaram satisfação com os resultados, destacando que o modelo consegue capturar bem a sazonalidade

existente no mercado. Esta observação é corroborada pela análise visual das Figuras 4 e 5, onde os valores previstos pelo modelo ARIMA se alinham estreitamente com os valores reais da série temporal. Além disso, os diretores confirmaram que os resultados estão dentro dos limites aceitáveis para a empresa. Esta afirmação é apoiada pelo MAPE do ARIMA, que indica que, em média, suas previsões estão cerca de 10,72% distantes dos valores reais. Esta margem de erro foi considerada aceitável no contexto do estudo.

De maneira destacada, os diretores veem potencial prático nos resultados, foi cogitada a hipótese de implementação do modelo como uma ferramenta para melhorar as revisões da previsão de demanda realizadas anualmente pela empresa. Com o auxílio deste modelo, a empresa pode ter um embasamento mais sólido ao realizar revisões, permitindo uma tomada de decisão mais informada e estratégica. O processo de revisão da previsão, dentro desse contexto, torna-se mais rápido e viável, visto que o modelo possibilita os diretores compararem os valores previstos do modelo com a previsão anual da empresa, desta forma, indicando se as ações devem ser tomadas para mitigar as perdas e otimizar a produção.

A combinação dos resultados técnicos obtidos e da opinião dos diretores da empresa reforça a relevância e acurácia do modelo ARIMA no contexto desta pesquisa. A capacidade do modelo de refletir com precisão a sazonalidade e tendência nos dados, juntamente com a validação prática dos diretores da empresa, destaca o potencial do ARIMA como uma ferramenta precisa para previsões de séries temporais no ambiente empresarial.

Além disso, é importante destacar o atingimento dos objetivos propostos na introdução do presente estudo, iniciando-se pela identificação na literatura dos métodos e técnicas de previsão de demanda – etapa cumprida a partir da discussão dos métodos de previsão da demanda ao longo da revisão bibliográfica, em que foram identificados e abordados os métodos de Regressão Linear, Séries Temporais, Médias Móveis, Ponderação Exponencial, ARIMA, LSTM e XG BOOST. Para cada modelo, foram apresentadas as principais vantagens e desvantagens, viabilizando a construção da base sólida de conhecimento do autor para o desenvolvimento das etapas seguintes do estudo.

O objetivo de elaboração das simulações com dados reais a partir da aplicação dos modelos selecionados cumpriu-se na etapa de modelagem e

avaliação por meio do desenvolvimento de simulações com a biblioteca Pycaret sobre a série temporal de demanda da empresa estudada. Com base nas simulações conduzidas, foi possível obter as medidas de acurácia para cada modelo simulado, viabilizando, também, o cumprimento do objetivo de análise dos resultados obtidos a partir das simulações.

Dado o atingimento dos três primeiros objetivos específicos propostos, fez-se possível, por fim, cumprir o objetivo de identificação do modelo mais adequado à série temporal de demanda da microcervejaria. De acordo com a análise dos resultados, constatou-se o modelo ARIMA, que apresentou MAPE de 10,72%, como o mais adequado para a série temporal da microcervejaria.

Finalmente, o objetivo geral da presente pesquisa foi atingido, uma vez que foi definido o modelo ARIMA como o modelo de previsão mais adequado para a empresa e foi proposta a utilização do mesmo para auxiliar a revisão da metodologia atualmente utilizada. Espera-se que, com a utilização do modelo proposto, a microcervejaria obtenha aumento de acuracidade em suas previsões e, conseqüentemente, aumento de eficiência na operação.

7 CONCLUSÃO

O presente trabalho apresentou o desenvolvimento de um modelo para prever a demanda mensal de uma microcervejaria situada no Rio Grande do Sul. Durante o desenvolvimento do TCC, foi analisada a fundamentação teórica em relação ao conceito dos métodos de previsão de demanda, analisando trabalhos e conceitos relacionados a esse tema, contribuindo com os estudos.

Em relação aos objetivos propostos, os resultados demonstraram que o modelo ARIMA foi aquele com maior acurácia, e também o melhor na opinião do gestor, respondendo, assim, a pergunta de pesquisa. Esse resultado foi obtido por meio dos dados disponibilizados pela empresa, a partir disso, foi criada a série temporal de demanda, a qual foram utilizados métodos preditivos da biblioteca Pycaret, dentro do Google Collab, e os resultados foram avaliados.

O modelo ARIMA foi o que obteve o melhor resultado em todas as métricas de acurácia entre todos os modelos testados, atingindo um MAPE de 10,72%. Como exposto no capítulo de análise dos resultados, de acordo com os diretores da empresa, o resultado foi satisfatório e foram levantadas hipóteses de utilização do modelo na empresa. Nesse sentido, definiu-se que ele seria o modelo a ser desenvolvido.

Uma sugestão de aprimoramento para o processo atual de previsão da empresa é instituir uma revisão mensal sistemática da previsão de demanda anual, ao invés de realizá-la apenas quando percebida a necessidade. Essa abordagem esporádica atual torna a empresa vulnerável a estimativas imprecisas, como foi exposto em entrevista com os diretores. É importante destacar que, para garantir uma acurácia consistente, o modelo ARIMA necessita de atualizações periódicas com os dados de demanda. Conforme apresentado no decorrer do projeto, a empresa estudada registrou um aumento significativo na demanda nos últimos anos e enfrenta expressiva sazonalidade em seu mercado. Isso ressalta a importância e o impacto desse estudo em suas operações.

O estudo possibilitou o entendimento teórico do autor sobre os métodos de previsão de demanda, permitindo uma análise mais profunda sobre a relevância e aplicação de cada técnica no contexto da empresa em que atua. Além disso, a pesquisa ofereceu uma introdução prática à linguagem de programação Python e às

ferramentas de Machine Learning. Essa experiência abre portas para futuras investigações, abrangendo outros produtos, bases de dados variadas e, possivelmente, a incorporação de novas variáveis.

Como principais limitações do presente estudo, destaca-se a impossibilidade de comparar diretamente o modelo desenvolvido com o método atualmente utilizado pela empresa. Tal comparação seria essencial para avaliar a eficácia relativa do novo modelo em relação ao método existente. A ausência de um histórico das previsões feitas pela empresa impede essa análise comparativa, o que pode limitar a compreensão completa do potencial e da precisão do modelo proposto em um contexto real de negócios.

Para trabalhos futuros, pode-se citar o desenvolvimento de um modelo de regressão multivariada, com a inclusão de variáveis externas à empresa que impactam o seu mercado. Além disso, pode-se desenvolver testes com outras dimensões de tempo, como dias, semanas ou anos, com a capacidade de prever um espaço de tempo maior que o atual.

REFERÊNCIAS BIBLIOGRÁFICAS

ABASTECIMENTO, 2021. Disponível em:
<<https://www.gov.br/agricultura/pt-br/assuntos/inspecao/produtos-vegetal/publicacoes/anuario-da-cerveja-2021.pdf>>. Acesso em: 11 jan. 2022.

ACKERMANN, Andres E. F.; SELLITTO, Miguel A. **Métodos de previsão de demanda: uma revisão da literatura**. Bogotá: Innovar, v. 32, n. 85, p. 83-99, set. 2022. Disponível em:
<http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0121-50512022000300083&lng=en&nrm=iso>. Acesso em: 26 Jan. 2023.

ARVAN, M.; FAHIMNIA, B.; REISI, M.; SIEMSEN, E. **Integrating human judgement into quantitative forecasting methods: A review**. Omega, [S.l.], v. 86, p. 237-252, 2019. Disponível em: <<https://doi.org/10.1016/j.omega.2018.07.012>>. Acesso em: 01 mar. 2023.

BALLOU, Ronald. **Gerenciamento da Cadeia de Suprimentos**. 5. ed. Porto Alegre: Bookman, 2006. Disponível em:
<<https://redeprocurtos.com.br/docs/T%C3%89CNICO%20EM%20LOG%C3%8DSTICA/Logistica%20Ballou.pdf>>. Acesso em: 02 mar. 2023.

BARTHHAAS. **BarthHaas Report 2021/2022**. Nuremberg: BARTHHAAS, 2022. Disponível em:
<<https://www.hops.com.au/downloads/news-events/BarthHaas-Hop-Report-2021-2022.pdf>>. Acesso em: 13 jan. 2022.

BESSA, Adriana Bezerra. Ensaio sobre previsão de vendas no varejo de moda. 2021. In: BESSA, Adriana Bezerra. Ensaio sobre previsão de vendas no varejo de moda. 2021. Monografia (Doutorado), FGV EESP, Escola de Economia de São Paulo, 2021. Disponível em:
<https://bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/31137/Tese_AdrianaBessa_Final.pdf?sequence=1&isAllowed=y>. Acesso em: 05 mar.2022.

BOUSDEKIS, A. *et al.* **A Review of Data-Driven Decision-Making Methods for Industry 4.0 Maintenance Applications**. Electronics, v. 10, n. 7, p. 828, 31 mar. 2021. Disponível em: <<http://dx.doi.org/10.3390/electronics10070828>>. Acesso em: 05. jul. 2023.

BOX, G. E. P.; JENKINS, G. M. **Time series analysis: forecasting and control**. San Francisco: Holden-Day, 1976.

BRENTAN, B. M.; LUVIZOTTO JR, E.; HERRERA, M.; IZQUIERDO, J.; PÉREZ-GARCIA, R. Hybrid regression model for near real-time urban water demand forecasting. **Journal of Computational and Applied Mathematics**, v. 309, p. 532-541, 2017. Disponível em: <<https://doi.org/10.1016/j.cam.2016.02.009>>. Acesso em: 01 mar. 2023.

CALÔBA, G. M. CALÔBA, L. P.; SALIBY, E.. **Cooperação entre redes neurais artificiais e técnicas 'clássicas' para previsão de demanda de uma série de vendas de cerveja na Austrália.** *Pesquisa Operacional*, v. 22, n. Pesqui. Oper., 2002 22(3), jul. 2002. Disponível em: <<https://www.scielo.br/j/pope/a/GvmZQVNPkXxtMBBpkGsSjyH/?lang=pt>>. Acesso em: 24 jan. 2023.

CAPITANIO, Sidimar. Desenvolvimento de algoritmo adaptável utilizando redes neurais artificiais para previsão de demanda.. *In: CAPITANIO, Sidimar. Desenvolvimento de algoritmo adaptável utilizando redes neurais artificiais para previsão de demanda.* 2019. Dissertação (Mestrado em Engenharia de Produção), Escola de Engenharia, Universidade de Caxias do Sul, Rio Grande do Sul. Disponível em: <<https://repositorio.ucs.br/11338/5144>>. Acesso em: 24 jan. 2023.

CERVBRASIL. **Anuário da Cerveja**, 2016. São Paulo: CERVBRASIL, 2017. Disponível em: <http://www.cervbrasil.org.br/novo_site/anuarios/CervBrasil-Anuario2016_WEB.pdf>. Acesso em: 13 jan. 2022.

CHASE, Richard B., AQUILANO, Nicholas J. and JACOBS, F. Robert. **Administração da produção para a vantagem competitiva.** 10.ed. Porto Alegre: Bookman, 2006.

CHEN, Tianqi; GUESTRIN, Carlos. **XGBoost: a scalable tree boosting system.** KDD'16: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, p. 785-794, ago. 2016. Disponível em: <<https://www.kdd.org/kdd2016/papers/files/rfp0697-chenAemb.pdf>>. Acesso em: 05 mar. 2023.

DEB, C.; ZHANG, F.; YANG, J.; LEE, S. E.; SHAH, K. W. A review on time series forecasting techniques for building energy consumption. **Renewable and Sustainable Energy Reviews**, v. 74, p. 902-924, 2017. Disponível em: <<https://doi.org/10.1016/j.rser.2017.02.085>>. Acesso em: 03 mar. 2023.

DIAS, Cristian da Silva. **A administração baseada em evidências ou evidence-based management é um procedimento que aumenta a qualidade do processo de tomada de decisões: uma revisão sistemática de literatura.** 2021. 89 f. Trabalho de Conclusão de Curso (Graduação em Administração) - Faculdade de Administração e Ciências Contábeis, Universidade Federal Fluminense, Niterói, 2021. Disponível em: <<https://app.uff.br/riuff/handle/1/25665>>. Acesso em: 05. jul. 2023.

DOS REIS, Josimar Vieira; *et al.* **A previsão de demanda como ferramenta para auxiliar na tomada de decisão no contexto organizacional das microempresas em tempos de pandemia.** *Revista Livre de Sustentabilidade e Empreendedorismo*, [S. l.] v. 7, n. 3, p. 145-162, 2022. Disponível em: <<http://relise.eco.br/index.php/relise/article/view/597/610>>. Acesso em: 25 jan. 2023.

EVANS, James; OLSON, David. **Statistics, Data Analysis and Decision Modeling**. New Jersey: Prentice-Hall, 2000.

FORTSCH, S. M.; KHAPALOVA, E. A. Reducing uncertainty in demand for blood. **Operations Research for Health Care**, v. 9, p. 16-28, 2016. Disponível em: <<https://doi.org/10.1016/j.orhc.2016.02.002>>. Acesso em: 03 mar. 2023.

GIL, P. A. V. D. F. **Unfolding the drivers for academic success: The case of ISCTE-IUL** [ISCTE]. n. 41, 2019. Disponível em: <<http://hdl.handle.net/10071/20069>>. Acesso em: 30 jun. 2023.

GIL, P. D., Martins, S. da C., Moro, S., & Costa, J. M. A data-driven approach to predict first-year students' academic success in higher education institutions. **Education and Information Technologies**, [s. l.], n. 26, v. 2, p. 2165–2190, 2021. Disponível em: <<https://doi.org/10.1007/s10639-020-10346-6>>. Acesso em: 30 jun. 2023.

HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. **Neural Computation**, v. 9, n. 8, p. 1735-1780, 1997. Disponível em: <<https://www.bioinf.jku.at/publications/older/2604.pdf>>. Acesso em: 01 mar. 2023.

HOFMANN, E.; RUTSCHMANN, E. **Big data analytics and demand forecasting in supply chains: a conceptual analysis**. The International Journal of Logistics Management, [S.l.], v. 29, n. 2, p. 739-766, 2018. Disponível em: <<https://doi.org/10.1108/IJLM-04-2017-0088>>. Acesso em: 25 jan. 2023.

JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. Data clustering: a review. **ACM Computing Surveys (CSUR)**, [S.l.], v. 31, n. 3, p. 264-323, Nova Iorque, 1999. Disponível em: <<https://dl.acm.org/doi/10.1145/331499.331504>>. Acesso em: 10 jul. 2023.

Kochański, A. (2003). Data preparation. *Informatyka w Technologii Materiałów*, 10 (Janeiro 2010), 24–137. Disponível em: <<https://doi.org/10.1017/9781107051386.003>>. Acesso em: 30 jun. 2023.

LEMOS, F. O. **Metodologia para seleção de métodos de previsão de demanda**. 2006. Dissertação de Mestrado (Mestrado em Engenharia de Produção) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2006. Disponível em: <<http://www.producao.ufrgs.br/arquivos/publicacoes/fernandooliveiralemos.pdf>>. Acesso em: 30 jun. 2023.

LIMBERGER, S. C.; ESPÍNDOLA, C. J. **A desnacionalização da indústria cervejeira no Brasil: da reestruturação produtiva aos movimentos de fusões e aquisições**. Goiânia: Ateliê Geográfico, v. 13, n. 2, p. 148–164, 2019. DOI: 10.5216/ag.v13i2.54257. Disponível em: <<https://revistas.ufg.br/atelie/article/view/54257>>. Acesso em: 3 fev. 2023.

MARTINS, Amanda Nunes; MAÇADA, Antonio Carlos Gastaud. Big data analytics na qualidade da tomada de decisão: uma revisão sistemática da literatura. *In*:

ENCONTRO NACIONAL DA ASSOCIAÇÃO NACIONAL DE PÓS-GRADUAÇÃO E PESQUISA EM ADMINISTRAÇÃO, 46., 2022, [s. l.]. **Anais [...]**. [S. l.]: ANPAD, 2022. Disponível em:

<<http://anpad.com.br/uploads/articles/120/approved/007ff380ee5ac49ffc34442f5c2a2b86.pdf>>. Acesso em: 05. jul. 2023.

Meijden, V. D. L. H., Nunen, J. A. E. E. V. & Ramondt, A. **Forecasting: bridging the gap between sales and manufacturing**. International Journal Production Economics, v. 37, n. 1, p. 101-114, 1994. Disponível em:

<<https://ideas.repec.org/a/eee/proeco/v37y1994i1p101-114.html>> Acesso em: 20. ago. 2023.

MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO. **Anuário da Cerveja**, 2021. Distrito Federal: MINISTÉRIO DA AGRICULTURA, PECUÁRIA E LIMBERGER, S. C. O setor cervejeiro no Brasil: gênese e evolução. CaderNAU, [S. l.], v. 6, n. 1, 2014. Disponível em:

<<https://periodicos.furg.br/cnau/article/view/4769>>. Acesso em: 25 jan. 2023.

Mircetic, D., Nikolicic, S., Maslaric, M., Ralevic, N., & Debelic, B. **Development of S-ARIMA Model for Forecasting Demand in a Beverage Supply Chain**. De Gruyter. Open Engineering, v. 6, n. 1, 2016. Disponível em:

<<https://www.degruyter.com/document/doi/10.1515/eng-2016-0056/html>>. Acesso em: 20 ago. 2023

MONTGOMERY, Douglas C.; JENNINGS, Cheryl L.; KULAHCI, Murat. **Introduction to Time Series Analysis and Forecasting**. 2. ed. Hoboken, NJ: John Wiley & Sons, 2015. 472 p. Disponível em: <<https://pedro.unifei.edu.br/download/Montgomery.pdf>>. Acesso em: 01 mar. 2023.

NAVES DINIZ, I. M.; NUNES, D. M.; ROSA, V. A. de O.; CALIFE, N. F. S. Aplicação de um modelo de previsão de demanda em uma empresa do setor cervejeiro.

Brazilian Journal of Production Engineering, [S. l.], v. 5, n. 3, p. 120–138, 2019. Disponível em: <https://periodicos.ufes.br/bjpe/article/view/V05N03_09>. Acesso em: 25 jan. 2023.

PETROPOULOS, Fotios *et al.* Judgmental selection of forecasting models. **Journal of Operations Management**, [s. l.], v. 60, p. 34-46, 2018. Disponível em:

<<https://doi.org/10.1016/j.jom.2018.05.005>>. Acesso em: 28 fev. 2023.

RAMOS, Jorge Luis Cavalcanti; RODRIGUES, Rodrigo Lins; SILVA, João Carlos Sedraz; OLIVEIRA, Pamela Letícia Silva de. **CRISP-EDM: uma proposta de adaptação do Modelo CRISP-DM para mineração de dados educacionais**. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 31. , 2020, Online. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2020. p. 1092-1101. Disponível em: <<https://doi.org/10.5753/cbie.sbie.2020.1092>>. Acesso em: 30 jun. 2023.

REN, Y.; SUGANTHAN, P. N.; SRIKANTH, N.; AMARATUNGA, G. **Random vector functional link network for short-term electricity load demand forecasting**.

Information Sciences, v. 367-368, p. 1078-1093, 2016. Disponível em: <https://doi.org/10.1016/j.ins.2015.11.039>. Acesso em: 02 mar. 2023.

SEMMELMANN, L.; HENNI, S.; WEINHARDT, C. **Load forecasting for energy communities: a novel LSTM-XGBoost hybrid model based on smart meter data.** *Energy Informatics*, [S.l.], v. 5, n. Suppl 1, p. 24, 2022. DOI: 10.1186/s42162-022-00212-9. Disponível em: <https://doi.org/10.1186/s42162-022-00212-9>. Acesso em: 01 mar. 2023.

SOARES JUNIOR, Edvan Joaquim. **Modelos de Redes Neurais Recorrentes para Previsão de Demanda de Curto Prazo.** 2020. 87 f. Universidade Federal de Pernambuco, Recife, 2020. Disponível em: <https://repositorio.ufpe.br/handle/123456789/37901>. Acesso em: 05 mar.2023.

SOUZA, Francisca Mendonça. **Modelos Box & Jenkins aplicados a previsão de demanda de leitos hospitalares.** 2006. Monografia (Especialização), Universidade Federal de Santa Maria, Centro de Ciências Naturais e Exatas, Curso de Especialização em Estatística e Modelagem Quantitativa, RS, 2006. Disponível em: <http://w3.ufsm.br/adriano/mon/fm.pdf>. Acesso em: 01 mar. 2023.

STÜKER, T. A. **Modelo para atualização da previsão de demanda em cadeia de suprimentos de moda rápida na indústria calçadista.** 2014. 117 f. Dissertação (Mestrado em Administração), Universidade do Vale do Rio dos Sinos, São Leopoldo, 2014. Disponível em: <http://www.repositorio.jesuita.org.br/handle/UNISINOS/3585>. Acesso em: 01 mar. 2023.

THOMAS, R. J. Estimating Demand for Services: Issues in Combining Sales Forecasts. *Journal of Retailing and Consumer Services*. v. 3, n. 4, p. 241-250, 1996. Disponível em: [https://doi.org/10.1016/0969-6989\(95\)00095-X](https://doi.org/10.1016/0969-6989(95)00095-X). Acesso em: 02 mar. 2023.

VAN DER LAAN, E.; VAN DALEN, J.; ROHRMOSER, M.; SIMPSON, R. **Demand forecasting and order planning for humanitarian logistics: An empirical assessment.** *Journal of Operations Management*, v. 45, p. 114-122, 2016. Disponível em: <https://doi.org/10.1016/j.jom.2016.05.004>. Acesso em: 03 mar. 2023.

Veiga, C. P., Veiga, C. P., Vieira, G. E. & Tortato, U. **Impacto financeiro dos erros de previsão: um estudo comparativo entre modelos de previsão lineares e redes neurais aplicados na gestão empresarial.** *Produção Online*, v. 12, p. 629-656, 2012. Disponível em: <https://producaoonline.org.br/rpo/article/view/959>. Acesso em: 20 de ago. 2023.

Wirth, R. (2000). CRISP-DM : Towards a Standard Process Model for Data Mining. In: INTERNATIONAL CONFERENCE ON THE PRACTICAL APPLICATION OF KNOWLEDGE DISCOVERY AND DATA MINING, 4., 2000, [s. l.]. *Anais [...]*. [S. l.; s. n.], 2000. p. 29-39. Disponível em:

<<http://cs.unibo.it/~danilo.montesi/CBD/Beatriz/10.1.1.198.5133.pdf>>. Acesso em: 30 jun. 2023.

WHEELWRIGHT, Steven C.; MAKRIDAKIS, Spyros; HYNDMAN, Rob J.
Forecasting: methods and applications. 3rd. ed. New York: Wiley, 1980.

ANEXO A - ROTEIRO DE ENTREVISTA SEMIESTRUTURADA PRÉ-MODELAGEM

Qualificação do respondente:

1. Qual é o seu nível de formação educacional e em que área é a sua formação?
2. Qual o seu cargo na empresa?
3. Há quantos anos você está trabalhando nesta empresa?
4. Que experiência anterior você tem em previsão de demanda e planejamento de produção?

Compreensão do Processo:

5. A empresa adota processo e planejamento de previsão de demanda?
6. Os objetivos do negócio estão alinhados ao planejamento de previsão de demanda?
7. Quantas pessoas estão envolvidas no processo de criar os cenários para o planejamento da demanda?
8. Quanto tempo, em média, leva para gerar as previsões de demanda que atualmente são utilizadas na empresa?
9. Quais áreas utilizam os cenários para tomada de decisão?

Fatores	Perguntas	Autores
Processo e Metodologia	10. Quais métodos sua empresa atualmente utiliza para a previsão de demanda? 11. Como esses métodos foram selecionados e implementados na empresa? 12. Quão frequentemente as previsões de demanda são atualizadas em sua empresa?	<ul style="list-style-type: none"> ● Makridakis, Wheelwright , Hyndman (1998) ● Armstrong (2001) ● Chopra, Meindl, 2016
Tempo (Lead-time)	13. Qual é o horizonte de tempo utilizado pela sua empresa para prever a demanda e organizar a produção?	<ul style="list-style-type: none"> ● Hyndman, R.J., Athanasopoulos, G (2018)
Acuracidade	14. Qual a importância da acuracidade do modelo de previsão de demanda para o planejamento do negócio? Pode ter alguma flexibilidade?	<ul style="list-style-type: none"> ● Davydenko, Fildes (2013)
Desafios	15. Quais são os principais desafios que sua empresa enfrenta ao prever a demanda? 16. Como a imprecisão na previsão de demanda impactou a empresa no passado?	<ul style="list-style-type: none"> ● Svoronos, Newbold (1997) ● Lee, Padmanabhan, Whang (2004)
Melhorias	17. Quais mudanças você gostaria de implementar no processo de previsão de demanda de sua empresa?	<ul style="list-style-type: none"> ● Syntetos, Babai, Davies (2016)

Respostas:

Nº. Pergunta	Resposta
5	Sim.
6	Sim, os objetivos do negócio, como a meta de crescimento, são levados em consideração para o desenvolvimento da previsão de demanda.
7	Três pessoas estão envolvidas no processo de previsão.
8	Atualmente o processo leva de 7 a 10 dias.
9	A previsão de demanda é gerada pela área comercial e é desdobrada nas áreas de produção e financeira.
10	O método foi desenvolvido pela própria empresa. É feito uma média de crescimento dos últimos 2 anos, junto a isto se distribui a demanda de acordo com a sazonalidade do trimestre.
11	Pela experiência e aprimoramento do método atual.
12	Não existe uma rotina para revisão da previsão feita anualmente, quando é observada a necessidade ela é desenvolvida.
13	30 dias.
14	É de extrema importância ter uma alta acuracidade na previsão de demanda, pois este planejamento tem um grande impacto em todas as áreas da empresa.
15	Conseguir prever o comportamento do mercado, é difícil prever a sazonalidade da demanda e as mudanças existentes dentro do ano.
16	Sim, ela vem reduzindo bastante mas existe um histórico de falta de produto para fazer venda, o que impactou na geração de receita da empresa, assim como houve períodos em que a previsão superou a demanda o que fez com que fosse gerado estoque desnecessário e por se tratar de produtos perecíveis houve perdas neste sentido.
17	Ferramentas que possam trazer insights sobre o comportamento do mercado e melhorar o processo de previsão atual, servindo como base para a revisão da previsão.

Anexo B - ROTEIRO DE ENTREVISTA PÓS-MODELAGEM

O objetivo desta entrevista é avaliar o interesse em utilizar o modelo de previsão como ferramenta no processo de previsão de vendas da microcervejaria, com base na apresentação dos resultados dos testes realizados neste trabalho.

1. Como você avalia os resultados das previsões mensais geradas pelo modelo?
2. Os resultados obtidos estão em conformidade com os limites estabelecidos pela empresa?
3. Na sua perspectiva, é viável que esse modelo contribua para aprimorar o processo de previsão de vendas mensais da empresa? Se sim, de que maneira você enxerga sua utilização? Caso contrário, quais seriam os motivos dessa visão?

Respostas:

Nro. Pergunta	Resposta
1	Os resultados são satisfatórios, é possível observar que o modelo consegue prever bem a sazonalidade existente no mercado e de acordo com as métricas de acurácia desempenhou bem nos testes.
2	Sim, os resultados estão dentro dos limites.
3	É possível utilizar o modelo para melhorar as revisões da previsão de demanda feita anualmente. Tendo esta ferramenta poderemos ter um melhor embasamento para realizar a revisão.