

**UNIVERSIDADE FEDERAL DE ALAGOAS  
UNIDADE ACADÊMICA CENTRO DE TECNOLOGIA  
CURSO DE ENGENHARIA QUÍMICA**

**ALVARO PESSOA QUINTILIANO SILVA**

**APLICAÇÃO DO MÉTODO ARIMA COMO FERRAMENTA DE GESTÃO DE  
CONTROLE DE QUALIDADE DE MATÉRIA-PRIMA EM INVENTÁRIO**

Maceió  
2022

ALVARO PESSOA QUINTLIANO SILVA

**APLICAÇÃO DO MÉTODO ARIMA COMO FERRAMENTA DE GESTÃO DE  
CONTROLE DE QUALIDADE DE MATÉRIA-PRIMA EM INVENTÁRIO**

Plano de Trabalho de Conclusão de Curso  
apresentado ao curso de Engenharia Química da  
Universidade Federal de Alagoas como  
requisito parcial para obtenção do título de  
Bacharel em Engenharia Química.

Orientador: Prof. Dr. João Inácio Soletti.

Maceió  
2022

**Catálogo na Fonte**  
**Universidade Federal de Alagoas**  
**Biblioteca Central**  
**Divisão de Tratamento Técnico**

Bibliotecário: Marcelino de Carvalho Freitas Neto – CRB-4 – 1767

- S586a Silva, Alvaro Pessoa Quintliano.  
Aplicação do método ARIMA como ferramenta de gestão de controle de qualidade de matéria-prima em inventário / Alvaro Pessoa Quintliano Silva. – 2022.  
36 f. il. : figs. ; tabs. color.
- Orientador: João Inácio Soletti.  
Monografia (Trabalho de Conclusão de Curso em Engenharia Química).  
Universidade Federal de Alagoas. Centro de Tecnologia. Maceió, 2022.
- Bibliografia: f. 35-36.
1. Controle de qualidade. 2. Otimização de processos. 3. Gestão da produção. 4. Matérias-primas. 5. Estatística. I. Título.

CDU:66.011

Dedico este trabalho à minha mãe e ao meu padrasto por todo incentivo, amor e exemplo durante toda minha vida.

## AGRADECIMENTOS

À minha mãe, Karyne, que sempre me apoiou, me educou e se dedicou para que eu pudesse ter a melhor educação possível. Ao meu irmão, Vinicius, que sempre foi um exemplo de força para mim desde pequeno. À minha avó, Lourdes, por todo amor e suporte constante.

Ao meu padrasto, Jairo Júnior, que é um dos grandes responsáveis pela minha formação e por ser, voluntariamente, um pilar para a minha família. A meu pai, Álvaro, e à toda minha família pelo amor e suporte durante esses anos de graduação.

À minha namorada, Thays, que, com muito amor e carinho, sempre esteve comigo nos momentos de dificuldades e alegrias tanto a graduação quanto na vida.

Aos meus amigos da vida na Hungria, por todos os momentos compartilhados e apoio mútuo. Em especial ao meu amigo Alisson, que se tornou um irmão na vida fora do Brasil.

À todas as pessoas que cruzaram meu caminho e que me fizeram ser melhor que a minha versão anterior.

Ao professor Jorge Brito por todas as boas conversas, troca de experiências, pelos conselhos sempre muito pertinentes e pelo incentivo constante à busca do conhecimento.

Por fim, ao meu orientador, professor João Inácio, pela paciência, incentivo, conselhos e pelo conhecimento compartilhado durante toda a graduação, contribuições estas que levarei para a minha vida profissional.

## RESUMO

O controle de qualidade da matéria prima é peça fundamental para a indústria. Materiais que permaneçam armazenados por muito tempo acabam saindo dos padrões de especificação requeridos e, como consequência, perdem seu valor agregado. Além disso, estes materiais geram custos relacionados de transporte e armazenagem, como também carregam um capital que poderia ser direcionado a outros investimentos. Neste trabalho, o objetivo é a gestão da matéria prima da produção de adesivos de uma empresa que atende a demanda de vários países da Europa, sendo líder de mercado neste setor. Para tal, foi realizada a modelagem matemática através do método Box-Jenkins (ARIMA), utilizando dados históricos de consumo dos materiais, a fim de prever possíveis materiais que sairão de especificação. Foram feitas as modelagens manual e automática, em linguagem Python, deste método a fim de comparar a eficácia de cada método. A metodologia do trabalho consistiu em uma revisão sistemática da literatura, seguida pelo estudo do método ARIMA. Logo após, foram implementadas as etapas manual e automática dos modelos de predição dos valores de depreciação dos 12 meses seguintes aos últimos meses presentes nos dados e os resultados obtidos foram submetidos a testes de correlação estatística com os valores reais. Na comparação, o método manual apresentou uma Raiz-Quadrada do Erro Médio (RMSE) de 11,18%, enquanto o método automático obteve 12,81%. Percebeu-se que ambos os modelos desenvolvidos se aproximaram da realidade e possivelmente permitiriam minimizar os custos associados ao controle da matéria prima utilizada, reduzindo a produção de produtos indesejáveis (fora de especificação).

Palavras-Chave: Controle; otimização; gestão da produção; matérias-primas; estatística.

## **ABSTRACT**

The quality control of raw materials is a fundamental part for the proper functioning of an industry. Materials that remain stored for a long time end up falling outside the required specification standards and, consequently, losing their added value. In addition, these materials generate costs related to transportation and storage, and carry a capital that could be directed to other investments. In this work, the objective is the management of the raw material production of adhesives in a company that meets the demand from several European countries, being a market leader in this sector. To this end, mathematical modeling was performed through the Box-Jenkins method (ARIMA), using historical data of material consumption, to predict possible materials that will go out of specification. Manual and automatic modeling, in Python language, of this method was done to compare the effectiveness of each method. The methodology of the work consisted of a systematic review of the literature, followed by the study of the ARIMA method. Soon after, the manual and automatic stages of the models for predicting the depreciation values for the 12 months following the last months present in the data were implemented, and the results obtained were submitted to statistical correlation tests with the original raw values. In the comparison, the manual method presented a Root Mean Square Error (RMSE) of 11.18%, while the automatic method obtained 12.81%. It was noticed that both models developed were close to reality and would possibly allow the minimization of the costs associated with the control of the raw material used, reducing the generation of undesirable products (out of specification).

Keywords: Control; optimization; production management; raw materials; statistics.

## LISTA DE FIGURAS

Figura 1: Fluxograma geral para obtenção do modelo matemático. ....	17
Figura 2: Exemplo de Plot PACF .....	21
Figura 3: Exemplo de Plot ACF. ....	22
Figura 4: Série Temporal – Dados reais de depreciação do material A. ....	25
Figura 5: Resultados da Decomposição Sazonal. ....	25
Figura 6: Média Móvel e Desvio Padrão em dados originais. ....	27
Figura 7: Média Móvel e Desvio Padrão após primeira diferenciação. ....	27
Figura 8: Teste ACF. ....	28
Figura 9: Teste PACF. ....	29
Figura 10: Resultados modelagem SARIMAX. ....	30
Figura 11: Gráfico dos valores de depreciação reais e modelados por mês. ....	31



## LISTA DE TABELAS

Tabela 1: Resultados do Teste ADF aplicado aos dados originais.....	26
Tabela 2: Resultados do Teste ADF após diferenciação dos dados. ....	26
Tabela 3: Resultados das Previsões. ....	31
Tabela 4: Análise Estatística para a Previsão Manual. ....	32
Tabela 5: Análise Estatística para a Previsão Manual. ....	32
Tabela 6: Viés Estatístico das Previsões Manual e Automática. ....	33
Tabela 7: Erro Médio Absoluto das Previsões Manual e Automática. ....	33
Tabela 8: RMSE das Previsões Manual e Automática. ....	34

## SUMÁRIO

1. INTRODUÇÃO.....	9
2. OBJETIVOS.....	10
2.1 Geral.....	10
2.2 Específicos.....	10
3. REVISÃO BIBLIOGRÁFICA.....	11
3.1 Tecnologia de Adesivos.....	11
3.1.1 Produção.....	11
3.2 Controle de Qualidade de Inventário.....	12
3.2.1 Desvalorização de Inventário ( <i>Write-Down</i> ).....	12
3.3 Modelagem Matemática e Séries Temporais.....	13
3.3.1 Metodologia Box-Jenkins (ARIMA).....	13
4. METODOLOGIA.....	17
4.1 Estudo da Série Temporal.....	17
4.1.1 Teste de Sazonalidade (Decomposição Sazonal).....	18
4.1.2 Estacionariedade da Série.....	18
4.1.3 Diferenciação dos Dados.....	20
4.2 Determinação dos Parâmetros $p$ , $d$ e $q$ do Modelo ARIMA.....	20
4.3 Aplicação do Método ARIMA e Geração do Modelo Matemático.....	22
4.4 Validação dos Resultados.....	23
4.4.1 Teste de Viés ( <i>BIAS</i> ).....	23
4.4.2 Erro Médio Absoluto ( <i>MAE</i> ).....	23
4.4.3 Raiz-Quadrada do Erro Médio ( <i>RMSE</i> ).....	24
5. RESULTADOS E DISCUSSÕES.....	25
5.1 Teste de Sazonalidade (Decomposição Sazonal).....	25
5.2 Teste de Estacionariedade.....	26

5.2.1	Teste ADF .....	26
5.2.2	Média Móvel e Desvio Padrão .....	27
5.3	Obtenção dos Parâmetros $p$ , $d$ , e $q$ do Modelo ARIMA .....	28
5.4	Geração automática do Modelo ARIMA .....	29
5.5	Validação dos Resultados .....	31
5.5.1	Teste de Viés ( <i>BIAS</i> ).....	32
5.5.2	Erro Médio Absoluto (MAE – <i>Mean Absolute Error</i> ) .....	33
5.5.3	Raiz-Quadrada do Erro Médio (RMSE – <i>Root Mean Square Deviation</i> ) .....	33
6.	CONCLUSÃO.....	34
7.	REFERÊNCIAS .....	35

## 1. INTRODUÇÃO

O controle, monitoramento e a análise dos materiais em uma empresa garantem o bom funcionamento de todas as operações associadas à mesma. A gestão é a peça-chave para o sucesso das empresas, visto que objetiva garantir a matéria prima necessária para sua atividade, bem como assegura o atendimento ao cliente. Por outro lado, uma má execução ou elaboração, de um plano de gestão de materiais pode levar a perdas significativa: todos os processos entram em estado de risco de perdas sem reparos para os produtos e processos de atendimento junto ao cliente final.

Atualmente, durante o processo de produção, são elaborados relatórios que reportam a qualidade da matéria-prima com a finalidade do controle da produção de forma mais eficiente, a fim de reduzir os custos de armazenamento e/ou descarte quando não são atingidas as especificações necessárias. Estes relatórios reportam eventos já ocorridos, produtos fora de especificação, sem a possibilidade de reutilização, sendo estas análises imprescindíveis para planejamentos futuros da empresa. Os dados coletados e utilizados nos fornecem informações sobre o estoque e desvalorizações de novembro de 1997 até dezembro de 2021 e têm características interessantes, haja visto que possibilitam a predição de ocorrências indesejadas. O presente trabalho tem como objetivo a utilização destes dados para possíveis predições de controle da qualidade, evitando a produção de produtos foras de especificação, aumentando a receita da empresa pela minimização de perdas.

## **2. OBJETIVOS**

### **2.1 Geral**

Desenvolver uma ferramenta para a o controle das matérias-primas em uma empresa de adesivos.

### **2.2 Específicos**

Elaborar de uma ferramenta capaz de estimar a depreciação do valor do material em estudo quando armazenado em estoque.

Modelar de forma analítica, utilizando ferramentas estatísticas, bem como utilizando a linguagem de programação Python, esta ferramenta de previsão.

Aumentar a receita da empresa ao minimizar os custos em relação a geração de produtos indesejados.

### **3. REVISÃO BIBLIOGRÁFICA**

#### **3.1 Tecnologia de Adesivos**

Por ser o setor que tem maior impacto na empresa para o mercado europeu, esta linha de negócios será aquela com os produtos analisados neste projeto. Atualmente, este setor abrange quatro áreas diferentes. Eles existem porque cada área tem divisões diferentes, que são tratadas de forma diferente pelas equipes responsáveis, como a equipe de vendas, a equipe de logística e a equipe de planejamento, que alocam os respectivos produtos em seus devidos usos. São estes os setores: Automotivo e Metalúrgico; Embalagens e Bens de consumo; Eletrônica e Industrial; Construção e Profissional.

##### **3.1.1 Produção**

A estrutura da cadeia de suprimentos da empresa possui 133 locais de produção, mais de 200 armazéns e cerca de 70.000 produtos, que juntos formam a espinha dorsal do setor de Tecnologia de Adesivos.

O processo produtivo é realizado em fábricas próprias, bem como em fabricantes contratados. Esta etapa inclui a produção de um ou vários lotes de um material para atingir os resultados de fabricação planejados:

- Qualidade: produtos finais que atendem a especificações rigorosas;
- Pontualidade: Os produtos finais são produzidos na data prevista, de acordo com o plano de produção

Os padrões de custo são atendidos por meio de rendimentos reais de materiais, consumos e tempo de recursos de acordo com a receita mestre e as versões de produção relevantes. Inclui os processos de conversão direta de materiais com recursos primários e secundários e todas as etapas de postagem associadas para refletir as alterações físicas e as alterações de valor.

As principais atividades estão descritas a seguir:

- Preparo das produções: Antes do início da produção, todos os processos na instalação de produção são concluídos. Consiste na verificação da especificação do produto, na configuração dos dados mestres de fabricação no sistema ERP principal e nos subsistemas relacionados.
- Gerenciamento dos materiais: descreve o recebimento de mercadorias e o processo de inventário físico em uma planta de produção. Os materiais de estágio

descrevem os processos para fornecer as matérias-primas, embalagens, granéis e materiais intermediários para as linhas de produção.

- Fabricação de produtos: inclui todos os processos durante os processos de mistura e enchimento de embalagens na produção.
- Gerenciamento de devolução: descreve o processo de recebimento, manuseio e tomada de decisão para as mercadorias ou materiais devolvidos.

### **3.2 Controle de Qualidade de Inventário**

O controle de inventário cuida especialmente do pedido de um cliente para o processo de reabastecimento, rastreando e controlando a extensão deste e alocando-o aos espaços físicos essenciais. Determinar o nível de estoque perfeito é crucial porque os objetos armazenados estão vinculados a uma taxa que influencia o desempenho geral das operações (BRENDEN, L., 2021).

O estoque é um dos cinco principais impulsionadores da cadeia de suprimentos (junto com produção, transporte, localização e informações) e responde a duas questões importantes:

- Quanto produzir, e
- quanto armazenar.

A quantidade de estoque que a empresa carrega pode ter consequências econômicas significativas. Como resultado, os estoques devem ser gerenciados de maneira que o tempo de vida útil seja crucial no processo de tomada de decisão (PRINIOTAKIS et al, 2018)

#### **3.2.1 Desvalorização de Inventário (*Write-Down*)**

O Estoque de Movimento Lento ou SMI (do inglês, *Slow Moving Inventory*) são os materiais com baixo índice de rotatividade e que ficam estocados no armazém por mais tempo, diferentemente dos materiais de movimento rápido que tem maior taxa de consumo, tendo assim um ciclo muito mais rápido (HGSC. 2018).

Quando o material não é utilizado por muito tempo, isso pode causar problemas porque geram custos de armazenamento, transporte e a perda de outros recursos que poderiam ser usados para investir no negócio (JOPPÉ, P. 2020). Estes custos gerados são refletidos no valor agregado do próprio material, que já não é mais lucrativo como antes. Essa depreciação é denominada *write-down*. A empresa em questão calcula a depreciação baseada em dois parâmetros: valor de inventário e consumo médio mensal. Estes são, portanto, os principais meios de ação para o gerenciamento da depreciação.

### 3.3 Modelagem Matemática e Séries Temporais

Um modelo matemático é uma construção matemática abstrata, simplificada, relacionada a uma parte da realidade e criada para um propósito específico. Quão bem qualquer objetivo específico é alcançado, depende tanto do estado de conhecimento sobre um sistema, quanto de quão bem a modelagem é feita. (BENDER, E. A., 1978).

A análise de séries temporais é uma maneira de analisar uma sequência de pontos de dados coletados em um intervalo de tempo. Em outras palavras, o tempo é uma variável crucial porque mostra como os dados se ajustam ao longo dos pontos, bem como os resultados. Ela fornece uma fonte adicional de informações e uma ordem definida de dependências entre os dados (TABLEAU, 2020). Neste trabalho, a série temporal em análise é a sequência de valores de depreciação de um certo material ao longo dos anos.

#### 3.3.1 Metodologia Box-Jenkins (ARIMA)

Para modelar séries temporais, também podemos trabalhar com os modelos estatísticos tradicionais, incluindo média móvel, suavização exponencial e ARIMA. Esses modelos são lineares, pois os valores futuros são limitados para serem funções lineares de dados passados (HAYES, A, 2021). Um modelo ARIMA pode ser entendido delineando cada um de seus componentes da seguinte forma:

- Autoregressão (AR): refere-se a um modelo que mostra uma variável que regride em seus próprios valores defasados ou anteriores. Os modelos tentam prever valores futuros com base em valores passados.
- Integrada (I): representa a diferenciação das observações brutas para permitir que a série temporal se torne estacionária (ou seja, os valores dos dados são substituídos pela diferença entre os valores dos dados e os valores anteriores).
- Média Móvel (MA – *Moving Average*): incorpora a dependência entre uma observação e um erro residual de um modelo de média móvel aplicado a observações defasadas. Os modelos tentam prever valores futuros com base em erros de previsão passados. Os modelos MA assumem que um modelo autorregressivo pode aproximar a série dada. Não deve ser confundido com a média móvel, que é um processo de suavização e não um modelo de previsão.



### ➤ **Parâmetros da ARIMA**

Cada componente do ARIMA funciona como um parâmetro com uma notação padrão. Para modelos ARIMA, uma notação padrão seria ARIMA com  $p$ ,  $d$  e  $q$ , onde valores inteiros substituem os parâmetros para indicar o tipo de modelo ARIMA usado (BROWNLEE, J, 2017). Os parâmetros podem ser definidos como:

- $p$ : o número de observações defasadas no modelo; também conhecida como ordem de atraso.
- $d$ : o número de vezes que as observações brutas são diferenciadas; também conhecido como o grau de diferenciação.
- $q$ : o tamanho da janela de média móvel; também conhecida como a ordem da média móvel.

Em um modelo de regressão linear, por exemplo, o número e o tipo de termos são incluídos. Um valor 0, que pode ser usado como parâmetro, significaria que o componente não deveria ser usado no modelo (BROWNLEE, J., 2017).

O modelo ARIMA (Auto Regressive Moving Average) é um modelo de previsão de séries temporais muito comum. É uma extensão mais sofisticada do modelo ARMA (Auto Regressive Moving Average) mais simples, que em si é apenas uma fusão de dois componentes ainda mais simples: AR e MA, já expostos acima.

### ➤ **Modelo AR**

O modelo AR pode ser definido como:

$$y_t = c + \varphi_1 y_{t-1} + \epsilon_t \quad (1)$$

Onde,  $y_t$  é o valor da função num dado tempo  $t$ ,  $c$  é uma constante,  $\varphi_1$  é um coeficiente e  $\epsilon_t$  é um termo de erro de ruído branco com  $\epsilon_t \sim N(0, \sigma^2)$ .

O modelo AR, no entanto, leva em consideração uma ordem  $p$ , que ira ditar quantas defasagens anteriores a serem usadas na regressão. Um modelo AR( $p$ ) pode ser expresso por:

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} \dots + \varphi_p y_{t-p} + \epsilon_t \quad (2)$$

Que pode ser finalmente representado por:

$$y_t = c + \sum_{i=1}^p \varphi_i y_{t-i} \quad (3)$$

Onde  $\varphi_i$  é o coeficiente correspondente a cada intervalo de tempo anterior  $y_{t-i}$  respectivo. Portanto:

$$\varphi = (\varphi_1, \varphi_2, \dots, \varphi_i) \quad (4)$$

### ➤ Modelo MA

O modelo MA pode ser definido como:

$$y_t = c + \theta_1 \epsilon_{t-1} \quad (5)$$

Assim, em contraste a um modelo AR, um modelo MA é uma regressão linear do valor atual da série contra termos de erro de ruído branco observados anteriormente. Da mesma forma que os modelos AR, os modelos MA também aceitam um termo de ordem,  $q$ , que ditará quantos erros anteriores serão considerados. Um modelo MA( $q$ ) pode ser expresso por:

$$y_t = c + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} \dots + \theta_q \epsilon_{t-q} + \epsilon_t \quad (6)$$

Onde,  $y_t$  é o valor da função num dado tempo  $t$ ,  $c$  é uma constante,  $\theta_1$  é um coeficiente e  $\epsilon_{t-1}$  é um erro de ruído branco do termo anterior.

Pode-se então se resumir em:

$$y_t = c + \sum_{j=1}^q \theta_j \epsilon_{t-j} \quad (7)$$

Onde  $\theta_j$  é o coeficiente correspondente a cada intervalo de tempo anterior  $y_{t-j}$  respectivo. Portanto:

$$\theta = (\theta_1, \theta_2, \dots, \theta_j) \quad (8)$$

Os parâmetros  $p$  e  $q$  são estimados. Os métodos utilizados para estimação serão descritos mais adiante neste trabalho.

➤ **Modelo ARMA**

Um modelo ARMA é simplesmente uma fusão dos dois modelos AR e MA descritos anteriormente. Recordando suas definições, podemos, portanto, expressar um modelo ARMA como:

$$y_t = c + \sum_{i=1}^p \varphi_i y_{t-i} + \sum_{j=1}^q \theta_j \epsilon_{t-j} + \epsilon_t \quad (9)$$

Onde  $p$  e  $q$  são as ordens dos modelos AR e MA, respectivamente.

➤ **Modelo ARIMA**

O modelo ARIMA é uma extensão do modelo ARMA, com a adição de um componente de integração.

Os modelos ARMA devem funcionar em séries temporais estacionárias. Uma série temporal estacionária é uma série cujas propriedades estatísticas, como média e variância, não mudam com o tempo. Infelizmente, a maioria das séries temporais do mundo real não são estacionárias e, portanto, muitas vezes devem ser transformadas para torná-las estacionárias. O processo de transformação é conhecido como integração. O processo de transformação empregado é chamado de diferenciação, onde tomamos uma dada  $d$ -ésima diferença da série até que a série seja estacionária. Este processo será descrito posteriormente neste trabalho.

Portanto, o modelo ARIMA pode ser finalmente descrito com a notação:  $ARIMA(p,d,q)$ .

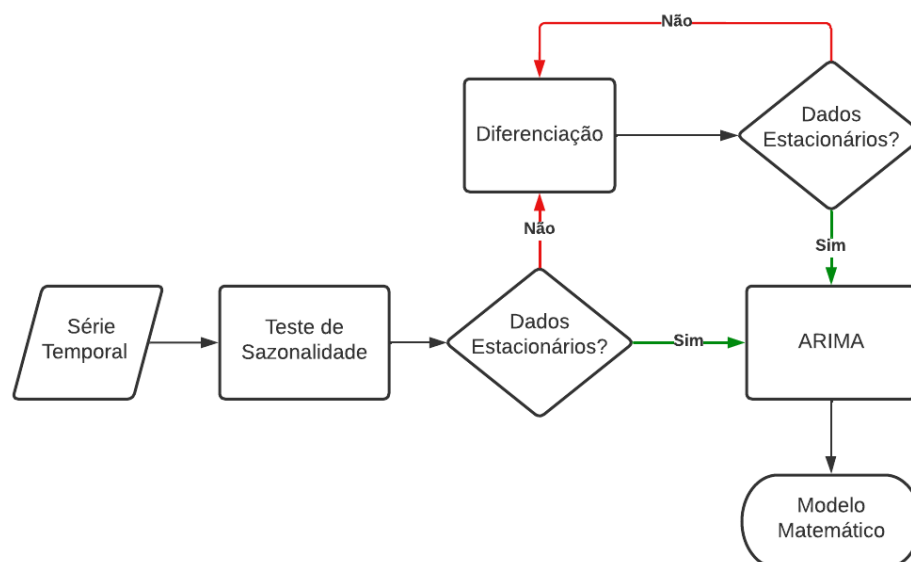
#### 4. METODOLOGIA

A metodologia foi constituída inicialmente por uma avaliação dos materiais mais problemáticos utilizados na produção de adesivos. Os critérios para avaliação destes materiais foram rotatividade, valor e desvalorização histórica. Decorrente do acordo de confidencialidade entre empresa e empregado, tais materiais não foram citados de forma explícita neste trabalho. Logo em seguida, o material com maior desvalorização proporcional ao seu preço unitário foi escolhido para ser objeto de estudo, nos resultados este será referido como material ‘A’. Dados de desvalorização em Euros foram coletados mensalmente a partir novembro de 1997 até dezembro de 2021, portanto, uma série temporal. Posteriormente, foram implementadas as etapas de utilização destes dados para a geração de um modelo matemático capaz de prever, em até 12 meses, a depreciação em Euros do material A.

##### 4.1 Estudo da Série Temporal

Neste cenário, a elaboração do modelo matemático constituiu-se em três etapas: teste de sazonalidade dos dados brutos, tratamento de dados (diferenciação) para que estes se adequassem ao método ARIMA e, em seguida, a aplicação deste método. Este processo está demonstrado na Figura 1:

**Figura 1:** Fluxograma geral para obtenção do modelo matemático.



**Fonte:** Autor (2022).

#### 4.1.1 Teste de Sazonalidade (Decomposição Sazonal)

A divisão de uma série temporal em componentes sistemáticos e não-sistemáticos é uma abstração útil para a escolha de técnicas de previsão (BROWNLEE, J, 2017).

- Sistemático: Componentes de séries temporais que podem ser caracterizados e modelados e ter consistência ou repetição.
- Não-Sistemáticos: Elementos de séries cronológicas que não podem ser modelados explicitamente.

Dois componentes sistemáticos - tendência e sazonalidade - com um componente não-sistemático - compõem uma série temporal. Estes componentes são definidos como:

- Tendência: O comportamento crescente, decrescente ou estável da série;
- Sazonalidade: A repetição do ciclo de curto prazo na série;
- Ruído (Resíduo): A variação aleatória na série.

É possível descrever uma série temporal através de um modelo aditivo destes componentes da seguinte forma:

$$y(t) = Tendência + Sazonalidade + Resíduo \quad (10)$$

Através da função `seasonal_decompose` da biblioteca `statsmodels` da linguagem Python, os valores destes componentes são detalhados e demonstrados graficamente, sendo assim possível determinar se há, ou não, sazonalidade na amostra.

#### 4.1.2 Estacionariedade da Série

O modelo ARIMA requer, para retornar resultados plausíveis, que a série temporal utilizada seja estacionária. Uma série temporal estacionária é aquela cujas propriedades não dependem do momento em que a série é observada, assim, as séries com tendências ou com sazonalidade, não são estacionárias - a tendência e a sazonalidade afetarão o valor das séries temporais em momentos diferentes (KWIATKOWSKI, D. et al, 1992). Neste trabalho, dois métodos foram utilizados para determinar a Estacionariedade da série temporal, o Teste de Dickey-Fuller Aumentado (Teste ADF) e estimadores Média Móvel (MM) e Desvio Padrão.

- Teste ADF (Dickey-Fuller Aumentado)

O teste ADF é fundamentalmente um teste de significância estatística. Isto significa que há um teste de hipótese envolvido com uma hipótese nula e como resultado, uma estatística de teste (FULLER, W. A, 1976). A hipótese nula é que uma raiz unitária está presente em uma

série temporal, ou seja, que esta série não é estacionária. A estatística ADF, usada no teste, é um número negativo, e quanto mais negativo, mais indicativo o teste se torna de rejeitar a hipótese nula de que uma raiz unitária existe na série.

Vejam primeiro a fórmula do Teste de Dickey Fuller (que é a origem do Teste ADF), e isto é:

$$y_t = c + \beta_t + \alpha Y_{t-1} + \varphi \Delta Y_{t-1} + e_t \quad (11)$$

- $y_t$  é o valor na série no tempo  $t$ ;
- $\Delta Y_{t-1}$  é a primeira diferença da série no tempo  $t-1$ .

O teste ADF, como o nome sugere, é uma expansão do Teste de Dickey-Fuller, e é dado por:

$$y_t = c + \beta_t + \alpha Y_{t-1} + \varphi \Delta Y_{t-1} + \varphi_2 \Delta Y_{t-2} + \dots + \varphi_p \Delta Y_{t-p} \quad (12)$$

Onde são adicionados mais termos de diferenciação, indicando uma série temporal maior. A hipótese nula é dada por  $p > 0.05$ . Caso  $p$  tenha valor inferior a este, a hipótese é rejeitada e a série temporal é tida como estacionária.

- Média Móvel (MM) e Desvio Padrão

A média móvel é um indicador/estimador que suaviza os dados de séries temporais a fim de filtrar ruídos representados por oscilações mais fortes, tornando mais fácil compreender o comportamento da série. A MM tem a capacidade de revelar tendências e, quando apresentada graficamente, estas tendências são mais bem observadas.

O cálculo da Média Móvel é dado por:

$$MM_n = \frac{\sum_{i=1}^n D_i}{n} \quad (13)$$

- $MM_n$  = média móvel de  $n$  períodos;
- $D_i$  = depreciação ocorrida no período  $i$ ;
- $n$  = número de períodos;
- $i$ : índice do período ( $i = 1, 2, 3, \dots$ )

Já o Desvio Padrão é um parâmetro que indica o grau de variação de um conjunto de elementos. Diferente da média móvel, este é afetado por ruídos de forma mais acentuada. Quanto mais próximo de 0 for o desvio padrão, mais homogêneo são os dados.

Este é calculado da seguinte forma:

$$DP = \sqrt{\frac{\sum_{i=1}^n (D_i - M_A)^2}{n}} \quad (14)$$

Sendo  $M_A$  a média aritmética dos  $n$  dados estudados.

#### 4.1.3 Diferenciação dos Dados

A diferenciação dos dados de uma série temporal é a transformação da série em uma nova onde os valores são as diferenças entre os valores consecutivos. Este procedimento pode ser aplicado mais de uma vez, dando origem às "primeiras diferenças", "segundas diferenças" e assim por diante (KWIATKOWSKI, D. et al, 1992). As primeiras diferenças são descritas pela equação a seguir:

$$d^{(1)}(t) = x(t) - x(t - 1) \quad (15)$$

A segunda diferença é calculada de acordo com a expressão:

$$d^{(2)}(t) = d^{(1)}(t) - d^{(1)}(t - 1) \quad (16)$$

E a expressão geral para as diferenças de ordem  $m$  é dada por:

$$d^{(m)}(t) = d^{(m-1)}(t) - d^{(m-1)}(t - 1) \quad (17)$$

Basicamente, a diferenciação torna uma série não-estacionária em estacionária. Deste modo, cumprindo um dos requisitos de utilização do modelo ARIMA.

#### 4.2 Determinação dos Parâmetros $p$ , $d$ e $q$ do Modelo ARIMA

A fim de determinar os parâmetros que constituem o modelo ARIMA, sendo estes  $p$  (AR),  $d$  (I) e  $q$  (MA), são aplicadas as funções ACF (Função de Autocorrelação) e PACF (Função de Autocorrelação Parcial).

Para a determinação do parâmetro  $p$ , é utilizada a função PACF. Esta pode ser imaginada como a correlação entre a série e a sua defasagem, após excluir as contribuições dos defasamentos intermediários. Assim, o PACF transmite a pura correlação entre uma defasagem e a série. Matematicamente, a autocorrelação parcial da defasagem ( $k$ ) de uma série é o coeficiente dessa defasagem na equação de autocorrelação de  $Y$ :

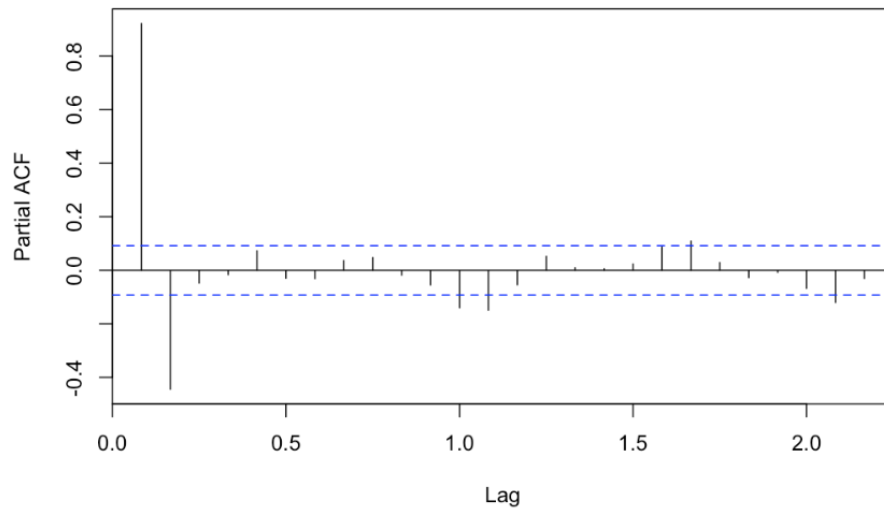
$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_k Y_{t-k} \quad (18)$$

Isto é,  $Y_t$  é a série atual,  $Y_{t-1}$  é a primeira defasagem da série e  $\alpha_k$  é o valor de PACF na  $k$ -ésima defasagem da série atual.

Já para a determinação do parâmetro  $q$ , a função ACF é utilizada. Esta, por sua vez, é utilizada para determinar quantos termos de MM são necessários para remover qualquer autocorrelação na série estacionarizada.

Na prática, tanto ACF quanto PACF são analisadas graficamente. Na linguagem Python, as funções `plot_acf()` e `plot_pacf()` retornam gráficos como os das Figuras 2 e 3 a seguir:

**Figura 2:** Exemplo de Plot PACF

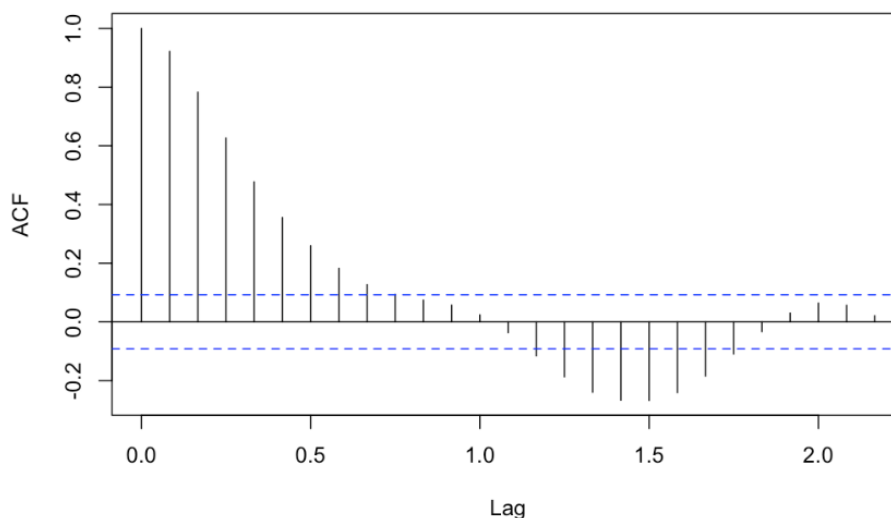


**Fonte:** TrainDataHub (2021)

Para determinar  $p$ , analisa-se na Figura 2 quantos picos existem antes da defasagem ( $lag$ ) estar dentro do intervalo de significância (limitado pelo pontilhado azul). Não consideramos o primeiro pico já que é a comparação da série com ela mesma, logo a correlação é 100% (ou seja, 1). Neste caso, portanto, verificamos que antes de algum valor de defasagem entrar no intervalo de significância existe apenas um pico, logo,  $p = 1$ .

A mesma análise é feita para a determinação do parâmetro  $q$ . Porém, desta vez é utilizado o ACF. Na Figura 3 a seguir, nota-se que existem oito picos antes da defasagem estar dentro do intervalo de significância, portanto,  $q = 8$ .



**Figura 3:** Exemplo de Plot ACF.

Fonte: TrainDataHub (2021)

A determinação do parâmetro de diferenciação ( $d$ ) é, por sua vez, a observação de alguns fatores que indicam se uma série está suficientemente diferenciada. A ordem correta de diferenciação é a diferença mínima necessária para obter uma série quase estacionária que gira em torno de uma média definida e o Plot ACF chega a zero razoavelmente rápido. Por exemplo, se as autocorrelações são positivas para muitas defasagens (10 ou mais), então a série precisa ser diferenciada ainda mais. Por outro lado, se a própria autocorrelação de defasagem 1 for muito negativa, então a série provavelmente é super diferenciada. Além disso, a análise do Valor  $P$  gerado pelo teste AD-Fuller também indica se a série está diferenciada o suficiente.

### 4.3 Aplicação do Método ARIMA e Geração do Modelo Matemático

Após o tratamento dos dados, o conjunto original é então aplicado na função *auto\_arima* da biblioteca *pmd.arima* da linguagem Python. Esta função tem o objetivo de descobrir automaticamente os parâmetros  $p$ ,  $d$ , e  $q$  ótimos para gerar um modelo ARIMA ajustado. Este processo funciona conduzindo testes de diferenciação (Kwiatkowski–Phillips–Schmidt–Shin, Dickey-Fuller Augmented ou Phillips–Perron) para determinar a ordem de diferenciação ( $d$ ) e, em seguida, ajusta o modelo dentro de intervalos definidos de  $p$ . Esta etapa do processo sumariza todos os pontos dos tópicos 4.1 e 4.2 em uma função automática. O modelo matemático é então gerado e usado para estimar dados futuros.

#### 4.4 Validação dos Resultados

Após a obtenção dos valores gerados pelo modelo matemático, estes são comparados com os valores reais através de três métodos de validação.

##### 4.4.1 Teste de Viés (*BIAS*)

O viés é uma medida de quão longe o valor da estimativa está do valor real do parâmetro que está sendo estimado. Ele é obtido pela fórmula:

$$BIAS = \frac{\sum_{i=1}^n (Dp_i - Dr_i)}{\sum_{i=1}^n Dr_i} \times 100\% \quad (2)$$

Onde,

$Dp_i$  = valor da depreciação prevista (pelo modelo) do material em um índice  $i$ ;

$Dr_i$  = valor da depreciação real do material em um índice  $i$ ;

$i$  = índice em meses;

$n$  = número de pontos estimados pelo modelo.

Este teste resulta em uma métrica transparente, tornando mais fácil e concreta a sua comunicação. Quando as estimativas são enviesadas, elas estão consistentemente erradas em uma direção devido a erros no sistema utilizado para as estimativas. Por exemplo, uma previsão meteorológica pode prever consistentemente temperaturas mais altas do que aquelas realmente observadas. Se o método de previsão for imparcial, ele ainda pode prever temperaturas que não são corretas, mas as temperaturas incorretas serão às vezes mais altas e às vezes mais baixas do que as temperaturas observadas. (MARKGRAF, B. 2018)

##### 4.4.2 Erro Médio Absoluto (*MAE*)

Do inglês *Mean Absolute Error*, o MAE é média dos erros entre observações que expressam o mesmo fenômeno. Neste caso, é o erro entre o valor de depreciação real e os valores estimados a cada mês.

$$MAE = \frac{\sum_{i=1}^n |Dp_i - Dr_i|}{n} \quad (20)$$

Onde  $|Dp_i - Dr_i|$  é o erro absoluto entre o valor previsto e o valor real, respectivamente, no ponto  $i$ .

#### 4.4.3 Raiz-Quadrada do Erro Médio (*RMSE*)

O RMSD representa a raiz quadrada das diferenças entre os valores preditos e os valores observados, normalizados pela média dos valores reais. É dado pela fórmula:

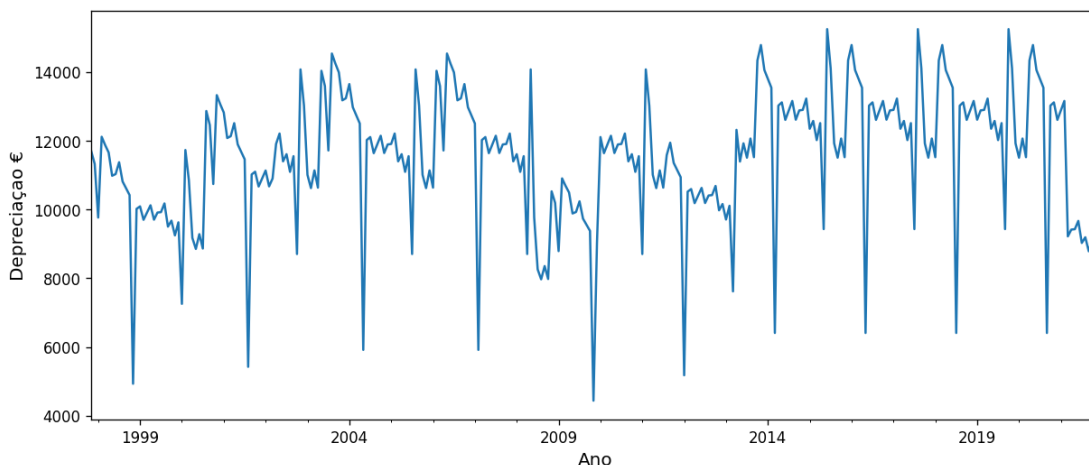
$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Dp_i - Dr_i)^2}{n}} \times 100\% \quad (213)$$

A diferença entre o RMSE e os outros métodos anteriores é que este penaliza fortemente maiores erros de previsão, já que utiliza a quadratura dos erros. Os resultados em porcentagem demonstram com clareza se há erros nos valores obtidos pelo modelo matemático gerado (KENNEY et al., 1962)

## 5. RESULTADOS E DISCUSSÕES

Tendo em vista que o objetivo deste trabalho foi gerar um modelo matemático a partir dos valores históricos de depreciação do material A fim de fazer previsões, os dados foram analisados em cada um dos processos propostos na metodologia e, por fim, os resultados previstos pelo modelo foram comparados com os valores reais.

**Figura 4:** Série Temporal – Dados reais de depreciação do material A.



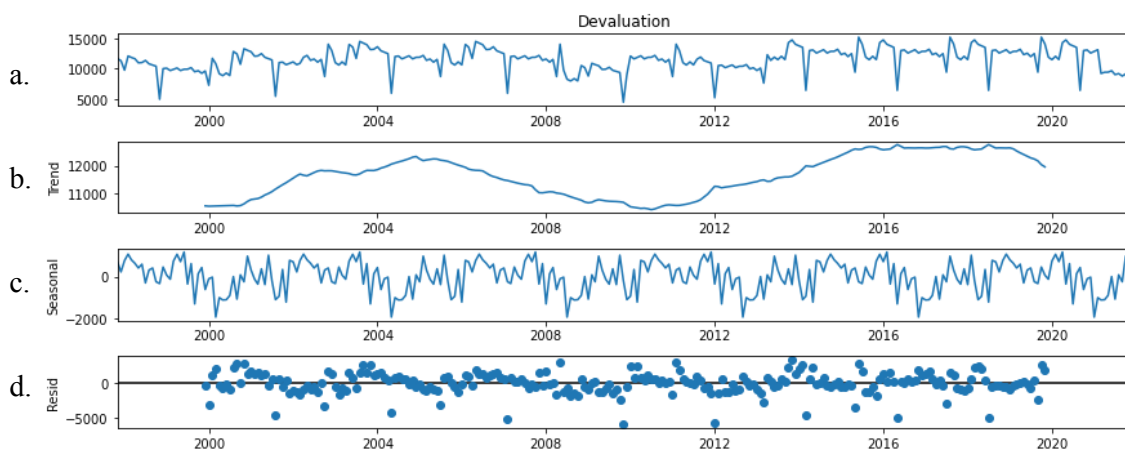
Fonte: Autor (2022).

### 5.1 Teste de Sazonalidade (Decomposição Sazonal)

A função *seasonal\_decompose()* foi aplicada aos dados brutos demonstrados na Figura 5. Foi utilizado o modelo aditivo de decomposição e escolheu-se um período de 48 meses para análise de sazonalidade.

**Figura 5:** Resultados da Decomposição Sazonal.

**5a.** Dados Originais; **5b.** Gráfico de Tendências; **5c.** Gráfico de Sazonalidade; **5d.** Gráfico de Resíduos/Ruídos.



Fonte: Autor (2022).

O gráfico 5b. aponta que há de fato um comportamento previsível em um período de 4 anos para a depreciação do material A. O gráfico 5c. demonstra que há alta sazonalidade nestes dados. Além disso, o gráfico de resíduos 5d. possui padrão pouco perceptível, sendo este um indicativo de que a escolha do modelo de decomposição foi acertada. Os resíduos também indicam que todos os anos, pelo menos uma vez ao ano, há uma oscilação mais brusca dos valores de depreciação. Tal fato não possui explicação direta e clara, podendo ser este o resultado de vários procedimentos internos da empresa, ou fatores externos que afetam a utilização do material A.

## 5.2 Teste de Estacionariedade

Para determinar a Estacionariedade da série, a fim de cumprir o pré-requisito do modelo ARIMA, foram feitos dois testes: Dickey-Fuller Aumentado (ADF) e Média-Móvel e Desvio Padrao.

### 5.2.1 Teste ADF

Aplicando-se a função *adfuller()*, também da biblioteca *statsmodels* da linguagem Python, aos dados originais, foram obtidos os resultados:

**Tabela 1:** Resultados do Teste ADF aplicado aos dados originais.

<b>Dados Originais</b>	
<b>Estatística ADF</b>	-2.839768
<b>Valor <math>p</math></b>	0.052814

**Fonte:** Autor (2022).

Nota-se que o valor  $p$  é ligeiramente maior que 0.05, portanto, a hipótese de que há uma raiz unitária na série temporal ainda não pode ser rejeitada, indicando que a série não é estacionária. Para resolver este problema, é aplicado o método de diferenciação na série. Após a diferenciação, estes são os resultados obtidos:

**Tabela 2:** Resultados do Teste ADF após diferenciação dos dados.

<b>Dados após Diferenciação</b>	
<b>Estatística ADF</b>	-5.703658
<b>Valor <math>p</math></b>	0.000001

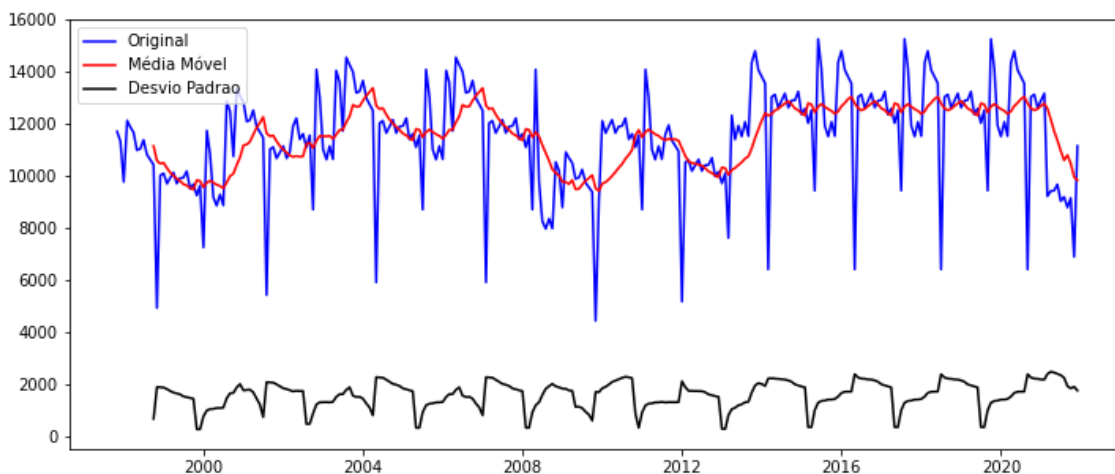
**Fonte:** Autor (2022).

Desta vez, o valor  $p$  é muito menor que 0.05 e a hipótese de que existe uma raiz unitária é rejeitada, ou seja, a série se tornou estacionária após apenas uma diferenciação. Além disso, nota-se que a Estatística ADF é mais negativa após a diferenciação, e quanto mais negativo este indicador, mais provável da série ser estacionária. É esta série diferenciada que servirá de base para a aplicação do modelo ARIMA.

### 5.2.2 Média Móvel e Desvio Padrão

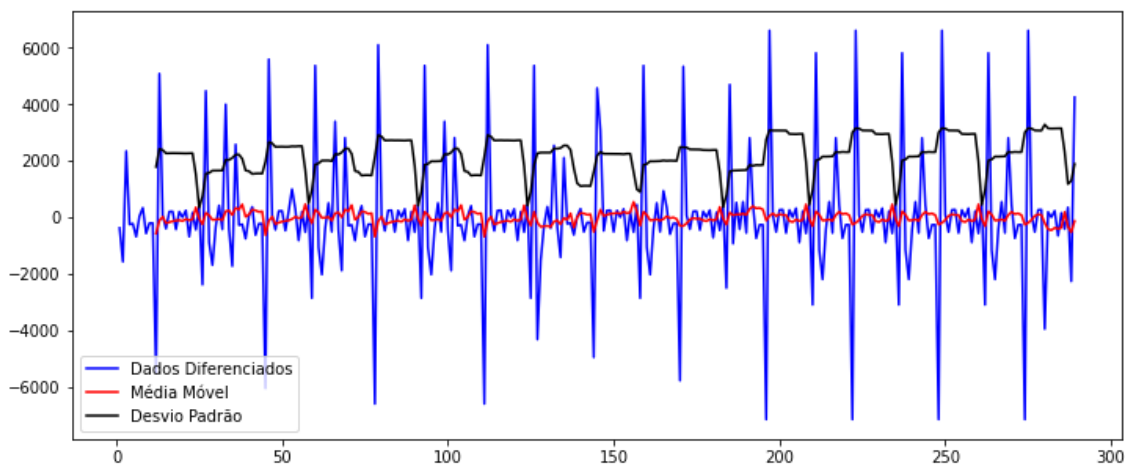
Outra maneira de observar a Estacionariedade de uma série é através do método gráfico de Média Móvel (MM) e Desvio Padrão (DP). Desta vez, os métodos foram aplicados aos dados originais e aos dados obtidos após primeira diferenciação.

**Figura 6:** Média Móvel e Desvio Padrão em dados originais.



**Fonte:** Autor (2022)

**Figura 7:** Média Móvel e Desvio Padrão após primeira diferenciação.



**Fonte:** Autor (2022)

Ao comparar a MM das figuras 4 e 5, nota-se que após a primeira diferenciação, esta se mantém muito próxima de zero, um indicativo de que as tendências foram removidas e a série se tornou de fato estacionária. Outra observação importante é de que o DP ficou praticamente inalterado, sugerindo que a natureza randômica da série original permaneceu, portanto, a série ainda é heterogênea.

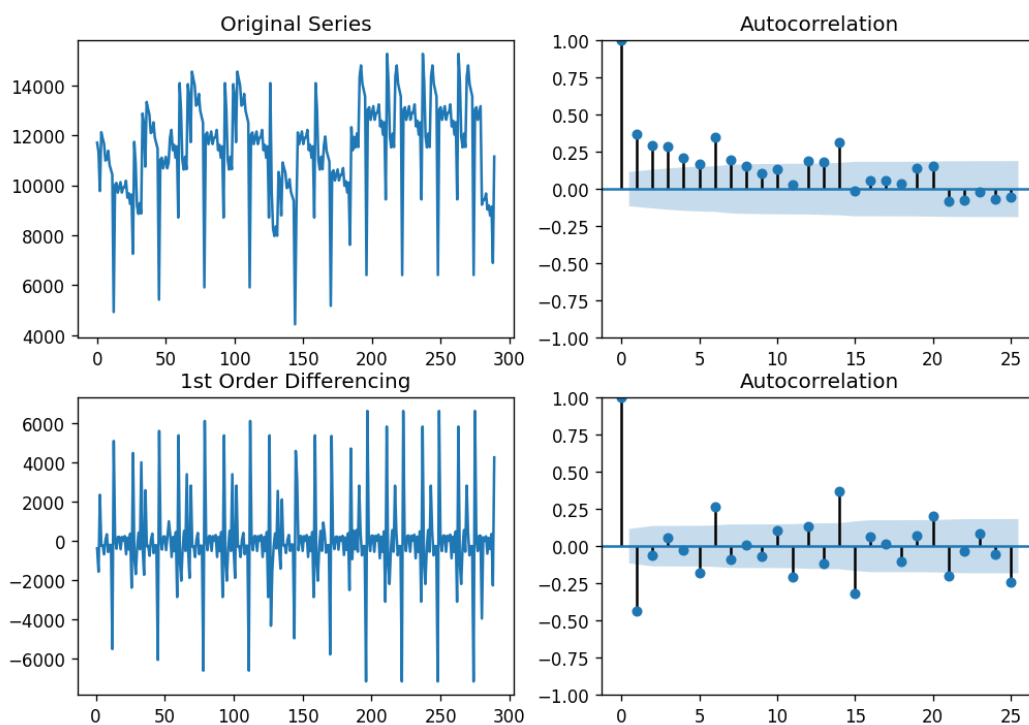
Após a aplicação de ambos os métodos para testar a Estacionariedade, é possível notar que apenas a primeira diferenciação é necessária para a aplicação do modelo ARIMA.

### 5.3 Obtenção dos Parâmetros $p$ , $d$ , e $q$ do Modelo ARIMA

Foram aplicadas as funções ACF e PACF nos dados de depreciação originais e após a primeira diferenciação.

De acordo com o tópico 4.2 e ao analisar o gráfico de autocorrelação dos dados após a primeira diferenciação, na Figura 8 a seguir, nota-se que o número de picos existentes antes de haver um pico dentro do intervalo de significância é 1. Portanto, tem-se que  $q = 1$ .

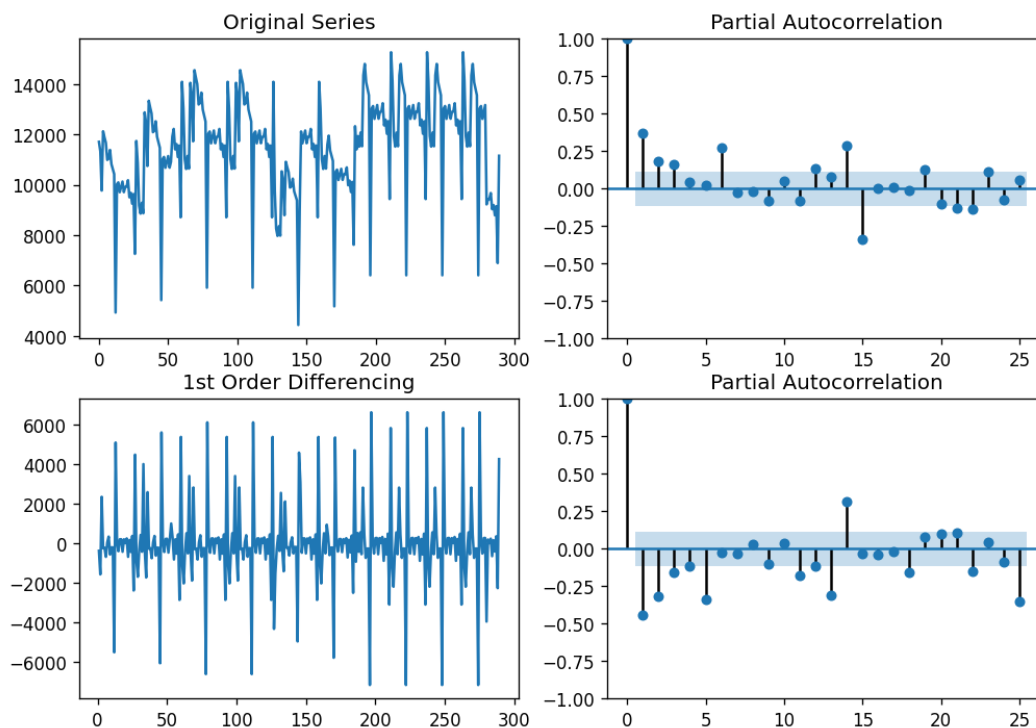
**Figura 8:** Teste ACF.



**Fonte:** Autor (2022)

Da mesma forma, analisando-se desta vez o gráfico de autocorrelação parcial dos dados, após a primeira diferenciação, na Figura 9 a seguir, nota-se que o número de picos existentes antes de haver um pico dentro do intervalo de significância é 5. Logo,  $p = 5$ .

**Figura 9:** Teste PACF.



Fonte: Autor (2022)

E como visto no tópico 4.2, apenas uma diferenciação foi necessária para tornar a série estacionária. Logo, temos que  $d = 1$ .

A partir desta análise, o modelo ARIMA encontrado é descrito como ARIMA(5, 1, 1).

#### 5.4 Geração automática do Modelo ARIMA

Utilizando a função `auto_arima()` da linguagem Python, foi obtido um modelo SARIMAX, que é um modelo ARIMA adicionado de um fator de sazonalidade. Seu formato é dado por SARIMAX( $p, d, q$ )x( $P, D, Q, s$ ), onde  $p, d,$  e  $q$  são os parâmetros já conhecidos do ARIMA e  $s$  é o parâmetro de sazonalidade adicionado. Neste caso, o parâmetro de sazonalidade ( $s$ ) foi inserido manualmente.

O resultado da modelagem automática é mostrado a seguir:



**Figura 10:** Resultados modelagem SARIMAX.

SARIMAX Results						
Dep. Variable:	y			No. Observations: 290		
Model:	SARIMAX(5, 1, 4)x(0, 0, [1], 48)			Log Likelihood	-2542.218	
Date:	Sat, 01 Oct 2022			AIC	5110.437	
Time:	11:48:19			BIC	5158.101	
Sample:	11-01-1997 - 12-01-2021			HQIC	5129.535	
Covariance Type: opg						
	coef	std err	z	P> z	[0.025	0.975]
ar.L1	-1.2934	0.103	-12.532	0.000	-1.496	-1.091
ar.L2	-1.0727	0.204	-5.258	0.000	-1.473	-0.673
ar.L3	-0.4685	0.257	-1.821	0.069	-0.973	0.036
ar.L4	0.2670	0.222	1.202	0.229	-0.168	0.703
ar.L5	-0.0373	0.113	-0.330	0.742	-0.259	0.184
ma.L1	0.6188	0.078	7.974	0.000	0.467	0.771
ma.L2	0.1582	0.105	1.510	0.131	-0.047	0.364
ma.L3	-0.3654	0.098	-3.710	0.000	-0.558	-0.172
ma.L4	-0.7912	0.081	-9.746	0.000	-0.950	-0.632
ar.S.L12	0.8337	0.143	5.845	0.000	0.554	1.113
ar.S.L24	-0.2099	0.093	-2.255	0.024	-0.392	-0.027
ma.S.L12	-0.8453	0.127	-6.632	0.000	-1.095	-0.595
sigma2	2.919e+06	1.81e-07	1.61e+13	0.000	2.92e+06	2.92e+06
Ljung-Box (L1) (Q):	0.10		Jarque-Bera (JB):	138.60		
Prob(Q):	0.75		Prob(JB):	0.00		
Heteroskedasticity (H):	0.99		Skew:	-1.15		
Prob(H) (two-sided):	0.97		Kurtosis:	5.49		

Fonte: Autor (2022)

Os parâmetros obtidos foram razoavelmente próximos das estimações feitas de forma manual, diferenciando-se apenas o parâmetro  $q$ , que neste caso é igual a 4. Os parâmetros  $p$  e  $d$  obtidos foram idênticos, indicando que a correlação entre os dados é de fato de ordem 5 e principalmente que a escolha do número de diferenciações necessárias foi acertada.

## 5.5 Validação dos Resultados

Após a obtenção dos modelos, estes foram aplicados para a previsão dos dados de depreciação do material A. Foram estimados os valores de depreciação dos meses de janeiro de 2022 até setembro de 2022.

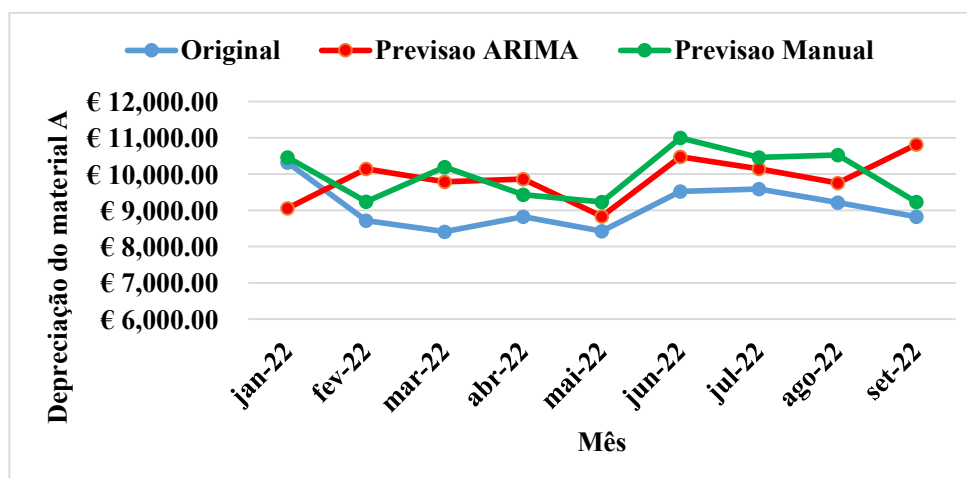
Tabela 3: Resultados das Previsões.

	Dados Reais de Depreciação	Previsão Manual	Previsão Automática
Jan-22	€ 10,313.80	€ 10,460.62	€ 9,057.71
Fev-22	€ 8,717.29	€ 9,241.37	€ 10,142.65
Mar-22	€ 8,412.45	€ 10,191.43	€ 9,786.54
Abr-22	€ 8,819.83	€ 9,432.37	€ 9,858.13
Mai-22	€ 8,422.76	€ 9,231.57	€ 8,832.51
Jun-22	€ 9,519.67	€ 10,995.70	€ 10,472.50
Jul-22	€ 9,589.01	€ 10,456.65	€ 10,140.52
Ago-22	€ 9,217.23	€ 10,526.73	€ 9,749.85
Set-22	€ 8,819.83	€ 9,227.66	€ 10,813.50
<b>Total:</b>	€ 81,831.87	€ 89,764.10	€ 88,853.91

Fonte: Autor (2022)

É possível notar a proximidade dos valores previstos por ambos os métodos aos dados reais. Graficamente também se nota esse comportamento:

Figura 11: Gráfico dos valores de depreciação reais e modelados por mês.



Fonte: Autor (2022)

Para analisar de forma efetiva estes resultados, são aplicados três testes estatísticos que tem como objetivo avaliar a proximidade dos valores previstos aos dados reais. Para tal, serão necessários os dados presentes nas seguintes tabelas:

**Tabela 4:** Análise Estatística para a Previsão Manual.

	<b>Original</b>	<b>Previsão Manual</b>	<b>Erro</b>	<b>Erro Absoluto %</b>	<b>Erro Absoluto</b>	<b>Erro<sup>2</sup></b>
<b>Jan-22</b>	€ 10,313.80	€ 10,460.62	€ 146.82	1.42%	€ 146.82	€ 21,555.38
<b>Fev-22</b>	€ 8,717.29	€ 9,241.37	€ 524.08	6.01%	€ 524.08	€ 274,662.92
<b>Mar-22</b>	€ 8,412.45	€ 10,191.43	€ 1,778.98	21.15%	€ 1,778.98	€ 3,164,784.66
<b>Abr-22</b>	€ 8,819.83	€ 9,432.37	€ 612.54	6.95%	€ 612.54	€ 375,203.60
<b>Mai-22</b>	€ 8,422.76	€ 9,231.57	€ 808.81	9.60%	€ 808.81	€ 654,167.29
<b>Jun-22</b>	€ 9,519.67	€ 10,995.70	€ 1,476.03	15.51%	€ 1,476.03	€ 2,178,656.65
<b>Jul-22</b>	€ 9,589.01	€ 10,456.65	€ 867.64	9.05%	€ 867.64	€ 752,802.45
<b>Ago-22</b>	€ 9,217.23	€ 10,526.73	€ 1,309.50	14.21%	€ 1,309.50	€ 1,714,787.09
<b>Set-22</b>	€ 8,819.83	€ 9,227.66	€ 407.83	4.62%	€ 407.83	€ 166,327.30
<b>Total:</b>	€ 81,831.87	€ 89,764.10	€ 7,932.23		€ 7,932.23	

**Fonte:** Autor (2022)

**Tabela 5:** Análise Estatística para a Previsão Manual.

	<b>Original</b>	<b>Previsão Automática</b>	<b>Erro</b>	<b>Erro Absoluto %</b>	<b>Erro Absoluto</b>	<b>Erro<sup>2</sup></b>
<b>Jan-22</b>	€ 10,313.80	€ 9,057.71	€ (1,256.09)	12.18%	€ 1,256.09	€ 1,577,759.99
<b>Fev-22</b>	€ 8,717.29	€ 10,142.65	€ 1,425.36	16.35%	€ 1,425.36	€ 2,031,661.56
<b>Mar-22</b>	€ 8,412.45	€ 9,786.54	€ 1,374.09	16.33%	€ 1,374.09	€ 1,888,128.19
<b>Abr-22</b>	€ 8,819.83	€ 9,858.13	€ 1,038.30	11.77%	€ 1,038.30	€ 1,078,070.31
<b>Mai-22</b>	€ 8,422.76	€ 8,832.51	€ 409.75	4.86%	€ 409.75	€ 167,896.52
<b>Jun-22</b>	€ 9,519.67	€ 10,472.50	€ 952.83	10.01%	€ 952.83	€ 907,886.49
<b>Jul-22</b>	€ 9,589.01	€ 10,140.52	€ 551.50	5.75%	€ 551.50	€ 304,155.74
<b>Ago-22</b>	€ 9,217.23	€ 9,749.85	€ 532.62	5.78%	€ 532.62	€ 283,683.48
<b>Set-22</b>	€ 8,819.83	€ 10,813.50	€ 1,993.67	22.60%	€ 1,993.67	€ 3,974,713.13
<b>Total:</b>	€ 81,831.87	€ 88,853.91	€ 7,022.04		€ 9,534.22	

**Fonte:** Autor (2022)

### 5.5.1 Teste de Viés (*BIAS*)

Aplicando a fórmula (10), foram obtidos os resultados do Viés estatístico. Estes podem ser visto na tabela a seguir.

**Tabela 6:** Viés Estatístico das Previsões Manual e Automática.

	<b>Dados Reais</b>	<b>Previsão Manual</b>	<b>Previsão Automática</b>
<b>Depreciação Total</b>	€ 81,831.87	€ 89,764.10	€ 88,853.91
<b>Erro Total</b>	-	€ 7,932.23	€ 7,022.04
<b>Viés</b>	-	9.69%	8.58%

Fonte: Autor (2022)

É possível notar que ambos os métodos de geração de modelo matemático possuem viés positivo, ou seja, os valores que eles apresentarão como estimativas provavelmente serão maiores que os valores reais. Isso fica claro ao se observar o gráfico da figura 11, onde a curva dos dados originais se apresenta abaixo dos valores estimados pelos modelos matemáticos. Quanto mais próximo de zero o viés for, mais preciso é o modelo matemático, logo, a previsão automática teve melhor performance neste teste.

### 5.5.2 Erro Médio Absoluto (MAE – *Mean Absolute Error*)

Tratando-se do MAE, o modelo de previsão com melhor performance foi o manual. Este fato se dá pelo comportamento do MAE não deixar com que erros negativos e positivos se anulem, dando a falsa impressão de que o modelo é preciso. Como houve um desvio negativo no mês de janeiro de 2022 para a modelagem automática, acarretou num aumento significativo do erro total e, portanto, do MAE para esta modelagem.

**Tabela 7:** Erro Médio Absoluto das Previsões Manual e Automática.

	<b>Dados Reais</b>	<b>Previsão Manual</b>	<b>Previsão Automática</b>
<b>Depreciação Total</b>	€ 81,831.87	€ 89,764.10	€ 88,853.91
<b>Erro Total Absoluto</b>	-	€ 7,932.23	€ 9,534.22
<b>MAE</b>	-	€ 881.36	€ 1,059.36

Fonte: Autor (2022)

### 5.5.3 Raiz-Quadrada do Erro Médio (RMSE – *Root Mean Square Deviation*)

Inicialmente calculou-se o MSE, que é a média da quadratura dos Erros, representada na fórmula (12) como  $\frac{\sum_{i=1}^n (Dp_i - Dr_i)^2}{n}$ . Os valores são apresentados na tabela 8 a seguir. Em seguida, calculou-se o RMSE.

**Tabela 8:** RMSE das Previsões Manual e Automática.

	<b>Dados Reais</b>	<b>Previsão Manual</b>	<b>Previsão Automática</b>
<b>Depreciação Total</b>	€ 81,831.87	€ 89,764.10	€ 88,853.91
<b>MSE</b>	-	€ 1,033,660.81	€ 1,357,106.16
<b>RMSE</b>	-	11.18%	12.81%

Fonte: Autor (2022)

É possível notar que os valores de RMSE obtidos são próximos em ambos os modelos. Além disso, vale salientar que o cálculo do RMSE penaliza com maior rigidez os erros por se tratar da quadratura destes e, ainda assim, a porcentagem obtida é aceitavelmente baixa (entre 11% e 13%). Outro fator importante é de que a correlação de uma série com ela mesma gera um RMSE de 0, o que indica que os dados obtidos pelos modelos também possuem boa qualidade.

## 6. CONCLUSÃO

Com o desenvolvimento do presente trabalho, implementou-se, em linguagem de programação Python, estratégias de estimação de valores futuros de depreciação de materiais guardados em inventário, utilizando técnicas de modelagem estatística. De um modo geral, pôde-se concluir que tanto a modelagem manual quanto a aplicação da função *ARIMA()* da linguagem Python, apresentaram resultado satisfatório para se trabalhar com séries temporais, o que foi ilustrado pela similaridade de performance entre elas. Além disso, foi possível observar a capacidade do método de estimar com relativa precisão, grandezas que são função de mais de uma variável, sendo estas dependentes de alguns fatores que são por muitas vezes pouco previsíveis, como demanda e valor de matéria-prima. Portanto, o desenvolvimento deste método pode acarretar ganhos significativos para a empresa.

Por fim, durante a realização deste trabalho, foram aplicados conhecimentos adquiridos durante a graduação nas áreas de modelagem matemática, estatística e tópicos de programação para desenvolver as estratégias de estimação propostas, aprofundando os conhecimentos nestas áreas e contribuindo na formação profissional em Engenharia e Tecnologia aplicada à indústria.

## 7. REFERÊNCIAS

MENTZER, J.T. DEWITT, W. KEEBLER, J.S. MIN, S. NIX, N.W. SMITH, C.D. ZACHARIA, Z.G. (2001), **Defining Supply Chain Management**, Journal of Business Logistics, Vol 22, issue 2, pp 1-25.

HENKEL (2022) **About Henkel**, Disponível em: <https://www.henkel.hu/vallalat>

JOPPÉ, P. (2020), **Raw and Pack Material Planning: R&P Planning – Roles Description**, Amsterdam.

DOEDEN, M. (2021), **Adhesive Technologies Organizational Structure**, Germany.

HGSC. (2018), **Supply Chain Operations Unified Toolbox (SCOUT) & Slow-Moving Inventory (SMI) Guide**, Subsequent Measurement of Inventory, version 2.2, Amsterdam.

LUN, Y.H.V. LAI, K.-H. CHENG, T.C.E. (2010) **Shipping and Logistics Management**, The Hong Kong Polytechnic University Department of Logistics and Maritime Studies, Springer.

PRINIOTAKIS, G. ARGYROPOULOS, P. (2018), **Inventory Management Concepts and Techniques**, IOP Conf. Ser.: Mater. Sci. Eng. 459 012060.

E. P. B. GEORGE, M. J. GWILYM, C. R. GREGORY, M. L. GRETA. **Time Series Analysis: Forecasting and Control**, vol.5, John Wiley & Sons Inc., 2016.

BOX, G.E.P. and G.M. JENKINS. **Time series analysis: Forecasting and control**, San Francisco: Holden-Day. 1980.

Fuller, W. A. (1976). **Introduction to Statistical Time Series**. New York: John Wiley and Sons. ISBN 0-471-28715-6.

TABLEU. **Time Series Analysis: Definition, Types, Techniques, and When It's Used**. Disponível em: <https://www.tableau.com/learn/articles/time-series-analysis>. Acessado em: 15 de maio de 2022.

MICHAEL, K. **Forecast with ARIMA in Python More Easily with Scalecast**. Disponível em: <https://towardsdatascience.com/forecast-with-arima-in-python-more-easily-with-scalecast-35125fc7dc2e>. Acessado em: 12 de maio de 2022.

HAYES, A. **Autoregressive Integrated Moving Average (ARIMA)**. Investopedia, 2021. Disponível em: <https://www.investopedia.com/terms/a/autoregressive-integrated-moving-average-arima.asp#:~:text=An%20autoregressive%20integrated%20moving%20average%2C%20or%20ARIMA%2C%20is%20a%20statistical,value%20based%20on%20past%20values>. Acessado em: 12 de maio de 2022.

BROWNLEE, J. **How to Create an ARIMA Model for Time Series Forecasting in Python**. Disponível em: <https://machinelearningmastery.com/arima-for-time-series-forecasting-with-python/>. Acessado em 17 de junho de 2022.

TRAINDATAHUB. **How to Interpret ACF and PACF plots for Identifying AR, MA, ARMA, or ARIMA Models**. Disponível em: <https://medium.com/@ooemma83/how-to->

interpret-acf-and-pacf-plots-for-identifying-ar-ma-arma-or-arima-models-498717e815b6t.  
Acessado em 5 de outubro de 2022.

KWIATKOWSKI, D., PHILLIPS, P. C. B., SCHMIDT, P., & SHIN, Y. (1992). **Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?** *Journal of Econometrics*, 54(1-3), 159–178.

MARKGRAF, B. **How to Calculate Bias.** Disponível em: <https://sciencing.com/how-to-calculate-bias-13710241.html>. Acessado em: 20 de setembro de 2022.

KENNEY, J. F. AND KEEPING, E. S. “**Root Mean Square.**” §4.15 in *Mathematics of Statistics*, Pt. 1, 3rd ed. Princeton, NJ: Van Nostrand, pp. 59-60, 1962.